

# **METHOD AND DEVICE FOR STORING LYRIC PHONETIC NOTATIONS**

[0001] This application claims priority to Chinese Patent Application No. 201810545601.9, filed on May 25, 2018 and entitled "METHOD AND DEVICE FOR STORING LYRIC PHONETIC NOTATIONS", the entire contents of which are incorporated herein by reference.

## **TECHNICAL FIELD**

[0002] The present disclosure relates to the field of computer technology, and in particular, to a method and device for storing lyric phonetic notations.

## **BACKGROUND**

[0003] With the development of computer technology, audio applications have gradually become globalized and can play Chinese songs, English songs, Japanese songs and the like. When playing songs, in order to make the songs sound more artistic, the pronunciations of some pronunciation characters are usually different from the most commonly used pronunciations. For example, in Japanese songs, the common pronunciation of the character "宿命" is "し ゆ く め い", and the pronunciation in the song is "さ だ め", so in order to display the correct pronunciations of the polyphonic characters the lyric, generally the polyphonic characters with special pronunciations are marked with the correct pronunciations.

[0004] In related technologies, when a song is being played, the pronunciation of the lyric is usually searched in real time and then marked. In this way, the pronunciation needs to be queried in a dictionary in real time and then marked (or "phonetic notation" in short), which may cause the display of the lyric phonetic notation to be slow.

## SUMMARY

[0005] To solve the problems of the prior art, embodiments of the present disclosure provides a method and device for storing lyric phonetic notations. The technical solutions are as follows:

[0006] In a first aspect, a method for storing lyric phonetic notations is provided. The method includes:

[0007] receiving a phonetic notation storage request for a target lyric;

[0008] obtaining pronunciations of polyphonic characters in the target lyric, and determining the playback duration of each furigana in the pronunciations of the polyphonic characters;

[0009] for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character is different, then obtaining a playback start identification of each furigana in the pronunciation of the polyphonic character and determining the same as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then obtaining a playback start identification of first furigana in the pronunciation of the polyphonic character and determining the same as display information of the pronunciation of the polyphonic character; and

[0010] correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

[0011] Optionally, correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric includes:

[0012] storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric.

[0013] Optionally, storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric includes:

[0014] determining pronunciations of polyphonic characters in each line of the target lyric and display information thereof; and

[0015] storing the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric respectively.

[0016] Optionally, correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric includes:

[0017] storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

[0018] In a second aspect, a method for storing lyric phonetic notations is provided. The method includes:

[0019] receiving a phonetic notation storage request for a target lyric;

[0020] obtaining pronunciations of polyphonic characters in the target lyric, and determining a playback duration of each furigana in the pronunciations of the polyphonic characters in the target lyric;

[0021] for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, then obtaining a number of Chinese characters in the polyphonic character and a playback start identification of each furigana in the pronunciation of the polyphonic character and determining the number of Chinese characters in the polyphonic character and the playback start identification as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then obtaining the number of Chinese characters in the polyphonic character and determining the same as display information of the pronunciation of the polyphonic character; and

[0022] correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

[0023] Optionally, correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric includes:

[0024] storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric.

[0025] Optionally, storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric includes:

[0026] determining pronunciations of polyphonic characters in each line of the target lyric and display information thereof; and

[0027] storing the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric respectively.

[0028] Optionally, correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric includes:

[0029] storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

[0030] In a third aspect, a device for storing lyric phonetic notations is provided. The device includes:

[0031] a receiving module configured to receive a phonetic notation storage request for a target lyric;

[0032] an obtaining module configured to obtain pronunciations of polyphonic characters in the target lyric and determine a playback duration of each furigana in the pronunciations of the polyphonic characters;

[0033] a determining module configured to, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, obtain a playback start identification of each furigana in the pronunciation of the polyphonic character and determine the same as display information thereof, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, obtain a playback start identification of first furigana in the pronunciation of the polyphonic character and determine the same as display information thereof; and

[0034] a storing module configured to correspondingly store the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

[0035] Optionally, the storing module is configured to:

[0036] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric.

[0037] Optionally, the storing module is configured to:

[0038] determine pronunciations of polyphonic characters in each line of the target lyric and display information thereof; and

[0039] store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric respectively.

[0040] Optionally, the storing module is configured to:

[0041] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

[0042] In a fourth aspect, a device for storing lyric phonetic notations is provided. The device includes:

[0043] a first receiving module configured to receive a phonetic notation storage request for a target lyric;

[0044] a first obtaining module configured to obtain pronunciations of polyphonic characters in the target lyric and determine a playback duration of each furigana in the pronunciations of the polyphonic characters in the target lyric;

[0045] a first determining module configured to, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, obtain a number of Chinese characters in the polyphonic character and a playback start identification of each furigana in the pronunciation of the polyphonic character and determine the number of Chinese characters in the polyphonic character and the playback start identification as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, obtain the number of Chinese characters in the polyphonic character and determine the same as display information thereof; and

[0046] a first storing module configured to correspondingly store the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

[0047] Optionally, the first storing module is configured to:

[0048] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric.

[0049] Optionally, the first storing module is configured to:

[0050] determine pronunciations of polyphonic characters in each line of the target lyric and display information thereof;

[0051] store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric respectively.

[0052] Optionally, the first storing module is configured to:

[0053] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

[0054] The technical solutions provided by embodiments of the present disclosure may include at least the following benefits.

[0055] In the embodiment of the present disclosure, the pronunciations of the polyphonic characters in the lyric and the display information thereof are correspondingly stored with the lyric. In this way, during subsequent playback of the song, the pronunciations of the polyphonic characters and the display information thereof can be obtained directly based on the lyric without real-time query, so the efficiency of displaying the phonetic notations of the lyric is improved.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

[0056] FIG. 1 is a flowchart of a method for storing lyric phonetic notations according to an embodiment of the present disclosure;

[0057] FIG. 2 is a schematic diagram of an additional part according to an embodiment of the present disclosure;

[0058] FIG. 3 is a schematic diagram of storing lyric phonetic notations according to an embodiment of the present disclosure;

[0059] FIG. 4 is a schematic diagram of storing lyric phonetic notations according to an embodiment of the present disclosure;

[0060] FIG. 5 is a flowchart of a method for storing lyric phonetic notations according to an embodiment of the present disclosure;

[0061] FIG. 6 is a schematic structural diagram of a device for storing lyric phonetic notations according to an embodiment of the present disclosure;

[0062] FIG. 7 is a schematic structural diagram of a device for storing lyric phonetic notations according to an embodiment of the present disclosure;

[0063] FIG. 8 is a schematic structural diagram of a server according to an embodiment of the present disclosure.

