



(11) **EP 2 485 029 A1**

(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 153(4) EPC

(43) Date of publication:
08.08.2012 Bulletin 2012/32

(51) Int Cl.:
G01L 19/00 ^(2006.01) **G01L 19/04** ^(2006.01)
G01L 19/12 ^(2006.01)

(21) Application number: **11774406.0**

(86) International application number:
PCT/CN2011/073479

(22) Date of filing: **28.04.2011**

(87) International publication number:
WO 2011/134415 (03.11.2011 Gazette 2011/44)

(84) Designated Contracting States:
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR**

(30) Priority: **28.04.2010 CN 201010163406**

(71) Applicant: **Huawei Technologies Co., Ltd.
Shenzhen, Guangdong 518129 (CN)**

(72) Inventors:
• **LIU, Zexin
Shenzhen
Guangdong 518129 (CN)**
• **MIAO, Lei
Shenzhen
Guangdong 518129 (CN)**

• **HU, Chen
Shenzhen
Guangdong 518129 (CN)**
• **WU, Wenhai
Shenzhen
Guangdong 518129 (CN)**
• **LANG, Yue
Shenzhen
Guangdong 518129 (CN)**
• **ZHANG, Qing
Shenzhen
Guangdong 518129 (CN)**

(74) Representative: **Isarpatent
Patent- und Rechtsanwälte
Postfach 44 01 51
80750 München (DE)**

(54) **AUDIO SIGNAL SWITCHING METHOD AND DEVICE**

(57) A method and an apparatus for switching speech or audio signals are disclosed. The method for switching speech or audio signals includes: when a switching of a speech or audio occurs, weighting a first high frequency band signal of a current frame of speech or audio signal and a second high frequency band signal

of the previous M frame of speech or audio signals to obtain a processed first high frequency band signal (101); and synthesizing the processed first high frequency band signal and a first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal (102).

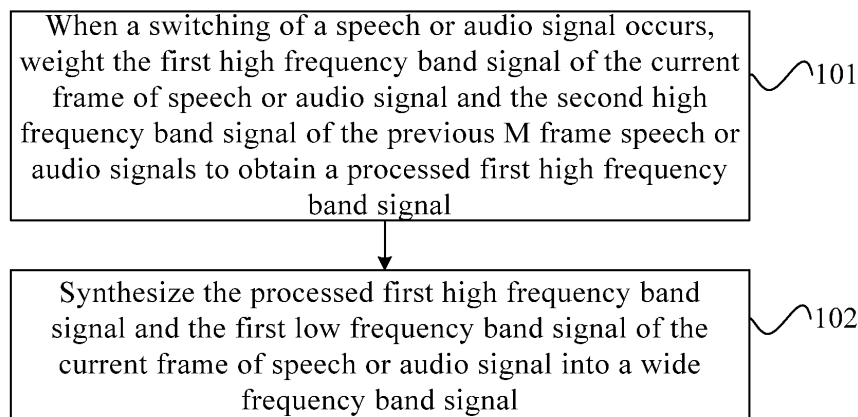


FIG. 1

EP 2 485 029 A1

Description

[0001] This application claims priority to Chinese Patent, titled as "METHOD AND APPARATUS FOR SWITCHING SPEECH OR AUDIO SIGNALS", Application No. 201010163406.3, filed on Apr. 28, 2010, which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present invention relates to communication technologies, and in particular, to a method and an apparatus for switching speech or audio signals.

BACKGROUND OF THE INVENTION

[0003] Currently, during the process of transmitting speech or audio signals on a network, because the network conditions may vary, the network may intercept the bit stream of the speech or audio signals transmitted from an encoder to the network with different bit rates, so that the decoder may decode the speech or audio signals with different bandwidths from the intercepted bit stream.

[0004] In the prior art, because the speech or audio signals transmitted on the network have different bandwidths, the bidirectional switching from/to a narrow frequency band speech or audio signal to/from a wide frequency band speech or audio signal may occur during the process of transmitting speech or audio signals. In embodiments of the present invention, the narrow frequency band signal is switched to a wide frequency band signal with only a low frequency band component through up-sampling and low-pass filtering; the wide frequency band speech or audio signal includes both a low frequency band signal component and a high frequency band signal component.

[0005] During the implementation of the present invention, the inventor discovers at least the following problems in the prior art: Because high frequency band signal information is available in wide frequency band speech or audio signals but is absent in narrow frequency band speech or audio signals, when speech or audio signals with different bandwidths are switched, a energy jump may occur in the speech or audio signals resulting in uncomfortable feeling in listening, and thus reducing the quality of audio signals received by a user.

SUMMARY OF THE INVENTION

[0006] Embodiments of the present invention provide a method and an apparatus for switching speech or audio signals to smoothly switch speech or audio signals between different bandwidths, thereby improving the quality of audio signals received by a user.

[0007] A method for switching speech or audio signals includes:

when a switching of a speech or audio signal occurs,

weighting a first high frequency band signal of a current frame of speech or audio signal and a second high frequency band signal of the previous M frame of speech or audio signals to obtain a processed first high frequency band signal, wherein, M is greater than or equal to 1; and synthesizing the processed first high frequency band signal and a first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal.

[0008] An apparatus for switching speech or audio signals includes:

a processing module, configured to: when a switching of a speech or audio occurs, weight a first high frequency band signal of a current frame of speech or audio signal and a second high frequency band signal of the previous M frame of speech or audio signals to obtain a processed first high frequency band signal, wherein, M is greater than or equal to 1; and

a first synthesizing module, configured to: synthesize the processed first high frequency band signal and a first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal.

[0009] By using the method and apparatus for switching speech or audio signals in embodiments of the present invention, the first high frequency band signal of the current frame of speech or audio signal is processed according to the second high frequency band signal of the previous M frame of speech or audio signals, so that the second high frequency band signal of the previous M frame of speech or audio signals can be smoothly switched to the processed first high frequency band signal; the processed first high frequency band signal and the first low frequency band signal are synthesized into a wide frequency band signal. In this way, during the process of switching between speech or audio signals with different bandwidths, these speech or audio signals can be smoothly switched, thus reducing the ill impact of the energy jump on the subjective audio quality of the speech or audio signals and improving the quality of speech or audio signals received by the user.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] To make the technical solution of the present invention clearer, the accompanying drawings for illustrating the embodiments of the present invention are outlined below. Apparently, the accompanying drawings are exemplary only, and those skilled in the art can derive other drawings from such accompanying drawings without creative efforts.

[0011] FIG. 1 is a flowchart of a first embodiment of a method for switching speech or audio signals;

[0012] FIG. 2 is a flowchart of a second embodiment of the method for switching speech or audio signals;

[0013] FIG.2 is a flowchart of an embodiment of step 201 shown in FIG. 2;

[0014] FIG.4 is a flowchart of an embodiment of step 302 shown in FIG. 3;

[0015] FIG.5 is a second flowchart of another embodiment of step 302 shown in FIG. 3;

[0016] FIG.6 is a flowchart of an embodiment of step 202 shown in FIG. 2;

[0017] FIG.7 is a second flowchart of another embodiment of step 201 shown in FIG. 2;

[0018] FIG.7 is a third flowchart of another embodiment of step 201 shown in FIG. 2;

[0019] FIG. 9 shows a structure of a first embodiment of an apparatus for switching speech or audio signals;

[0020] FIG. 10 shows a structure of a second embodiment of the apparatus for switching speech or audio signals;

[0021] FIG. 11 is a first schematic diagram illustrating a structure of a processing module in the second embodiment of the apparatus for switching speech or audio signals;

[0022] FIG. 12 is a schematic diagram illustrating a structure of a first module in the second embodiment of the apparatus for switching speech or audio signals;

[0023] FIG. 13a is a second schematic diagram illustrating a structure of the processing module in the second embodiment of the apparatus for switching speech or audio signals; and

[0024] FIG. 13b is a third schematic diagram illustrating a structure of the processing module in the second embodiment of the apparatus for switching speech or audio signals.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0025] To facilitate the understanding of the object, technical solution, and merit of the present invention, the following describes the present invention in detail with reference to embodiments and accompanying drawings. Apparently, the embodiments are exemplary only and the present invention is not limited to such embodiments. Persons having ordinary skill in the related art can derive other embodiments from the embodiments given herein without making remarkable creative effort, and all such embodiments are covered in the scope of the present invention.

[0026] FIG. 1 is a flowchart of the first embodiment of a method for switching speech or audio signals. As shown in FIG. 1, by using the method for switching speech or audio signals, when a switching of a speech or audio occurs, each frame after a switching frame is processed according to the following steps:

[0027] Step 101: When a switching of a speech or audio occurs, weight the first high frequency band signal of the current frame of speech or audio signal and the second high frequency band signal of the previous M frame

of speech or audio signals to obtain a processed first high frequency band signal, where M is greater than or equal to 1.

[0028] Step 102: Synthesize the processed first high frequency band signal and the first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal.

[0029] In this embodiment, the previous M frame of speech or audio signals refer to M frame of speech or audio signals before the current frame. The L frame of speech or audio signals before the switching refer to L frame of speech or audio signals before the switching frame. When a switching of a speech or audio occurs. If the current speech frame is a wide frequency band signal but the previous speech frame is a narrow frequency band signal or if the current speech frame is a narrow frequency band signal but the previous speech frame is a wide frequency band signal, the speech or audio signal is switched and the current speech frame is the switching frame.

[0030] By using the method for switching speech or audio signals in this embodiment, the first high frequency band signal of the current frame of speech or audio signal is processed according to the second high frequency band signal of the previous M frame of speech or audio signals, so that the second high frequency band signal of the previous M frame of speech or audio signals can be smoothly switched to the processed first high frequency band signal. In this way, during the process of switching between speech or audio signals with different bandwidths, the high frequency band signal of these speech or audio signals can be smoothly switched. Finally, the processed first high frequency band signal and the first low frequency band signal are synthesized into a wide frequency band signal; the wide frequency band signal is transmitted to a user terminal, so that the user enjoys a high quality speech or audio signal. By using the method for switching speech or audio signals in this embodiment, speech or audio signals with different bandwidths can be switched smoothly, thus reducing the impact of the sudden energy change on the subjective audio quality of the speech or audio signals and improving the quality of speech or audio signals received by the user.

[0031] FIG. 2 is a flowchart of the second embodiment of the method for switching speech or audio signals. As shown in FIG. 2, the method includes the following steps:

[0032] Step 200: When a switching of the speech or audio signal does not occur, synthesize the first high frequency band signal of the current frame of speech or audio signal and the first low frequency band signal into a wide frequency band signal.

[0033] Specifically, the first frequency band speech or audio signal in this embodiment may be a wide frequency band speech or audio signal or a narrow frequency band speech or audio signal. When the first frequency band speech or audio signal is not switched during the transmission of the speech or audio signal, the operation may be executed according to the following two cases: 1. If

the first frequency band speech or audio signal is a wide frequency band speech or audio signal, the low frequency band signal and high frequency band signal of the wide frequency band speech or audio signals are synthesized into a wide frequency band signal. 2. If the first frequency band speech or audio signal is a narrow frequency band speech or audio signal, the low frequency band signal and the high frequency band signal of the narrow frequency band speech or audio signal are synthesized into a wide frequency band signal. In this case, although the signal is a wide frequency band signal, the high frequency band is null.

[0034] Step 201: When the speech or audio signal is switched, weight the first high frequency band signal of the current frame of speech or audio signal and the second high frequency band signal of the previous M frame of speech or audio signals to obtain a processed first high frequency band signal. M is greater than or equal to 1.

