



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
17.03.2021 Bulletin 2021/11

(51) Int Cl.:
G10L 21/0364^(2013.01) G10L 21/0208^(2013.01)

(21) Application number: **20195635.6**

(22) Date of filing: **11.09.2020**

(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
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(72) Inventors:
• **HE, Zhihui**
Xiamen, Fujian 361009 (CN)
• **LIN, Lifeng**
Xiamen, Fujian 361009 (CN)
• **KANG, Yuanxun**
Xiamen, Fujian 361009 (CN)

(30) Priority: **11.09.2019 CN 201910860075**

(74) Representative: **Michalski Hüttermann & Partner**
Patentanwälte mbB
Speditionstraße 21
40221 Düsseldorf (DE)

(71) Applicant: **Yealink (Xiamen) Network Technology Co., Ltd.**
Xiamen, Fujian 361009 (CN)

(54) **DIGITAL AUTOMATIC GAIN CONTROL METHOD AND APPARATUS**

(57) A digital automatic gain control (AGC) method and apparatus are disclosed in this application. A specific implementation of the method includes: calculating a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model; performing a speech envelope estimation on the signal data to obtain a speech envelope amplitude; calculating a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude; classifying the signal data based on the speech probability P_n , and counting the number of noise occurrences in the signal data based on a classification result, to calculate a second gain; and performing signal gain amplitude adjustment on the signal data by using the first gain and the second gain. This implementation helps achieve an overall amplification effect and reduce distortion caused by automatic gain processing. Moreover, in the case of user change, the signal gain amplitude can also be adjusted quickly in response to different users. Meanwhile, in the case of excessive ambient noise, the apparatus can also adjust the gain automatically, to reduce amplification of the noise.

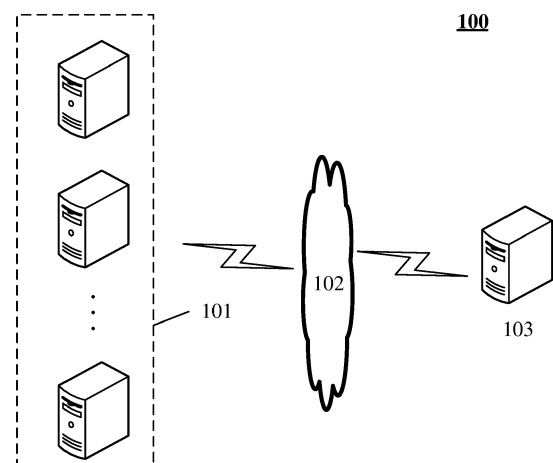


FIG. 1

Description**TECHNICAL FIELD**

5 **[0001]** This application relates to the field of signal processing, and in particular, to a digital automatic gain control (AGC) method and apparatus.

BACKGROUND

10 **[0002]** In use of a conventional video conference system, speech data picked up by a microphone depends on a distance between a speaker and the microphone and a volume of the speaker. If the speech data is directly transmitted, sound heard by a receiver is unstable, which severely affects call quality. A main solution to this problem is to control a transmit volume by using an AGC algorithm, so that signal amplitudes of input signals with different amplitudes are adjusted by using different gains. For example, a greater gain is used for a weak signal, and a smaller gain is used for a strong signal, so that amplitudes of transmitted signals are finally stabilized within a certain range.

15 **[0003]** In existing digital automatic control processing solutions at present, a gain is generally calculated according to an input amplitude of a module and an anticipated amplitude, and then an output amplitude of a signal is adjusted. Since amplitudes of speech signals change quickly, instead of directly using a current input signal as an input of an AGC module, an envelope of the input signal is generally used as the input of the module. A common envelope is obtained by smoothing a peak signal. Because the envelope changes with the speech signal, the calculated gain also changes correspondingly, which inevitably results in distortion of the processed speed signal. Moreover, an envelope value of a non-speech signal is relatively small, and if the gain calculated accordingly is excessively large, it also increases output noise. In addition, in some solutions, a user is identified by using a voiceprint recognition technology. When an apparatus recognizes that a user speaks for the first time, a steady-state gain is calculated by using standard AGC and stored into a memory. When the apparatus recognizes that the user has a historical use record, a historical gain may be directly retrieved from the memory. This method needs to use voiceprint recognition for calculation, which results in a great amount of calculation, and the method is easily affected by the accuracy of voiceprint recognition.

SUMMARY

30 **[0004]** An objective of this application is to provide an improved digital AGC method and apparatus, to solve the technical problem mentioned in the background.

[0005] According to a first aspect, the present invention provides a digital AGC method, including steps of: calculating a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model; performing a speech envelope estimation on the signal data to obtain a speech envelope amplitude; calculating a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude; classifying the signal data based on the speech probability P_n , and counting the number of noise occurrences in the signal data based on a classification result, to calculate a second gain; and performing signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

40 **[0006]** In some embodiments, the step of performing a speech envelope estimation on the signal data to obtain a speech envelope amplitude specifically includes: performing speech envelope estimation processing on each segment of signal data respectively, a calculation formula of the speech envelope estimation being as follows:

$$45 \quad A[n] = \begin{cases} A[n-1] + \alpha[n] * (E[n] - A[n-1]) & P_n > P_{T1} \\ A[n-1] & \text{others} \end{cases}$$

where $A[n]$ is a speech envelope estimation in n^{th} processing, $E[n]$ is a maximum value of the piece of signal data, α is an update coefficient, and P_{T1} is a speech update threshold;

50 a calculation formula of the update coefficient α is as follows:

$$\alpha[n] = \begin{cases} \alpha[n-1] + \Delta & P_n < P_{T2} \\ \alpha[n-1] - \Delta & \text{others} \end{cases}$$

55 where Δ is an adjustment amount of the update coefficient, and P_{T2} is an update threshold value; and after the update coefficient α is obtained, restricting a maximum value and a minimum value of the update coefficient:

$$\alpha[n] = \begin{cases} 1 & \alpha[n] > 1 \\ \alpha_0 & \alpha[n] < \alpha_0 \end{cases}$$

where α_0 is a lower-limit threshold, and $\alpha_0 < 1$.

[0007] In some embodiments, the step of calculating a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude specifically includes: calculating the first gain according to the deviation between the speech envelope amplitude and the anticipated amplitude, a calculation formula of the first gain being as follows:

$$g_1 = \frac{A_T}{A[n]}$$

where A_T is the anticipated amplitude.

[0008] In some embodiments, the step of classifying the signal data based on the speech probability P_n , and counting the number of noise occurrences in the signal data based on a classification result, to calculate a second gain includes: classifying the signal data according to the speech probability P_n , a classification method being as follows:

$$T[n] = \begin{cases} 1 & P_n > P_{T3} \\ 0 & \text{others} \end{cases}$$

where P_{T3} is a classification threshold, $T[n]$ is the classification result, 1 represents a speech, and 0 represents noise; and

[0009] calculating the second gain according to the number c of noise occurrences in the speech data counted based on the classification result:

$$g_2 = \begin{cases} 1 & c < c_1 \\ g_{min} & c > c_2 \end{cases}$$

where c_1 and c_2 are thresholds, and $c_1 < c_2$; g_{min} is a minimum gain and $g_{min} < 1$.

[0010] In some embodiments, the step of performing signal gain amplitude adjustment on the signal data by using the first gain and the second gain includes: performing the signal gain amplitude adjustment on the signal data based on the first gain and the second gain, a formula of the signal gain amplitude adjustment being as follows:

$$x_o = g_1 * g_2 * x$$

where x is the signal data, and x_o is an output of the gain amplitude adjustment.