## **DETAILED DESCRIPTION**

[0064] The present disclosure will be described in further detail with reference to the accompanying drawings, to present the objects, technical solutions, and advantages of the present disclosure more clearly.

[0065] An embodiment of the present disclosure provides a method for storing lyric phonetic notations. The method is executed by a server, and the server may be a background server of an audio application. The server may be used to store an audio file of a target audio. The audio file includes audio data and lyric data and the like. The lyric data includes lyric text and an additional part. The additional part is used to store authors, singers, and other content.

[0066] The server may be provided with a processor, a memory, a transceiver, and the like. The processor may be used to process the process of storing the lyric. The memory may be used to store the data required during the process of storing the lyric and the data generated. The transceiver may be used for receiving and sending data.

[0067] First, an application scenario of the embodiment of the present disclosure is introduced. Generally, an audio application displays lyric when playing a song. Generally, the staff of the audio application uploads lyric data to a background server (referring to as a server in the following) of the audio application. The server may correspondingly store the lyric data with the song's audio data. Since certain characters in the lyrics of some songs are polyphonic characters, when displaying the lyric, the polyphonic information of

the polyphonic characters will be found by looking up a dictionary, and the polyphonic characters will be marked, which may cause the display of the lyric to be slow. For this purpose, a way to store lyric phonetic notations in advance is provided. For example, in some Japanese songs, in order to achieve artistic effects, the pronunciations of certain characters in the lyric are different from the most commonly used pronunciations (the most commonly used pronunciation of the character "宿命" is "し ゅ く め い", but the pronunciation in the lyric is "さ だ め").

[0068] In addition, the storage manner of the lyric text in the embodiment of the present disclosure is introduced:

[0069] [First line playback start identification] [Playback start identification] First line first character <Playback start identification> First line second character <Playback start identification> ... First line last character.

[0070] An embodiment of the present disclosure provides a method for storing lyric phonetic notations. Japanese lyric is used as an example. As shown in FIG. 1, the processing steps of the method may be as follows.

[0071] In a step 101, a phonetic notation storage request for a target lyric is received.

[0072] The phonetic notation includes a pronunciation and display information thereof.

[0073] In implementation, a user (generally a staff member of an audio application) may install an audio application in a terminal. The audio application may provide an interface for the user to edit lyric and pronunciations of polyphonic characters. For the lyric of a song (which may be referred to as target lyric later), the user may mark pronunciations of polyphonic characters in the target lyric and start timestamps of furigana in each pronunciation, and then click a submit button. The terminal will generate a phonetic notation storage request which carries an identification of the target lyric, the pronunciations of the polyphonic characters, and the start timestamps of furigana in each pronunciation, and then send the phonetic notation storage request for the target lyric to a server. After receiving the phonetic notation storage request sent by the terminal, the server may parse the same to obtain the identification of the target lyric, the pronunciations of the polyphonic characters, and the start timestamps of furigana in each pronunciation.

[0074] Alternatively, the audio application provides the user with a phonetic notation adding option. If the user wants to add a phonetic notation to the lyric of a song (which may be referred to as target lyric later), he/she may click the phonetic notation adding option. The terminal will detect a click instruction of the phonetic notation adding option, generate a phonetic notation storage request which carries an identification of the target lyric. After receiving the phonetic notation storage request, the server may parse the same to obtain the identification of the target lyric.

[0075] In a step 102, pronunciations of polyphonic characters in the target lyric are obtained, and a playback duration of each furigana in the pronunciations of the polyphonic characters is determined.

[0076] The playback duration is the coloring duration when rendering the lyric. The playback start identification is used to indicate that the furigana starts to play, which may be a playback start timestamp and the like. Furigana is a kana symbol noted above or around a polyphonic character in Japanese. For example, the furigana of the pronunciation corresponding to "絆" (meaning "牽絆" in Chinese) is "き ず な".

[0077] In implementation, the server may obtain the pronunciations of the polyphonic characters in the target lyric and the start timestamps of furigana in each pronunciation from the phonetic notation storage request. Alternatively, after the server parses the phonetic notation storage request to obtain the identification of the target lyric, the server may use the identification of the target lyric to find the target lyric, and then input the target lyric into a phonetic notation determination program to obtain the start timestamps of furigana in each pronunciation in the polyphonic characters of the target lyric (detailed later).

[0078] For any polyphonic character in the target lyric, the start timestamps of two consecutive furigana in the pronunciation of the polyphonic character may be subtracted to obtain the timestamps of other furigana in the pronunciation of the polyphonic character except the last furigana. The start timestamp of the character after the polyphonic character and the start timestamp of the last furigana may be subtracted, the difference of which is the playback duration of the last furigana. In this way, the playback duration of each furigana in the pronunciation of the polyphonic character may be

obtained. For example, the furigana of the pronunciation corresponding to "絆" (meaning "牽絆" in Chinese) is "き ず な". The start timestamp of "き" is 0ms, the start timestamp of "ず" is 10ms, and the start timestamp of "な" is 30ms, the playback duration of "き" is 10ms-0ms, which is 10ms, the playback duration of "ず" is 30ms-10ms, which is 20ms, and the start timestamp of the character after "絆" is 55ms, then the playback duration of "な" is 55ms-30ms, which is 25ms.

**[0079]** In step a 103, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, then a playback start identification of each furigana in the pronunciation of the polyphonic character is obtained and determined as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then a playback start identification of first furigana in the pronunciation of the polyphonic character is obtained and determined as display information of the pronunciation of the polyphonic character.

**[0080]** In implementation, for any of the polyphonic characters, whether the playback durations of different furigana are the same may be determined. If they are different, the playback start identification of each furigana in the pronunciation of the polyphonic character may be determined as the display information of the pronunciation of the polyphonic character. If they are the same, the playback start identification of the first furigana in the pronunciation of the polyphonic character may be determined as the display information of the pronunciation of the polyphonic character. Each polyphonic character in the target lyric may be determined in this way.