[0035] Specifically, when the switching between speech or audio signals with different bandwidths occurs, the first high frequency band signal of the current frame of speech or audio signal is processed according to the second high frequency band signal of the previous M frame of speech or audio signals, so that the second high frequency band signal of the previous M frame of speech or audio signals can be smoothly switched to the processed first high frequency band signal. For example, when the wide frequency band speech or audio signal is switched to the narrow frequency band speech or audio signal, because the high frequency band signal information corresponding to the narrow frequency band speech or audio signal is null, the component of the high frequency band signal corresponding to the narrow frequency band speech or audio signal needs to be restored to enable the wide frequency band speech or audio signal to be smoothly switched to the narrow frequency band speech or audio signal. However, when the narrow frequency band speech or audio signal is switched to the wide frequency band speech or audio signal, because the high frequency band signal of the wide frequency band speech or audio signal is not null, the energy of the high frequency band signals of consecutive multiple-frame wide frequency band speech or audio signals after the switching must be weakened to enable the narrow frequency band speech or audio signal to be smoothly switched to the wide frequency band speech or audio signal, so that the high frequency band signal of the wide frequency band speech or audio signal is gradually switched to a real high frequency band signal. By processing the current frame of speech or audio signal in step 201, high frequency band signals in speech or audio signals with different bandwidths can be smoothly switched, which avoids uncomfortable listening of the user due to the sudden energy change in the process of switching between the wide frequency band speech or audio signal and the narrow frequency band speech or audio signal, enabling the user to receive high quality audio signals. To simplify the process of obtaining the

processed first high frequency band signal, the first high frequency band signal and the second high frequency band signal of the previous M frame of speech or audio signals may be directly weighted. The weighted result is the processed first high frequency band signal.

[0036] Step 202: Synthesize the processed first high frequency band signal and the first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal.

[0037] Specifically, after the current frame of speech or audio signal is processed in step 201, the second high frequency band signal of the previous M frame of speech or audio signals can be smoothly switched to the processed first high frequency band signal of the current frame; then, in step 202, the processed first high frequency band signal and the first low frequency band signal of the current frame of speech or audio signal are synthesized into a wide frequency band signal, so that the speech or audio signals received by the user are always wide frequency band speech or audio signals. In this way, speech or audio signals with different bandwidths are smoothly switched, which helps improve the quality of audio signals received by the user.

[0038] By using the method for switching speech or audio signals in this embodiment, the first high frequency band signal of the current frame of speech or audio signal is processed according to the second high frequency band signal of the previous M frame of speech or audio signals, so that the second high frequency band signal of the previous M frame of speech or audio signals can be smoothly switched to the processed first high frequency band signal. In this way, during the process of switching between speech or audio signals with different bandwidths, the high frequency band signal of these speech or audio signals can be smoothly switched. Finally, the processed first high frequency band signal and the first low frequency band signal are synthesized into a wide frequency band signal; the wide frequency band signal is transmitted to a user terminal, so that the user enjoys a high quality speech or audio signal. By using the method for switching speech or audio signals in this embodiment, speech or audio signals with different bandwidths can be switched smoothly, thus reducing the impact of the sudden energy change on the subjective audio quality of the speech or audio signals and improving the quality of audio signals received by the user. In addition, when speech or audio signals with different bandwidths are not switched, the first high frequency band signal and the first low frequency band signal of the current frame of speech or audio signal are synthesized into a wide frequency band signal, so that the user can obtain high quality audio signal.

[0039] According to the preceding technical solution, optionally, as shown in FIG. 3, when a switching from wide frequency band speech or audio signal to a narrow frequency band speech or audio signal occurs, step 201 includes the following steps:

[0040] Step 301: Predict fine structure information and

envelope information corresponding to the first high frequency band signal.

[0041] Specifically, the speech or audio signal may be divided into fine structure information and envelope information, so that the speech or audio signal can be restored according to the fine structure information and envelope information. In the process of switching from a wide frequency band speech or audio signal to a narrow frequency band speech or audio signal, because only a low frequency band signal is available in the narrow frequency band speech or audio signal and the high frequency band signal is null, to enable the wide frequency band speech or audio signal to be smoothly switched to the narrow frequency band speech or audio signal, the high frequency band signal needed by the current narrow frequency band speech or audio signal needs to be restored so as to implement smooth switching between speech or audio signals. In step 301, the predicted fine structure information and envelope information corresponding to the first high frequency band signal of the narrow frequency band speech or audio signal are predicted.

[0042] To predict the fine structure information and envelope information corresponding to the current frame of speech or audio signal more accurately, the first low frequency band signal of the current frame of speech or audio signal may be classified in step 301, and then the predicted fine structure information and envelope information corresponding to the first high frequency band signal are predicted according to the signal type of the first low frequency band signal. For example, the narrow frequency band speech or audio signal of the current frame may be a harmonic signal, or a non-harmonic signal or a transient signal. In this case, the fine structure information and envelope information corresponding to the type of the narrow frequency band speech or audio signal can be obtained, so that the fine structure information and envelope information corresponding to the high frequency band signal can be predicted more accurately. The method for switching speech or audio signals in this embodiment does not limit the signal type of the narrow frequency band speech or audio signal.

[0043] Step 302: Weight the predicted envelope information and the previous M frame envelope information corresponding to the second high frequency band signal of the previous M frame of speech or audio signals to obtain first envelope information corresponding to the first high frequency band signal.

[0044] Specifically, after the predicted fine structure information and envelope information corresponding to the first high frequency band signal of the current frame are predicted in step 301, the first envelope information corresponding to the first high frequency band signal may be generated according to the predicted envelope information and the previous M frame envelope information corresponding to the second high frequency band signal of the previous M frame of speech or audio signals.

[0045] Specifically, the process of generating the first

envelope information corresponding to the first high frequency band signal in step 302 may be implemented by using the following two modes:

[0046] 1. As shown in FIG. 4, an embodiment of obtaining the first envelope information through step 302 may include the following steps:

[0047] Step 401: Calculate a correlation coefficient between the first low frequency band signal and the low frequency band signal of the previous N frame of speech or audio signals according to the first low frequency band signal and the low frequency band signal of the previous N frame of speech or audio signals, where N is greater than or equal to 1.

[0048] Specifically, the first low frequency band signal of the current frame of speech or audio signal is compared with the low frequency band signal of the previous N frame of speech or audio signals to obtain a correlation coefficient between the first low frequency band signal of the current frame of speech or audio signal and the low frequency band signal of the previous N frame of speech or audio signals. For example, the correlation between the first low frequency band signal of the current frame of speech or audio signal and the low frequency band signal of the previous N frame of speech or audio signals may be determined by judging the difference between a frequency band of the first low frequency band signal of the current frame of speech or audio signal and the same frequency band of the low frequency band signal of the previous N frame of speech or audio signals in terms of the energy size or the information type, so that the desired correlation coefficient can be calculated. The previous N frame of speech or audio signals may be narrow frequency band speech or audio signals, wide frequency band speech or audio signals, or hybrid signals of narrow frequency band speech or audio signals and wide frequency band speech or audio signals.

[0049] Step 402: Judge whether the correlation coefficient is within a given first threshold range.

[0050] Specifically, after the correlation coefficient is calculated in step 401, whether the correlation coefficient is within the given threshold range is judged. The purpose of calculating the correlation coefficient is to judge whether the current frame of speech or audio signal is gradually switched from the previous N frame of speech or audio signals or suddenly switched from the previous N frame of speech or audio signals. That is, the purpose is to judge whether their characteristics are the same and then determine the weight of the high frequency band signal of the previous frame in the process of predicting the high frequency band signal of the current speech or audio signal. For example, if the first low frequency band signal of the current frame of speech or audio signal has the same energy as the low frequency band signal of the previous frame of speech or audio signal and their signal types are the same, it indicates that the previous frame of speech or audio signal is highly correlated with the current frame of speech or audio signal. Therefore, to accurately restore the first envelope information corre-

sponding to the current frame of speech or audio signal, the high frequency band envelope information or transitional envelope information corresponding to the previous frame of speech or audio signal occupies a larger weight; otherwise, if there is a huge difference between the first low frequency band signal of the current frame of speech or audio signal and the low frequency band signal of the previous frame of speech or audio signal in terms of energy and their signal types are different, it indicates that the previous speech or audio signal is lowly correlated with the current frame of speech or audio signal. Therefore, to accurately restore the first envelope information corresponding to the current frame of speech or audio signal, the high frequency band envelope information or transitional envelope information corresponding to the previous frame of speech or audio signal occupies a smaller weight.

[0051] Step 403: If the correlation coefficient is not within the given first threshold range, weight according to a set first weight 1 and a set first weight 2 to calculate the first envelope information. The first weight 1 refers to the weight value of the previous frame envelope information corresponding to the high frequency band signal of the previous frame of speech or audio signal, and the first weight 2 refers to the weight value of the envelope information.

[0052] Specifically, if the correlation coefficient is determined to be not within the given first threshold range in step 402, it indicates that the current frame of speech or audio signal is slightly correlated with the previous N frame of speech or audio signals. Therefore, the previous M frame envelope information or transitional envelope information corresponding to the first frequency band speech or audio signal of the previous M frames or the high frequency band envelope information corresponding to the previous frame of speech or audio signal has a slight impact on the first envelope information. When the first envelope information corresponding to the current frame of speech or audio signal is restored, the previous M frame envelope information or transitional envelope information corresponding to the first frequency band speech or audio signal of the previous M frames or the high frequency band envelope information corresponding to the previous frame of speech or audio signal occupies a smaller weight. Therefore, the first envelope information of the current frame may be calculated according to the set first weight 1 and the first weight 2. The first weight 1 refers to the weight value of the envelope information corresponding to the high frequency band signal of the previous frame of speech or audio signal. The previous frame of speech or audio signal may be a wide frequency band speech or audio signal or a processed narrow frequency band speech or audio signal. In the case of first switching, the previous frame of speech or audio signal is the wide frequency band speech or audio signal, while the first weight 2 refers to the weight value of the predicted envelope information. The product of the predicted envelope information and the first weight

2 is added to the product of the previous frame envelope information and the first weight 1, and the weighted sum is the first envelope information of the current frame. In addition, subsequently transmitted speech or audio signals are processed according to this method and weight. The first envelope information corresponding to the speech or audio signal is restored until a speech or audio signal is switched again.

[0053] Step 404: If the correlation coefficient is within the given first threshold range, weight according to a set second weight 1 and a set second weight 2 to calculate the transitional envelope information. The second weight 1 refers to the weight value of the envelope information before the switching, and the second weight 2 refers to the weight value of the previous M frame envelope information, where M is greater than or equal to 1.

[0054] Specifically, if the correlation coefficient is determined to be within the given threshold range in step 402, the current frame of speech or audio signal has characteristics similar to those of the previous consecutive N frame of speech or audio signals, and the first envelope information corresponding to the current frame of speech or audio signal is greatly affected by the envelope information of the previous consecutive N frame of speech or audio signals. In view of the authenticity of the previous M frame envelopes, the transitional envelope information corresponding to the current frame of speech or audio signal needs to be calculated according to the previous M frame envelope information and the envelope information before the switching. When the first envelope information of the current frame of speech or audio signal is restored, the previous M frame envelope information and the previous L frame envelope information before the switching should occupy a larger weight. Then, the first envelope information is calculated according to the transitional envelope information. The second weight 1 refers to the weight value of the envelope information before the switching, and the second weight 2 refers to the weight value of the previous M frame envelope information. In this case, the product of the envelope information before the switching and the second weight 1 is added to the product of the previous M frame envelope information and the second weight 2, and the weighted value is the transitional envelope information.