[0011] In some embodiments, the digital AGC method further includes: dynamically compressing the signal data by using a dynamic compressor, a formula of the dynamic compression being as follows:

$$y = \begin{cases} x_o & x_o < T - W/2 \\ T + \frac{(x_o - T)}{R} & x_o > T + W/2 \\ x_o + \frac{(\frac{1}{R} - 1)(x_o - T + \frac{W}{2})^2}{2W} & T - \frac{W}{2} \leq x_o \leq T + \frac{W}{2} \end{cases}$$

where T is a compression threshold, R is a compression coefficient, W is a compression transition range, x_o is a signal after gain adjustment, and y is an output signal.

[0012] According to a second aspect, the present invention provides a digital AGC apparatus, including: a speech probability calculation module, configured to calculate a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model; a speech envelope estimation module, configured to perform a speech envelope estimation on the signal data to obtain a speech envelope amplitude; a first gain calculation module, configured to calculate a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude; a second gain calculation module, configured to classify the signal data based on the speech

probability P_n , and count the number of noise occurrences in the signal data based on a classification result, to calculate a second gain; and a signal gain amplitude adjustment module, configured to perform signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

[0013] In some embodiments, the digital AGC apparatus further includes: a signal data classification module, configured to classify the signal data according to the speech probability P_n , a classification method being as follows:

$$T[n] = \begin{cases} 1 & P_n > P_{T3} \\ 0 & \text{others} \end{cases}$$

where P_{T3} is a classification threshold, $T[n]$ is the classification result, 1 represents a speech, and 0 represents noise.

[0014] In some embodiments, the digital AGC apparatus further includes: a dynamic compression module, configured to dynamically compress the signal data by using a dynamic compressor, a formula of the dynamic compression being as follows:

$$y = \begin{cases} x_o & x_o < T - W/2 \\ T + \frac{(x_o - T)}{R} & x_o > T + W/2 \\ x_o + \frac{(\frac{1}{R} - 1)(x_o - T + \frac{W}{2})^2}{2W} & T - \frac{W}{2} \leq x_o \leq T + \frac{W}{2} \end{cases}$$

where T is a compression threshold, R is a compression coefficient, W is a compression transition range, x_o is a signal after gain adjustment, and y is an output signal.

[0015] According to a third aspect, the present invention provides a computer readable storage medium, storing a computer program, where the program implements the method of the first aspect when being executed by a processor.

[0016] This application provides a digital AGC method and apparatus. In the method, a first gain and a second gain are calculated for signal data through speech probability calculation, speech envelope estimation and signal data classification, thereby adjusting a signal gain amplitude, which helps achieve an overall amplification effect and reduce distortion caused by automatic gain processing. Moreover, in the case of user change, the signal gain amplitude can also be adjusted quickly in response to different users. Meanwhile, in the case of excessive ambient noise, the apparatus can also adjust the gain automatically, to reduce amplification of the noise.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Other features, objectives and advantages of this application will become more apparent upon reading the detailed description of the non-restrictive embodiments with reference to the following accompanying drawings.

FIG. 1 is a diagram of an exemplary system architecture to which this application can be applied;

FIG. 2 is a flowchart of an embodiment of a digital AGC method according to this application;

FIG. 3 is a flowchart of another embodiment of a digital AGC method according to this application;

FIG. 4 is a schematic structural diagram of an embodiment of a digital AGC apparatus according to this application; and

FIG. 5 is a schematic structural diagram of a computer system of an electronic device suitable for implementing the embodiments of this application.

DETAILED DESCRIPTION

[0018] This application will be further described in detail below in conjunction with the accompanying drawings and embodiments. It may be understood that the specific embodiments described herein are merely intended to explain the related invention, rather than to limit the present invention. It should also be noted that, for convenience of description, only the parts related to the present invention are shown in the accompany drawings.

[0019] It should be noted that the embodiments in this application and features in the embodiments may be combined with each other if no conflict occurs. This application will be described in detail below with reference to the accompanying drawings and the embodiments.

[0020] FIG. 1 shows an exemplary system architecture 100 to which a digital AGC method according to an embodiment of this application can be applied.

[0021] As shown in FIG. 1, the system architecture 100 may include a data server 101, a network 102, and a primary

server 103. The network 102 is configured to provide a medium of a communication link between the data server 101 and the primary server 103. The network 102 may include various connection types, for example, a wired or wireless communication link, a fiber-optic cable, and the like.

[0022] The primary server 103 may be a server providing various services, for example, a data processing server for processing information uploaded by the data server 101. The data processing server may process received event information, and store processing results (such as an element information set and a label) into an event information library in an associated manner.

[0023] It should be noted that, the digital AGC method provided in the embodiments of this application is generally executed by the primary server 103. Correspondingly, the digital AGC apparatus is generally set in the primary server 103.

[0024] It should be noted that, the data server and the primary server may be hardware or software. When being hardware, the data server and the primary server each may be implemented as a distributed server cluster including multiple servers, or may be implemented as a single server. When being software, the data server and the primary server each may be implemented as multiple pieces of software or multiple software modules (for example, software or software modules for providing distributed services), or may be implemented as a single piece of software or a single software module.

[0025] It should be appreciated that the quantities of the data servers, the networks, and the primary servers in FIG. 1 are merely an example. Any quantities of terminal devices, network devices and servers may be provided according to an implementation requirement.

[0026] Further refer to FIG. 2, which shows a process 200 of an embodiment of a digital AGC method according to this application. The method includes the following steps:

[0027] Step 201. calculate a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model.

[0028] In an example, each segment of signal data is set to be speech data with a length of 10 ms, or may be set otherwise according to an actual situation and an application scenario.

[0029] Step 202. perform a speech envelope estimation on the signal data to obtain a speech envelope amplitude.

[0030] In this embodiment, speech envelope estimation processing is performed on each segment of signal data respectively, and a calculation formula of the speech envelope estimation is as follows:

$$A[n] = \begin{cases} A[n-1] + \alpha[n] * (E[n] - A[n-1]) & P_n > P_{T1} \\ A[n-1] & \text{others} \end{cases}$$

where $A[n]$ is a speech envelope estimation in n^{th} processing, $E[n]$ is a maximum value of the piece of signal data, α is an update coefficient, and P_{T1} is a speech update threshold.

[0031] In an example, when the speech probability of the signal data is greater than the threshold, the input signal data has a high probability of being a speech, and in this case, a speech envelope is updated, to avoid an excessively small estimated value of the envelope due to a noise signal, thus preventing amplification of the noise signal. The value of the update coefficient α controls the stability of the envelope. A greater value of the update coefficient allows the envelope to fluctuate more drastically, so that the estimated value approaches the maximum value of the current input signal more quickly; otherwise, the envelope is more stable.

[0032] A calculation formula of the update coefficient α is as follows:

$$\alpha[n] = \begin{cases} \alpha[n-1] + \Delta & P_n < P_{T2} \\ \alpha[n-1] - \Delta & \text{others} \end{cases}$$

where Δ is an adjustment amount of the update coefficient, and P_{T2} is an update threshold value.

[0033] After the update coefficient α is obtained, a maximum value and a minimum value of the update coefficient are then restricted:

$$\alpha[n] = \begin{cases} 1 & \alpha[n] > 1 \\ \alpha_0 & \alpha[n] < \alpha_0 \end{cases}$$

where α_0 is a lower-limit threshold, and $\alpha_0 < 1$.