**[0081]** It should be noted that, in the pronunciation of any of the polyphonic characters above, if the playback durations of different furigana are the same, the playback start identification of each furigana may be known by only using the playback start identification of the first furigana. For example, the furigana of the pronunciation corresponding to "絆" is "き ず な" and the playback duration of each furigana of the pronunciation corresponding to "絆" is the same, so only the playback start timestamp of

"き" may be stored. When rendering the furigana during the subsequent playback of the song, the total duration of "きずな" may be determined according to the playback start identification of "き" and the playback start timestamp of the character after "絆", which may be divided by 3 to obtain the playback duration of each furigana. In this way, it is not necessary to store the playback start identification of each furigana, so storage resources can be saved. If the playback durations of different furigana are different, only when the playback start identification of each furigana is marked, when to render the furigana may be determined.

**[0082]** In a step 104, the pronunciations of the polyphonic characters in the target lyric and the display information thereof are correspondingly stored with the target lyric.

**[0083]** In implementation, after determining the phonetic notations of the target lyric, the server may correspondingly store the pronunciations of the polyphonic characters and the display information thereof with the target lyric based on the identification of the target lyric. In this way, the pronunciations of the target lyric and the display information thereof may be obtained subsequently based on the target lyric.

**[0084]** Optionally, there may be multiple methods of corresponding storage, and two feasible methods are given below:

**[0085]** Method I: the pronunciations of the polyphonic characters and the display information thereof are stored in an additional part of the text of the target lyric.

**[0086]** The additional part of the text of the target lyric may be a lyric Meta segment, which refers to an additional part for recording the lyric that has nothing to do with the rendered content of the lyric. For example, as shown in FIG. 2, in the lyric in an LRC (lyric) format, before the lyric text, there are the following contents: al: the album of the song, ar: performer-singer, au: lyric author-composer, by: creator of the LRC (lyric) file, etc.

**[0087]** In implementation, the server may obtain the additional part of the text of the target lyric based on the identification of the target lyric, and then store the pronunciations of the polyphonic characters and the display information thereof in the additional part of the text of the target lyric according to the recording format of the

additional part, for example, as shown in FIG. 3, in a format of "kana (meaning "假名" in Chinese): [playback start identification] furigana [playback start identification] furigana ...".

**[0088]** In this way, the pronunciation-related information is added in the original Meta segment. Since the Meta-segment parsing mechanism is relatively loose, even if the pronunciation-related information is added in the original Meta segment, it will not affect the format change, and will not cause an exception in the audio application.

**[0089]** Optionally, each line of lyric in the lyric text may correspond to the pronunciation, and the corresponding processing may be as follows.

**[0090]** Pronunciations of polyphonic characters in each line of the target lyric and display information thereof are determined; and the pronunciations of the polyphonic characters in each line of the target lyric and the display information thereof are respectively stored in the additional part of the text of the target lyric.

**[0091]** In implementation, the server may find each line of lyric in the text of the target lyric based on the identification of the target lyric, determine pronunciations of polyphonic characters in each line of lyric and display information thereof, and then store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof respectively in the additional part of the text of the target lyric. For example, as shown in FIG. 4, it may be stored in the following format:

**[0092]** "[kana2: pronunciation, and display information thereof]

**[0093]** [kana6: pronunciation and display information thereof]

**[0094]** [...]"

**[0095]** where 2 represents phonetic notations of polyphonic characters corresponding to the second line of lyric, and 6 represents phonetic notations of polyphonic characters corresponding to the sixth line of lyric.

**[0096]** Method II: the pronunciations of the polyphonic characters in the target lyric and the display information thereof are correspondingly stored with the target lyric in a phonetic notation file.

**[0097]** The phonetic notation file is a file for storing pronunciations of polyphonic characters and display information thereof. Each lyric corresponds to one phonetic notation file, which is distinguished by the identification of the lyric. When the

pronunciation and the display information thereof are not stored, the phonetic notation file is still stored, but is empty. Alternatively, when the pronunciation and the display information thereof are not stored, there is no phonetic notation file, which is newly created upon storage.

**[0098]** In implementation, the server may obtain the phonetic notation file of the target lyric based on the identification of the target lyric, and then store the pronunciations of the polyphonic characters and the display information thereof in the phonetic notation file according to the recording format of the text, for example, in a format of "kana: [playback start identification] furigana [playback start identification] furigana ...".

**[0099]** It should be noted that, in the step 103, the display information is the playback start identification of each furigana. The polyphonic character is "雨 の 空港" (meaning "雨中的机场" in Chinese), and "(39112) 雨(39803) の (40356) 空(40606) 港 (41176)" is stored in the lyric text. The pronunciation of "雨 の 空港" is "あ め の く う こ う", and "kana: (39112) あ (39921) め (40356) く う (40606) こ う" is stored in the additional part or phonetic notation file. The content in "()" indicates the playback start identification of the immediately following content. For example, (39921) indicates the playback start identification of "あ". In this way, the pronunciation may be obtained based on the playback start identification. (39112) corresponds to "雨" and "あ め". (40356) corresponds to "空" and "く う", and so on. In this way, during the subsequent playback of the song, the pronunciation can be obtained directly based on the playback start identification.

**[00100]** For the step 103, the display information is the playback start identification of the first furigana. The polyphonic character is "絆" (meaning "牵绊" in Chinese), and "(40576) 絆" is stored in the lyric text. The pronunciation of "絆" is "き ず な", and "kana: (40576) き ず な" may be stored in the additional part or phonetic notation file. Since the playback duration of the three furigana is the same, the total playback duration of the three furigana may be determined based on (40576) and the playback start identification of the character after "絆" (if the playback start identification

is a timestamp, the absolute value of the difference between the two timestamps is the total playback duration of the three furigana), which is then divided by three to obtain the playback duration of each of the three furigana. In this way, during the subsequent playback of the song, the pronunciation may be obtained based on the playback start identification of the first furigana. For example, it can be found that the pronunciation is "き す な" by using (40576) to make "絆" correspond to "き す な".

**[00101]** It should also be noted that, when the display information is the playback start identification of each furigana or the playback start identification of the first furigana, during the subsequent playback of the song, when obtaining the lyric, the playback start identification of the polyphonic character in the lyric text may be used to find the pronunciation corresponding to the playback start identification in the phonetic notation file or additional information for display. In this way, the correspondence between the polyphonic character and the pronunciation can be realized.