[0055] Step 405: Decrease the second weight 1 as per the first weight step, and increase the second weight 2 as per the first weight step.

[0056] Specifically, as the speech or audio signals are transmitted, the impact of the wide frequency band speech or audio signals before the switching on the subsequent narrow frequency band speech or audio signals is gradually decreased. To calculate the first envelope information more accurately, adaptive adjustment needs to be performed on the second weight 1 and the second weight 2. Because the impact of the L frame wide frequency band speech or audio signals before the switching on the subsequent speech or audio signals is decreased gradually, the value of the second weight 1 turns

smaller gradually, while the value of the second weight 2 turns larger gradually, thus weakening the impact of the envelope information before the switching on the first envelope information. In step 405, the second weight 1 and the second weight 2 may be modified according to the following formulas: New second weight 1 = Old second weight 1 - First weight step; New second weight 2 = Old second weight 2 + First weight step, where the first weight step is a set value.

[0057] Step 406: Judge whether a set third weight 1 is greater than the first weight 1.

[0058] Specifically, the third weight 1 refers to the weight value of the transitional envelope information. The impact of the transitional envelope information on the first envelope information of the current frame may be determined by comparing the third weight 1 with the second weight 1. The transitional envelope information is calculated according to the previous M frame envelope information and the envelope information before the switching. Therefore, the third weight 1 actually represents the degree of the impact that the first envelope information suffers from the envelope information before the switching.

[0059] Step 407: If the third weight 1 is not greater than the first weight 1, weight according to the set first weight 1 and the first weight 2 to calculate the first envelope information.

[0060] Specifically, when the third weight 1 is determined to be smaller than or equal to the first weight 1 in step 406, it indicates that the current frame of speech or audio signal is a little far from the L frame of speech or audio signals before the switching and that the first envelope information is mainly affected by the previous M frame envelope information. Therefore, the first envelope information of the current frame may be calculated according to the set first weight 1 and the first weight 2.

[0061] Step 408: If the third weight 1 is greater than the first weight 1, weight according to the set third weight 1 and the third weight 2 to calculate the first envelope information. The third weight 1 refers to the weight value of the transitional envelope information, and the third weight 2 refers to the weight value of the predicted envelope information.

[0062] Specifically, if the third weight 1 is determined to be greater than the first weight 1 in step 406, it indicates that the current frame of speech or audio signal is closer to the L frame of speech or audio signals before the switching and that the first envelope information is greatly affected by the envelope information before the switching. Therefore, the first envelope information of the current frame needs to be calculated according to the transitional envelope information. The third weight 1 refers to the weight value of the transitional envelope information, and the third weight 2 refers to the weight value of the predicted envelope information. In this case, the product of the transitional envelope information and the third weight 1 is added to the product of the predicted envelope information and the third weight 2, and the weighted value

is the first envelope information.

[0063] Step 409: Decrease the third weight 1 as per the second weight step, and increase the third weight 2 as per the second weight step until the third weight 1 is equal to 0.

[0064] Specifically, the purpose of modifying the third weight 1 and the third weight 2 in step 409 is the same as that of modifying the second weight 1 and the second weight 2 in step 405, that is, the purpose is to perform adaptive adjustment on the third weight 1 and the third weight 2 to calculate the first envelope information more accurately when the impact of the L frame of speech or audio signals before the switching on the subsequently transmitted speech or audio signals is decreased gradually. Because the impact of the L frame of speech or audio signals before the switching on the subsequent speech or audio signals is decreased gradually, the value of the third weight 1 turns smaller gradually, while the value of the third weight 2 turns larger gradually, thus weakening the impact of the envelope information before the switching on the first envelope information. In step 409, the third weight 1 and the third weight 2 may be modified according to the following formulas: New third weight 1 = Old third weight 1 - Second weight step; New third weight 2 = Old third weight 2 + Second weight step, where the second weight step is a set value.

[0065] The sum of the first weight 1 and the first weight 2 is equal to 1; the sum of the second weight 1 and the second weight 2 is equal to 1; the sum of the third weight 1 and the third weight 2 is equal to 1; the initial value of the third weight 1 is greater than the initial value of the first weight 1; and the first weight 1 and the first weight 2 are fixed constants. Specifically, the weight 1 and the weight 2 in this embodiment actually represent the percentages of the envelope information before the switching and the previous M frame envelope information in the first envelope information of the current frame. If the current frame of speech or audio signal is close to the L frame of speech or audio signals before the switching and their correlation is high, the percentage of the envelope information before the switching is high, while the percentage of the previous M frame envelope information is low. If the current frame of speech or audio signal is a little far from the L frame of speech or audio signals before the switching, it indicates that the speech or audio signal is stably transmitted on the network; or if the current frame of speech or audio signal is slightly correlated with the L frame of speech or audio signals before the switching, it indicates that the characteristics of the current frame of speech or audio signal are already changed. Therefore, if the current frame of speech or audio signal is slightly affected by the L frame of speech or audio signals before the switching, the percentage of the envelope information before the switching is low.

[0066] In addition, step 404 may be executed after step 405. That is, the second weight 1 and the second weight 2 may be modified firstly, and then the transitional envelope information is calculated according to the second

weight 1 and the second weight 2. Similarly, step 408 may be executed after step 409. That is, the third weight 1 and the third weight 2 may be modified firstly, and then the first envelope information is calculated according to the third weight 1 and the third weight 2.

[0067] 2. As shown in FIG. 5, another embodiment of obtaining the first envelope information through step 302 may further include the following steps:

[0068] Step 501: Calculate a correlation coefficient between the first low frequency band signal and the low frequency band signal of the previous frame of speech or audio signal according to the first low frequency band signal of the current frame of speech or audio signal and the low frequency band signal of the previous frame of speech or audio signal.

[0069] Specifically, to obtain more accurate first envelope information, the relationship between a frequency band of the first low frequency band signal of the current frame of speech or audio signal and the same frequency band of the low frequency band signal of the previous frame of speech or audio signal is calculated. In this embodiment, "corr" may be used to indicate the correlation coefficient. This correlation coefficient is obtained according to the energy relationship between the first low frequency band signal of the current frame of speech or audio signal and the low frequency band signal of the previous frame of speech or audio signal. If the energy difference is small, the "corr" is large; otherwise, the "corr" is small. For the specific process, see the calculation about the correlation of the previous N frame of speech or audio signals in step 401.

[0070] Step 502: Judge whether the correlation coefficient is within a given second threshold range.

[0071] Specifically, after the value of the "corr" is calculated in step 501, whether the calculated "corr" value is within the given second threshold is judged. For example, the second threshold range may be represented by c1 to c2 in this embodiment.

[0072] Step 503: If the correlation coefficient is not within the given second threshold range, weight according to the set first weight 1 and the first weight 2 to calculate the first envelope information. The first weight 1 refers to the weight value of the previous frame envelope information corresponding to the high frequency band signal of the previous frame of speech or audio signal, and the first weight 2 refers to the weight value of the predicted envelope information. The first weight 1 and the second weight 2 are fixed constants.

[0073] Specifically, when the "corr" value is determined to be smaller than c1 or greater than c2, it is determined that the first envelope information corresponding to the current frame of speech or audio signal is slightly affected by the envelope information of the previous frame of speech or audio signal before the switching. Therefore, the first envelope information of the current frame is calculated according to the set first weight 1 and the first weight 2. The product of the predicted envelope information and the first weight 2 is added to the product

of the previous frame envelope information and the first weight 1, and the weighted sum is the first envelope information of the current frame. In addition, subsequently transmitted narrowband speech or audio signals are processed according to this method and weight. The first envelope information corresponding to the narrowband speech or audio signal is restored until the speech or audio signals with different bandwidths are switched again. For example, the first weight 1 in this embodiment may be represented by a1; the first weight 2 may be represented by b1; the previous frame envelope information may be represented by pre_fenv; the predicted envelope information may be represented by fenv; and the first envelope information may be represented by cur_fenv. In this case, step 503 may be represented by the following formula: $\text{cur_fenv} = \text{pre_fenv} \times a1 + \text{fenv} \times b1$.

[0074] Step 504: If the correlation coefficient is within the second threshold range, judge whether the set second weight 1 is greater than the first weight 1. The second weight 1 refers to the weight value of the envelope information before the switching that corresponds to the high frequency band signal of the previous frame of speech or audio signal before the switching.

[0075] Specifically, if $c1 < \text{corr} < c2$, the degree of the impact of the envelope information before the switching and the previous frame envelope information on the first envelope information of the current frame may be obtained by comparing the second weight 1 with the first weight 1.

[0076] Step 505: If the second weight 1 is not greater than the first weight 1, weight according to the set first weight 1 and the first weight 2 to calculate the first envelope information.

[0077] Specifically, when the second weight 1 is determined to be smaller than the first weight 1 in step 504, it indicates that the current frame of speech or audio signal is a little far from the previous frame of speech or audio signal before the switching and that the first envelope information is slightly affected by the previous frame envelope information before the switching. Therefore, the first envelope information of the current frame may be calculated according to the set first weight 1 and the first weight 2. In this case, step 505 may be represented by the following formula: $\text{cur_fenv} = \text{pre_fenv} \times a1 + \text{fenv} \times b1$.

[0078] Step 506: If the second weight 1 is greater than the first weight 1, weight according to the second weight 1 and the set second weight 2 to calculate the first envelope information. The second weight 2 refers to the weight value of the predicted envelope information. For example, the second weight 1 may be represented by a2, and the second weight 2 may be represented by b2.

[0079] Specifically, when the second weight 1 is determined to be greater than the first weight 1 in step 504, it indicates that the current frame of speech or audio signal is closer to the first frequency band speech or audio signal of the previous frame before the switching and that the first envelope information is greatly affected by the

envelope information before the switching that corresponds to the previous frame of speech or audio signal before the switching. Therefore, the first envelope information of the current frame may be calculated according to the set second weight 1 and the second weight 2. In this case, the product of the predicted envelope information and the second weight 2 is added to the product of the envelope information before the switching and the second weight 1, and the weighted sum is the first envelope information of the current frame. The envelope information before the switching may be represented by con_fenv. In this case, step 506 may be represented by the following formula: $\text{cur_fenv} = \text{con_fenv} \times a2 + \text{fenv} \times b2$.

[0080] Step 507: Decrease the second weight 1 as per the second weight step, and increase the second weight 2 as per the second weight step.

[0081] Specifically, as the speech or audio signals are transmitted, the impact of a speech or audio signal before the switching on the subsequent frame of speech or audio signal is gradually decreased. To calculate the first envelope information more accurately, adaptive adjustment needs to be performed on the second weight 1 and the second weight 2. The impact of the speech or audio signal before the switching on the subsequent frame of speech or audio signal is gradually decreased, while the impact of the previous frame of speech or audio signal close to the current frame of speech or audio signal turns larger gradually. Therefore, the value of the second weight 1 turns smaller gradually, while the value of the second weight 2 turns larger gradually. In this way, the impact of the envelope information before the switching on the first envelope information is weakened, while the impact of the predicted envelope information on the first envelope information is enhanced. In step 507, the second weight 1 and the second weight 2 may be modified according to the following formulas: New second weight 1 = Old second weight 1 - First weight step; New second weight 2 = Old second weight 2 + First weight step, where the first weight step is a set value.