[0034] In an example, when the speech probability of the input signal is relatively low, the value of α increases; when the speech probability of the input signal is relatively high, the value of α decreases. When the input signal is a continuous speech, the value of α decreases gradually, an update speed of the envelope decreases gradually, and the estimated

envelope is more stable; when non-speech input signals appear many times, the value of α increases gradually, the update speed of the envelope increases gradually, and the value thereof is closer to the maximum value of the current input.

[0035] In some optional implementations of this embodiment, when the signal data comes from the same person, gain variations during processing are preferred to be as small as possible. Therefore, the envelope estimation value is preferred to be as stable as possible. When the user changes, the gain needs to be adjusted quickly. In this case, the envelope estimation value also needs to change quickly. In general cases, signal data is highly continuous when the same person speaks, and a relatively stable envelope can be obtained according to the foregoing method. After a person finishes speaking, the update coefficient increases gradually. If another person speaks again, the envelope estimation value may track the input signal quickly, and the envelope can also change quickly. Thereafter, the envelope tends to be stable as the user speaks.

[0036] Step 203. calculate a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude.

[0037] In this embodiment, the first gain is calculated according to the deviation between the speech envelope amplitude and the anticipated amplitude, and a calculation formula of the first gain is as follows:

$$g_1 = \frac{A_T}{A[n]}$$

where A_T is the anticipated amplitude.

[0038] Step 204. classify the signal data based on the speech probability P_n , and count the number of noise occurrences in the signal data based on a classification result, to calculate a second gain.

[0039] In this embodiment, the signal data is classified according to the speech probability P_n , and a classification method is as follows:

$$T[n] = \begin{cases} 1 & P_n > P_{T3} \\ 0 & \text{others} \end{cases}$$

where P_{T3} is a classification threshold, $T[n]$ is the classification result, 1 represents a speech, and 0 represents noise.

[0040] The second gain is calculated according to the number c of noise occurrences in the speech data counted based on the classification result:

$$g_2 = \begin{cases} 1 & c < c_1 \\ g_{min} & c > c_2 \end{cases}$$

where c_1 and c_2 are thresholds, and $c_1 < c_2$; g_{min} is a minimum gain and $g_{min} < 1$.

[0041] Step 205. perform signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

[0042] In this embodiment, the signal gain amplitude adjustment is performed on the signal data based on the first gain and the second gain, and a formula of the signal gain amplitude adjustment is as follows:

$$x_o = g_1 * g_2 * x$$

where x is the signal data, and x_o is an output of the gain amplitude adjustment.

[0043] In an example, if noise occurs frequently in a period of time, it is highly probable that no one is speaking locally; in this case, g_2 is less than 1, and a final gain value is also less than g_1 , which can reduce a noise gain amplitude when no one speaks locally. On the contrary, if the number of noise occurrences counted in a period of time is less than the threshold, it is highly probable that someone is speaking locally; in this case, g_2 is equal to 1, the final gain value is g_1 , and the amplitude of the output signal may be close to the anticipated amplitude.

[0044] The foregoing embodiments of this application provide a digital AGC method and apparatus. In the method, a first gain and a second gain are calculated for signal data through speech probability calculation, speech envelope estimation and signal data classification, thereby adjusting a signal gain amplitude, which helps achieve an overall amplification effect and reduce distortion caused by automatic gain processing. Moreover, in the case of user change, the signal gain amplitude can also be adjusted quickly in response to different users. Meanwhile, in the case of excessive

ambient noise, the apparatus can also adjust the gain automatically, to reduce amplification of the noise.

[0045] Further refer to FIG. 3, which shows a process 300 of another embodiment of the digital AGC method according to this embodiment. The method includes the following steps:

[0046] Step 301. calculate a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model.

[0047] In this embodiment, step 301 is basically the same as step 201 in the embodiment corresponding to FIG. 2, and is not described in detail herein again.

[0048] Step 302. perform a speech envelope estimation on the signal data to obtain a speech envelope amplitude.

[0049] In this embodiment, step 302 is basically the same as step 202 in the embodiment corresponding to FIG. 2, and is not described in detail herein again.

[0050] Step 303. calculate a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude.

[0051] In this embodiment, step 303 is basically the same as step 203 in the embodiment corresponding to FIG. 2, and is not described in detail herein again.

[0052] Step 304. classify the signal data based on the speech probability P_n , and count the number of noise occurrences in the signal data based on a classification result, to calculate a second gain.

[0053] In this embodiment, step 304 is basically the same as step 204 in the embodiment corresponding to FIG. 2, and is not described in detail herein again.

[0054] Step 305. perform signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

[0055] In this embodiment, step 305 is basically the same as step 205 in the embodiment corresponding to FIG. 2, and is not described in detail herein again.

[0056] Step 306. dynamically compress the signal data by using a dynamic compressor.

[0057] In this embodiment, the signal data is dynamically compressed by using the dynamic compressor, and a formula of the dynamic compression is as follows:

$$y = \begin{cases} x_o & x_o < T - W/2 \\ T + \frac{(x_o - T)}{R} & x_o > T + W/2 \\ x_o + \frac{(\frac{1}{R} - 1)(x_o - T + \frac{W}{2})^2}{2W} & T - \frac{W}{2} \leq x_o \leq T + \frac{W}{2} \end{cases}$$

where T is a compression threshold, R is a compression coefficient, W is a compression transition range, x_o is a signal after gain adjustment, and y is an output signal.

[0058] In an example, when a user speaks stably, the envelope estimation obtained during the period is also stable, and therefore an overall amplification effect can be achieved. However, when a user makes a speech, amplitudes of some tones are often obviously higher than other tones, and such high-amplitude tones may cause cracking voice during output if the overall amplification effect is still used. In addition, when a user quickly switches from a far distance to a near distance, if the update coefficient fails to be adjusted to 1 in time, the envelope estimation value may be less than the maximum value of the input, which may also result in an excessively high amplitude of the output signal. Therefore, the dynamic compressor may be used for processing, to reduce voice cracking.

[0059] It can be seen from FIG. 3 that, compared with the embodiment corresponding to FIG. 2, the process 300 of the digital AGC method in this embodiment highlights the step of dynamically compressing the signal data by using the dynamic compressor. Therefore, the solution described in this embodiment may reduce voice cracking caused by an excessively large gain, which facilitates stable output of the speech signal.

[0060] Further refer to FIG. 4, as an implementation of the methods shown in the foregoing figures, this application provides an embodiment of a digital AGC apparatus. The apparatus embodiment corresponds to the method embodiment shown in FIG. 2, and the apparatus may be specifically applied to various electronic devices.

[0061] As shown in FIG. 4, a digital AGC apparatus 400 of this embodiment includes: a speech probability calculation module 401, configured to calculate a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model; a speech envelope estimation module 402, configured to perform a speech envelope estimation on the signal data to obtain a speech envelope amplitude; a first gain calculation module 403, configured to calculate a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude; a second gain calculation module 404, configured to classify the signal data based on the speech probability P_n , and count the number of noise occurrences in the signal data based on a classification result, to calculate a second gain; and a signal gain amplitude adjustment module 405, configured to perform signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

[0062] In this embodiment, the speech probability calculation module 401 may calculate the speech probability P_n of

each segment of signal data in a piece of read signal data by using the speech probability calculation model.