**[00102]** In the embodiment of the present disclosure, the pronunciations of the polyphonic characters in the lyric and the display information thereof are correspondingly stored with the lyric. In this way, during the subsequent playback of the song, the pronunciations of the polyphonic characters and the display information thereof can be obtained directly based on the lyric without real-time query, so the efficiency of displaying the phonetic notations of the lyric is improved.

**[00103]** Another embodiment of the present disclosure provides a method for implicitly storing phonetic notations. As shown in FIG. 5, the corresponding steps may be as follows.

**[00104]** In a step 501, a phonetic notation storage request for a target lyric is received.

**[00105]** The phonetic notation includes a pronunciation and display information thereof.

**[00106]** In implementation, a user may install an audio application in a terminal. The audio application may provide an interface for the user to edit lyric and pronunciations of polyphonic characters. For the lyric of a song (which may be referred to as target lyric later), the user may mark pronunciations of polyphonic characters in the target lyric and start timestamps of furigana in each pronunciation, and then click a

submit button. The terminal will generate a phonetic notation storage request which carries an identification of the target lyric, the pronunciations of the polyphonic characters, and the start timestamps of furigana in each pronunciation, and then send the phonetic notation storage request for the target lyric to a server. After receiving the phonetic notation storage request sent by the terminal, the server may parse the same to obtain the identification of the target lyric, the pronunciations of the polyphonic characters, and the start timestamps of furigana in each pronunciation.

**[00107]** Alternatively, the audio application provides the user with a phonetic notation adding option. If the user wants to add a phonetic notation to the lyric of a song (which may be referred to as target lyric later), he/she may click the phonetic notation adding option. The terminal will detect a click instruction of the phonetic notation adding option, generate a phonetic notation storage request which carries the identification of the target lyric. After receiving the phonetic notation storage request, the server may parse the same to obtain the identification of the target lyric.

**[00108]** In a step 502, pronunciations of polyphonic characters in the target lyric are obtained, and a playback duration of each furigana in the pronunciations of the polyphonic characters is determined.

**[00109]** The playback duration is the coloring duration when rendering the lyric. The playback start identification is used to indicate that the furigana starts to play, which may be a playback start timestamp and the like. Furigana is a kana symbol noted above or around a polyphonic character in Japanese. For example, the furigana of the pronunciation corresponding to "絆" (meaning "牵绊" in Chinese) is "き ず な".

**[00110]** For the processing in the step 502, reference may be made to the processing in step 102, and details are not described herein again.

**[00111]** In a step 503, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, then the number of Chinese characters in the polyphonic character and a playback start identification of each furigana in the pronunciation of the polyphonic character are obtained and determined as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then the number of Chinese

characters in the polyphonic character is obtained and determined as display information of the pronunciation of the polyphonic character.

**[00112]** In implementation, for any of the polyphonic characters, whether the playback durations of different furigana are the same may be determined. If they are different, the number of Chinese characters in the polyphonic character and the playback start identification of each furigana in the pronunciation of the polyphonic character may be determined as the display information of the pronunciation of the polyphonic character. If they are the same, the number of Chinese characters in the polyphonic character may be determined as the display information of the pronunciation of the polyphonic character.

**[00113]** It should be noted that this is implicit phonetic notation. For example, in Japanese, "欠片の絆" (meaning "碎片的牵绊" in Chinese), the pronunciation may be "2 か け ら 1 き ず な", where "2" means the following is the pronunciations of two characters, that is, "欠片", and "1" means that the following is the pronunciation of one character, that is, "絆". Since the playback duration of three furigana in the pronunciation of the polyphonic character "欠片" is different, the phonetic notation of "欠片" may be "2 か (1000) け (1500) ら (2000) 1 き ず な", where "1000", "1500", and "2000" indicate playback start identifications. Since the playback duration of three furigana in the pronunciation of the polyphonic character "絆" is the same, merely the number of Chinese characters may be marked. In addition, when the playback duration of the furigana is different, the playback duration may also be marked. For example, the display information of "欠片" may be "2 か (1000,500) け (1500,500) ら (2000,1000)", where 500 , 500, and 1000 respectively represent the playback duration of "か", "け", and "ら". In this way, for the furigana with the same playback duration, the playback time identification can be omitted and storage resources can be saved.

**[00114]** In a step 504, the pronunciations of the polyphonic characters in the target lyric and the display information thereof are correspondingly stored with the target lyric.

**[00115]** For the processing in the step 504, reference may be made to the processing in step 104, and details are not described herein again.

[00116] Optionally, for the case where the display information of the pronunciation of the polyphonic character includes the number of Chinese characters in the polyphonic character, during storage, the pronunciation of the polyphonic character needs to be stored in the order of the polyphonic character in the text. For example, the polyphonic characters are "欠片の絆" (meaning "碎片的牽絆" in Chinese). When the pronunciations are stored, because "欠片" precedes "絆", the storage order is that the pronunciation of "欠片" precedes "絆", that is, "2 か け ら 1 き ず な ".

[00117] Optionally, there may be multiple methods of corresponding storage, and two feasible methods are given below:

[00118] Method I: the pronunciations of the polyphonic characters and the display information thereof are stored in an additional part of the text of the target lyric.

[00119] The additional part of the text of the target lyric may be a lyric Meta segment, which refers to an additional part for recording the lyric that has nothing to do with the rendered content of the lyric. For example, as shown in FIG. 2, in the lyric in an LRC (lyric) format, before the lyric text, there are the following contents: al: the album of the song, ar: performer-singer, au: lyric author-composer, by: creator of the LRC (lyric) file, etc.

[00120] In implementation, the server may obtain the additional part of the text of the target lyric based on the identification of the target lyric, and then store the pronunciations of the polyphonic characters and the display information thereof in the additional part of the text of the target lyric according to the recording format of the additional part, for example, as shown in FIG. 3, in a format of "kana (meaning "假名" in Chinese): [playback start identification] furigana [playback start identification] furigana ...".

[00121] In this way, the pronunciation-related information is added in the original Meta segment. Since the Meta-segment parsing mechanism is relatively loose, even if the pronunciation-related information is added in the original Meta segment, it will not affect the format change, and will not cause an exception in the audio application.

**[00122]** Optionally, each line of lyric in the lyric text may correspond to the pronunciation, and the corresponding processing may be as follows.