[0082] The sum of the first weight 1 and the first weight 2 is equal to 1; the sum of the second weight 1 and the second weight 2 is equal to 1; the initial value of the second weight 1 is greater than the initial value of the first weight 1.

[0083] Step 303: Generate a processed first high frequency band signal according to the first envelope information and the predicted fine structure information.

[0084] Specifically, after the first envelope information of the current frame is obtained in step 302, the processed first high frequency band signal may be generated according to the first envelope information and predicted fine structure information, so that the second high frequency band signal can be smoothly switched to the processed first high frequency band signal.

[0085] By using the method for switching speech or audio signals in this embodiment, in the process of switching a speech or audio signal from a wide frequency

band speech or audio signal to a narrow frequency band speech or audio signal, the processed first high frequency band signal of the current frame is obtained according to the predicted fine structure information and the first envelope information. In this way, the second high frequency band signal of the wide frequency band speech or audio signal before the switching can be smoothly switched to the processed first high frequency band signal corresponding to the narrow frequency band speech or audio signal, thus improving the quality of audio signals received by the user.

[0086] Based on the preceding technical solution, step 202 shown in FIG. 6 includes the following steps:

[0087] Step 601: Judge whether the processed first high frequency band signal needs to be attenuated according to the current frame of speech or audio signal and the previous frame of speech or audio signal before the switching.

[0088] Specifically, the first high frequency band signal of the narrowband speech or audio signal is null. In the process of switching the wide frequency band speech or audio signal to the narrow frequency band speech or audio signal, to prevent the negative impact of the processed first high frequency band signal corresponding to the restored narrow frequency band speech or audio signal, the energy of the processed first high frequency band signal is attenuated by frames until the attenuation coefficient reaches a given threshold after the number of frames of the wide frequency band signal extended from the narrow frequency band speech or audio signal reaches a given number of frames. The interval between the current frame of speech or audio signal and the speech or audio signal of a frame before the switching may be obtained according to the current frame of speech or audio signal and the speech or audio signal of the frame before the switching. For example, the number of frames of the narrow frequency band speech or audio signal may be recorded by using a counter, where the number of frames may be a predetermined value greater than or equal to 0.

[0089] Step 602: If the processed first high frequency band signal does not need to be attenuated, synthesize the processed first high frequency band signal and the first low frequency band signal into a wide frequency band signal.

[0090] Specifically, if it is determined that the processed first high frequency band signal does not need to be attenuated in step 601, the processed first high frequency band signal and the first low frequency band signal are directly synthesized into a wide frequency band signal.

[0091] Step 603: If the processed first high frequency band signal needs to be attenuated, judge whether the attenuation factor corresponding to the processed first high frequency band signal is greater than the threshold.

[0092] Specifically, the initial value of the attenuation factor is 1, and the threshold is greater than or equal to 0 and smaller than 1. If it is determined that the processed

first high frequency band signal needs to be attenuated in step 601, whether the attenuation factor corresponding to the processed first high frequency band signal is greater than a given threshold is judged in step 603.

[0093] Step 604: If the attenuation factor is not greater than the given threshold, multiply the processed first high frequency band signal by the threshold, and synthesize the product and the first low frequency band signal into the wide frequency band signal.

[0094] Specifically, if the attenuation factor is determined to be not greater than the given threshold in step 603, it indicates that the energy of the processed first high frequency band signal is already attenuated to a certain degree and that the processed first high frequency band signal may not cause negative impacts. In this case, this attenuation ratio may be kept. Then, the processed first high frequency band signal is multiplied by the threshold, and then the product and the first low frequency band signal are synthesized into a wide frequency band signal.

[0095] Step 605: If the attenuation factor is greater than the given threshold, multiply the processed first high frequency band signal by the attenuation factor, and synthesize the product and the first low frequency band signal into the wide frequency band signal.

[0096] Specifically, if the attenuation factor is greater than the given threshold in step 603, it indicates that the processed first high frequency band signal may cause poor listening at the attenuation factor and needs to be further attenuated until it reaches the given threshold. Then, the processed first high frequency band signal is multiplied by the attenuation factor, and then the product and the first low frequency band signal are synthesized into a wide frequency band signal.

[0097] Step 606: Modify the attenuation factor to decrease the attenuation factor.

[0098] Specifically, as the speech or audio signals are transmitted, the impact of the speech or audio signals before the switching on subsequent narrowband speech or audio signals gradually turns smaller, and the attenuation factor also turns smaller gradually.

[0099] Optionally, based on the preceding technical solution, when a switching from a narrow frequency band speech or audio signal a wide frequency band speech or audio signal occurs, an embodiment of obtaining the processed first high frequency band signal through step 201 includes the following steps, as shown in FIG. 7:

[0100] Step 701: Weight according to the set fourth weight 1 and the fourth weight 2 to calculate a processed first high frequency band signal. The fourth weight 1 refers to the weight value of the second high frequency band signal, and the fourth weight 2 refers to the weight value of the first high frequency band signal of the current frame of speech or audio signal.

[0101] Specifically, in the process of switching the narrow frequency band speech or audio signal to the wide frequency band speech or audio signal, because the high frequency band signal of the wide frequency band

speech or audio signal is not null but the high frequency band signal corresponding to the narrow frequency band speech or audio signal is null, the energy of the high frequency band signal of the wide frequency band speech or audio signal needs to be attenuated to ensure that the narrow frequency band speech or audio signal can be smoothly switched to the wide frequency band speech or audio signal. The product of the second high frequency band signal and the fourth weight 1 is added to the product of the first high frequency band signal and the fourth weight 2; the weighted value is the processed first high frequency band signal.

[0102] Step 702: Decrease the fourth weight 1 as per the third weight step, and increase the fourth weight 2 as per the third weight step until the fourth weight 1 is equal to 0. The sum of the fourth weight 1 and the fourth weight 2 is equal to 1.

[0103] Specifically, as the speech or audio signals are transmitted, the impact of the narrow frequency band speech or audio signals before the switching on subsequent wide frequency band speech or audio signals gradually turns smaller. Therefore, the fourth weight 1 gradually turns smaller, while the fourth weight 2 gradually turns larger until the fourth weight 1 is equal to 0 and the fourth weight 2 is equal to 1. That is, the transmitted speech or audio signals are always wide frequency band speech or audio signals.

[0104] Similarly, as shown in FIG. 8, another embodiment of obtaining the processed first high frequency band signal through step 201 may further include the following steps:

[0105] Step 801: Weight according to the set fifth weight 1 and the fifth weight 2 to calculate a processed first high frequency band signal. The fifth weight 1 is the weight value of a set fixed parameter, and the fifth weight 2 is the weight value of the first high frequency band signal of the current frame of speech or audio signal.

[0106] Specifically, because the first high frequency band signal of the narrow frequency band speech or audio signal is null, a fixed parameter may be set to replace the high frequency band signal of the narrow frequency band speech or audio signal, where the fixed parameter is a constant greater than or equal to 0 and smaller than the energy of the first high frequency band signal. The product of the fixed parameter and the fifth weight 1 is added to the product of the first high frequency band signal and the fifth weight 2; the weighted value is the processed first high frequency band signal.

[0107] Step 802: Decrease the fifth weight 1 as per the fourth weight step, and increase the fifth weight 2 as per the fourth weight step until the fifth weight 1 is equal to 0. The sum of the fifth weight 1 and the fifth weight 2 is equal to 1.

[0108] Specifically, as the speech or audio signals are transmitted, the impact of the narrow frequency band speech or audio signals before the switching on subsequent wide frequency band speech or audio signals gradually turns smaller. Therefore, the fifth weight 1 gradually

turns smaller, while the fifth weight 2 gradually turns larger until the fifth weight 1 is equal to 0 and the fifth weight 2 is equal to 1. That is, the transmitted speech or audio signals are always real wide frequency band speech or audio signals.

[0109] By using the method for switching speech or audio signals in this embodiment, in the process of switching a speech or audio signal from a narrow frequency band speech or audio signal to a wide frequency band speech or audio signal, the high frequency band signal of the wide frequency band speech or audio signal is attenuated to obtain a processed high frequency band signal. In this way, the high frequency band signal corresponding to the narrow frequency band speech or audio signal before the switching can be smoothly switched to the processed high frequency band signal corresponding to the wide frequency band speech or audio signal, thus helping to improve the quality of audio signals received by the user.

[0110] In this embodiment, the envelope information may also be replaced by other parameters that can represent the high frequency band signal, for example, a linear predictive coding (LPC) parameter or an amplitude parameter.

[0111] Those skilled in the art may understand that all or a part of the steps of the method according to the embodiments of the present invention may be implemented by a program instructing relevant hardware. The program may be stored in a computer readable storage medium. When the program runs, the steps of the method according to the embodiments of the present invention are performed. The storage medium may be a read only memory (ROM), a random access memory (RAM), a magnetic disk, or a compact disk-read only memory (CD-ROM).

[0112] FIG. 9 shows a structure of the first embodiment of an apparatus for switching speech or audio signals. As shown in FIG. 9, the apparatus for switching speech or audio signals includes a processing module 91 and a first synthesizing module 92.

[0113] The processing module 91 is adapted to weight the first high frequency band signal of the current frame of speech or audio signal and the second high frequency band signal of the previous M frame of speech or audio signals to obtain a processed first high frequency band signal. When a switching of a speech or audio occurs, M is greater than or equal to 1.

[0114] The first synthesizing module 92 is adapted to synthesize the processed first high frequency band signal and the first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal.

[0115] In the apparatus for switching speech or audio signals in this embodiment, the processing module processes the first high frequency band signal of the current frame of speech or audio signal according to the second high frequency band signal of the previous M frame of speech or audio signals, so that the second high frequen-

cy band signal can be smoothly switched to the processed first high frequency band signal. In this way, during the process of switching between speech or audio signals with different bandwidths, the high frequency band signal of these speech or audio signals can be smoothly switched. Finally, the first synthesizing module synthesizes the processed first high frequency band signal and the first low frequency band signal into a wide frequency band signal; the wide frequency band signal is transmitted to a user terminal, so that the user enjoys a high quality speech or audio signal. By using the method for switching speech or audio signals in this embodiment, speech or audio signals with different bandwidths can be switched smoothly, thus reducing the impact of the sudden energy change on the subjective audio quality of the speech or audio signals and improving the quality of audio signals received by the user.

[0116] FIG. 10 shows a structure of the second embodiment of the apparatus for switching speech or audio signals. As shown in FIG. 10, the apparatus for switching speech or audio signals in this embodiment is based on the first embodiment, and further includes a second synthesizing module 103.

[0117] The second synthesizing module 103 is adapted to synthesize the first high frequency band signal and the first low frequency band signal into the wide frequency band signal when a switching of the speech or audio signal does not occur.

[0118] In the apparatus for switching speech or audio signal in this embodiment, the second synthesizing module is set to synthesize the first low frequency band signal and the first high frequency band signal of the first frequency band speech or audio signals of the current frame into a wide frequency band signal when a switching between speech or audio signals with different bandwidths occurs. In this way, the quality of speech or audio signals received by the user is improved.

[0119] According to the preceding technical solution, optionally, when a switching from wide frequency band speech or audio signal to a narrow frequency band speech or audio signal occurs, the processing module 101 includes the following modules, as shown in FIG. 10 and FIG. 11:

a predicting module 1011, adapted to predict fine structure information and envelope information corresponding to the first high frequency band signal; a first generating module 1012, adapted to weight the predicted envelope information and the previous M frame envelope information corresponding to the second high frequency band signal of the previous M frame of speech or audio signals to obtain first envelope information corresponding to the first high frequency band signal; and a second generating module 1013, adapted to generate a processed first high frequency band signal according to the first envelope information and the predicted fine structure information.

