



(11) **EP 2 096 629 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention
of the grant of the patent:
24.10.2012 Bulletin 2012/43

(21) Application number: **07855800.4**

(22) Date of filing: **26.12.2007**

(51) Int Cl.:
G10L 11/02 (2006.01)

(86) International application number:
PCT/CN2007/003798

(87) International publication number:
WO 2008/067735 (12.06.2008 Gazette 2008/24)

(54) **METHOD AND APPARATUS FOR CLASSIFYING SOUND SIGNALS**

VERFAHREN UND GERÄT ZUR KLASSIFIZIERUNG VON TONSIGNALEN

PROCÉDÉ ET APPAREIL POUR LE CLASSEMENT DE SIGNAUX SONORES

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IS IT LI LT LU LV MC MT NL PL PT RO SE
SI SK TR**

(30) Priority: **05.12.2006 CN 200610164456**

(43) Date of publication of application:
02.09.2009 Bulletin 2009/36

(73) Proprietor: **Huawei Technologies Co., Ltd.
Shenzhen, Guangdong Province 518129 (CN)**

(72) Inventors:

- **LI, Wei**
Guangdong 518129 (CN)
- **XU, Lijing**
Guangdong 518129 (CN)
- **ZHANG, Qing**
Guangdong 518129 (CN)
- **XU, Jianfeng**
Guangdong 518129 (CN)
- **SANG, Shenghu**
Guangdong 518129 (CN)
- **DU, Zhengzhong**
Guangdong 518129 (CN)
- **YAN, Qin**
Guangdong 100080 (CN)
- **DENG, Haojiang**
Guangdong 100080 (CN)
- **WANG, Jun**
Guangdong 100080 (CN)

(74) Representative: **Pfenning, Meinig & Partner GbR**
Patent- und Rechtsanwälte
Theresienhöhe 13
80339 München (DE)

(56) References cited:

WO-A1-96/05592 WO-A2-02/065457
CN-A- 1 296 258 CN-A- 1 331 825
CN-A- 1 354 455 CN-A- 1 430 778
CN-A- 1 624 766

- **JELINEK M ET AL: "Robust signal/noise discrimination for wideband speech and audio coding", SPEECH CODING, 2000. PROCEEDINGS. 2000 IEEE WORKSHOP ON SEPTEMBER 17-20, 2000, PISCATAWAY, NJ, USA, IEEE, 17 September 2000 (2000-09-17), pages 151-153, XP010520072, ISBN: 978-0-7803-6416-5**
- **"Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions (3GPP TS 26.290 version 6.3.0 Release 6); ETSI TS 126 290", ETSI STANDARDS, LIS, SOPHIA ANTIPOLIS CEDEX, FRANCE, vol. 3-SA4, no. V6.3.0, 1 June 2005 (2005-06-01), XP014030612, ISSN: 0000-0001**
- **QI F. AND BAO C.: 'A new method to voiced/unvoiced/silence of speech classification using Support Vector Machine' ACTA ELECTRONICA SINICA vol. 34, no. 4, April 2006, pages 605 - 611, XP008109092**
- **BAI L. ET AL.: 'Feature Analysis and Extraction for Audio Automatic Classification' MINI-MICRO SYSTEMS vol. 26, no. 11, November 2005, pages 2029 - 2034, XP008109115**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 2 096 629 B1

Description

FIELD OF THE INVENTION

5 **[0001]** The present invention relates to speech coding technologies, and in particular, to a method and apparatus for classifying sound signals.

BACKGROUND

10 **[0002]** In speech communication, only about 40% signals include speech, and the others are mute or background noise. In order to save transmission bandwidth, a Voice Activity Detection (VAD) technique is applied in speech coding in the speech signal processing field. Therefore, the coder may encode the background noise and active speech at different rates. That is, the coder encodes the background noise at a lower rate, and encodes the active speech at a higher rate, thus reducing the average code rate and enhancing the variable-rate speech coding technology greatly.

15 **[0003]** The VAD in the related art is developed for speech signals only, and categorizes input audio signals into only two types: noise and non-noise. Later coders such as AMR-WB+ and SMV covers detection of music signals, serving as a correction and supplement to the VAD decision. The AMR-WB+ coder is characterized that after VAD, the coding mode varies between a speech signal and a music signal, and depends on whether the input audio signal is a speech signal or music signal, thus minimizing the code rate and ensuring the coding quality.

20 **[0004]** The two different coding modes in the AMR-WB+ are: Algebraic Code Excited Linear Prediction (ACELP)-based coding algorithm, and Transform Coded eXcitation (TCX)-based coding algorithm. The ACELP sets up a speech phonation model, makes the most of the speech characteristics, and is highly efficient in encoding speech signals. Moreover, the ACELP technology is so mature that the ACELP may be extended on a universal audio coder to improve the speech coding quality massively. Likewise, the TCX may be extended on the low-bit-rate speech coder to improve the quality

25 of encoding broadband music.

[0005] Depending on complexity, the ACELP mode selection algorithm and the TCX mode selection algorithm of the AMR-WB+ coding algorithm come in two types: open loop selection algorithm, and closed loop selection algorithm. Closed-loop selection corresponds to high complexity, and is default option. It is a traversal search selection mode based on a perceptive weighted Signal-to-Noise Ratio (SNR). Evidently, such a selection method is rather accurate, but involves

30 rather complicated operation and a huge amount of codes.

[0006] The open-loop selection includes the following steps.

[0007] In step 101, the VAD module judges whether the signal is a non-usable signal or usable signal according to the Tone_flag and the sub-band energy parameter (Level[n]).

[0008] In step 102, primary mode selection (EC) is performed.

35 **[0009]** In step 103, the mode primarily determined in step 102 is corrected, and refined mode selection is performed to determine the coding mode to be selected. Specifically, this step is performed based on open loop pitch parameters and Immittance Spectral Frequency (ISF) parameters.

[0010] In step 104, TCXS processing is performed. That is, when the number of times of selecting the speech signal coding mode continuously is less than three times, a small-sized closed-loop traversal search is performed to determine

40 the coding mode finally, where the speech signal coding mode is ACELP and the music signal coding mode is TCX.

[0011] In the process of implementing the present invention, the inventor finds that the AMR-WB+ speech signal selection algorithm in the related art involves the following defects:

45 1. The VAD module in the related art underperforms in identifying noise and some music signals in the process of classifying signals, thus reducing accuracy of classifying sound signals.

2. Calculation of the open pitch parameters is necessary to the ACELP coding mode, but unnecessary to the TCX coding mode. According to the AMR-WB+ structure design, the VAD and the open-loop mode selection algorithm involve use of the open loop pitch parameters. Therefore, the open loop pitch needs to be calculated for all frames. However, as for other non-ACELP coding modes (such as TCX), the calculation of such parameters is redundant

50 complexity, increases the calculation load of coding mode selection, and reduces the efficiency.

3. Although the VAD algorithm is superior in speech detection and noise immunity among the coders currently available, it may mistake music signals for noise at the hangover of some special music signals, thus intercepting the music hangover and making the music unnatural.

4. The AMR-WB+ mode selection algorithm disregards the Signal Noise Ratio (SNR) environment of the signal, and

55 its performance of identifying speech and music in the case of a low SNR is further deteriorated.

[0012] JELINEK M ET AL: "Robust signal/noise discrimination for wideband speech and audio coding" discloses a robust discrimination method to separate information carrying signals from ambient noise coding applications. The

method consists of a two-stage procedure. First, a local decision is made based on a set of extracted parameters to update the estimated noise level. The parameters have been chosen to detect reliably speech as well as music signals. Then, the final decision is made based only on a frequency dependent Signal to Noise Ratio (SNR). The noise level update does not depend on the final decision to prevent the discriminator from locking when noise level changes suddenly.

The performance is compared with the performance of the Voice Activity Detector (VAD) of G. 729, Annex G.

[0013] ETSI TS 126290 V6.3.0 (2005-06) discloses an extended adaptive multi-rate wideband (AMR-WB+) codec with ACELP/TCX open-loop mode selection.

[0014] WO 02/065457 A2 discloses a speech coding system with a music classifier. An encoder is disposed to receive an input signal and provides a bitstream based upon a speech coding of a portion of the input signal. The encoder provides a classification of the input signal as one of noise, speech and music. The music classifier analyzes or determines signal properties of the input signal. The music classifier compares the signal properties to thresholds to determine the classification of the input signal.

SUMMARY

[0015] A method according to claim 1 and an apparatus according to claim 4 for classifying sound signals are provided by the present invention to improve accuracy of sound signal classification. Further improvements of the method and the apparatus are provided in the respective dependent claims.

[0016] In the embodiments of the present invention, the update rate of the background noise is determined, the noise parameters are updated according to the update rate, the signals are classified primarily according to the sub-band energy parameters and the updated noise parameters, and the non-useful signals and the useful signals in the received speech signals are determined, thus reducing the probability of mistaking useful signals for noise signals and improving accuracy of classifying sound signals.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Figure 1 shows open loop selection of AMR-WB+ coding algorithm in the related art;

[0018] Figure 2 is a general flowchart of a method for classifying sound signals in an embodiment of the present invention;

[0019] Figure 3 is a schematic diagram showing an apparatus for classifying sound signals in an embodiment of the present invention;

[0020] Figure 4 is a schematic diagram showing a system in an embodiment of the present invention;

[0021] Figure 5 is a flowchart of calculating various parameters on a coder parameter extracting module in an embodiment of the present invention;

[0022] Figure 6 is a flowchart of calculating various parameters on another coder parameter extracting module in an embodiment of the present invention;

[0023] Figure 7 shows composition of a PSC module in an embodiment of the present invention;

[0024] Figure 8 shows how a signal type judging module determines characteristic parameters in an embodiment of the present invention;

[0025] Figure 9 shows how a signal type judging module performs speech judgment in an embodiment of the present invention;

[0026] Figure 10 shows how a signal type judging module performs music judgment in an embodiment of the present invention;

[0027] Figure 11 shows how a signal type judging module corrects a primary judgment result in an embodiment of the present invention;

[0028] Figure 12 shows how a signal type judging module performs primary type correction for uncertain signals in an embodiment of the present invention;

[0029] Figure 13 shows how a signal type judging module performs final type correction for signals in an embodiment of the present invention; and

[0030] Figure 14 shows how a signal type judging module performs parameter update in an embodiment of the present invention.

DETAILED DESCRIPTION

[0031] In order to make the technical solution, objectives and merits of the present invention clearer, a detailed description of the present invention is given below by reference to the accompanying drawings and preferred embodiments.

[0032] In the embodiments of the present invention, the update rate of the background noise is determined according to the spectral distribution parameters of the current sound signal and the background noise, and the noise parameters

are updated according to the update rate. Therefore, the useful signals and the non-useful signals in the received speech signals are determined according to the updated noise parameters, thus improving the accuracy of the noise parameters in determining the useful signals and non-useful signals, and improving the accuracy of classifying sound signals.

[0033] Figure 2 shows a method for classifying sound signals in an embodiment of the present invention, including the following process:

[0034] Block 201: Sound signals are received, and the update rate of background noise is determined according to the spectral distribution parameters of the background noise and the sound signals.

[0035] Block 202: The noise parameters are updated according to the update rate, and the sound signals are classified according to sub-band energy parameters and updated noise parameters.