**[00123]** Pronunciations of polyphonic characters in each line of the target lyric and display information thereof are determined; and the pronunciations of the polyphonic characters in each line of the target lyric and the display information thereof are respectively stored in the additional part of the text of the target lyric.

**[00124]** In implementation, the server may find each line of lyric in the text of the target lyric based on the identification of the target lyric, determine pronunciations of polyphonic characters in each line of lyric and display information thereof, and then store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof respectively in the additional part of the text of the target lyric. For example, as shown in FIG. 4, it may be stored in the following format:

```
" [kana2: pronunciation, and display information thereof]
[kana6: pronunciation and display information thereof]
[...]"
```

**[00125]** where 2 represents phonetic notations of polyphonic characters corresponding to the second line of lyric, and 6 represents phonetic notations of polyphonic characters corresponding to the sixth line of lyric.

**[00126]** Method II: the pronunciations of the polyphonic characters in the target lyric and the display information thereof are correspondingly stored with the target lyric in a phonetic notation file.

**[00127]** The phonetic notation file is a file for storing pronunciations of polyphonic characters and display information thereof. Each lyric corresponds to one phonetic notation file, which is distinguished by the identification of the lyric. When the pronunciation and the display information thereof are not stored, the phonetic notation file is still stored, but is empty. Alternatively, when the pronunciation and the display information thereof are not stored, there is no phonetic notation file, which is newly created upon storage.

**[00128]** In implementation, the server may obtain the phonetic notation file of the target lyric based on the identification of the target lyric, and then store the

pronunciations of the polyphonic characters and the display information thereof in the phonetic notation file according to the recording format of the text, for example, in a format of "kana: [playback start identification] furigana [playback start identification] furigana ...".

**[00129]** It should be noted that, in the step 503, the display information is the number of Chinese characters in the polyphonic character. The polyphonic character is "絆" (meaning "牽絆" in Chinese), and "(40576) 絆" is stored in the lyric text". The pronunciation of "絆" is "き ず な", and "kana: 1 き ず な" may be stored in the additional part or phonetic notation file. "1" represents the pronunciation of one character, that is, the pronunciation of "絆". Since the playback duration of the three furigana is the same, the total playback duration of the three furigana may be determined based on the playback start identification of "絆" in the lyric text and the playback start identification of the character after "絆" (if the playback start identification is a timestamp, the absolute value of the difference between the two timestamps is the total playback duration of the three furigana), which is then divided by three to obtain the playback duration of each of the three furigana.

**[00130]** For the step 503, the display information is the number of Chinese characters in the polyphonic character and the playback start identification of each furigana in the pronunciation of the polyphonic character. The polyphonic character is "欠片" (meaning "碎片" in Chinese), and "(40569) 欠片" is stored in the lyric text. The pronunciation of "欠片" is "か け ら", and "kana: 2 : (1000) け (1500) ら (2000)" may be stored in the additional part or phonetic notation file. "2" represents the pronunciations of two characters, that is, the pronunciation of "欠片". Since the playback duration of the three furigana is different, a playback start identification may be marked for each of the furigana. "1000", "1500", and "2000" indicate playback start identifications. In addition, "kana: 2 か (1000,500) け (1500,500) ら (2000,1000)" may be stored in the additional

part or phonetic notation file, where 500, 500, and 1000 respectively represent the playback duration of "か" and "け", "ら".

**[00131]** It should be noted that the display information is the number of Chinese characters in the polyphonic character the playback start identification of each furigana in the pronunciation of the polyphonic character, or the number of Chinese characters in the polyphonic character. In the process of use, when obtaining the target lyric, the server may obtain the encoding of the lyric text of the target lyric, and then determine which characters in the target lyric text are polyphonic characters based on the obtained encoding and need to be marked with pronunciations. The server determines the first polyphonic character in the target lyric text, and then determines the number of Chinese characters in the first polyphonic character to be the first number in "kana:", and makes the first polyphonic character correspond to the first pronunciation in "kana:"; and determines the second polyphonic character in the target lyric text, and then determines the number of Chinese characters in the second polyphonic character as the second number in "kana:", and makes the second polyphonic character correspond to the two pronunciations in "kana :", and so on. For example, the first polyphonic character identified in the target lyric text is "欠片", the number of Chinese characters is 2, and the first pronunciation in phonetic notation is 2 か け ら, then "か け ら" may be determined as the pronunciation of "欠片".

**[00132]** In the embodiment of the present disclosure, the pronunciations of the polyphonic characters in the lyric and the display information thereof are correspondingly stored with the lyric. In this way, during the subsequent playback of the song, the pronunciations of the polyphonic characters and the display information thereof can be obtained directly based on the lyric without real-time query, so the efficiency of displaying the phonetic representation of the lyric is improved.

**[00133]** In addition, in the above two embodiments of the present disclosure, a method for determining a pronunciation of a polyphonic character by a server is also provided. The corresponding processing may be as follows.

**[00134]** A dictionary is provided in the server which records pronunciations, prototypes, and part-of-speeches of multiple polyphonic characters and different

pronunciations thereof when being adjacent to different characters. For example, for "思い出" in Japanese, the pronunciation of "思" in "思い出を" (meaning "把回忆" in Chinese) is "おも", and the pronunciation of "出" is "で". The pronunciation of "思" in "思い出し" (meaning "请想起回忆" in Chinese) is "おも", and the pronunciation of "出" is "だ". Therefore, when actually looking up the dictionary, the longest matching principle may be adopted to look at the character or characters after a polyphonic character to obtain the pronunciation of the polyphonic character. In addition, other methods may also be used to determine the polyphonic character, which is only a feasible method here.

**[00135]** In addition, in the above two embodiments of the present disclosure, in the additional part, when the pronunciation is stored, it may not necessarily be "kana: [playback start identification] furigana [playback start identification] furigana ...". "[syllables (meaning "音节" in Chinese): [playback start identification] furigana [playback start identification] furigana..." or "[syll (meaning "音节" in Chinese): [playback start identification] furigana [playback start identification] furigana...", or other methods may be adopted. The additional part of the lyric text is generally provided before the lyric text, and may also be provided behind the lyric text.