[0063] In an example, each segment of signal data is set to be speech data with a length of 10 ms, or may be set otherwise according to an actual situation and an application scenario.

[0064] In this embodiment, the speech envelope estimation module 402 may perform the speech envelope estimation on the signal data to obtain the speech envelope amplitude.

[0065] Specifically, the speech envelope estimation module 402 performs the speech envelope estimation processing on each segment of signal data respectively, and a calculation formula of the speech envelope estimation is as follows:

$$A[n] = \begin{cases} A[n-1] + \alpha[n] * (E[n] - A[n-1]) & P_n > P_{T1} \\ A[n-1] & \text{others} \end{cases}$$

where $A[n]$ is a speech envelope estimation in n^{th} processing, $E[n]$ is a maximum value of the piece of signal data, α is an update coefficient, and P_{T1} is a speech update threshold.

[0066] In an example, when the speech probability of the signal data is greater than the threshold, the input signal data has a high probability of being a speech, and in this case, a speech envelope is updated, to avoid an excessively small estimated value of the envelope due to a noise signal, thus preventing amplification of the noise signal. The value of the update coefficient α controls the stability of the envelope. A greater value of the update coefficient allows the envelope to fluctuate more drastically, so that the estimated value approaches the maximum value of the current input signal more quickly; otherwise, the envelope is more stable.

[0067] A calculation formula of the update coefficient α is as follows:

$$\alpha[n] = \begin{cases} \alpha[n-1] + \Delta & P_n < P_{T2} \\ \alpha[n-1] - \Delta & \text{others} \end{cases}$$

where Δ is an adjustment amount of the update coefficient, and P_{T2} is an update threshold value.

[0068] After the update coefficient α is obtained, a maximum value and a minimum value of the update coefficient are then restricted:

$$\alpha[n] = \begin{cases} 1 & \alpha[n] > 1 \\ \alpha_0 & \alpha[n] < \alpha_0 \end{cases}$$

where α_0 is a lower-limit threshold, and $\alpha_0 < 1$.

[0069] In an example, when the speech probability of the input signal is relatively low, the value of α increases; when the speech probability of the input signal is relatively high, the value of α decreases. When the input signal is a continuous speech, the value of α decreases gradually, an update speed of the envelope decreases gradually, and the estimated envelope is more stable; when non-speech input signals appear many times, the value of α increases gradually, the update speed of the envelope increases gradually, and the value thereof is closer to the maximum value of the current input.

[0070] In some optional implementations of this embodiment, when the signal data comes from the same person, gain variations during processing are preferred to be as small as possible. Therefore, the envelope estimation value is preferred to be as stable as possible. When the user changes, the gain needs to be adjusted quickly. In this case, the envelope estimation value also needs to change quickly. In general cases, signal data is highly continuous when the same person speaks, and a relatively stable envelope can be obtained according to the foregoing method. After a person finishes speaking, the update coefficient increases gradually. If another person speaks again, the envelope estimation value may track the input signal quickly, and the envelope can also change quickly. Thereafter, the envelope tends to be stable as the user speaks.

[0071] In this embodiment, the first gain calculation module 403 may calculate the first gain according to the deviation between the speech envelope amplitude and the anticipated amplitude.

[0072] Specifically, the first gain calculation module 403 calculates the first gain according to the deviation between the speech envelope amplitude and the anticipated amplitude, and a calculation formula of the first gain is as follows:

$$g_1 = \frac{A_T}{A[n]}$$

where A_T is the anticipated amplitude.

[0073] In this embodiment, the second gain calculation module 404 may classify the signal data based on the speech probability P_n , and count the number of noise occurrences in the signal data based on the classification result, to calculate the second gain.

[0074] Specifically, the second gain calculation module 404 classifies the signal data based on the speech probability P_n , and a classification method is as follows:

$$T[n] = \begin{cases} 1 & P_n > P_{T3} \\ 0 & \text{others} \end{cases}$$

where P_{T3} is a classification threshold, $T[n]$ is the classification result, 1 represents a speech, and 0 represents noise.

[0075] The second gain is calculated according to the number c of noise occurrences in the speech data counted based on the classification result:

$$g_2 = \begin{cases} 1 & c < c_1 \\ g_{min} & c > c_2 \end{cases}$$

where c_1 and c_2 are thresholds, and $c_1 < c_2$; g_{min} is a minimum gain and $g_{min} < 1$; in this embodiment, a first storage unit 405 may store target event information, an element information set and a label into a preset event information library in an associated manner.

[0076] In this embodiment, the signal gain amplitude adjustment module 405 may perform the signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

[0077] Specifically, the signal gain amplitude adjustment module 405 performs the signal gain amplitude adjustment on the signal data based on the first gain and the second gain, and a formula of the signal gain amplitude adjustment is as follows:

$$x_o = g_1 * g_2 * x$$

where x is the signal data, and x_o is an output of the gain amplitude adjustment.

[0078] In an example, if noise occurs frequently in a period of time, it is highly probable that no one is speaking locally; in this case, g_2 is less than 1, and a final gain value is also less than g_1 , which can reduce a noise gain amplitude when no one speaks locally. On the contrary, if the number of noise occurrences counted in a period of time is less than the threshold, it is highly probable that someone is speaking locally; in this case, g_2 is equal to 1, the final gain value is g_1 , and the amplitude of the output signal may be close to the anticipated amplitude.

[0079] In some optional implementations of this embodiment, the apparatus 400 may further include: a dynamic compression module, configured to dynamically compress the signal data by using a dynamic compressor, and a formula of the dynamic compression is as follows:

$$y = \begin{cases} x_o & x_o < T - W/2 \\ T + \frac{(x_o - T)}{R} & x_o > T + W/2 \\ x_o + \frac{(\frac{1}{R} - 1)(x_o - T + \frac{W}{2})^2}{2W} & T - \frac{W}{2} \leq x_o \leq T + \frac{W}{2} \end{cases}$$

where T is a compression threshold, R is a compression coefficient, W is a compression transition range, x_o is a signal after gain adjustment, and y is an output signal.

[0080] In an example, when a user speaks stably, the envelope estimation obtained during the period is also stable, and therefore an overall amplification effect can be achieved. However, when a user makes a speech, amplitudes of some tones are often obviously higher than other tones, and such high-amplitude tones may cause cracking voice during output if the overall amplification effect is still used. In addition, when a user quickly switches from a far distance to a near distance, if the update coefficient fails to be adjusted to 1 in time, the envelope estimation value may be less than the maximum value of the input, which may also result in an excessively high amplitude of the output signal. Therefore, the dynamic compressor may be used for processing, to reduce voice cracking.

[0081] The apparatus provided in the foregoing embodiment of this application calculates a first gain and a second gain for signal data through speech probability calculation, speech envelope estimation and signal data classification, thereby adjusting a signal gain amplitude, which helps achieve an overall amplification effect and reduce distortion caused by automatic gain processing. Moreover, in the case of user change, the signal gain amplitude can also be adjusted quickly in response to different users. Meanwhile, in the case of excessive ambient noise, the apparatus can also adjust the gain automatically, to reduce amplification of the noise.

[0082] Refer to FIG. 5 below, which is a schematic structural diagram of a computer system 500 of an electronic device suitable for implementing the embodiments of this application. The electronic device shown in FIG. 5 is merely an example, and should not cause any limitation to the functions and application range of the embodiments of this application.