[0036] In block 202, the sound signals are classified into two types: useful signals, and non-useful signals. Afterward, the useful signals may be subdivided into speech signals and music signals, depending on whether the noise converges. The subdividing may be based on open loop pitch parameters, ISF parameters, and sub-band energy parameters, or based on ISF parameters and sub-band energy parameters.

[0037] Besides, in order to prevent mistaking music signal hangovers for non-useful signals and reducing the sound effect, a determined useful signal type is obtained in an embodiment of the present invention. The signal hangover length is determined according to the useful signal type, and the useful signals and the non-useful signals in the received speech signals are further determined according to the signal hangover length. Here the music signal hangover may be set to a relatively great value to improve the sound effect of the music signal.

[0038] In the process of determining a useful signal as a speech signal or music signal, it is appropriate to set the signal not accurately identifiable to an uncertain type first, and then correct the uncertain type according to other parameters, and finally determine the type of useful signals.

[0039] Calculation of ISF parameters is not necessarily involved in the coding mode of non-useful signals. Therefore, no ISF parameters are calculated for the determined noise signals if the corresponding coding mode needs no calculation of ISF parameters, with a view to reducing the calculation load in the classification process and improving the classification efficiency.

[0040] As shown in Figure 3, an apparatus for classifying sound signals in an embodiment of the present invention includes: a background noise parameter updating module, configured to: determine the update rate of background noise according to the spectral distribution parameters of the background noise and the current sound signals, and send the determined update rate to a PSC module; and a PSC module, configured to: update the noise parameters according to the update rate received from the background noise parameter updating module, perform primary classification for the signals according to the sub-band energy parameters and updated noise parameters, and determine the received speech signal to be a useful signal or non-useful signal.

[0041] The apparatus for classifying sound signals may further include a signal type judging module. The PSC module transfers the determined signal type to the signal type judging module. The signal type judging module determines the type of a useful signal based on the open loop pitch parameters, ISF parameters, and sub-band energy parameters, or based on ISF parameters and sub-band energy parameters, where the type of the useful signal includes speech and music.

[0042] The apparatus for classifying sound signals may further include a classification parameter extracting module. The PSC module transfers the determined signal type to the signal type judging module through the classification parameter extracting module. The classification parameter extracting module is further configured to: obtain ISF parameters and sub-band energy parameters, or further obtain open loop pitch parameters, process the obtained parameters into signal type characteristic parameters, and send the parameters to the signal type judging module; and process the obtained parameters into spectral distribution parameters of sound signals and background noise, and transfer the spectral distribution parameters to the background noise parameter updating module. Therefore, the signal type judging module determines the type of useful signals according to the foregoing signal type characteristic parameter and the signal type determined by the PSC module, where the type of useful signals includes speech and music.

[0043] The PSC module may be further configured to transfer the sound signal SNR calculated in the process of determining the signal type to the signal type judging module. The signal type judging module determines the useful signal to be a speech signal or music signal according to the SNR.

[0044] The apparatus for classifying sound signals may further include a coder mode and rate selecting module. The signal type judging module transfers the determined signal type to the coder mode and rate selecting module, and the coder mode and rate selecting module determines the coding mode and rate of sound signals according to the received signal type.

[0045] The apparatus for classifying sound signals may further include a coder parameter extracting module, which is configured to extract ISF parameters and sub-band energy parameters or additionally open loop pitch parameters, transfer the extracted parameters to the classification parameter extracting module, and transfer the extracted sub-band energy parameters to the PSC module.

[0046] The method for classifying sound signals and the apparatus for classifying sound signals in an embodiment of

the present invention are detailed below.

[0047] Figure 4 is a schematic diagram showing a system in an embodiment of the present invention. The system includes a Sound Activity Detector (SAD). As required by the coder, the SAD sorts the audio digital signals into three types: non-useful signal, speech, and music, thus forming a basis for the coder to select the coding mode and rate.

[0048] As shown in Figure 4, the SAD module includes: a background noise estimation control module, a PSC module, a classification parameter extracting module, and a signal type judging module. As a signal classifier used inside the coder, the SAD makes the most of the parameters of the coder in order to reduce resource occupation and calculation complexity. Therefore, the coder parameter extracting module in the coder is used to calculate the sub-band energy parameters and coder parameters, and provide the calculated parameters for the SAD module. Moreover, the SAD module finally outputs a determined signal type (namely, non-useful signal, speech, or music), and provides the determined signal type for the coder mode and rate selecting module to select the coder mode and rate.

[0049] The SAD-related modules in the coder, sub-modules in the SAD, and the interaction processes between the sub-modules are detailed below.

[0050] The coder parameter extracting module in the coder calculates the sub-band energy parameters and coder parameters, and provides the calculated parameters for the SAD module. The sub-band energy parameters may be calculated through filtering of a filter group. The specific quantity of sub-bands (for example, 12 sub-bands in this embodiment) is determined according to the calculation complexity requirement and classification accuracy requirement.

[0051] Figure 5 or Figure 6 shows how a coder parameter extracting module calculates various parameters required by the SAD module in this embodiment.

[0052] The process shown in Figure 5 includes the following process:

[0053] Block 501: The coder parameter extracting module calculates the sub-band energy parameters first.

[0054] Block 502: The coder parameter extracting module decides whether it is necessary to perform ISF calculation according to the primary signal judgment result (Vad_flag) received from the PSC module, and performs block 503 if necessary; or performs block 504 if not necessary.

[0055] The decision about whether to perform ISF calculation in this block includes: If the current frame is composed of non-useful signal signals, the mechanism of the coder applies. The mechanism of the coder is: If ISF parameters are required when the coder encodes non-useful signals, the ISF calculation needs to be performed; otherwise, the operation of the coder parameter extracting module is finished. If the current frame is composed of useful signals, the ISF calculation needs to be performed. Most coding modes require calculation of ISF parameters for useful signals. Therefore, the calculation brings no redundant complexity to the coder. The technical solution to calculation of ISF parameters is detailed in the instruction manuals of coders, and is not repeated here any further.

[0056] Block 503: The coder parameter extracting module calculates the ISF parameters and then performs block 504.

[0057] Block 504: The coder parameter extracting module calculates the open loop pitch parameters.

[0058] The sub-band energy parameters calculated through the process in Figure 5 are provided for the PSC module and the classification parameter extracting module in the SAD, and other parameters are provided for the classification parameter extracting module in the SAD.

[0059] In the process shown in Figure 6, a block is added on the basis of the process in Figure 5, where the added block is to decide whether to calculate the open-loop pitch parameters depending on whether the primary noise converges. Blocks 601-603 are basically identical to blocks 501-503 in Figure 5. In block 604, a judgment is made about whether the primary noise parameter (namely, noise estimation) converges. If the primary noise parameter converges, the open loop pitch parameters are calculated in block 60; otherwise, no open loop pitch parameter is calculated.

[0060] The calculation of open-loop pitch parameters is redundant to some coding modes such as TCX. In order to simplify calculation, it is basically certain that the corresponding coding mode of the signal does not need to calculate open loop pitch parameters after the noise estimation converges. Therefore, the open loop pitch parameters are not calculated any more.

[0061] Before convergence of the noise estimation, the open loop pitch parameters need to be calculated in order to ensure convergence of the noise estimation and the convergence speed. However, such calculation occurs at the startup stage, and the complexity of calculation is ignorable. The technical solution to calculation of open loop pitch parameters is detailed in the instruction about ACELP-based coding, and is not repeated here any further. The basis for judging whether the noise estimation converges may be: The count of determining as noise frames continuously exceeds the noise convergence threshold (THR1). In an example in this embodiment, the value of THR1 is 20.

[0062] The foregoing extracted sub-band energy parameter is: level[i], where i represents a member index of the vector, and its value falls within 1...12 in this embodiment, corresponding to 0-200 Hz, 200-400 Hz, 400-600 Hz, 600-800 Hz, 800-1200 Hz, 1200-1600 Hz, 1600-2000 Hz, 2000-2400 Hz, 2400-3200 Hz, 3200-4000 Hz, 4000-4800 Hz, and 4800-6400 Hz, respectively.

[0063] The foregoing extracted ISF parameter is $1/sf_n[i]$, where n represents a frame index, and the value of i falls within 1...16, representing a member index in the vector.

[0064] The foregoing extracted open loop pitch parameters include: open_loop pitch gain (ol_gain), open_loop pitch

lag (ol_lag), and tone_flag. If the value of ol_gain is greater than the value of tone threshold (TONE_THR), the tone_flag is set to 1.

[0065] The PSC module may be implemented through various VAD algorithms in the related art, and includes: background noise estimating sub-module, SNR calculating sub-module, useful signal estimating sub-module, judgment threshold adjusting sub-module, comparing sub-module, and hangover protective useful signal sub-module. In this embodiment, as shown in Figure 7, the implementation of the PSC module may differ from the VAD algorithm module in the related art in the following aspects:

[0066] I. The SNR calculating sub-module calculates the SNR according to sub-band energy estimation parameters of background noise and the sub-band energy parameters. The calculated SNR parameter is not only applied inside the PSC module, but also transferred to the signal type judging module so that the signal type judging module identifies the speech and music more accurately in the case of low SNR.

[0067] II. The VAD in the related art underperforms in identifying noise and some types of music, and improvement is made for the VAD in this embodiment: First, the calculation of the background noise parameter is controlled by the update rate (ACC) provided by the background noise parameter updating module. The background noise estimating sub-module receives the update rate from the background noise parameter updating module, updates the noise parameter, and transfers the sub-band energy estimation parameters of background noise calculated out according to the updated noise parameter to the SNR calculating sub-module. The calculation of the update rate is detailed in the instruction about the background noise parameter updating module hereinafter. In an example of this embodiment, the update rate comes in 4 levels: acc1, acc2, acc3, and acc4. For different update rates, different upward update parameters (update_up) and downward update parameters (update_down) are determined, where update_up corresponds to the upward update rate of background noise, and update_down corresponds to the downward update rate of background noise.

[0068] Afterwards, the solution to updating the noise parameter may be the solution in the AMR_WB+:

$$\begin{aligned} & \text{If } (bckr_est_m[n] < level_{m-1}[n]) \\ & \quad \text{update} = \text{update_up} \\ & \text{else} \\ & \quad \text{update} = \text{update_down} \end{aligned}$$

[0069] Therefore, the formula for updating noise estimation is:

$$bckr_est_{m+1}[n] = (1 - \text{update}) * bckr_est_m[n] + \text{update} * level_{m-1}[n]$$

[0070] Therefore, the formula for updating the spectral distribution parameter vector of noise is:

$$\tilde{p}_{m+1}[i] = (1 - \text{update}) * \tilde{p}_m[i] + \text{update} * p_m[i]$$

where,

m: frame index

n: sub-band index

i: element index of spectral distribution parameter vector, i = 1,2,3,4 bckr_est: sub-band energy of background noise estimation

$\tilde{p}$: estimation of spectral distribution parameter vector of background noise

p: spectral distribution parameter vector of the current signal

[0071] III. In the VAD in the related art, hangover is used to prevent useful signals from being mistaken for noise. The hangover length should be tradeoff between signal protection and transmission efficiency. For traditional speech coders, the hangover length may be a constant after learning. A multi-rate coder is oriented to audio signals such as music. Such signals tend to have a long low-energy hangover. It is difficult for a conventional VAD to detect such a hangover. Therefore, a relatively long hangover is required for protection. In this embodiment, the hangover length in the hangover protective useful signal sub-module is designed to be adaptive according to the SAD signal judgment result. If the

judgment result is a music signal (SAD_flag = MUSIC), a long hangover (hang_len = HANG_LONG) is set; if the judgment result is a speech signal (SAD_flag = SPEECH), a short hangover (hang_len = HANG_SHORT) is set. The detailed setting mode is as follows:

```

5      If (SAD_flag = MUSIC)
        hang_len = HANG_LONG
      else if (SAD_flag = SPEECH)
        hang_len = HANG_SHORT
      else
10     hang_len = 0
      where,
      SAD_flag: SAD judgment flag
      hang_len: protective hangover length

```

[0072] In an example of this embodiment, HANG_LONG = 100, and HANG_SHORT = 20, which may be measured in frames.

