AUDIO SIGNAL PROCESSING DEVICE, AUDIO SIGNAL PROCESSING METHOD, AND AUDIO SIGNAL PROCESSING PROGRAM

Technical Field

[0001] The present invention relates to an audio signal processing device, an audio signal processing method, and an audio signal processing program for processing an audio signal.

Background Art

[0002] In transmission of a coded and packetized audio signal through an Internet network with an IP (Internet Protocol) phone, a packet can be lost because of a network congestion or the like (this phenomenon will be referred to hereinafter as "packet loss"). With an occurrence of a packet loss, necessary audio codes are lost resulting in a failure in decoding of audio, thereby causing an audio discontinuity. A technology for preventing an audio discontinuity caused by a packet loss is an audio packet loss concealment technology. The audio packet loss concealment technology is designed to detect a packet loss and generate a pseudo audio signal corresponding to the lost packet (which will be referred to hereinafter as "concealment signal").

[0003] When an audio encoding technique used is a technique of performing audio encoding while updating internal states of encoder/decoder, encoding parameters to be originally received are not obtained and thus the audio packet loss concealment technology includes performing an update of the internal states of the decoder by

use of artificially-generated parameters as well.

[0004] The CELP (Code Excited Linear Prediction) encoding is widely used as a technique for performing the audio encoding while updating the internal states of encoder/decoder. In the CELP encoding, an autoregressive model is assumed, and an excitation signal e(n) is filtered by an all-pole synthesis filter a(i) to synthesize an audio signal. Namely, the audio signal s(n) is synthesized according to the below equation. In the equation below, a(i) represents linear prediction coefficients (LP (Linear Prediction) coefficients) and the degree to be used is a value such as P=16.

[Mathematical Equation 1]

$$s(n) = e(n) - \sum_{i=1}^{P} a(i) \cdot s(n-i)$$

[0005] In the CELP encoding, the internal states stored include ISF (Immittance Spectral Frequency) parameters as mathematically equivalent representation of the linear prediction coefficients, and a past excitation signal. With an occurrence of a packet loss, these are artificially generated, and there arises a deviation from the original parameters that would be obtained by decoding. An inconsistency of a synthesized audio caused by a deviation of the parameters is perceived as a noise by a listener, which significantly degrades the subjective quality.

[0006] The paragraphs below will describe a configuration and an operation of an audio decoder to perform the audio packet loss concealment, using an example where the CELP encoding is used as the audio encoding technique.

[0007] A configuration diagram and an operation of the audio decoder are shown in Fig. 1 and Fig. 2. As shown in Fig. 1, an audio decoder 1 has a packet loss detector 11, an audio code decoder 12, a concealment signal generator 13, and an internal state buffer 14.

[0008] The packet loss detector 11, when receiving an audio packet correctly , sends a control signal, and audio codes included in the audio packet, to the audio code decoder 12 (normal reception: YES in step S100 in Fig. 2). Thereafter, the audio code decoder 12 performs decoding of the audio codes and updating of the internal states as described below (steps S200 and S400 in Fig. 2). On the other hand, the packet loss detector 11, when failing to receive an audio packet correctly, sends a control signal to the concealment signal generator 13 (packet loss: NO in step S100 in Fig. 2). Thereafter, the concealment signal generator 13 generates a concealment signal and updates the internal states as described below (steps S300 and S400 in Fig. 2). The processes of steps S100 to S400 in Fig. 2 are repeated to the end of communication (or until step S500 results in a determination of YES).

[0009] The audio codes include at least encoded ISF parameters

[Mathematical Equation 2]

$$\dot{\omega}_i$$,

encoded pitch lags T^j_p of the first to fourth subframes, encoded adaptive codebook gains g^j_p of the first to fourth subframes, encoded fixed codebook gains g^j_c of the first to fourth subframes, and encoded fixed codebook vectors $c^j(n)$ of the first to fourth subframes. The ISF parameters may be replaced by LSF (line spectral frequency) parameters which are mathematically equivalent representation thereof.