[0083] As shown in FIG. 5, the computer system 500 includes a central processing unit (CPU) 501, which can perform various suitable actions and processing according to a program stored in a read-only memory (ROM) 502 or a program loaded from a storage part 508 to a random access memory (RAM) 503. The RAM 503 further stores various programs and data required for operations of the system 500. The CPU 501, the ROM 502, and the RAM 503 are mutually connected through a bus 504. An input/output (I/O) interface 505 is also connected to the bus 504.

[0084] The following components are connected to the I/O interface 505: an input part 506 including a keyboard, a mouse, and the like; an output part 507 including a liquid crystal display (LCD), a loudspeaker, and the like; the storage part 508 including a hard disk; and a communication part 509 including a network interface card such as a local area network (LAN) card or a modem. The communication part 509 executes communication processing via a network such as the Internet. A driver 510 may also be connected to the I/O interface 505 as required. A removable medium 511, such as a magnetic disk, an optical disc, a magneto-optical disc, a semiconductor memory, or the like, is installed on the driver 510 as required, so that a computer program read therefrom can be installed in the storage part 508 as required.

[0085] Particularly, according to the embodiments of the present disclosure, the process described above with reference to the flowchart may be implemented as a computer software program. For example, the embodiment of the present disclosure includes a computer program product, which includes a computer program carried by a computer readable medium. The computer program includes program code for executing the method shown in the flowchart. In this embodiment, the computer program may be downloaded from a network by means of the communication part 509 and installed, and/or be installed from the removable medium 511. When the computer program is executed by the CPU 501, the functions defined in the method of this application are executed.

[0086] It should be noted that, the computer readable medium in this application may be a computer readable signal medium or a computer readable storage medium or a combination thereof. For example, the computer readable storage medium may be, but is not limited to, electric, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus or device, or any combination of the above. More specific examples of the computer readable storage medium may include, but are not limited to: an electric connector with one or more wires, a portable computer magnetic disk, a hard drive, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or flash drive), an optical fiber, a compact disk read-only memory (CD-ROM), an optical storage device, a magnetic storage device, or any proper combination of the above. In this application, the computer readable storage medium may be any tangible medium containing or storing a program, and the program may be used by or used in combination with an instruction execution system, apparatus or device. In this application, the computer readable signal medium may include a data signal propagated in a baseband or propagated as a part of a carrier, and carries computer readable program code. Such a propagated data signal may be in multiple forms, including, but not limited to an electromagnetic signal, an optical signal, or any proper combination of the above. The computer readable signal medium may also be any computer readable medium except the computer readable storage medium. The computer readable storage medium may send, propagate or transmit a program used by or used in combination with an instruction execution system, apparatus or device. The program code contained in the computer readable storage medium may be transmitted by using any suitable medium, including, but not limited to radio, an electric wire, an optical fiber, RF, and the like, or any proper combination of the above.

[0087] Computer program code for executing the operations in this application may be compiled by using one or more programming languages or a combination thereof. The programming languages include object oriented programming languages, such as Java, Smalltalk, and C++, and also include conventional procedural programming languages, such as C or similar programming languages. The program code can be executed fully on a user computer, executed partially on a user computer, executed as an independent software package, executed partially on a user computer and partially on a remote computer, or executed fully on a remote computer or a server. In a circumstance in which a remote computer is involved, the remote computer may be connected to a user computer via any type of network, including a local area network (LAN) or a wide area network (WAN), or may be connected to an external computer (for example, connected via the Internet by using an Internet service provider).

[0088] The flowcharts and block diagrams in the accompanying drawings illustrate system architectures, functions and operations that may be implemented by the system, method and computer program product according to the embodiments of this application. In this regard, each block in the flowcharts or block diagrams may represent a module,

a program segment or a part of code, and the module, the program segment or the part of code includes one or more executable instructions for implementing specified logic functions. It should also be noted that, in some alternative implementations, functions marked in the blocks may also take place in an order different from the order marked in the accompanying drawings. For example, two consecutive blocks can actually be executed substantially in parallel, and they may sometimes be executed in a reverse order, which depends on involved functions. It should also be noted that each block in the flowcharts and/or block diagrams and combinations of the blocks in the flowcharts and/or block diagrams may be implemented by a dedicated hardware-based system for executing specified functions or operations, or may be implemented by a combination of dedicated hardware and computer instructions.

[0089] Related units described in the embodiments of this application may be implemented by software, or may be implemented by hardware. The described units may also be set in a processor, which, for example, may be described as a processor including a speech probability calculation unit, a speech envelope estimation unit, a first gain calculation unit, a second gain calculation unit, and a signal gain amplitude adjustment unit.

[0090] In another aspect, this application further provides a computer readable storage medium. The computer readable storage medium may be contained in the electronic device described in the foregoing embodiment, or may exist alone without being assembled in the electronic device. The computer readable storage medium carries one or more programs. When executed by the electronic device, the one or more programs cause the electronic device to: calculate a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model; perform a speech envelope estimation on the signal data to obtain a speech envelope amplitude; calculate a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude; classify the signal data based on the speech probability P_n , and count the number of noise occurrences in the signal data based on a classification result, to calculate a second gain; and perform signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

[0091] The above description is merely an illustration of preferred embodiments of this application and the technical principle in use. Those skilled in the art should understand that, the scope of invention of this application is not limited to the technical solution formed by a specific combination of the foregoing technical features, but should cover other technical solutions formed by any combination of the foregoing technical features or equivalent features thereof without departing from the foregoing inventive concept, for example, a technical solution formed by replacing the foregoing feature with a technical feature having a similar function disclosed in (but not limited to) this application.

Claims

1. A digital automatic gain control (AGC) method, comprising steps of:

calculating a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model;
performing a speech envelope estimation on the signal data to obtain a speech envelope amplitude;
calculating a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude;
classifying the signal data based on the speech probability P_n , and counting a number of noise occurrences in the signal data based on a result of the classifying, to calculate a second gain; and
performing signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

2. The digital AGC method according to the previous claim, whereby the step of performing a speech envelope estimation on the signal data to obtain a speech envelope amplitude specifically comprises:

performing speech envelope estimation processing on the each segment of signal data respectively, a calculation formula of the speech envelope estimation being as follows:

$$A[n] = \begin{cases} A[n-1] + \alpha[n] * (E[n] - A[n-1]) & P_n > P_{T1} \\ A[n-1] & \text{others} \end{cases}$$

wherein $A[n]$ is a speech envelope estimation in n^{th} processing, $E[n]$ is a maximum value of the piece of signal data, α is an update coefficient, and P_{T1} is a speech update threshold;
wherein a calculation formula of the update coefficient α is as follows:

$$\alpha[n] = \begin{cases} \alpha[n-1] + \Delta & P_n < P_{T2} \\ \alpha[n-1] - \Delta & \text{others} \end{cases}$$

wherein Δ is an adjustment amount of the update coefficient, and P_{T2} is an update threshold value; and after the update coefficient α is obtained, restricting a maximum value and a minimum value of the update coefficient:

$$\alpha[n] = \begin{cases} 1 & \alpha[n] > 1 \\ \alpha_0 & \alpha[n] < \alpha_0 \end{cases}$$

wherein α_0 is a lower-limit threshold, and $\alpha_0 < 1$.

3. The digital AGC method according to any of the previous claims, whereby the step of calculating a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude specifically comprises:

calculating the first gain according to the deviation between the speech envelope amplitude and the anticipated amplitude, a calculation formula of the first gain being as follows:

$$g_1 = \frac{A_T}{A[n]}$$

wherein A_T is the anticipated amplitude.