[0073] The classification parameter extracting module is configured to: calculate the parameters required by the signal type judging module and the background noise parameter updating module according to the Vad_flag parameter determined by the PSC module and the sub-band energy parameters, ISF parameters, and open loop pitch parameters provided by the coder parameter extracting module; and provide the sub-band energy parameters, ISF parameters, open loop pitch parameters, and calculated parameters for the signal type judging module and the background noise parameter updating module. The parameters calculated by the classification parameter extracting module include:

1. Pitch parameter

[0074] Difference of continuous open loop pitch lags is compared. If the increment of the open loop pitch lag is less than a set threshold, the lag count accrues; if the sum of the lag counts of two continuous frames is great enough, the pitch is set to 1; otherwise, the pitch is set to 0. The formula for calculating the open loop pitch lag is specified in the AMR-WB+/AMR-WB standard document.

2. Longtime signal correlation value parameter (meangain)

[0075] The meangain is a moving average of tones of three adjacent frames, where tone = 1000*tone_flg. The definition of tone_flg is the same as that in the AMR-WB+.

[0076] 3. Zero Cross Rate (zcr)

$$zcr = \frac{1}{T} \sum_{i=1}^{T-1} \Pi\{x(i)x(i-1) < 0\}$$

$\Pi\{A\}$ is 1 when **A** is "truth", and is 0 when **A** is false.

[0077] 4. Time domain fluctuation of sub-band energy (t_flux)

$$t_flux = \frac{\sum_{i=1}^{12} |level_m(i) - level_{m-1}(i)|}{short_mean_level_energy}$$

where short_mean_level_energy represents short-time average energy.

[0078] 5. Ratio of high sub-band energy to low sub-band energy (ra)

$$ra = \frac{sublevel_high_energy}{sublevel_low_energy}$$

Given below is an instance of the present invention:

sublevel_high_energy = level[10]+ level[11];

sublevel_low_energy = level[0]+ level[1]+ level[2]+ level[3]+ level[4]+
level[5]+ level[6]+ level[7] + level[8]+ level[9];

[0079] 6. Frequency domain fluctuation of sub-band energy (f_flux)

$$f_flux = \frac{\sum_{i=2}^{12} |level_m(i) - level_m(i-1)|}{short_mean_level_energy}$$

[0080] 7. ISF mean short-time distance (isf_meanSD): average of ISF distance (Isf_SD) of five adjacent frames, where

$$Isf_SD = \sum_{i=1}^{16} |Isf_m(i) - Isf_{m-1}(i)|$$

[0081] 8. Sub-band energy standard deviation mean (level_meanSD) parameter: average of the sub-band energy standard deviation (level_SD) of two adjacent frames, where the calculation method of the level_SD parameter is similar to the calculation method of the Isf_SD described above.

[0082] In the foregoing 8 parameters, the parameters provided for the background noise parameter updating module include: zcr, ra, f_flux, and t_flux; the parameters provided for the signal type judging module include: pitch, meangain, isf_meanSD, and level_meanSD.

[0083] The signal type judging module is configured to sort the signals into non-useful(such as noise), speech, and music according to the snr and Vad_flag parameters received from the PSC module and the sub-band energy parameter, pitch, meangain, Isf_meansD, and level_meanSD parameters received from the classification parameter extracting module. The signal type judging module may include:

a parameter updating sub-module, configured to: update the threshold in the signal type judgment process according to the SNR, and provide the updated threshold for a judging sub-module; and

a judging sub-module, configured to: receive the sound signal type from the PSC module, determine the type of the useful signals in the sound signals based on the open loop pitch parameter, ISF parameter, sub-band energy parameter, and updated threshold, or based on the ISF parameter and sub-band energy parameter and the updated threshold, and send the determined type of the useful signals to the coder mode and rate selecting module.

[0084] The process of determining a useful signal to be a speech signal or music signal includes:

firstly, setting both the speech flag bit and the music flag bit to 0, sorting the signals into speech, music and uncertain signals primarily according to the pitch parameter flag, longtime signal correlation value, isf_meansD, and level_meanSD, and modifying the value of the speech flag bit or music flag bit according to the primarily determined speech or music;

secondly, correcting the primarily determined speech, music, and uncertain signals according to: sub-band energy, longtime signal correlation value, level_meanSD, speech_flag, music_flag, whether the number of continuous frames whose pitch value is 1 exceeds the preset hangover frame threshold, number of continuous music frames, number of continuous speech frames, and type of the previous frame; and determining the type of useful signals, where the type of a useful signal includes speech signal and music signal.

[0085] The process of determining a useful signal to be a speech signal or music signal is detailed below.

[0086] In order to ensure stability of judging signals and avoid frequent conversion of judgment results, this embodiment

provides a parameter flag hangover mechanism. The characteristic parameter values such as pitch_flag, level_meanSD_high_flag, ISF_meanSD_high_flag, ISF_meanSD_low_flag, level_meanSD_low_flag, and meangain_flag are determined according to the hangover mechanism, as shown in Figure 8.

[0087] In Figure 8, the length of the hangover period is determined according to the hangover parameter flag value. This embodiment provides two types of hangover settings (namely, two solutions to determining the hangover parameter flag value).

[0088] In the first hangover setting solution, when the parameter value is higher or lower than a threshold, the corresponding parameter hangover counter value increases by one; otherwise, the corresponding parameter hangover counter value is set to 0, and different parameter hangover flags are set according to the value of the parameter hangover counter. If the value of the parameter hangover counter is higher, the parameter hangover flag value is greater. The specific value is determined as required at the time of setting the parameter hangover flag value according to the parameter counter, and is not described here any further.

[0089] In the second hangover setting solution, the hangover length is controlled according to the Error Rate (ER) of the internal nodes of the decision tree corresponding to the training parameter. If the ER is lower, the hangover is shorter; if the ER is higher, the hangover is longer.

[0090] Afterwards, if the current signal is classified as a useful signal, the signal is primarily sorted into either speech or music:

[0091] Firstly, primary speech judgment is performed. As shown in Figure 9, in block 901, the speech flag bit is set to 0, and then in block 902, a judgment is made about whether the Isf_meanSD is greater than the first ISF speech threshold (such as 1500). If the Isf_meanSD is greater than the first ISF speech threshold, the speech flag bit is set to 1; otherwise, in block 903, a judgment is made about whether the pitch value is 1 and the pitch lag value (t_top_mean) obtained switching on and switching off the pitch search is less than the pitch speech threshold (such as 40). If yes, the speech flag bit is set to 1; otherwise, in block 904, a judgment is made about whether the number of continuous frames whose pitch value is 1 exceeds the preset threshold of the number of hangover frames (such as 2 frames). If yes, the speech flag bit is set to 1; otherwise:

in block 905, a judgment is made about whether the meangain exceeds the preset threshold of the longtime correlation speech (such as 8000). If yes, the speech flag bit is set to 1; otherwise,

in block 906, a judgment is made about whether either or both of the level_meanSD_high_flag value and the ISF_meanSD_high_flag value are 1. If yes, the speech flag bit is set to 1; otherwise, the value of the speech flag bit remains unchanged.

[0092] Afterwards, primary music judgment is performed, as shown in Figure 10:

[0093] In block 1001, the music flag bit is set to 0 first, and then in block 1002, a judgment is made about whether the signal fulfills both ISF_meanSD_low_flag = 1 and level_meanSD_low_flag = 1. If yes, the music signal flag (music_flag) is set; otherwise, the value of the music flag bit remains unchanged.

[0094] Afterwards, as shown in Figure 11, the primary judgment result is corrected:

[0095] In block 1101, a judgment is made about whether the instant energy of the sub-band is less than the sub-band energy threshold (such as 5000). If yes, the process proceeds to block 1102; otherwise, the signal is determined to be of the uncertain type.

[0096] In block 1102, a judgment is made about whether meangain_flag is 1 and the continuous count of music is less than the speech judgment threshold (such as 3) of continuous music count. If yes, the signal is determined to be a speech signal; otherwise,

[0097] in block 1103, a judgment is made about whether the ISF_meanSD value exceeds the preset threshold of the second ISF speech (such as 2000). If yes, the signal is determined to be a speech signal; otherwise,

[0098] in block 1104, a judgment is made about whether the level_energy is less than 10000 and more than five frames are previously determined to be noise. If yes, the current signal type is set to the uncertain type, with a view to reducing the probability of mistaking noise for music; otherwise,

[0099] in block 1105, a judgment is made about whether both the music flag bit and the speech flag bit are 1s. If yes, the current signal type is determined to be the uncertain type; otherwise,

[0100] in block 1106, a judgment is made about whether both the music flag bit and the speech flag bit are 0s. If yes, the current signal type is determined to be the uncertain type; otherwise,

[0101] in block 1107, a judgment is made about whether the music flag bit is 0 and the speech flag bit is 1. If yes, the current signal type is determined to be the speech type; otherwise,

[0102] in block 1108, because the music flag bit is 1 and the speech flag bit is 0, the current signal type is determined to be the music type.

[0103] After the signal is determined to be of the uncertain type in the foregoing blocks 1104, 1105 and 1106, block 1109 is performed to judge whether pitch_flag is 1, the ISF_meanSD is less than the ISF music threshold (such as 900), and the number of continuous speech frames is less than 3. If yes, the signal is determined to be of the music type; otherwise, the signal is still determined to be of the uncertain type.

[0104] After the signal is determined to be of the speech type in the foregoing blocks 1103 and 1107, block 1110 is performed to judge whether the number of continuous music frames is greater than 3 and the ISF_meanSD is less than the ISF music threshold. If yes, the signal is determined to be a music signal; otherwise, the signal is determined to be a speech signal.

[0105] After the speech signals and music signals are determined through the foregoing process, the signals of the uncertain type undergo the primary corrective classification process shown in Figure 12, including:

[0106] In block 1201, a judgment is made about whether the level_energy is less than the threshold (such as 5000) of the uncertain type of sub-band energy. If yes, the signal type is still determined to be the uncertain class; otherwise,

[0107] in block 1202, a judgment is made about whether the number of continuous music frames is greater than 1 and ISF_meanSD is less than the ISF music threshold. If yes, the signal is determined to be of the music class; otherwise,

[0108] the speech and music hangover flags are cleared. If the signals before this frame are continuous speech signals and the continuity is strong, the speech is judged according to the characteristic parameters of the speech. If the speech conditions are fulfilled, the speech_hangover_flag is set to 1, as illustrated in blocks 1203 to 1206 in Figure 12. If the signals before this frame are continuous music signals and the continuity is strong, the music is judged according to the characteristic parameters of the music. If the music conditions are fulfilled, the music_hangover_flag is set to 1, as illustrated in blocks 1207 to 1210 in Figure 12.