Although the discussion below uses the ISF parameters, the same discussion may also be true for the case using the LSF parameters.

[0010] The internal state buffer includes past ISF parameters

[Mathematical Equation 3]

$\dot{\omega}_i^{-1}$

and, as equivalent representation of

[Mathematical Equation 4]

$\dot{\omega}_i^{-1}$,

ISP (Immittance Spectral Pair) parameters

[Mathematical Equation 5]

$\dot{\mathbf{q}}_i^{-1}$,

ISF residual parameters

[Mathematical Equation 6]

$\dot{r}_i^{-1}$,

past pitch lags T^j_p, past adaptive codebook gains g^j_p, past fixed codebook gains g^j_c, and an adaptive codebook u(n). It is determined, depending upon a design principle, how many subframes of the past parameters should be included. It is assumed in the present specification that one frame includes four subframes, but another value may be adopted depending upon the design principle.

[0011] <Case of Normal Reception>

Fig. 3 shows an exemplary functional configuration of the audio code decoder 12. As shown in this Fig. 3, the audio code decoder 12

has an ISF decoder 120, a stability processor 121, an LP coefficient calculator 122, an adaptive codebook calculator 123, a fixed codebook decoder 124, a gain decoder 125, an excitation vector synthesizer 126, a post-filter 127, and a synthesis filter 128. It should be noted, however, that the post-filter 127 is not an indispensable constitutive element. In Fig. 3, for convenience of explanation, the internal state buffer 14 is indicated by a double-dot line inside the audio code decoder 12. However, the internal state buffer 14 is not included inside the audio code decoder 12, but is indeed the internal state buffer 14 itself shown in Fig. 1. The same is also true in the configuration diagrams of the audio code decoder hereinafter.

[0012] A configuration diagram of the LP coefficient calculator 122 is shown in Fig. 4 and a processing flow of calculation of LP coefficients from the encoded ISF parameters is shown in Fig. 5. As shown in Fig. 4, the LP coefficient calculator 122 has an ISF-ISP converter 122A, an ISP interpolator 122B, and an ISP-LPC converter 122C.

[0013] First described are a functional configuration and its operation associated with the process of calculating the LP coefficients from the encoded ISF parameters (Fig. 5).

[0014] The ISF decoder 120 decodes the encoded ISF parameters to obtain the ISF residual parameters

[Mathematical Equation 7]

$$\dot{r}_i^0$$

and calculates the ISF parameters

[Mathematical Equation 8]

$\dot{\omega}_i$

in accordance with the following equation (step S1 in Fig. 5). Here, $mean_i$ represents mean vectors obtained in advance by learning or the like.

[Mathematical Equation 9]

$$\dot{\omega}_i = mean_i + \dot{r}_i^0 + \frac{1}{3}\dot{r}_i^{-1}$$

[0015] The example of using an MA prediction for the calculation of the ISF parameters is described herein, but it is also possible to adopt a configuration to perform calculation of the ISF parameters using an AR prediction as described below. Here, the ISF parameters of the immediately preceding frame are denoted by

[Mathematical Equation 10]

$$\dot{\omega}_i^{-1}$$

and weight factors of the AR prediction by ρ_i.

[Mathematical Equation 11]

$$\dot{\omega}_i = mean_i + \rho_i(\dot{\omega}_i^{-1} - mean_i)$$

[0016] The stability processor 121 performs a process according to the below equation so as to place a distance of not less than 50 Hz between elements of the ISF parameters in order to secure stability of the filter (step S2 in Fig. 5). The ISF parameters are indicative of a line spectrum representing the shape of an audio spectrum envelope, and as the distance between them becomes shorter, peaks of the spectrum

become larger, causing resonance. For this reason, the process for securing stability becomes necessary to prevent gains from becoming too large at the peaks of the spectrum. Here, min_dist represents a minimum ISF distance, and isf_min represents a minimum of ISF necessary for securing the distance of min_dist. isf_min is successively updated by adding the distance of min_dist to a value of neighboring ISF. On the other hand, isf_max represents a maximum of ISF necessary for securing the distance of min_dist. isf_max is successively updated by subtracting the distance of min_dist from a value of neighboring ISF.