4. The digital AGC method according to any of the previous claims, whereby the step of classifying the signal data based on the speech probability P_n , and counting a number of noise occurrences in the signal data based on a result of the classifying, to calculate a second gain comprises:

classifying the signal data according to the speech probability P_n , a classification method being as follows:

$$T[n] = \begin{cases} 1 & P_n > P_{T3} \\ 0 & \text{others} \end{cases}$$

wherein P_{T3} is a classification threshold, $T[n]$ is the result of the classifying, 1 represents a speech, and 0 represents noise; and

calculating the second gain according to a number c of noise occurrences in the speech data counted based on the result of the classifying:

$$g_2 = \begin{cases} 1 & c < c_1 \\ g_{min} & c > c_2 \end{cases}$$

wherein c_1 and c_2 are thresholds, and $c_1 < c_2$; g_{min} is a minimum gain and $g_{min} < 1$.

5. The digital AGC method according to any of the previous claims, whereby the step of performing signal gain amplitude adjustment on the signal data by using the first gain and the second gain comprises:

performing the signal gain amplitude adjustment on the signal data based on the first gain and the second gain, a formula of the signal gain amplitude adjustment being as follows:

$$x_o = g_1 * g_2 * x$$

wherein x is the signal data, and x_o is an output of the gain amplitude adjustment.

6. The digital AGC method according to any of the previous claims, further comprising:

dynamically compressing the signal data by using a dynamic compressor, a formula of the dynamic compression being as follows:

5

10

$$y = \begin{cases} x_o & x_o < T - W/2 \\ T + \frac{(x_o - T)}{R} & x_o > T + W/2 \\ x_o + \frac{(\frac{1}{R} - 1)(x_o - T + \frac{W}{2})^2}{2W} & T - \frac{W}{2} \leq x_o \leq T + \frac{W}{2} \end{cases}$$

15

wherein T is a compression threshold, R is a compression coefficient, W is a compression transition range, x_o is a signal after gain adjustment, and y is an output signal.

7. A digital AGC apparatus, comprising:

20

a speech probability calculation module, configured to calculate a speech probability P_n of each segment of signal data in a piece of read signal data by using a speech probability calculation model;

a speech envelope estimation module, configured to perform a speech envelope estimation on the signal data to obtain a speech envelope amplitude;

25

a first gain calculation module, configured to calculate a first gain according to a deviation between the speech envelope amplitude and an anticipated amplitude;

a second gain calculation module, configured to classify the signal data based on the speech probability P_n , and count a number of noise occurrences in the signal data based on a classification result, to calculate a second gain; and

30

a signal gain amplitude adjustment module, configured to perform signal gain amplitude adjustment on the signal data by using the first gain and the second gain.

8. The digital AGC apparatus according to the previous claim, further comprising:

35

a signal data classification module, configured to classify the signal data according to the speech probability P_n , a classification method being as follows:

40

$$T[n] = \begin{cases} 1 & P_n > P_{T3} \\ 0 & \text{others} \end{cases}$$

wherein P_{T3} is a classification threshold, $T[n]$ is the classification result, 1 represents a speech, and 0 represents noise.

9. The digital AGC apparatus according to any of the two previous claims, further comprising:

45

a dynamic compression module, configured to dynamically compress the signal data by using a dynamic compressor, a formula of the dynamic compression being as follows:

50

55

$$y = \begin{cases} x_o & x_o < T - W/2 \\ T + \frac{(x_o - T)}{R} & x_o > T + W/2 \\ x_o + \frac{(\frac{1}{R} - 1)(x_o - T + \frac{W}{2})^2}{2W} & T - \frac{W}{2} \leq x_o \leq T + \frac{W}{2} \end{cases}$$

wherein T is a compression threshold, R is a compression coefficient, W is a compression transition range, x_o

is a signal after gain adjustment, and y is an output signal.

10. A computer readable storage medium storing a computer program, whereby the program implements the method according to any one of claims 1 to 6 when being executed by a processor.