[0109] Afterwards, as illustrated in blocks 1211 to 1216 in Figure 12, if the speech hangover flag is 1 and the music hangover flag is 0, the current signal type is set to the speech class. If the music hangover flag is 1 and the speech hangover flag is 0, the current signal type is set to the music class. If both the music hangover flag and the speech hangover flag are 1 or both are 0, the signal type is set to the uncertain class. In this case, if more than 20 previous music frames are continuous, the signal is determined to be of the music class; if more than 20 previous speech frames are continuous, the signal is determined to be of the speech class.

[0110] After the foregoing primary correction is performed, the useful signal type is corrected finally in Figure 13. The type is further corrected according to the current context. In block 1301, if the current context is music and the continuity is longer than 3 seconds, namely, the current continuous music frames are more than 150 frames, mandatory correction may be performed according to the ISF_meanSD value to determine the music signal. In block 1302, if the current context is speech and the continuity is longer than 3 seconds, namely, the current continuous speech frames are more than 150 frames, mandatory correction may be performed according to the ISF_meanSD value to determine the speech signal class. Afterwards, if the signal type is still uncertain, the signal type is corrected according to the previous context in block 1303, namely, the current uncertain signal type is sorted into the previous signal type.

[0111] After the type of useful signals is determined in the foregoing process, the three type counters and the threshold values in the signal type judging module need to be updated. For the three type counters, if the current type is music (signal_sort = music), the music counter (music_continue_counter) increases by 1; otherwise, the music counter is cleared. Other type counters are processed similarly as shown in Figure 14, and are not detailed here any further. The threshold values are updated according to the SNR output by the PSC module. The threshold examples given in the embodiments herein are the values learned in the case that the SNR is 20 dB.

[0112] The background noise parameter updating module uses some spectral distribution parameters calculated in the classification parameter extracting module in the SAD to control the update rate of the background noise. In the actual application environment, the energy level of the background noise may surge abruptly. In this case, it is probable that the background noise estimation remains non-updated because the signals are continuously determined to be useful signals. Such a problem is solved by the background noise parameter updating module.

[0113] The background noise parameter updating module calculates the vector of relevant spectral distribution parameters according to the parameters received from the classification parameter extracting module. The vector includes the following elements:

- zero cross rate short-time mean (zcr_mean)
- short-time mean of ratio of high sub-band energy to low sub-band energy (RA)
- short-time mean of frequency domain fluctuation (f_flux) of sub-band energy
- short-time mean of time domain fluctuation (t_flux) of sub-band energy

where the zcr_mean is calculated in the following way, and other elements are calculated similarly:

$$zcr_mean_m = ALPHA \cdot zcr_mean_{m-1} + (1 - ALPHA) \cdot zcr_m$$

where ALPHA = 0.96 and m represents a frame index.

[0114] This embodiment makes use of the stable spectral features of the background noise. The elements of the

spectral distribution parameter vector are not limited to the 4 elements listed above. The update rate of the current background noise is controlled by a difference (d_{cb}) between the current spectral distribution parameter and the spectral distribution parameter estimation of the background noise. The difference may be implemented through the algorithms such as Euclidean distance and Manhattan distance. An instance of the present invention adopts the Manhattan distance (a distance calculation method similar to Euclidean distance):

$$d_{cb} = \sum_{i=1}^4 |p(i) - \tilde{p}(i)|$$

where p is the spectral distribution parameter vector of the current signal, and $\tilde{p}$ is the spectral distribution parameter vector estimation of background noise.

[0115] In an example of this embodiment, if $d_{cb} < \text{TH1}$, the module outputs an update rate acc1 , which represents the fastest update rate; otherwise, if $d_{cb} < \text{TH2}$, the module outputs an update rate acc2 ; otherwise, if $d_{cb} < \text{TH3}$, the module outputs an update rate acc3 ; otherwise, the module outputs an update rate acc4 . TH1, TH2, TH3 and TH4 are update thresholds, and the specific threshold values depend on the actual environment conditions.

[0116] In the embodiments of the present invention, the update rate of the background noise is determined, the noise parameters are updated according to the update rate, the signals are classified primarily according to the sub-band energy parameters and the updated noise parameters, and the non-useful signals and the useful signals in the received speech signals are determined, thus reducing the probability of mistaking useful signals for noise signals and improving accuracy of classifying sound signals.

[0117] It is understandable to those skilled in the art that the embodiments of the present invention may be implemented through software in addition to a universal hardware platform or through hardware only. In most cases, however, software in addition to a universal hardware platform is preferred. Therefore, the technical solution under the present invention or contributions to the related art may be embodied by a software product. The software product is stored in a storage medium and incorporates several instructions so that a computer device (for example, PC, server, or network device) may execute the method in each embodiment of the present invention.

[0118] Described above are preferred embodiments of the present invention. In practice, those skilled in the art may make modifications to the method under the present invention to meet the specific requirements. Although the invention has been described through some exemplary embodiments, the invention is not limited to such embodiments.

Claims

1. A method for classifying sound signals, comprising:

receiving the sound signals, and determining an update rate of background noise according to spectral distribution parameters of the background noise and spectral distribution parameters of the sound signals;
 updating noise parameters according to the update rate and classifying the sound signals into non-useful signals and useful signals according to sub-band energy parameters and the updated noise parameters;
 obtaining an Immittance Spectral Frequency, ISF, parameter and a sub-band energy parameter; and
for the useful signals obtained through classification, determining the type of these useful signals based on the Immittance Spectral Frequency parameter and the sub-band energy parameter, wherein the type of the useful signals comprises speech and music.

2. The method according to the preceding claim, **characterized by** determining a signal hangover length according to the determined type of the useful signals and classifying the sound signals further according to the signal hangover length.

3. The method according to one of the preceding claims, **characterized by** calculating a difference between the spectral distribution parameters of the sound signals and the spectral distribution parameters of the background noise and determining the update rate according to this difference.

4. An apparatus for classifying sound signals, comprising:

a background noise parameter updating module configured to determine an update rate of background noise according to spectral distribution parameters of the background noise and spectral distribution parameters of

the sound signals;

a primary signal classification, PSC, module configured to update noise parameters according to the updating rate received from the background noise parameter updating module, perform primary classification for the sound signals according to sub-band energy parameters and the updated noise parameters and determine the received sound signals to be useful signals or non-useful signals; and

a signal type judging module configured to determine, for the useful signals classified by the primary signal classification module, the type of these useful signals based on Immittance Spectral Frequency, ISF, parameters and the sub-band energy parameters wherein the type of the useful signals comprises speech and music.

5. The apparatus according to the preceding claim, **characterized by** a classification parameter extracting module configured to obtain ISF parameters and sub-band energy parameters or further obtain open loop pitch parameters, process the obtained parameters into signal type characteristic parameters and send the parameters to the signal type judging module, and process the obtained parameters into spectral distribution parameters of sound signals and background noise and transfer the spectral distribution parameters to the background noise parameter updating module, the PSC module being further configured to transfer the determined signal type to the signal type judging module through the classification parameter extracting module, and the signal type judging module being further configured to determine the type of useful signals according to the foregoing signal type characteristic parameter and the signal type determined by the PSC module.

6. The apparatus according to the preceding claim, **characterized by** a coder mode and rate selecting module configured to determine the coding mode and rate of sound signals according to the signal type transferred by and received from the signal type judging module.

7. The apparatus according to the claims 4 to 6, **characterized by** a coder parameter extracting module, which is configured to extract ISF parameters and sub-band energy parameters or additionally open loop pitch parameters, transfer the extracted parameters to the classification parameter extracting module and transfer the extracted sub-band energy parameters to the PSC module.

Patentansprüche

1. Verfahren zum Klassifizieren von Tonsignalen, umfassend:

Empfangen der Tonsignale und Bestimmen einer Aktualisierungsrate von Hintergrundgeräusch gemäß Spektralverteilungsparametern des Hintergrundgeräuschs und Spektralverteilungsparametern der Tonsignale; Aktualisieren von Geräuschparametern gemäß der Aktualisierungsrate und Klassifizieren der Tonsignale in nichtnützliche Signale und nützliche Signale gemäß Subbandenergieparametern und den aktualisierten Geräuschparametern; Erhalten eines Immittanzspektralfrequenz- bzw. ISF-Parameters und eines Subbandenergieparameters; und für die durch Klassifikation erhaltenen nützlichen Signale: Bestimmen des Typs dieser nützlichen Signale auf der Basis des Immittanzspektralfrequenzparameters und des Subbandenergieparameters, wobei der Typ der nützlichen Signale Sprache und Musik umfasst.

2. Verfahren nach dem vorhergehenden Anspruch, **gekennzeichnet durch** Bestimmen einer Signalüberhanglänge gemäß dem bestimmten Typ der nützlichen Signale und weiteres Klassifizieren der Tonsignale gemäß der Signalüberhanglänge.

3. Verfahren nach einem der vorhergehenden Ansprüche, **gekennzeichnet durch** Berechnen einer Differenz zwischen den Spektralverteilungsparametern der Tonsignale und den Spektralverteilungsparametern des Hintergrundgeräuschs und Bestimmen der Aktualisierungsrate gemäß dieser Differenz.

4. Vorrichtung zum Klassifizieren von Tonsignalen, umfassend:

ein Hintergrundgeräusch-Parameteraktualisierungsmodul, das dafür ausgelegt ist, eine Aktualisierungsrate von Hintergrundgeräusch gemäß Spektralverteilungsparametern des Hintergrundgeräuschs und Spektralverteilungsparametern der Tonsignale zu bestimmen;

ein Modul zur Primärsignalklassifikation PSC, das dafür ausgelegt ist, Geräuschparameter gemäß der von dem Hintergrundgeräusch-Parameteraktualisierungsmodul empfangenen Aktualisierungsrate zu aktualisieren, Pri-

märklassifikation für die Tonsignale gemäß den Subbandenergieparametern und den aktualisierten Geräuschparametern durchzuführen und die empfangenen Tonsignale als nützliche Signale oder nichtnützliche Signale zu bestimmen; und

ein Signaltyp-Beurteilungsmodul, das dafür ausgelegt ist, für die durch das Primärsignalklassifikationsmodul klassifizierten nützlichen Signale den Typ dieser nützlichen Signale auf der Basis von Immittanzspektralfrequenz- bzw. ISF-Parametern und den Subbandenergieparametern zu bestimmen, wobei der Typ der nützlichen Signale Sprache und Musik umfasst.

5. Vorrichtung nach dem vorhergehenden Anspruch, **gekennzeichnet durch** ein Klassifikationsparameter-Extraktionsmodul, das dafür ausgelegt ist, ISF-Parameter und Subbandenergieparameter zu erhalten oder ferner Open-Loop-Pitch-Parameter zu erhalten, die erhaltenen Parameter zu Signaltypcharakteristikparametern zu verarbeiten und die Parameter zu dem Signaltyp-Beurteilungsmodul zu senden und die erhaltenen Parameter zu Spektralverteilungsparametern von Tonsignalen und Hintergrundgeräusch zu verarbeiten und die Spektralverteilungsparameter zu dem Hintergrundgeräusch-Parameteraktualisierungsmodul zu transferieren, wobei das PSC-Modul ferner dafür ausgelegt ist, den bestimmten Signaltyp **durch** das Klassifikationsparameter-Extraktionsmodul zu dem Signaltyp-Beurteilungsmodul zu transferieren und das Signaltyp-Beurteilungsmodul ferner dafür ausgelegt ist, den Typ nützlicher Signale gemäß dem obigen Signaltypcharakteristikparameter und dem **durch** das PSC-Modul bestimmten Signaltyp zu bestimmen.