[0120] Further, the apparatus for switching speech or audio signals in this embodiment may include a classifying module 1010 adapted to classify the first low frequency band signal of the current frame of speech or audio signal. The predicting module 1011 is further adapted to predict the fine structure information and envelope information corresponding to the first low frequency band signal of the current frame of speech or audio signal.

[0121] In the apparatus for switching speech or audio signals in this embodiment, the predicting module predicts the fine structure information and envelope information corresponding to the first high frequency band signal, so that the processed first high frequency band signal can be accurately generated by the first generating module and the second generating module. In this way, the first high frequency band signal can be smoothly switched to the processed first high frequency band signal, thus improving the quality of speech or audio signals received by the user. In addition, the classifying module classifies the first low frequency band signal of the current frame of speech or audio signal; the predicting module obtains the predicted fine structure information and predicted envelope information according to the signal type. In this way, the predicted fine structure information and predicted envelope information are more accurate, thus improving the quality of speech or audio signals received by the user.

[0122] Based on the preceding technical solution, optionally, the first synthesizing module 102 includes the following modules, as shown in FIG. 10 and FIG. 12:

a first judging module 1021, adapted to judge whether the processed first high frequency band signal needs to be attenuated according to the current frame of speech or audio signal and the previous frame of speech or audio signal before the switching; a third synthesizing module 1022, adapted to synthesize the processed first high frequency band signal and the first low frequency band signal into a wide frequency band signal when the first judging module 1021 determines that the processed first high frequency band signal does not need to be attenuated;

a second judging module 1023, adapted to judge whether the attenuation factor corresponding to the processed first high frequency band signal is greater than the given threshold when the first judging module 1021 determines that the processed first high frequency band signal needs to be attenuated;

a fourth synthesizing module 1024, adapted to: if the second judging module 1023 determines that the attenuation factor is not greater than the given threshold, multiply the processed first high frequency band signal by the threshold, and synthesize the product and the first low frequency band signal into a wide frequency band signal;

a fifth synthesizing module 1025, adapted to: if the

second judging module 1023 determines that the attenuation factor is greater than the given threshold, multiply the processed first high frequency band signal by the attenuation factor, and synthesize the product and the first low frequency band signal into a wide frequency band signal; and a first modifying module 1026, adapted to modify the attenuation factor to decrease the attenuation factor.

[0123] The initial value of the attenuation factor is 1, and the threshold is greater than or equal to 0 and smaller than 1.

[0124] By using the apparatus for switching speech or audio signals, the processed first high frequency band signal is attenuated, so that the wide frequency band signal obtained by processing the current frame of speech or audio signal is more accurate, thus improving the quality of audio signals received by the user.

[0125] According to the preceding technical solution, optionally, when a switching from a narrow frequency band speech or audio signal a wide frequency band speech or audio signal occurs, the processing module 101 in this embodiment includes the following modules, as shown in FIG. 10 and FIG. 13a:

a first calculating module 1011a, adapted to weight according to a set fourth weight 1 and a fourth weight 2 to calculate the processed first high frequency band signal, where the fourth weight 1 refers to the weight value of the second high frequency band signal and the fourth weight 2 refers to the weight value of the first high frequency band signal; and a second modifying module 1012a, adapted to: decrease the fourth weight 1 as per the third weight step, and increase the fourth weight 2 as per the third weight step until the fourth weight 1 is equal to 0, where the sum of the fourth weight 1 and the fourth weight 2 is equal to 1.

[0126] Similarly, when a switching from a narrow frequency band speech or audio signal a wide frequency band speech or audio signal occurs, the processing module 101 in this embodiment may further include the following modules, as shown in FIG. 10 and FIG. 13b:

a second calculating module 1011b, adapted to weight according to a set fifth weight 1 and a fifth weight 2 to calculate the processed first high frequency band signal, where the fifth weight 1 refers to the weight value of a set fixed parameter, and the fifth weight 2 refers to the weight value of the first high frequency band signal; and a third modifying module 1012b, adapted to: decrease the fifth weight 1 as per the fourth weight step, and increase the fifth weight 2 as per the fourth weight step until the fifth weight 1 is equal to 0, where the sum of the fifth weight 1 and the fifth weight 2 is equal to 1, where the fixed parameter is a fixed con-

stant greater than or equal to 0 and smaller than the energy value of the first high frequency band signal.

[0127] By using the apparatus for switching speech or audio signals in this embodiment, in the process of switching a speech or audio signal from a narrow frequency band speech or audio signal to a wide frequency band speech or audio signal, the high frequency band signal of the wide frequency band speech or audio signal is attenuated to obtain a processed high frequency band signal. In this way, the high frequency band signal corresponding to the narrow frequency band speech or audio signal before the switching can be smoothly switched to the processed high frequency band signal corresponding to the wide frequency band speech or audio signal, thus helping to improve the quality of audio signals received by the user.

[0128] It should be noted that the above embodiments are merely provided for describing the technical solution of the present invention, but not intended to limit the present invention. Although the present invention has been described in detail with reference to the foregoing embodiments, it is apparent that those skilled in the art can make various modifications and variations to the invention without departing from the spirit and scope of the invention. The invention shall cover the modifications and variations provided that they fall in the scope of protection defined by the following claims or their equivalents.

Claims

1. A method for switching speech or audio signals, comprising:

when a switching of a speech or audio occurs, weighting a first high frequency band signal of a current frame of speech or audio signal and a second high frequency band signal of previous M frame of speech or audio signals to obtain a processed first high frequency band signal, wherein M is greater than or equal to 1; and synthesizing the processed first high frequency band signal and a first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal.

2. The method of claim 1, further comprising:

when a switching of the speech or audio signal does not occur, synthesizing the first high frequency band signal and the first low frequency band signal into the wide frequency band signal.

3. The method of claim 1 or 2, wherein when a switching from a wide frequency band speech or audio signal to a narrow frequency band speech or audio signal occurs, the step of weighting the first high frequency

band signal of the current frame of speech or audio signal and the second high frequency band signal of the previous M frame of speech or audio signals to obtain the processed first high frequency band signal comprises:

predicting the fine structure information and the envelope information corresponding to the first high frequency band signal of the current frame of speech or audio signal;
weighting the predicted envelope information and previous M frame envelope information corresponding to the second high frequency band signal of the previous M frame of speech or audio signals to obtain first envelope information corresponding to the first high frequency band signal; and
generating the processed first high frequency band signal according to the first envelope information and the predicted fine structure information.

4. The method of claim 3, wherein the step of predicting the fine structure information and envelope information corresponding to the first high frequency band signal of the current frame of speech or audio signal comprises:

classifying the first low frequency band signal of the current frame of speech or audio signal; and predicting the fine structure information and envelope information according to the signal type of the first low frequency band signal.

5. The method of claim 3, wherein the step of weighting the predicted envelope information and the previous M frame envelope information corresponding to the second high frequency band signal of the previous M frame of speech or audio signals to obtain the first envelope information corresponding to the first high frequency band signal comprises:

calculating a correlation coefficient between the first low frequency band signal and a low frequency band signal of previous N frame of speech or audio signals according to the first low frequency band signal and the low frequency band signal of the previous N frame of speech or audio signals, wherein N is greater than or equal to 1;
judging whether the correlation coefficient is within a given first threshold range;
if the correlation coefficient is not within the first threshold range, weighting according to a set first weight 1 and a set first weight 2 to calculate the first envelope information, wherein the first weight 1 refers to a weight value of previous frame envelope information corresponding to a

high frequency band signal of a previous frame of speech or audio signal and the first weight 2 refers to a weight value of the envelope information;

if the correlation coefficient is within the first threshold range, weighting according to a set second weight 1 and a set second weight 2 to calculate transitional envelope information, wherein the second weight 1 refers to a weight value of envelope information corresponding to a high frequency band signal of L frame of speech or audio signals before the switching and the second weight 2 refers to the weight value of the previous M frame envelope information, wherein L is greater than or equal to 1; decreasing the second weight 1 as per a first weight step, and increasing the second weight 2 as per the first weight step; judging whether a set third weight 1 is greater than the first weight 1; if the third weight 1 is not greater than the first weight 1, weighting according to the set first weight 1 and the first weight 2 to calculate the first envelope information; if the third weight 1 is greater than the first weight 1, weighting according to the set third weight 1 and a third weight 2 to calculate the first envelope information, wherein the third weight 1 refers to a weight value of the transitional envelope information and the third weight 2 refers to a weight value of the predicted envelope information; and decreasing the third weight 1 as per a second weight step, and increasing the third weight 2 as per the second weight step until the third weight 1 is equal to 0; wherein:

a sum of the first weight 1 and the first weight 2 is equal to 1; a sum of the second weight 1 and the second weight 2 is equal to 1; a sum of the third weight 1 and the third weight 2 is equal to 1; an initial value of the third weight 1 is greater than an initial value of the first weight 1; and the first weight 1 and the first weight 2 are fixed constants.

6. The method of claim 3, wherein the step of weighting the predicted envelope information and the previous M frame envelope information corresponding to the second high frequency band signal of the previous M frame of speech or audio signals to obtain the first envelope information corresponding to the first high frequency band signal comprises:

calculating a correlation coefficient between the first low frequency band signal of a current frame and a low frequency band signal of a previous frame of speech or audio signal according to the

first low frequency band signal of the current frame and the low frequency band signal of the previous frame of speech or audio signal; judging whether the correlation coefficient is within a given second threshold range; if the correlation coefficient is not within the second threshold range, weighting according to a set first weight 1 and a set first weight 2 to calculate the first envelope information, wherein the first weight 1 refers to a weight value of previous frame envelope information corresponding to a high frequency band signal of the previous frame of speech or audio signal and the first weight 2 refers to a weight value of the predicted envelope information; and the first weight 1 and the first weight 2 are fixed constants; if the correlation coefficient is within the second threshold range, judging whether a set second weight 1 is greater than the first weight 1, wherein the second weight 1 refers to a weight value of envelope information corresponding to the high frequency band signal of the previous frame of speech or audio signal before the switching; if the second weight 1 is not greater than the first weight 1, weighting according to the set first weight 1 and the first weight 2 to calculate the first envelope information; if the second weight 1 is greater than the first weight 1, weighting according to the second weight 1 and a set second weight 2 to calculate the first envelope information, wherein the second weight 2 refers to a weight value of the predicted envelope information; and decreasing the second weight 1 as per a second weight step, and increasing the second weight 2 as per the second weight step; wherein:

a sum of the first weight 1 and the first weight 2 is equal to 1; a sum of the second weight 1 and the second weight 2 is equal to 1; an initial value of the second weight 1 is greater than an initial value of the first weight 1.

7. The method of claim 3, wherein the step of synthesizing the processed first high frequency band signal and the first low frequency band signal of the current frame of speech or audio signal into the wide frequency band signal comprises:

judging whether the processed first high frequency band signal needs to be attenuated according to the current frame of speech or audio signal and a previous frame of speech or audio signal before the switching; if attenuation is not required, synthesizing the processed first high frequency band signal and the first low frequency band signal into the wide frequency band signal;

if attenuation is required, judging whether an attenuation factor corresponding to the first high frequency band signal is greater than a given threshold;

if the attenuation factor is not greater than the given threshold, multiplying the processed first high frequency band signal by the threshold, and synthesizing the product of the processed first high frequency band signal and the threshold and the first low frequency band signal into the wide frequency band signal;

if the attenuation factor is greater than the given threshold, multiplying the processed first high frequency band signal by the attenuation factor, and synthesizing the product of the processed first high frequency band signal and the attenuation factor and the first low frequency band signal into the wide frequency band signal; and modifying the attenuation factor to decrease the attenuation factor; wherein:

an initial value of the attenuation factor is 1, and the threshold is greater than or equal to 0 and smaller than 1.