**[00136]** It should also be noted that the storage manners in the above two embodiments of the present disclosure may be applied to lyric using karaoke format. In the embodiment of the present disclosure, the description of the scheme using songs in Japanese may also be used for rendering polyphonic characters in lyric in other languages, such as French, English, and the like.

**[00137]** Based on the same technical concept, an embodiment of the present disclosure further provides a device for storing lyric phonetic notations, as shown in FIG. 6, the device includes:

**[00138]** a receiving module 610 configured to receive a phonetic notation storage request for a target lyric;

[00139] an obtaining module 620 configured to obtain pronunciations of polyphonic characters in the target lyric and determine a playback duration of each furigana in the pronunciations of the polyphonic characters; and

[00140] a determining module 630 configured to, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, then obtain a playback start identification of each furigana in the pronunciation of the polyphonic character and determine the same as display information thereof, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then obtain a playback start identification of first furigana in the pronunciation of the polyphonic character and determine the same as display information thereof; and

[00141] a storing module 640 configured to correspondingly store the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

[00142] Optionally, the storing module 640 is configured to:

[00143] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric.

[00144] Optionally, the storing module 640 is configured to:

[00145] determine pronunciations of polyphonic characters in each line of the target lyric and display information thereof; and

[00146] store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric respectively.

[00147] Optionally, the storing module 640 is configured to:

[00148] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

[00149] Based on the same technical concept, an embodiment of the present disclosure further provides a device for storing lyric phonetic notations, as shown in FIG. 7, the device includes:

- [00150] a first receiving module 710 configured to receive a phonetic notation storage request for a target lyric;
- [00151] a first obtaining module 720 configured to obtain pronunciations of polyphonic characters in the target lyric and determine a playback duration of each furigana in the pronunciations of the polyphonic characters in the target lyric;
- [00152] a first determining module 730 is configured to obtain, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, obtain a number of Chinese characters in the polyphonic character and a playback start identification of each furigana in the pronunciation of the polyphonic character and determine the same as display information thereof, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then obtain the number of Chinese characters in the polyphonic character and determine the same as display information thereof; and
- [00153] a first storing module 740 is configured to correspondingly store the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.
- [00154] Optionally, the first storing module 740 is configured to:
- [00155] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric.
- [00156] Optionally, the first storing module 740 is configured to:
- [00157] determine pronunciations of polyphonic characters in each line of the target lyric and display information thereof;
- [00158] store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric respectively.
- [00159] Optionally, the first storing module 740 is configured to
- [00160] store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.
- [00161] In the embodiment of the present disclosure, the pronunciations of the polyphonic characters in the lyric and the display information thereof are correspondingly

stored with the lyric. In this way, during subsequent playback of the song, the pronunciations of the polyphonic characters and the display information thereof can be obtained directly based on the lyric without real-time query, so the efficiency of displaying the phonetic notations of the lyric is improved.

**[00162]** It should be noted that the device for storing lyric phonetic notations provided by the present embodiment only takes division of all the functional modules as an example for explanation when storing lyric phonetic notations. In practice, the above functions can be finished by the different functional modules as required. That is, the internal structure of the apparatus is divided into different functional modules to finish all or part of the functions described above. In addition, the device for storing lyric phonetic notations provided by the present embodiment has the same concept as the method for storing lyric phonetic notations provided by the foregoing embodiment. Refer to the method embodiment for the specific implementation process of the apparatus, which will not be repeated herein.

**[00163]** FIG. 8 is a structural block diagram of a terminal 800 in accordance with one embodiment of the present disclosure. The terminal 800 may be a smart phone, a tablet computer, an MP3 (Moving Picture Experts Group Audio Layer III) player, an MP4 (Moving Picture Experts Group Audio Layer IV) player, a laptop or desk computer. The terminal 800 may also be called a UE (User Equipment), a portable terminal, a laptop terminal, a desk terminal, etc.

**[00164]** Generally, the terminal 800 includes a processor 801 and a memory 802.

**[00165]** The processor 801 may include one or more processing cores, such as a 4-core processor and an 8-core processor. The processor 801 may be formed by at least one hardware of a DSP (Digital Signal Processing), an FPGA (Field-Programmable Gate Array), and a PLA (Programmable Logic Array). The processor 801 may also include a main processor and a coprocessor. The main processor is a processor for processing the data in an awake state, and is also called a CPU (Central Processing Unit). The coprocessor is a low-power-consumption processor for processing the data in a standby state. In some embodiments, the processor 801 may be integrated with a GPU (Graphics Processing Unit), which is configured to render and draw the content that needs to be displayed by a display screen. In some embodiments, the processor 801 may also include

an AI (Artificial Intelligence) processor configured to process computational operations related to machine learning.

**[00166]** The memory 802 may include one or more computer-readable storage mediums, which can be non-transitory. The memory 802 may also include a high-speed random access memory, as well as a non-volatile memory, such as one or more disk storage devices and flash storage devices. In some embodiments, the non-transitory computer-readable storage medium in the memory 802 is configured to store at least one instruction. The at least one instruction is configured to be executed by the processor 801 to implement the method for storing lyric phonetic notations provided by the method embodiments of the present disclosure.

**[00167]** In some embodiments, the terminal 800 also optionally includes a peripheral device interface 803 and at least one peripheral device. The processor 801, the memory 802, and the peripheral device interface 803 may be connected by a bus or a signal line. Each peripheral device may be connected to the peripheral device interface 803 by a bus, a signal line or a circuit board. For example, the peripheral device includes at least one of a radio frequency circuit 804, a touch display screen 805, a camera 806, an audio circuit 807, a positioning component 808 and a power source 809.

**[00168]** The peripheral device interface 803 may be configured to connect at least one peripheral device associated with an I/O (Input/Output) to the processor 801 and the memory 802. In some embodiments, the processor 801, the memory 802 and the peripheral device interface 803 are integrated on the same chip or circuit board. In some other embodiments, any one or two of the processor 801, the memory 802 and the peripheral device interface 803 may be implemented on a separate chip or circuit board, which is not limited in the present embodiment.