6. Vorrichtung nach dem vorhergehenden Anspruch, **gekennzeichnet durch** ein Codierermodus- und Ratenauswahlmodul, das dafür ausgelegt ist, den Codierungsmodus und die Rate von Tonsignalen gemäß dem **durch** das Signaltyp-Beurteilungsmodul transferierten und von diesem empfangenen Signaltyp zu bestimmen.

7. Vorrichtung nach Anspruch 4 bis 6, **gekennzeichnet durch** ein Codiererparameter-Extraktionsmodul, das dafür ausgelegt ist, ISF-Parameter und Subbandenergieparameter oder zusätzlich Open-Loop-Pitch-Parameter zu extrahieren, die extrahierten Parameter zu dem Klassifikationsparameter-Extraktionsmodul zu transferieren und die extrahierten Subbandenergieparameter zu dem PSC-Modul zu transferieren.

Revendications

1. Procédé de classement de signaux sonores, comprenant :

la réception des signaux sonores et la détermination d'un débit de mise à jour du bruit de fond en fonction de paramètres de distribution spectrale du bruit de fond et de paramètres de distribution spectrale des signaux sonores,

la mise à jour de paramètres de bruit en fonction du débit de mise à jour et le classement des signaux sonores en signaux inutiles et en signaux utiles en fonction de paramètres d'énergie de sous-bande et des paramètres de bruit mis à jour,

la récupération d'un paramètre de fréquence spectrale d'immittance, ISF, et d'un paramètre d'énergie de sous bande, et

pour les signaux utiles obtenus grâce au classement, la détermination du type de ces signaux utiles sur la base du paramètre de fréquence spectrale d'immittance et du paramètre d'énergie de sous bande, le type des signaux utiles comprenant de la parole et de la musique.

2. Procédé selon la revendication précédente, **caractérisé par** la détermination d'une longueur de traînage de signal en fonction du type déterminé des signaux utiles et par le classement des signaux sonores plus avant en fonction de la longueur de traînage de signal.

3. Procédé selon l'une des revendications précédentes, **caractérisé par** le calcul d'une différence entre les paramètres de distribution spectrale des signaux sonores et les paramètres de distribution spectrale du bruit de fond, et par la détermination du débit de mise à jour en fonction de cette différence.

4. Appareil de classement de signaux sonores, comprenant :

un module de mise à jour de paramètres de bruit de fond configuré pour déterminer un débit de mise à jour du bruit de fond en fonction de paramètres de distribution spectrale du bruit de fond et de paramètres de distribution spectrale des signaux sonores,

un module de classement primaire de signaux, PSC, configuré pour mettre à jour les paramètres de bruit en fonction du débit de mise à jour reçu du module de mise à jour de paramètres de bruit de fond, pour exécuter un classement primaire des signaux sonores en fonction des paramètres d'énergie de sous bande et des paramètres de bruit mis à jour, et pour déterminer que les signaux sonores reçus sont des signaux utiles ou des signaux inutiles, et

un module d'évaluation de type de signal configuré pour déterminer, pour les signaux utiles classés par le module de classement primaire de signaux, le type de ces signaux utiles sur la base de paramètres de fréquence spectrale d'immitance, ISF, et des paramètres d'énergie de sous bande, le type des signaux utiles comprenant de la parole et de la musique.

5. Appareil selon la revendication précédente, **caractérisé par** :

un module d'extraction de paramètres de classement, configuré pour récupérer les paramètres de fréquence ISF et les paramètres d'énergie de sous bande ou pour récupérer en outre des paramètres de hauteur tonale en boucle ouverte, pour traiter les paramètres obtenus en paramètres de caractéristiques du type de signal et pour envoyer les paramètres au module d'évaluation de type de signal, ainsi que pour traiter les paramètres obtenus en paramètres de distribution spectrale de signaux sonores et de bruit de fond et pour transférer les paramètres de distribution spectrale vers le module de mise à jour des paramètres de bruit de fond,

le module de classement PSC étant en outre configuré pour transférer le type de signal déterminé vers le module d'évaluation de type de signal par l'intermédiaire du module d'extraction de paramètres de classement, et le module d'évaluation de type de signal étant en outre configuré pour déterminer le type de signaux utiles en fonction du paramètre de caractéristiques de type de signal précédent et du type de signal déterminé par le module de classement PSC.

6. Appareil selon la revendication précédente, **caractérisé par** un module de sélection de débit et de mode de codeur configuré pour déterminer le mode de codage et le débit de signaux sonores en fonction du type de signal transféré par le module d'évaluation de type de signal et reçu de celui-ci

7. Appareil selon les revendications 4 à 6, **caractérisé par** un module d'extraction de paramètres de codeur qui est configuré pour extraire les paramètres de fréquence ISF et les paramètres d'énergie de sous bande ou en plus les paramètres de hauteur tonale en boucle ouverte, pour transférer les paramètres extraits vers le module d'extraction de paramètres de classement et pour transférer les paramètres d'énergie de sous bande extraits vers le module de classement PSC.