8. The method of claim 1 or 2, wherein a switching from a narrow frequency band speech or audio signal to a wide frequency band speech or audio signal occurs, the step of weighting the first high frequency band signal of the current frame of speech or audio signal and the second high frequency band signal of the previous M frame of speech or audio signals to obtain the processed first high frequency band signal comprises:

weighting according to a set fourth weight 1 and a set fourth weight 2 to calculate the processed first high frequency band signal, wherein the fourth weight 1 refers to a weight value of the second high frequency band signal and the fourth weight 2 refers to a weight value of the first high frequency band signal; and decreasing the fourth weight 1 as per a third weight step, and increasing the fourth weight 2 as per the third weight step until the fourth weight 1 is equal to 0, wherein a sum of the fourth weight 1 and the fourth weight 2 is equal to 1.

9. The method of claim 1 or 2, wherein when a switching from a wide frequency band speech or audio signal to a narrow frequency band speech or audio signal occurs" the step of weighting the first high frequency band signal of the current frame of speech or audio signal and the second high frequency band signal of the previous M frame of speech or audio signals to obtain the processed first high frequency band signal comprises:

weighting according to a set fifth weight 1 and a set fifth weight 2 to calculate the processed first high frequency band signal, wherein the fifth weight 1 refers to a weight value of a set fixed parameter, and the fifth weight 2 refers to a weight value of the first high frequency band signal; and

reducing the fifth weight 1 as per a fourth weight step, and increasing the fifth weight 2 as per the fourth weight step until the fifth weight 1 is equal to 0, wherein a sum of the fifth weight 1 and the fifth weight 2 is equal to 1; wherein:

the fixed parameter is a constant greater than or equal to 0 and smaller than an energy value of the first high frequency band signal.

10. An apparatus for switching speech or audio signals, comprising:

a processing module, adapted to: when a switching of a speech or audio occurs, weight a first high frequency band signal of a current frame of speech or audio signal and a second high frequency band signal of previous M frame of speech or audio signals to obtain a processed first high frequency band signal, wherein M is greater than or equal to 1; and a first synthesizing module, adapted to synthesize the processed first high frequency band signal and a first low frequency band signal of the current frame of speech or audio signal into a wide frequency band signal.

11. The apparatus of claim 10, further comprising:

a second synthesizing module, adapted to synthesize the first high frequency band signal and the first low frequency band signal into the wide frequency band signal when a switching of the speech or audio signal does not occur.

12. The apparatus of claim 10 or 11, wherein when a switching from a wide frequency band speech or audio signal to a narrow frequency band speech or audio signal occurs, the processing module comprises:

a predicting module, adapted to predict the fine structure information and the envelope information corresponding to the first high frequency band signal of the current frame of speech or audio signal; a first generating module, adapted to weight the predicted envelope information and previous M frame envelope information corresponding to the second high frequency band signal of the previous M frame of speech or audio signals to

obtain first envelope information corresponding to the first high frequency band signal; and a second generating module, adapted to generate the processed first high frequency band signal according to the first envelope information and the predicted fine structure information.

13. The apparatus of claim 12, further comprising a classifying module adapted to classify the first low frequency band signal of the current frame of speech or audio signal, wherein:

the predicting module is further adapted to predict the fine structure information and the envelope information according to the signal type of the first low frequency band signal.

14. The apparatus of claim 12, wherein the first synthesizing module comprises:

a first judging module, adapted to judge whether the processed first high frequency band signal needs to be attenuated according to the current frame of speech or audio signal and a previous frame of speech or audio signal before the switching;

a third synthesizing module, adapted to synthesize the processed first high frequency band signal and the first low frequency band signal into the wide frequency band signal when the first judging module determines that the processed first high frequency band signal does not need to be attenuated;

a second judging module, adapted to judge whether an attenuation factor corresponding to the processed first high frequency band signal is greater than a given threshold when the first judging module determines that the processed first high frequency band signal needs to be attenuated;

a fourth synthesizing module, adapted to: if the second judging module determines that the attenuation factor is not greater than the given threshold, multiply the processed first high frequency band signal by the threshold, and synthesize the product and the first low frequency band signal into the wide frequency band signal; a fifth synthesizing module, adapted to: if the second judging module determines that the attenuation factor is greater than the given threshold, multiply the processed first high frequency band signal by the attenuation factor, and synthesize the product and the first low frequency band signal into the wide frequency band signal; and

a first modifying module, adapted to modify the attenuation factor to decrease the attenuation factor; wherein:

an initial value of the attenuation factor is 1, and the threshold is greater than or equal to 0 and smaller than 1.

15. The apparatus of claim 10 or 11, wherein, when a switching from a narrow frequency band speech or audio signal to a wide frequency band speech or audio signal occurs, the processing module comprises:

a first calculating module, adapted to weight according to a set fourth weight 1 and a set fourth weight 2 to calculate the processed first high frequency band signal, wherein the fourth weight 1 refers to a weight value of the second high frequency band signal and the fourth weight 2 refers to a weight value of the first high frequency band signal; and

a second modifying module, adapted to: decrease the fourth weight 1 as per a third weight step, and increase the fourth weight 2 as per the third weight step until the fourth weight 1 is equal to 0, wherein a sum of the fourth weight 1 and the fourth weight 2 is equal to 1.

16. The apparatus of claim 13, wherein, when a switching from a narrow frequency band speech or audio signal to a wide frequency band speech or audio signal occurs, the processing module comprises:

a second calculating module, adapted to weight according to a set fifth weight 1 and a set fifth weight 2 to calculate the processed first high frequency band signal, wherein the fifth weight 1 refers to a weight value of a set fixed parameter and the fifth weight 2 refers to a weight value of the first high frequency band signal; and a third modifying module, adapted to: decrease the fifth weight 1 as per a fourth weight step, and increase the fifth weight 2 as per the fourth weight step until the fifth weight 1 is equal to 0, wherein a sum of the fifth weight 1 and the fifth weight 2 is equal to 1, wherein the fixed parameter is a constant greater than or equal to 0 and smaller than an energy value of the first high frequency band signal.