**[00169]** The radio frequency circuit 804 is configured to receive and transmit an RF (Radio Frequency) signal, which is also referred to as an electromagnetic signal. The radio frequency circuit 804 communicates with a communication network and other communication devices via the electromagnetic signal. The radio frequency circuit 804 converts the electrical signal into the electromagnetic signal for transmission, or converts the received electromagnetic signal into the electrical signal. Optionally, the radio frequency circuit 804 includes an antenna system, an RF transceiver, one or more

amplifiers, a tuner, an oscillator, a digital signal processor, a codec chipset, a subscriber identity module card, and the like. The radio frequency circuit 804 can communicate with other terminals via at least one wireless communication protocol. The wireless communication protocol includes, but not limited to, the World Wide Web, a metropolitan area network, an intranet, various generations of mobile communication networks (2G, 3G, 4G, and 5G), a wireless local area network, and/or a WiFi (Wireless Fidelity) network. In some embodiments, the RF circuit 804 may also include NFC (Near Field Communication) related circuits, which is not limited in the present disclosure.

**[00170]** The display screen 805 is configured to display a UI (User Interface). The UI may include graphics, text, icons, videos, and any combination thereof. When the display screen 805 is a touch display screen, the display screen 805 also has the capacity to acquire touch signals on or over the surface of the display screen 805. The touch signal may be input into the processor 801 as a control signal for processing. At this time, the display screen 805 may also be configured to provide virtual buttons and/or virtual keyboards, which are also referred to as soft buttons and/or soft keyboards. In some embodiments, one display screen 805 may be disposed on the front panel of the terminal 800. In some other embodiments, at least two display screens 805 may be disposed respectively on different surfaces of the terminal 800 or in a folded design. In further embodiments, the display screen 805 may be a flexible display screen disposed on the curved or folded surface of the terminal 800. Even the display screen 805 may have an irregular shape other than a rectangle; that is, the display screen 805 may be an irregular-shaped screen. The display screen 805 may be an LCD (liquid Crystal Display) screen, an OLED (Organic Light-Emitting Diode) screen or the like.

**[00171]** The camera component 806 is configured to capture images or videos. In some embodiments of the present disclosure, the camera component 806 includes a front camera and a rear camera. Usually, the front camera is placed on the front panel of the terminal, and the rear camera is placed on the back of the terminal. In some embodiments, at least two rear cameras are disposed, and are at least one of a main camera, a depth-of-field camera, a wide-angle camera, and a telephoto camera respectively, so as to realize a background blurring function achieved by fusion of the main camera and the depth-of-field camera, panoramic shooting and VR (Virtual Reality) shooting functions achieved

by fusion of the main camera and the wide-angle camera or other fusion shooting functions. In some embodiments, the camera component 806 may also include a flashlight. The flashlight may be a mono-color temperature flashlight or a two-color temperature flashlight. The two-color temperature flash is a combination of a warm flashlight and a cold flashlight and can be used for light compensation at different color temperatures.

**[00172]** The audio circuit 807 may include a microphone and a speaker. The microphone is configured to collect sound waves of users and environments, and convert the sound waves into electrical signals which are input into the processor 801 for processing, or input into the RF circuit 804 for voice communication. For the purpose of stereo obtaining or noise reduction, there may be a plurality of microphones respectively disposed at different locations of the terminal 800. The microphone may also be an array microphone or an omnidirectional obtaining microphone. The speaker is then configured to convert the electrical signals from the processor 801 or the radio frequency circuit 804 into the sound waves. The speaker may be a conventional film speaker or a piezoelectric ceramic speaker. When the speaker is the piezoelectric ceramic speaker, the electrical signal can be converted into not only human-audible sound waves but also the sound waves which are inaudible to humans for the purpose of ranging and the like. In some embodiments, the audio circuit 807 may also include a headphone jack.

**[00173]** The positioning component 808 is configured to locate the current geographic location of the terminal 800 to implement navigation or LBS (Location Based Service). The positioning component 808 may be a positioning component based on the American GPS (Global Positioning System), the Chinese Beidou system, the Russian Grenas System or the European Galileo system.

**[00174]** The power source 809 is configured to power up various components in the terminal 800. The power source 809 may be alternating current, direct current, a disposable battery, or a rechargeable battery. When the power source 809 includes the rechargeable battery, the rechargeable battery may be a wired rechargeable battery or a wireless rechargeable battery. The rechargeable battery may also support the fast charging technology.

**[00175]** In some embodiments, terminal 800 also includes one or more sensors 810. The one or more sensors 810 include, but not limited to, an acceleration sensor 811, a gyro sensor 812, a pressure sensor 813, a fingerprint sensor 814, an optical sensor 815 and a proximity sensor 816.

**[00176]** The acceleration sensor 811 may detect magnitudes of accelerations on three coordinate axes of a coordinate system established by the terminal 800. For example, the acceleration sensor 811 may be configured to detect components of a gravitational acceleration on the three coordinate axes. The processor 801 may control the touch display screen 805 to display a user interface in a landscape view or a portrait view according to a gravity acceleration signal collected by the acceleration sensor 811. The acceleration sensor 811 may also be configured to collect motion data of a game or a user.

**[00177]** The gyro sensor 812 can detect a body direction and a rotation angle of the terminal 800, and can cooperate with the acceleration sensor 811 to collect a 3D motion of the user on the terminal 800. Based on the data collected by the gyro sensor 812, the processor 801 can serve the following functions: motion sensing (such as changing the UI according to a user's tilt operation), image stabilization during shooting, game control and inertial navigation.

**[00178]** The pressure sensor 813 may be disposed on a side frame of the terminal 800 and/or a lower layer of the touch display screen 805. When the pressure sensor 813 is disposed on the side frame of the terminal 800, a user's holding signal to the terminal 800 can be detected. The processor 801 can perform left-right hand recognition or quick operation according to the holding signal collected by the pressure sensor 813. When the pressure sensor 813 is disposed on the lower layer of the touch display screen 805, the processor 801 controls an operable control on the UI according to a user's pressure operation on the touch display screen 805. The operable control includes at least one of a button control, a scroll bar control, an icon control and a menu control.

**[00179]** The fingerprint sensor 814 is configured to collect a user's fingerprint. The processor 801 identifies the user's identity based on the fingerprint collected by the fingerprint sensor 814, or the fingerprint sensor 814 identifies the user's identity based on the collected fingerprint. When the user's identity is identified as trusted, the processor

801 authorizes the user to perform related sensitive operations, such as unlocking the screen, viewing encrypted information, downloading software, paying, and changing settings. The fingerprint sensor 814 may be provided on the front, back, or side of the terminal 800. When the terminal 800 is provided with a physical button or a manufacturer's Logo, the fingerprint sensor 814 may be integrated with the physical button or the manufacturer's Logo.