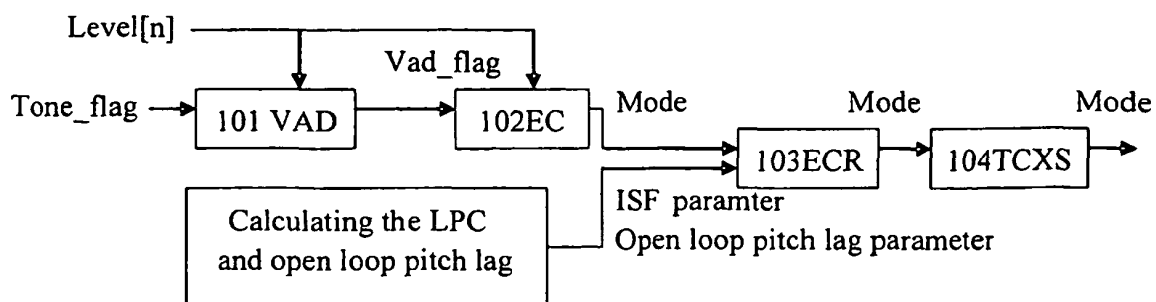


FIG 1

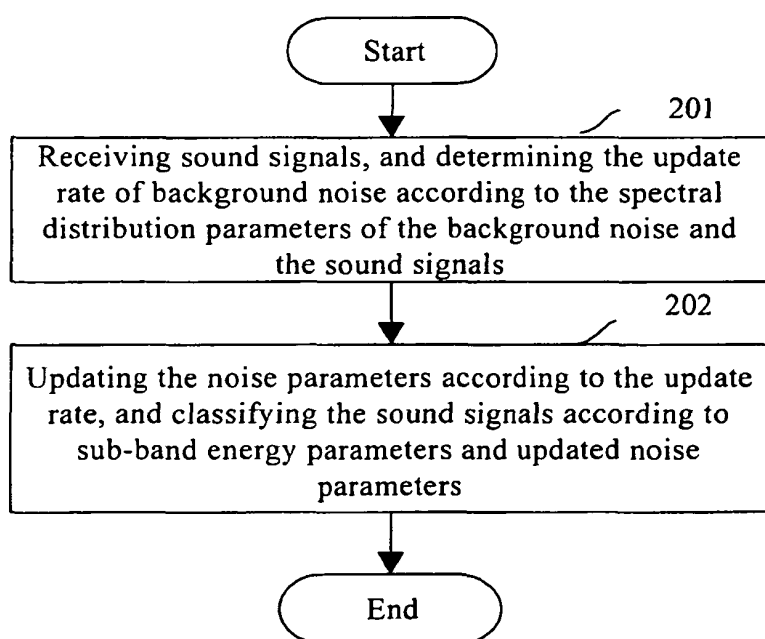


FIG 2

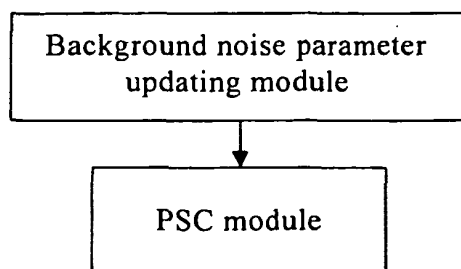


FIG 3

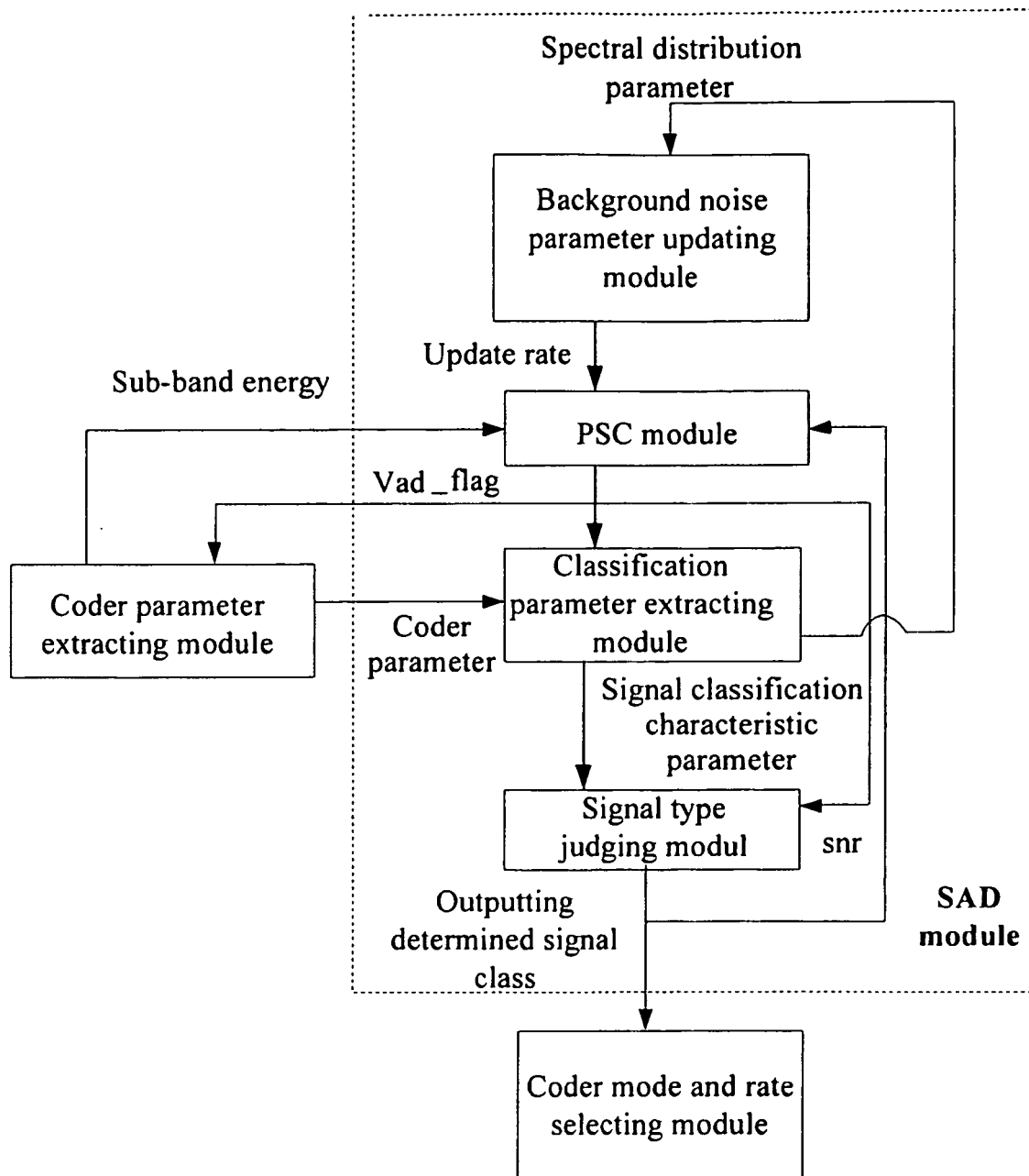


FIG 4

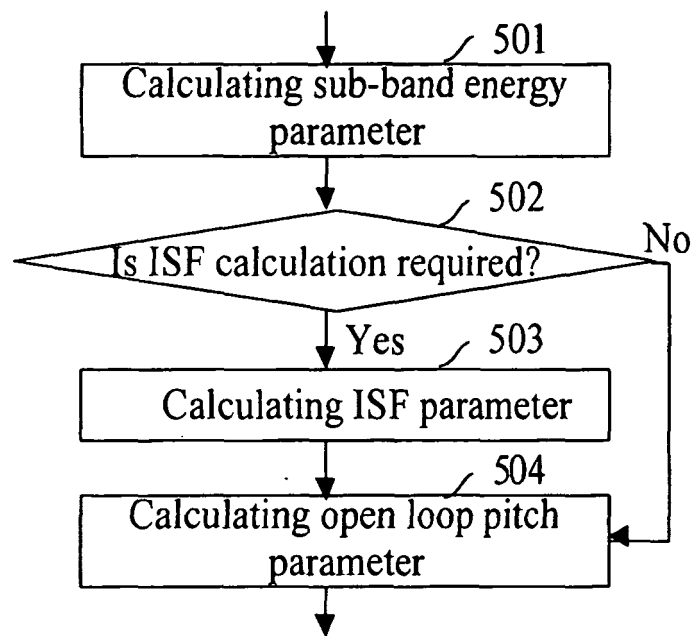


FIG 5

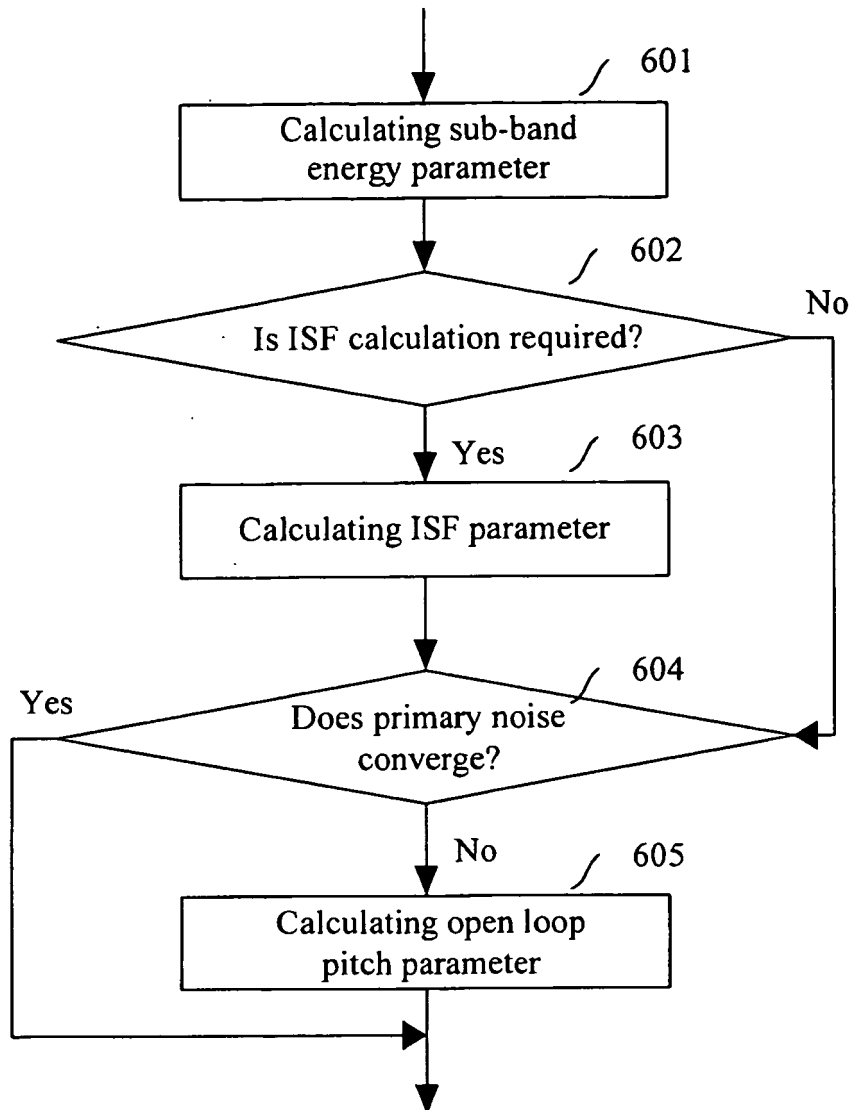


FIG 6

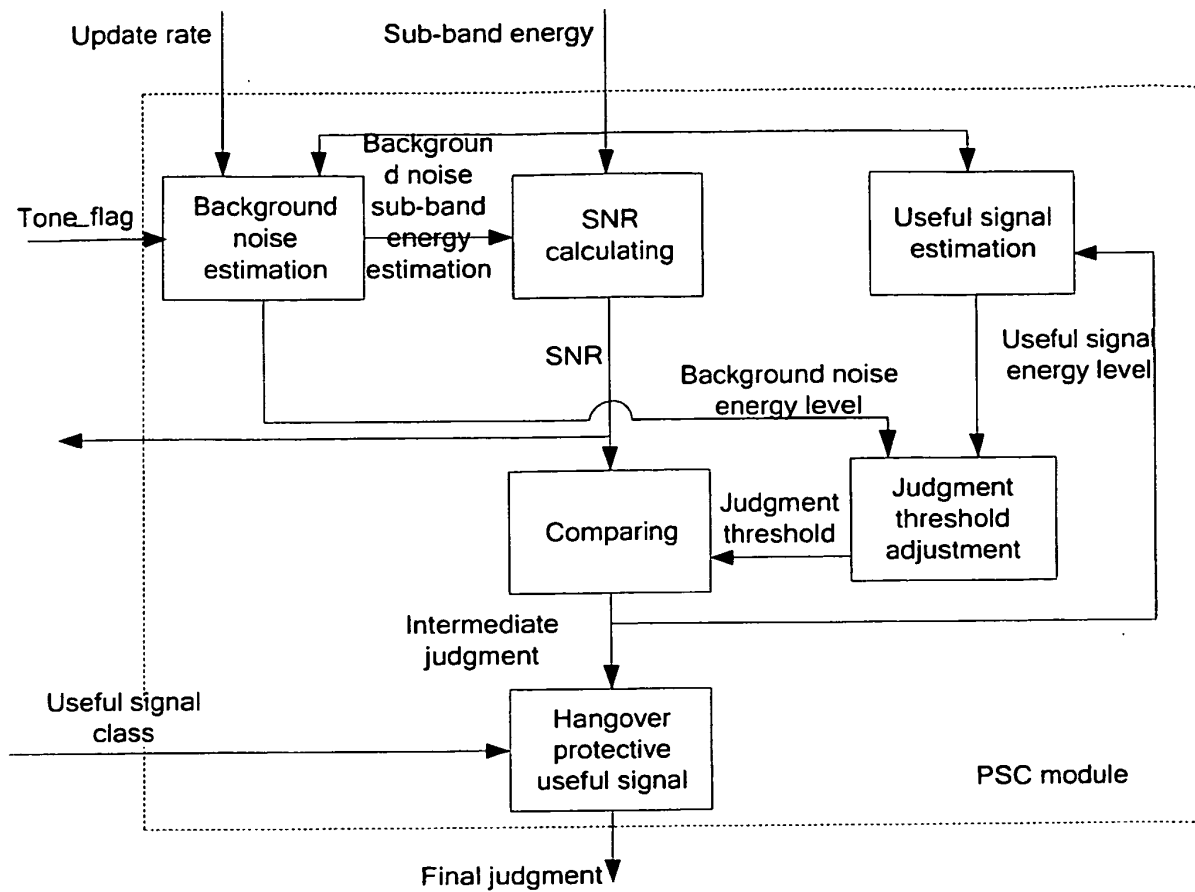