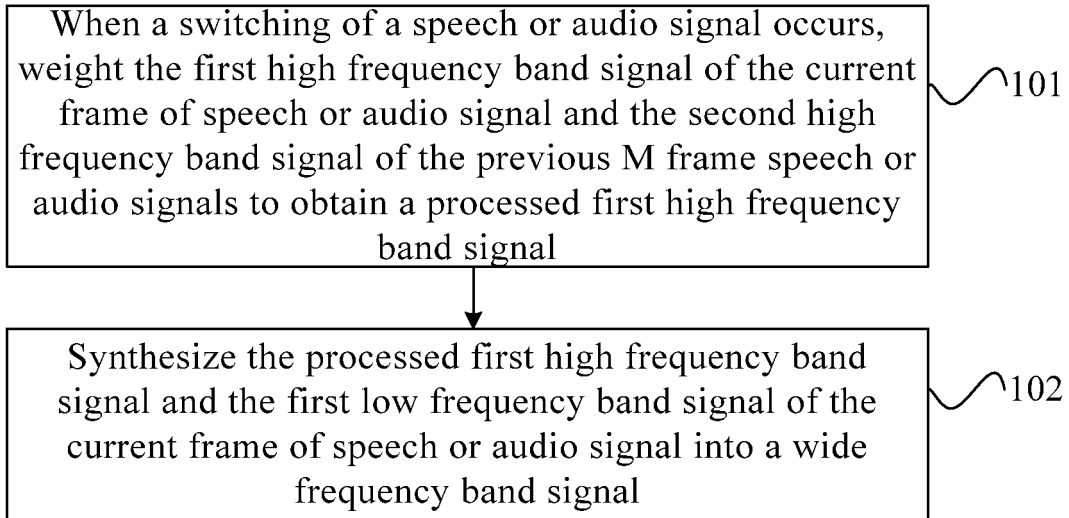


FIG. 1

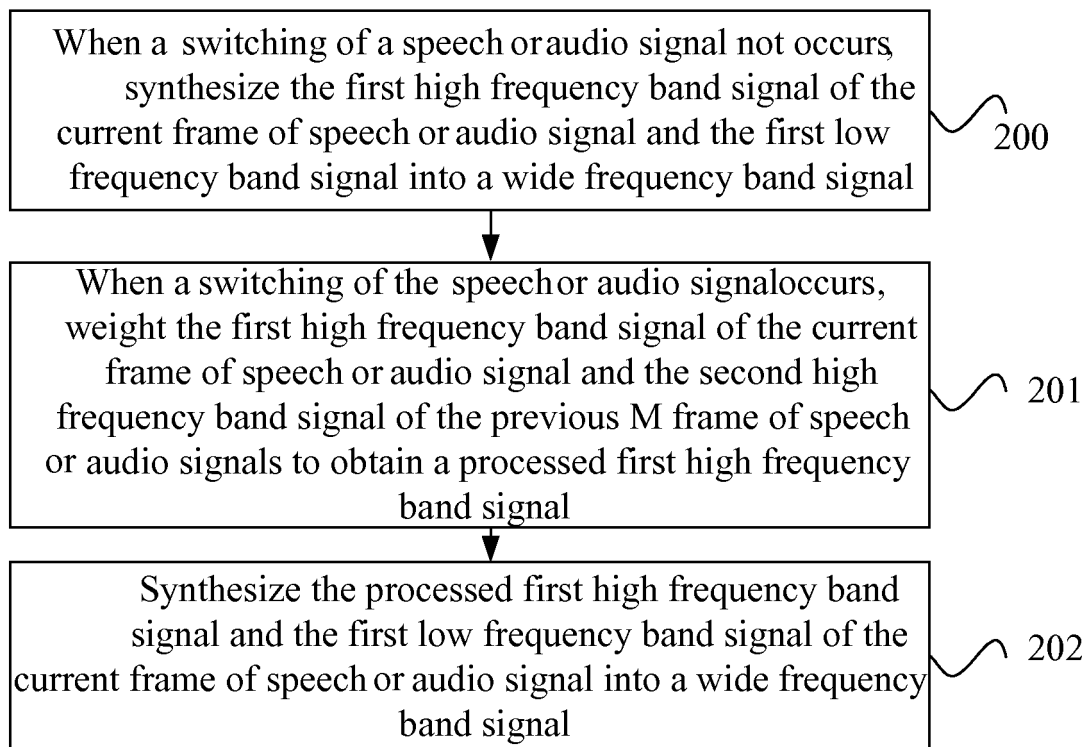


FIG. 2

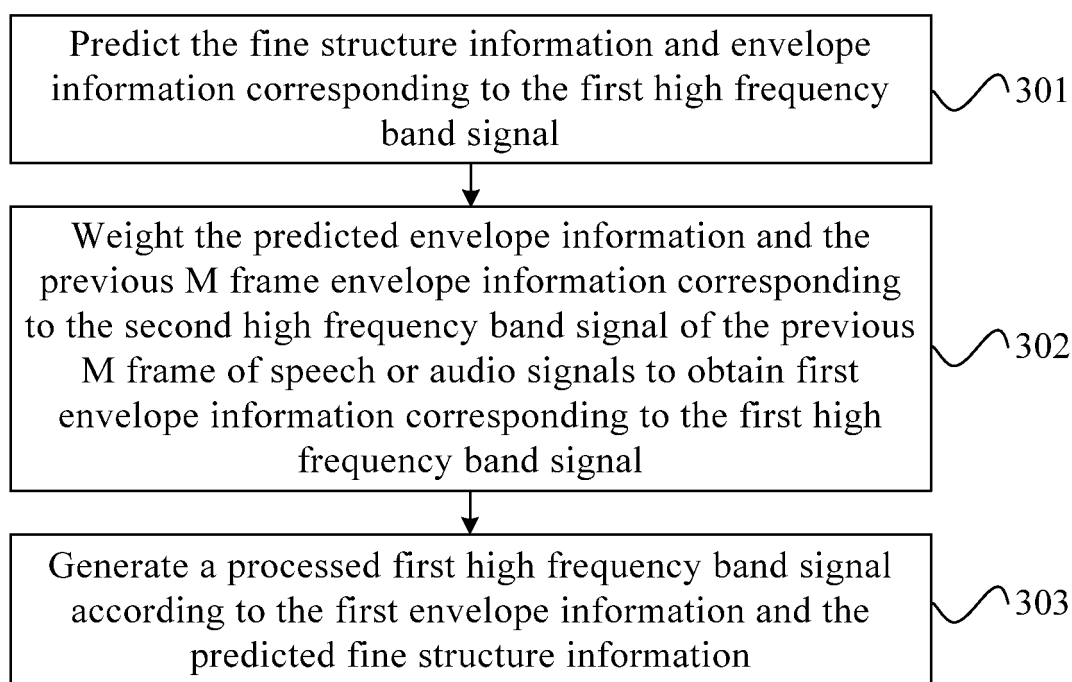


FIG. 3

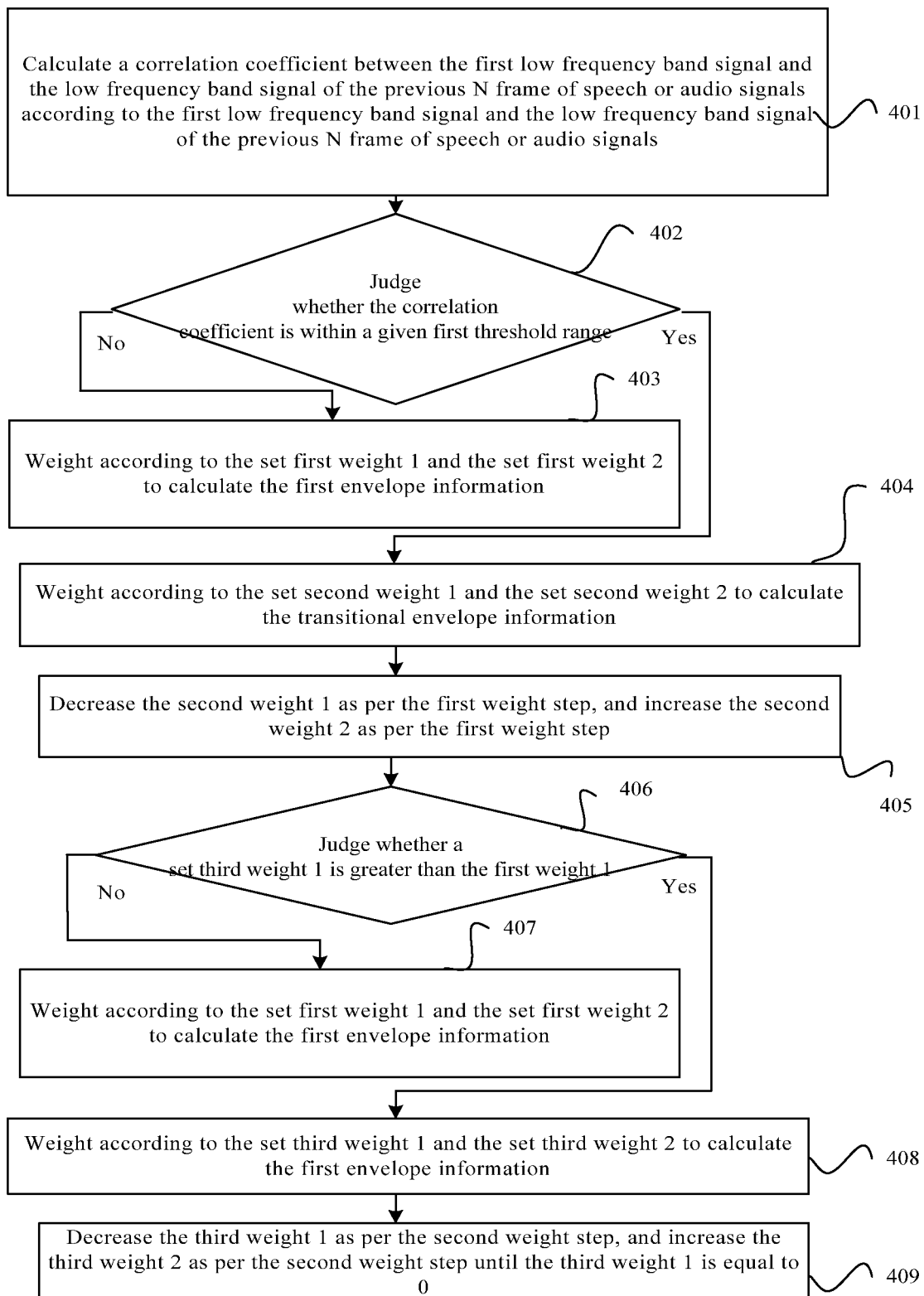


FIG. 4

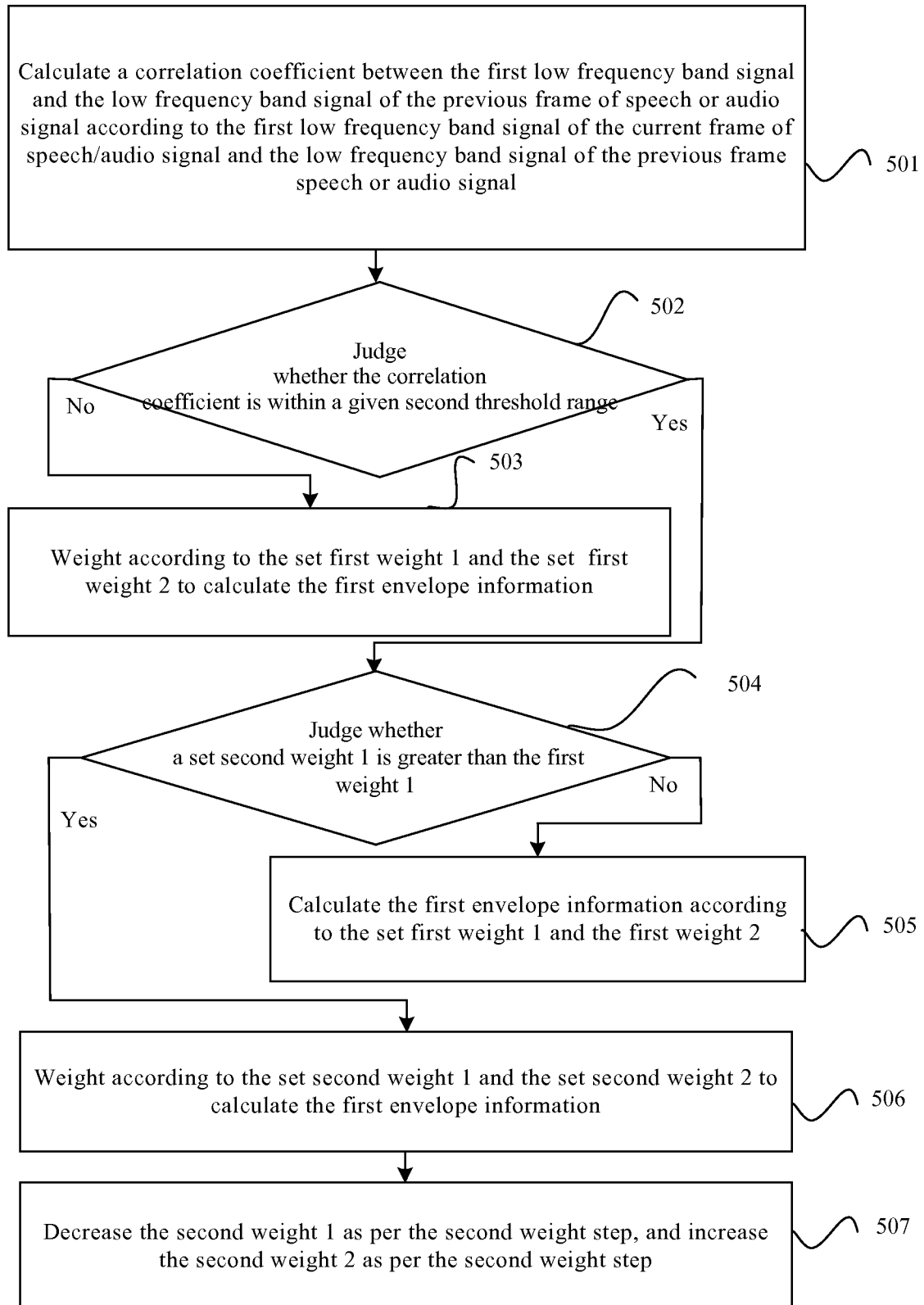


FIG. 5

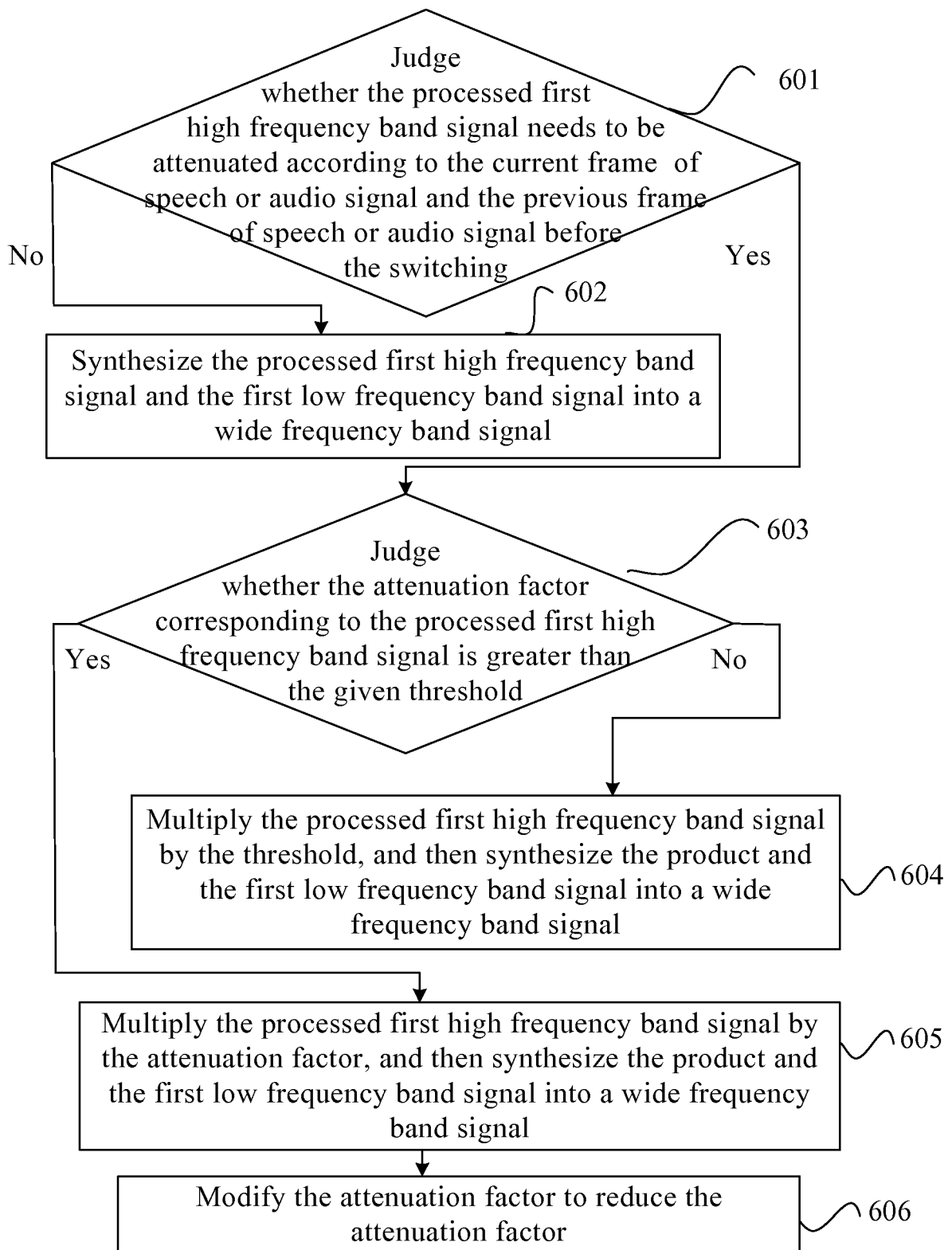


FIG. 6

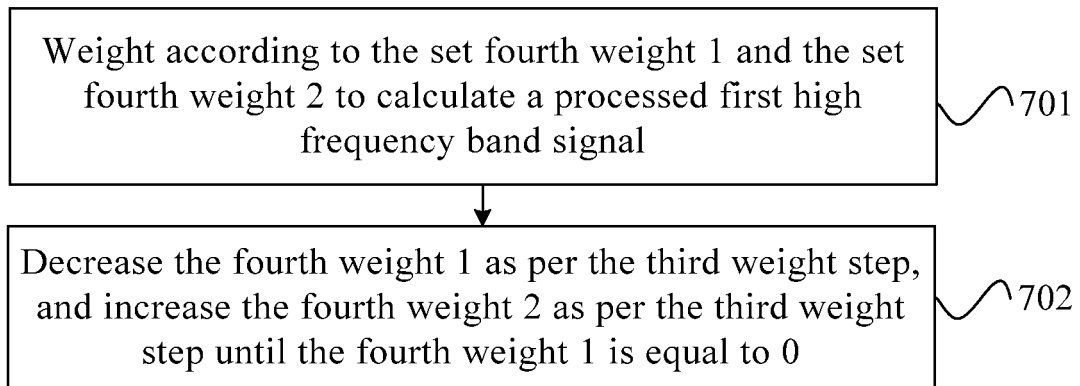


FIG. 7

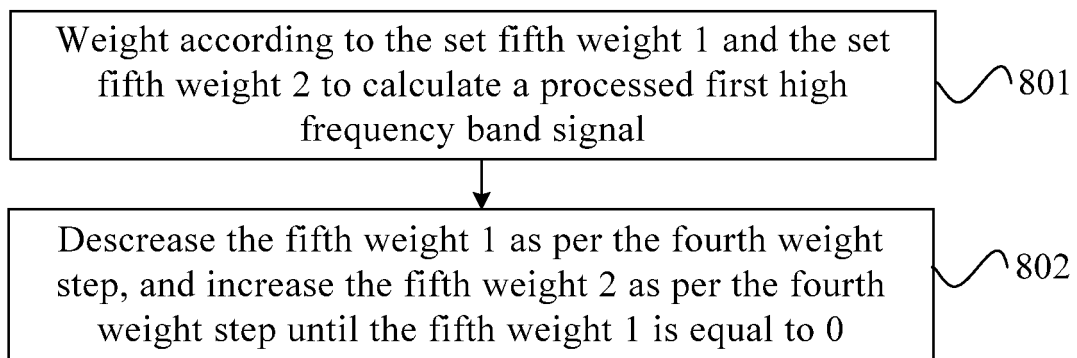


FIG. 8

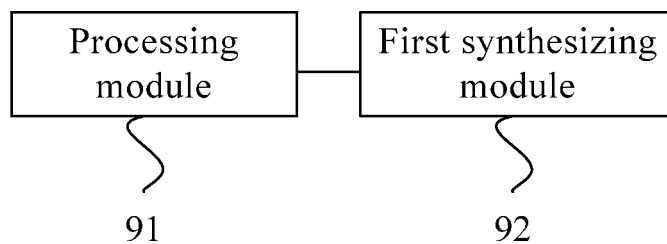


FIG. 9

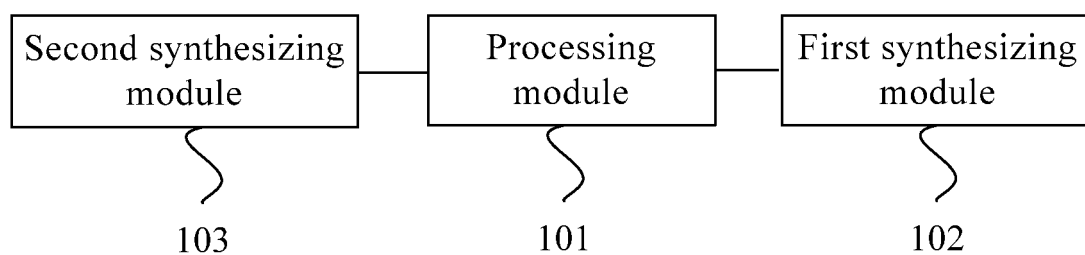


FIG. 10

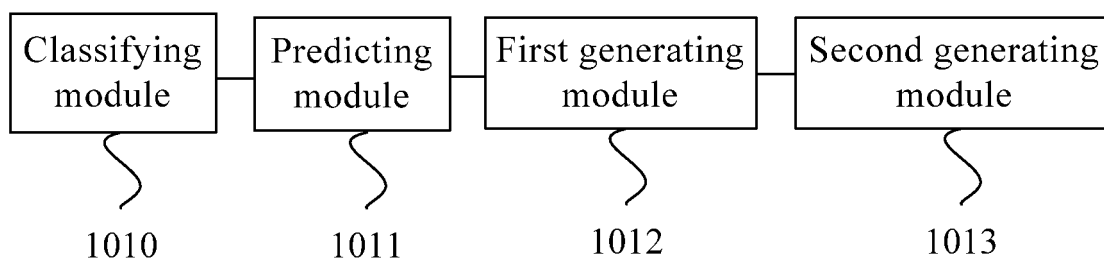


FIG. 11

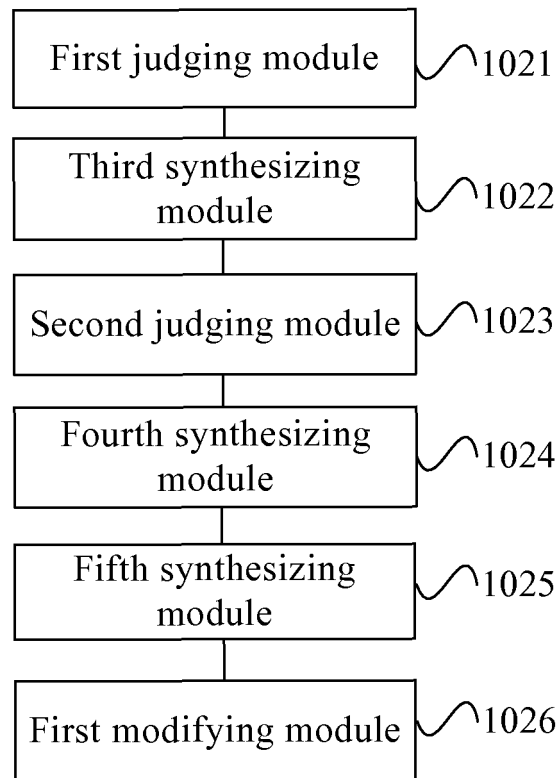


FIG. 12

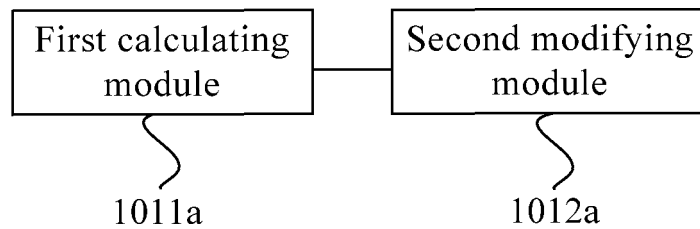


FIG. 13a

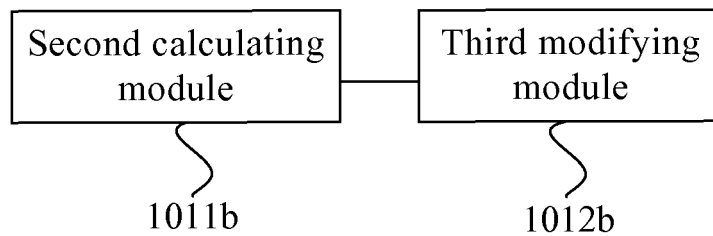


FIG. 13b

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2011/073479

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNA BS, VEN: Narrow, band, wide, frequency, audio, speech, signal?, switch+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN101964189A(HUAWEI TECHNOLOGIES CO. LTD.),02 Feb. 2011(02.02.2011), claims 1-16	1-16
X	CN101335002A(HUAWEI TECHNOLOGIES CO. LTD.), 31 Dec.2008(31.12.2008), Abstract, pages 4-12 and 31-32 of the description, figures 3-13	1-2, 10-11
A		3-9, 12-16
A	CN101425292A(HUAWEI TECHNOLOGIES CO. LTD.),06 May 2009(06.05.2009), the whole document	1-16
A	US2005027516A1(SAMSUNG ELECTRONICS CO., LTD.),03 Feb. 2005(03.02.2005), the whole document	1-16