**[00180]** The optical sensor 815 is configured to collect ambient light intensity. In one embodiment, the processor 801 may control the display brightness of the touch display screen 805 according to the ambient light intensity collected by the optical sensor 815. For example, when the ambient light intensity is high, the display brightness of the touch display screen 805 is increased; and when the ambient light intensity is low, the display brightness of the touch display screen 805 is decreased. In another embodiment, the processor 801 may also dynamically adjust shooting parameters of the camera component 806 according to the ambient light intensity collected by the optical sensor 815.

**[00181]** The proximity sensor 816, also referred to as a distance sensor, is usually disposed on the front panel of the terminal 800. The proximity sensor 816 is configured to capture a distance between the user and a front surface of the terminal 800. In one embodiment, when the proximity sensor 816 detects that the distance between the user and the front surface of the terminal 800 becomes gradually smaller, the processor 801 controls the touch display screen 805 to switch from a screen-on state to a screen-off state. When it is detected that the distance between the user and the front surface of the terminal 800 gradually increases, the processor 801 controls the touch display screen 805 to switch from the screen-off state to the screen-on state.

**[00182]** It will be understood by those skilled in the art that the structure shown in FIG. 8 does not constitute a limitation to the terminal 800, and may include more or less components than those illustrated, or combine some components or adopt different component arrangements.

**[00183]** Persons of ordinary skill in the art can understand that all or part of the steps described in the above embodiments can be completed through hardware, or

through relevant hardware instructed by applications stored in a non-transitory computer readable storage medium, such as a read-only memory, a disk or a CD, etc.

**[00184]** The foregoing descriptions are merely preferred embodiments of the present disclosure, and are not intended to limit the present disclosure. Within the spirit and principles of the disclosure, any modifications, equivalent substitutions, improvements, etc., are within the protection scope of the present disclosure.

## CLAIMS

What is claimed is:

1. A method for storing lyric phonetic notations, comprising:  
receiving a phonetic notation storage request for a target lyric;  
obtaining pronunciations of polyphonic characters in the target lyric, and  
determining a playback duration of each furigana in the pronunciations of the polyphonic characters;  
for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, then obtaining a playback start identification of each furigana in the pronunciation of the polyphonic character and determining the playback start identification as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then obtaining a playback start identification of first furigana in the pronunciation of the polyphonic character and determining the playback start identification as display information of the pronunciation of the polyphonic character; and  
correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.
2. The method according to claim 1, wherein correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric comprises:  
storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of a text of the target lyric.
3. The method according to claim 2, wherein storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric comprises:  
determining pronunciations of polyphonic characters in each line of the target lyric and display information thereof; and

storing the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric, respectively.

4. The method according to claim 1, wherein correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric comprises:

storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

5. A method for storing lyric phonetic notations, comprising:

receiving a phonetic notation storage request for a target lyric;

obtaining pronunciations of polyphonic characters in the target lyric, and determining a playback duration of each furigana in the pronunciations of the polyphonic characters in the target lyric;

for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, then obtaining a number of Chinese characters in the polyphonic character and a playback start identification of each furigana in the pronunciation of the polyphonic character and determining the number of Chinese characters in the polyphonic character and the playback start identification as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, then obtaining the number of Chinese characters in the polyphonic character and determining the number of Chinese characters in the polyphonic character as display information of the pronunciation of the polyphonic character; and

correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

6. The method according to claim 5, wherein correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric comprises:

storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of a text of the target lyric.

7. The method according to claim 6, wherein storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of the text of the target lyric comprises:

determining pronunciations of polyphonic characters in each line of the target lyric and display information thereof; and

storing the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric, respectively.

8. The method according to claim 5, wherein correspondingly storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric comprises:

storing the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

9. A device for storing lyric phonetic notations, comprising:

a receiving module configured to receive a phonetic notation storage request for a target lyric;

an obtaining module configured to obtain pronunciations of polyphonic characters in the target lyric and determine a playback duration of each furigana in the pronunciations of the polyphonic characters;

a determining module configured to, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, obtain a playback start identification of each furigana in the pronunciation of the polyphonic character and determine the playback start identification as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, obtain a playback start identification of first furigana in the pronunciation of the

polyphonic character and determine the same as display information of the pronunciation of the polyphonic character; and

a storing module configured to correspondingly store the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

10. The device according to claim 9, wherein the storing module is configured to:  
store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of a text of the target lyric.

11. The device according to claim 10, wherein the storing module is configured to:  
determine pronunciations of polyphonic characters in each line of the target lyric and display information thereof; and

store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric, respectively.

12. The device according to claim 9, wherein the storing module is configured to:  
store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.

13. A device for storing lyric phonetic notations, comprising:  
a first receiving module configured to receive a phonetic notation storage request for a target lyric;

a first obtaining module configured to obtain pronunciations of polyphonic characters in the target lyric and determine a playback duration of each furigana in the pronunciations of the polyphonic characters in the target lyric;

a first determining module configured to, for any of the polyphonic characters, if the playback durations of different furigana in the pronunciation of the polyphonic character are different, obtain a number of Chinese characters in the polyphonic character and a playback start identification of each furigana in the pronunciation of the polyphonic

character and determine the number of Chinese characters in the polyphonic character and the playback start identification as display information of the pronunciation of the polyphonic character, and if the playback durations of different furigana in the pronunciation of the polyphonic character are the same, obtain the number of Chinese characters in the polyphonic character and determine the number of Chinese characters in the polyphonic character as display information of the pronunciation of the polyphonic character; and

a first storing module configured to correspondingly store the pronunciations of the polyphonic characters in the target lyric and the display information thereof with the target lyric.

14. The device according to claim 13, wherein the first storing module is configured to:

store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in an additional part of a text of the target lyric.

15. The device according to claim 14, wherein the first storing module is configured to:

determine pronunciations of polyphonic characters in each line of the target lyric and display information thereof;

store the pronunciations of the polyphonic characters in each line of lyric and the display information thereof in the additional part of the text of the target lyric, respectively.

16. The device according to claim 13, wherein the first storing module is configured to:

store the pronunciations of the polyphonic characters in the target lyric and the display information thereof in a phonetic notation file corresponding to the target lyric.