FIG 7

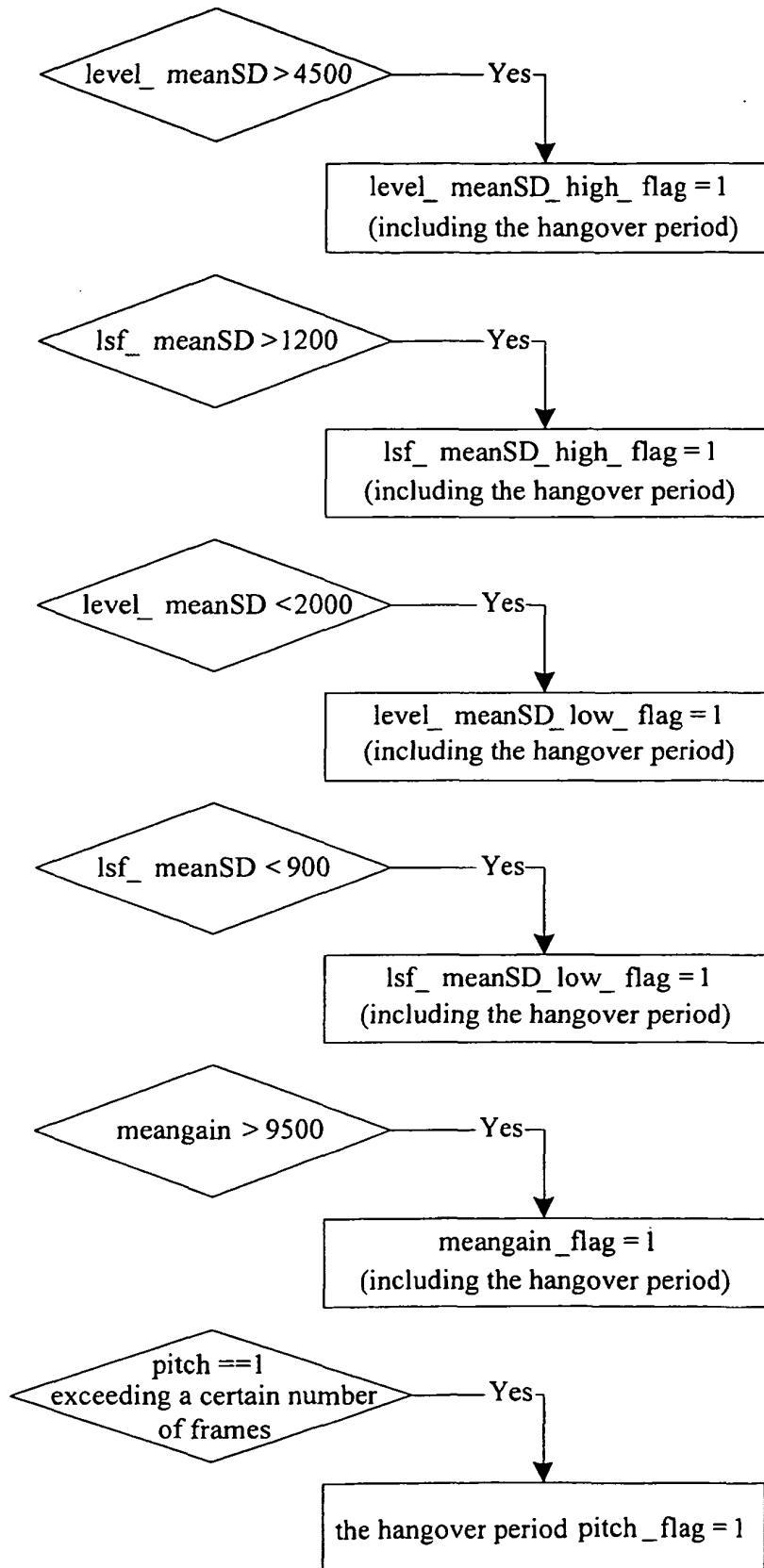


FIG 8

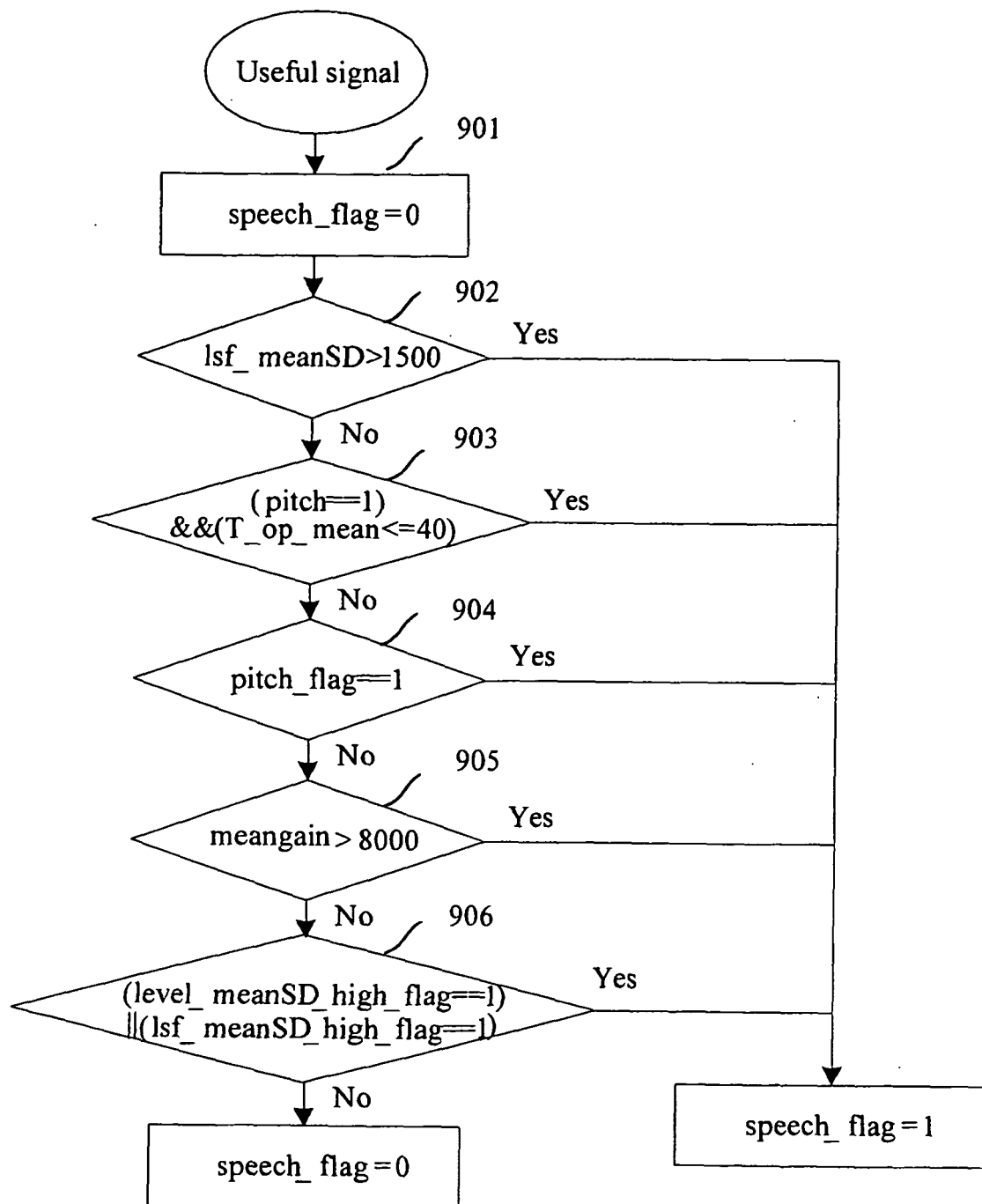


FIG 9

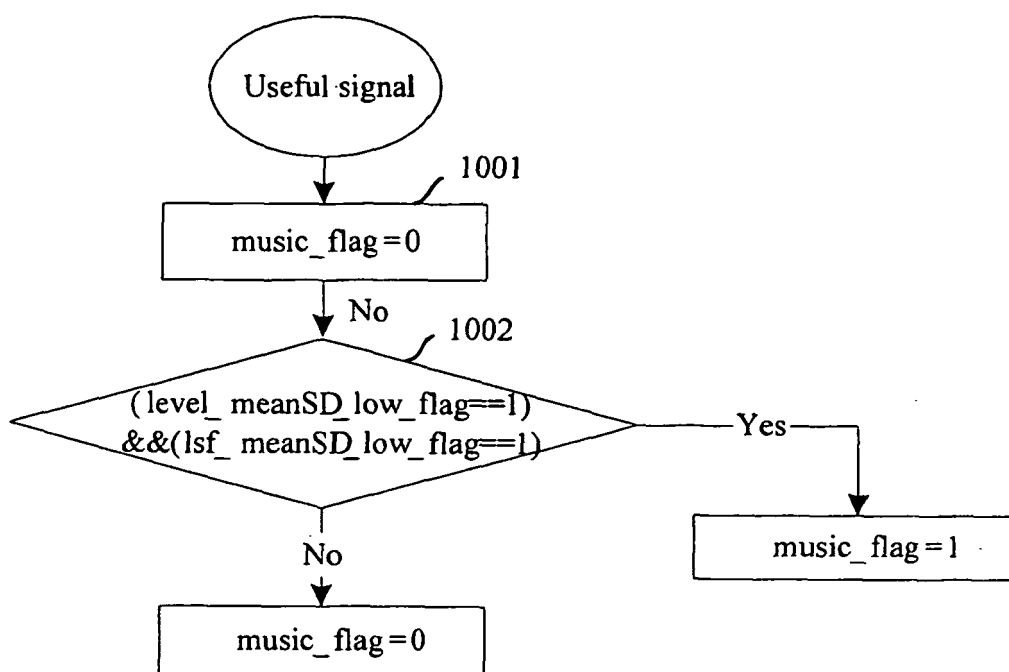


FIG 10

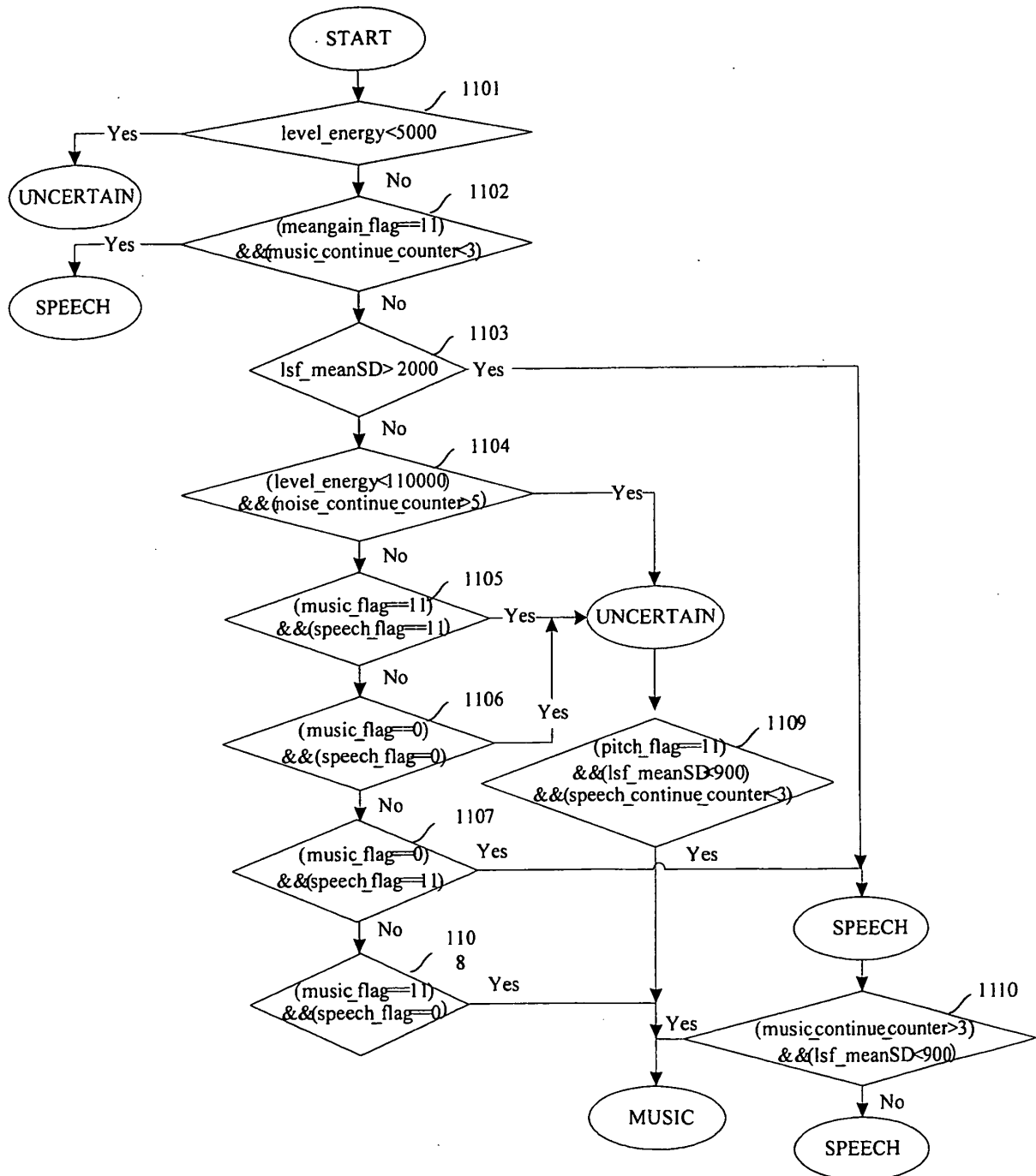


FIG 11

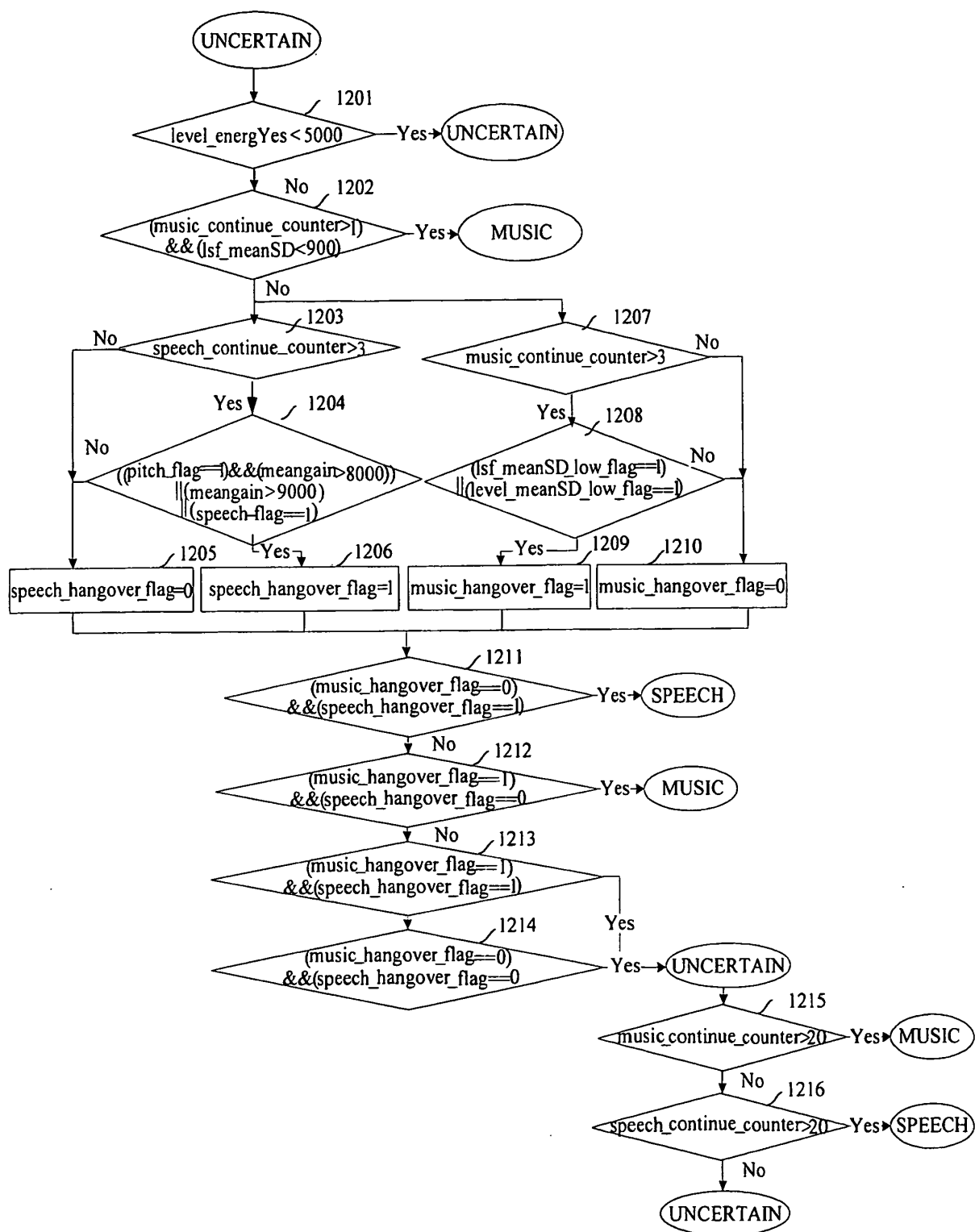


FIG 12

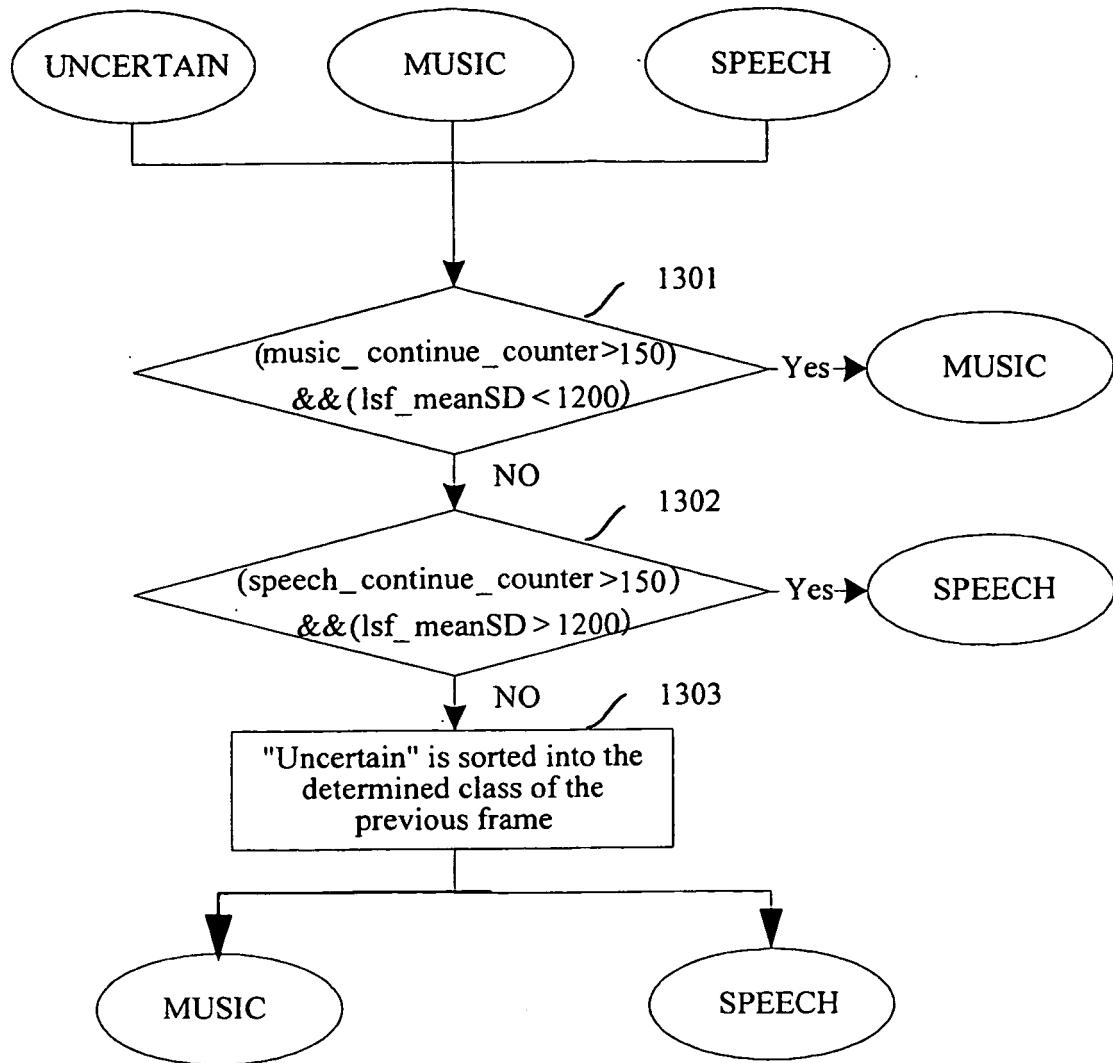


FIG 13