[Mathematical Equation 12]

```
isf_min = min_dist = 50
for i=0 to 14
        if ω̇_i<isf_min then ω̇_i=isf_min
        isf_min = ω̇_i + min_dist
isf_max = 6400 - min_dist
if ω̇_14>isf_max
        for i=14 down to 1
                if ω̇_i > isf_max then ω̇_i=isf_max
                isf_max = ω̇_i - min_dist
```

[0017] The ISF-ISP converter 122A in the LP coefficient calculator 122 converts

[Mathematical Equation 13]

$\dot{\omega}_i$

into ISP parameters

[Mathematical Equation 14]

$$\dot{\mathbf{q}}_i$$

in accordance with the following equation (step S3 in Fig. 5). Here, C is a constant determined in advance.

[Mathematical Equation 15]

$$\dot{\mathbf{q}}_i = \cos\left(C \cdot \dot{\omega}_i\right)$$

[0018] The ISP interpolator 122B calculates the ISP parameters for the respective subframes from the past ISP parameters

[Mathematical Equation 16]

$$\dot{\mathbf{q}}_i^{-1}$$

included in the internal state buffer 14 and the foregoing ISP parameters

[Mathematical Equation 17]

$$\dot{\mathbf{q}}_i$$

in accordance with the below equation (step S4 in Fig. 5). Other coefficients may be used for the interpolation.

[Mathematical Equation 18]

$$\mathbf{q}_i^{(1)} = 0.75 \cdot \dot{\mathbf{q}}_i^{-1} + 0.25 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(2)} = 0.5 \cdot \dot{\mathbf{q}}_i^{-1} + 0.5 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(3)} = 0.25 \cdot \dot{\mathbf{q}}_i^{-1} + 0.75 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(4)} = \dot{\mathbf{q}}_i$$

[0019] The ISP-LPC converter 122C converts the ISP parameters for the respective subframes into LP coefficients

[Mathematical Equation 19]

$$\dot{a}_i^j \quad (0 < i \le P \text{、} \ 0 \le j < 4)$$

(step S5 in Fig. 5). A specific conversion procedure to be used can be

the processing procedure described in Non Patent Literature 1. The number of subframes included in a look-ahead signal is assumed to be 4 herein, but the number of subframes may differ, depending upon the design principle.

[0020] Next described are other configurations and operations in the audio code decoder 12.

[0021] The adaptive codebook calculator 123 decodes encoded pitch lags to calculate the pitch lags T^j_P of the first to fourth subframes. Then, the adaptive codebook calculator 123 uses the adaptive codebook u(n) to calculate adaptive codebook vectors for the respective subframes in accordance with the below equation. The adaptive codebook vectors are calculated by interpolating the adaptive codebook u(n) by a FIR filter Int(i). Here, the length of the adaptive codebook is denoted by N_{adapt}. The filter Int(i) used for the interpolation is an FIR filter with a predetermined length $2l + 1$, and L' presents the sample number of the subframes. By using the interpolation filter Int(i), the pitch lags can be utilized to the accuracy of decimal places. For the details of the interpolation filter, the method described in Non Patent Literature 1 can be referred to.

[Mathematical Equation 20]

$$v^j(n) = \sum_{i=-l}^{l} Int(i) \cdot u(n + N_{adapt} - \hat{T}_p^{(j)} + i) \qquad (0 \le n < L')$$

[0022] The fixed codebook decoder 124 decodes the encoded fixed codebook vectors to acquire the fixed codebook vectors $c^j(n)$ of the first to fourth subframes.