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&”document member of the same patent family

Date of the actual completion of the international search
15 Jul. 2011(15.07.2011)Date of mailing of the international search report
04 Aug. 2011 (04.08.2011)Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451

Authorized officer

ZHANG Yaling

Telephone No. (86-10)62085722

INTERNATIONAL SEARCH REPORT
 Information on patent family members

International application No.

PCT/CN2011/073479

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN101964189A	02.02.2011	none	
CN101335002A	31.12.2008	CN100585699C	27.01.2010
		EP2207166 A1	14.07.2010
		WO2009056027 A1	07.05.2009
		US2010228557 A1	09.09.2010
		KR20100085991 A	29.07.2010
		JP2011502287T	20.01.2011
CN101425292A	06.05.2009	none	
US2005027516A1	03.02.2005	EP1498874B1	07.06.2006
		EP1498874 A1	19.01.2005
		JP2005037949 A	10.02.2005
		DE602004001101T2	14.06.2007
		DE602004001101 D1	20.07.2006
		KR20050009384 A	25.01.2005
		KR940531 B1	10.02.2010

Form PCT/ISA /210 (patent family annex) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2011/073479

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/00(2006.01)i

G10L 19/04(2006.01)i

G10L 19/12(2006.01)i

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- CN 201010163406 [0001]