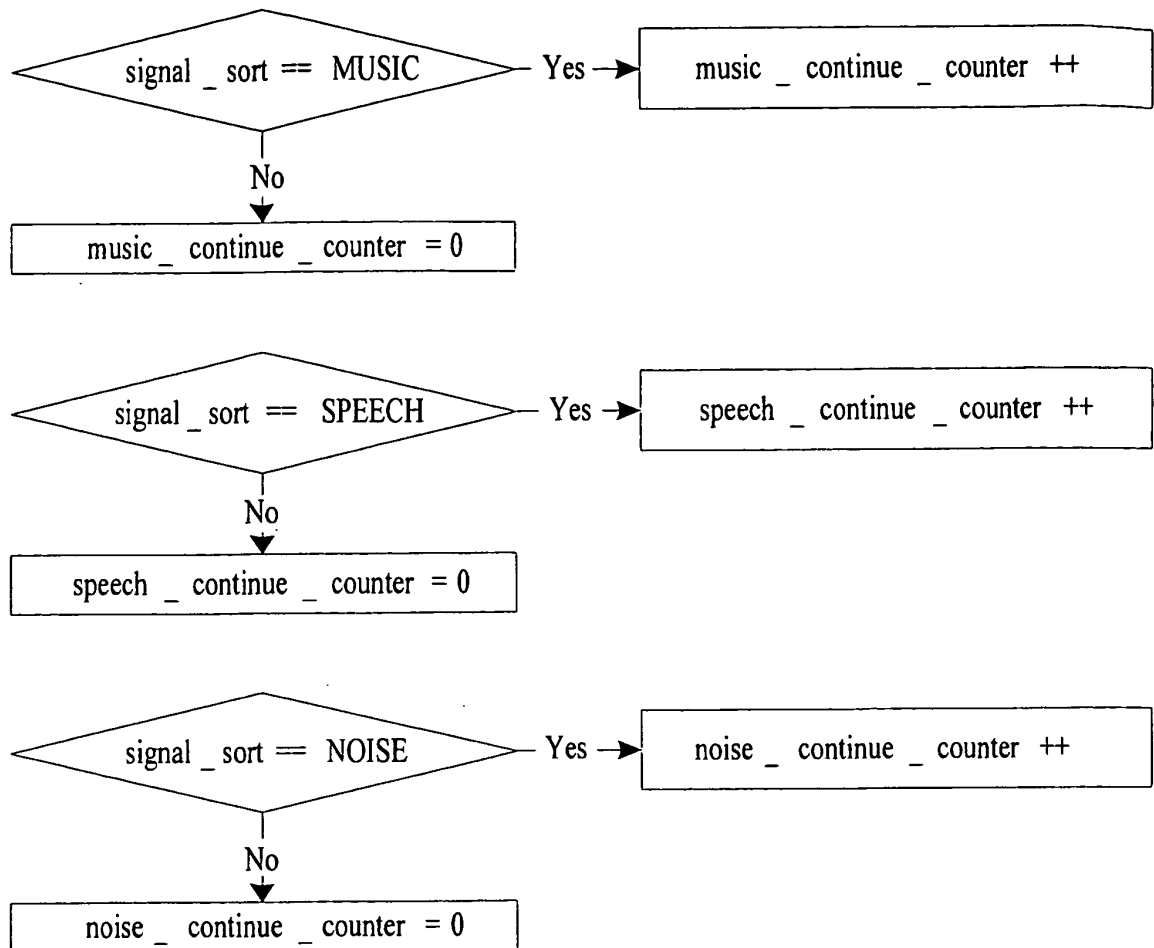


FIG 14

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

- WO 02065457 A2 [0014]

Non-patent literature cited in the description

- **JELINEK M et al.** *Robust signal/noise discrimination for wideband speech and audio coding* [0012]
- ETSI TS 126290 V6.3.0, June 2005 [0013]