5

10

15

20

25

30

35

40

45

50

55

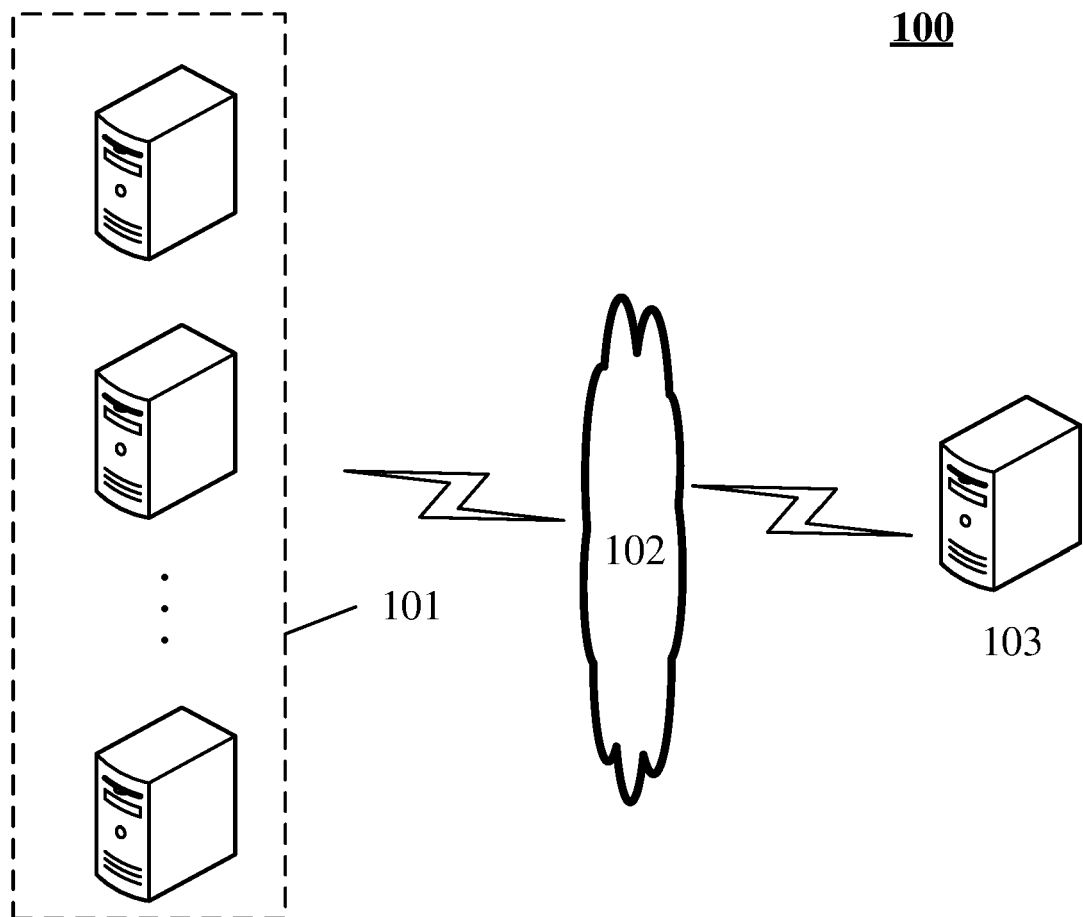


FIG. 1

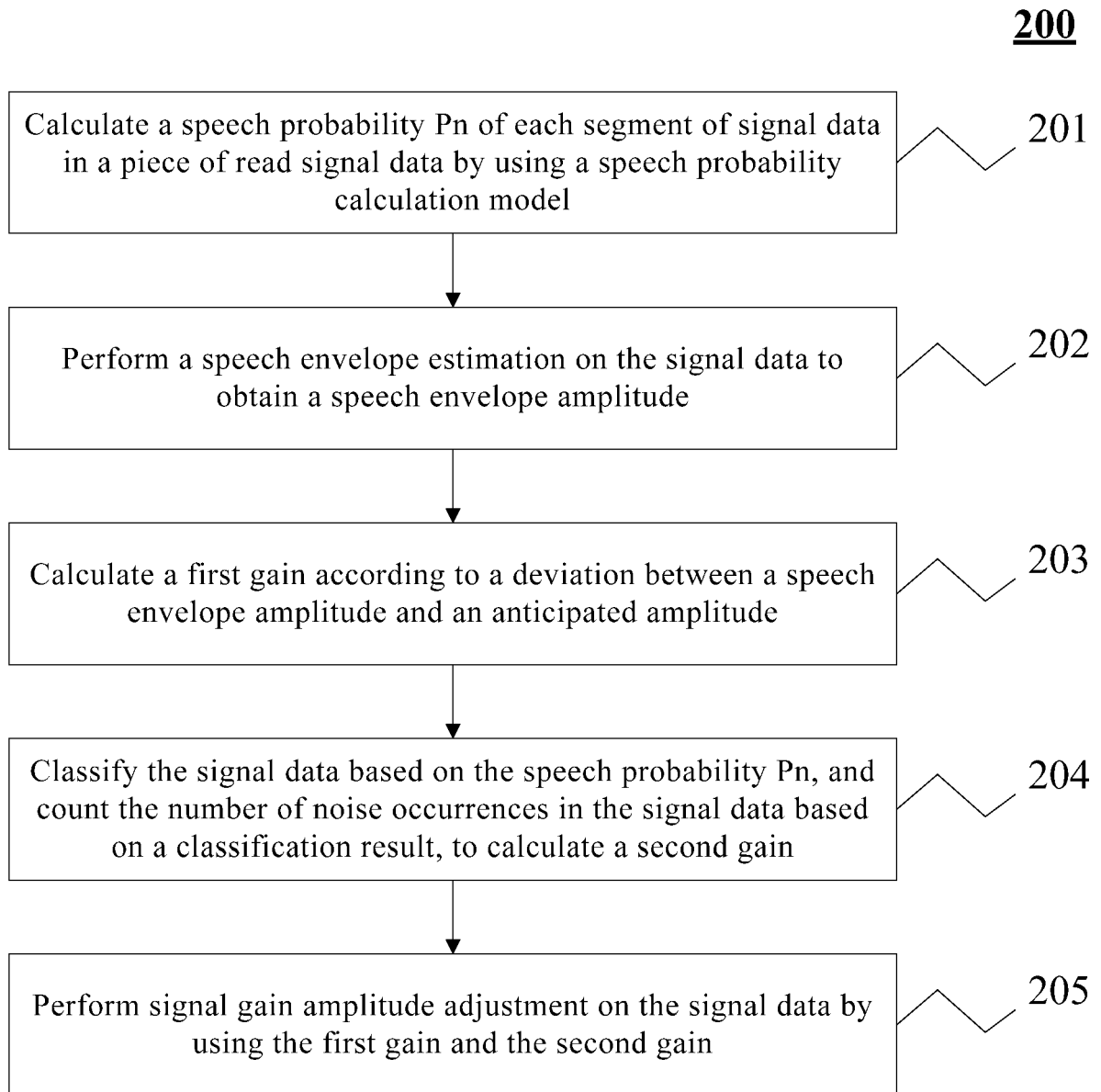


FIG. 2

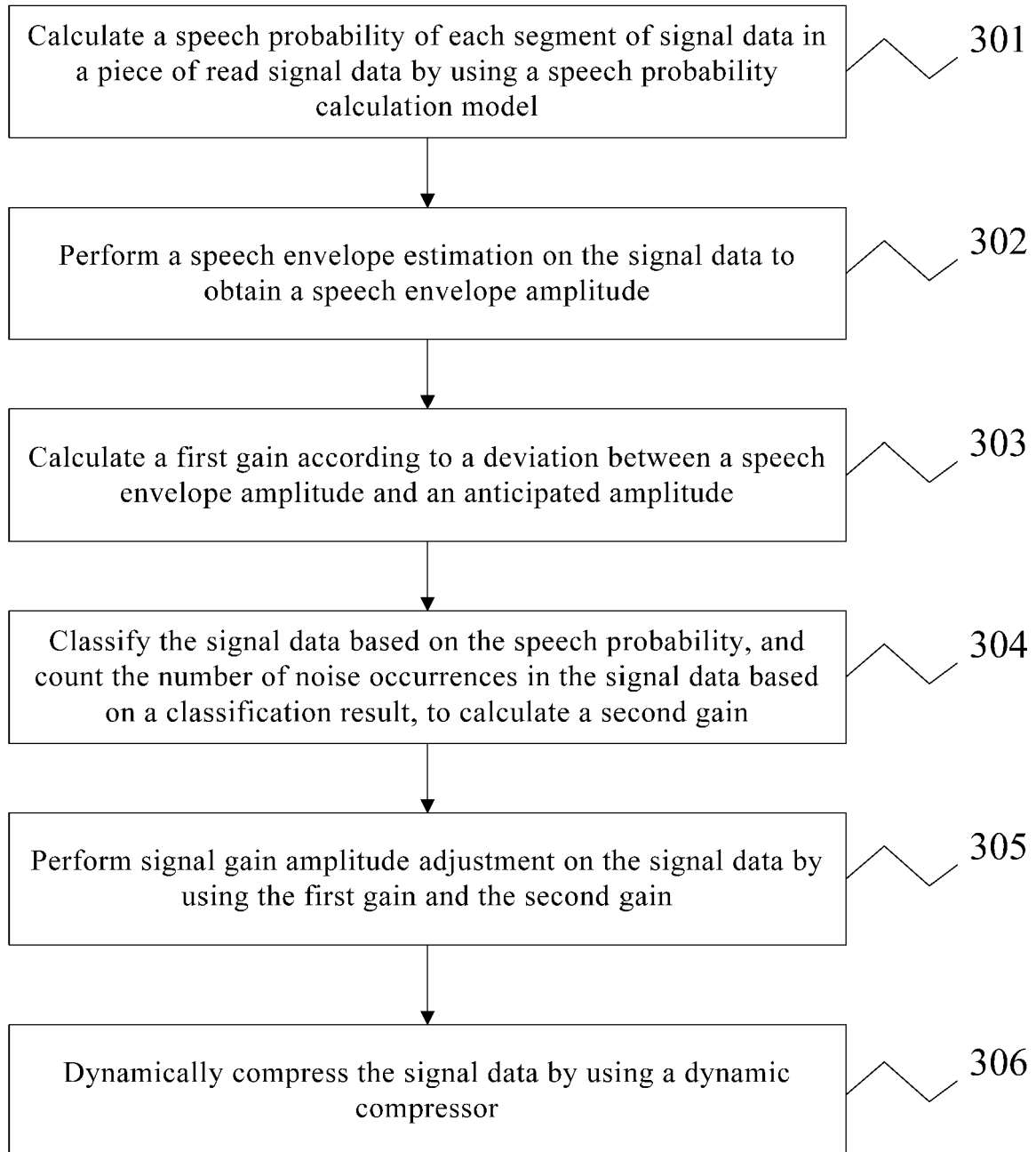
300

FIG. 3

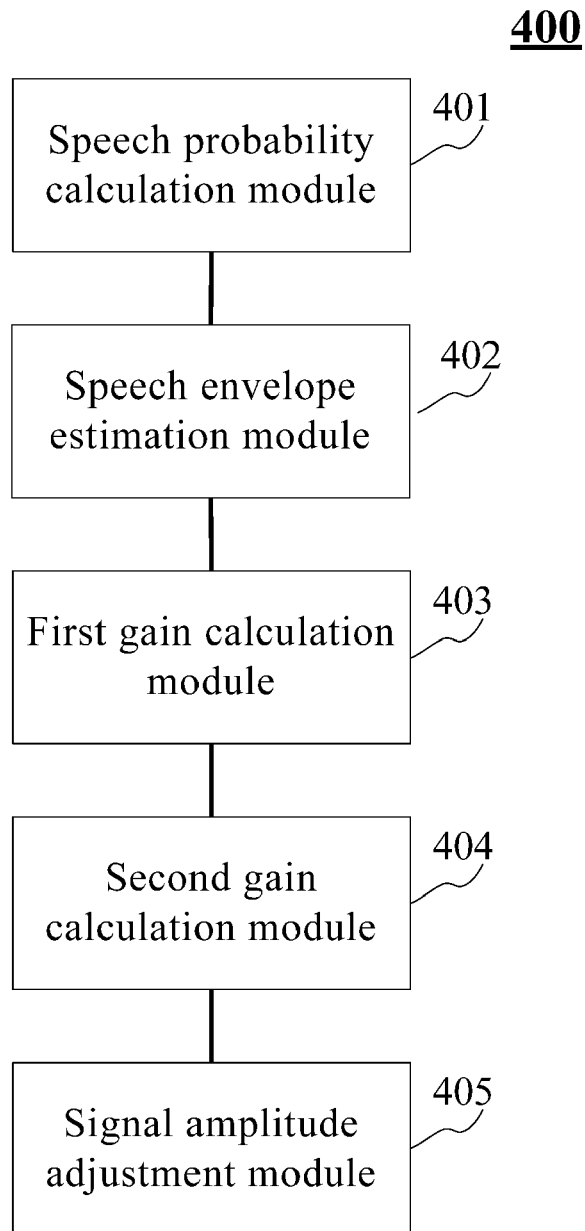


FIG. 4

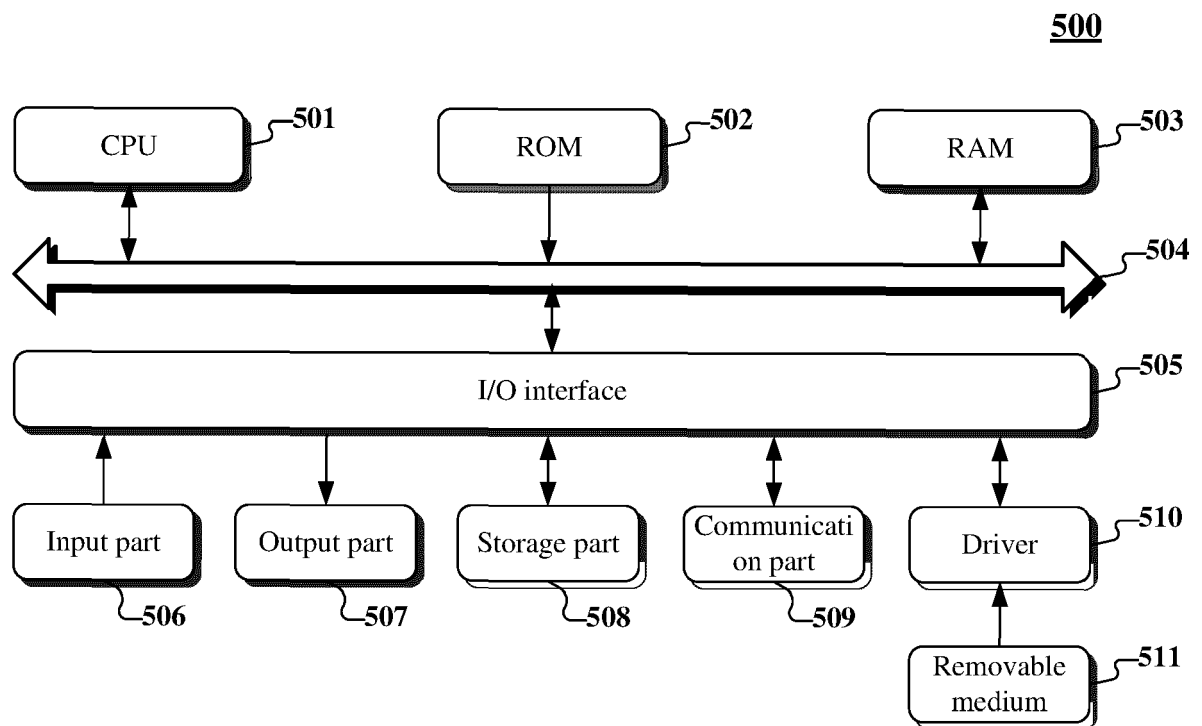


FIG. 5



EUROPEAN SEARCH REPORT

 Application Number
 EP 20 19 5635

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	US 2018/068670 A1 (LU LIE [CN] ET AL) 8 March 2018 (2018-03-08)	1,3-10	INV. G10L21/0364 G10L21/0208
A	* paragraphs [0131], [0145], [0264], [0266], [0275] *	2	
A	----- US 5 937 377 A (HARDIMAN BUDI AGUNG [US] ET AL) 10 August 1999 (1999-08-10) * column 8, line 66 - column 9, line 9 *	2,4,8	
A	----- US 9 685 921 B2 (DTS INC [US]) 20 June 2017 (2017-06-20) * column 7, line 50 - column 7, line 62 *	2	
			TECHNICAL FIELDS SEARCHED (IPC)
			G10L
The present search report has been drawn up for all claims			
Place of search		Date of completion of the search	Examiner
The Hague		20 January 2021	Taddei, Hervé
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

 1
 EPO FORM 1503 03.02 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 20 19 5635

5 This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

20-01-2021

10

15

20

25

30

35

40

45

50

55

CFM P0459

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2018068670 A1	08-03-2018	CN 104078050 A	01-10-2014
		CN 105074822 A	18-11-2015
		CN 109616142 A	12-04-2019
		EP 2979267 A2	03-02-2016
		EP 3598448 A1	22-01-2020
		HK 1212808 A1	17-06-2016
		JP 6573870 B2	11-09-2019
		JP 2016519784 A	07-07-2016
		JP 2019194742 A	07-11-2019
		US 2016078879 A1	17-03-2016
		US 2018068670 A1	08-03-2018
		WO 2014160678 A2	02-10-2014

US 5937377 A	10-08-1999	NONE	

US 9685921 B2	20-06-2017	CN 104471855 A	25-03-2015
		HK 1208290 A1	26-02-2016
		JP 6328627 B2	23-05-2018
		JP 2015523606 A	13-08-2015
		KR 20150036581 A	07-04-2015
		US 2014016791 A1	16-01-2014
		WO 2014011959 A2	16-01-2014