[0023] The gain decoder 125 decodes the encoded adaptive codebook

gains and the encoded fixed codebook gains to acquire the adaptive codebook gains and fixed codebook gains of the first to fourth subframes. For example, the decoding of the adaptive codebook gains and the fixed codebook gains can be carried out, for example, by the below technique described in Non Patent Literature 1. Since the below technique described in Non Patent Literature 1 does not use the interframe prediction as used in gain encoding of AMR-WB, it can enhance packet loss resistance.

[0024] For example, the gain decoder 125 acquires the fixed codebook gain in accordance with the below processing flow.

[0025] First, the gain decoder 125 calculates the power of the fixed codebook vector. Here, the length of the subframe is defined as N_s.

[Mathematical Equation 21]

$$E_c = 10\log\left(\frac{1}{N_s}\sum_{i=0}^{N_s-1} c^2(i)\right)$$

[0026] Next, the gain decoder 125 decodes the vector-quantized gain parameter to acquire the adaptive codebook gain

[Mathematical Equation 22]

$$\hat{g}_p$$

and the quantized fixed codebook gain

[Mathematical Equation 23]

$$\hat{E}_i$$

.

It then calculates a predictive fixed codebook gain as described below from the quantized fixed codebook gain and the aforementioned power of the fixed codebook vector.

[Mathematical Equation 24]

$$g_c' = 10^{0.05\left(\hat{E}_i - E_c\right)}$$

[0027] Finally, the gain decoder 125 decodes the prediction coefficient

[Mathematical Equation 25]

$\hat{\gamma}$

and multiplies it to the prediction gain to acquire the fixed codebook gain.

[Mathematical Equation 26]

$$\hat{g}_c = \hat{\gamma} \cdot g_c'$$

[0028] The excitation vector synthesizer 126 multiplies the adaptive codebook vector by the adaptive codebook gain and multiplies the fixed codebook vector by the fixed codebook gain and calculates a sum of them to acquire an excitation signal, as expressed by the following equation.

[Mathematical Equation 27]

$$e^j(n) = g_p^j \cdot v^j(n) + g_c^j \cdot c^j(n)$$

[0029] The post-filter 127 subjects the excitation signal vectors, for example, to post-processes such as processes of pitch enhancement, noise enhancement, and low-frequency enhancement. The pitch enhancement, the noise enhancement, and the low-frequency enhancement can be effected by use of the techniques described in Non Patent Literature 1.

[0030] The synthesis filter 128 synthesizes a decoded signal with the excitation signal as a drive audio source, by linear prediction inverse

filtering.

[Mathematical Equation 28]

$$\hat{s}(n) = e^{j}(n) - \sum_{i=1}^{P} \dot{a}_i^j \cdot \hat{s}(n-i)$$

[0031] If a pre-emphasis is done in the encoder, a de-emphasis is carried out.

[Mathematical Equation 29]

$$\hat{s}_{de-emph}(n) = \hat{s}(n) + \beta \cdot \hat{s}(n-1)$$

[0032] On the other hand, if a pre-emphasis is not done in the encoder, a de-emphasis is not carried out.

[0033] The paragraphs below will describe the operation concerning an internal state update.

[0034] In order to interpolate parameter upon an occurrence of packet loss, the LP coefficient calculator 122 updates the internal states of the ISF parameters by vectors calculated by the following equation.

[Mathematical Equation 30]

$$\vec{\omega}_i = \beta\omega_i^C + (1-\beta)\frac{\omega_i^{(-3)} + \omega_i^{(-2)} + \omega_i^{(-1)}}{3}$$

[0035] Here, $\omega_i^{(-j)}$ represents the ISF parameters j frames prior, which are stored in the buffer. ω_i^C represents the ISF parameters in speech intervals obtained in advance by learning or the like. β is a constant and can be a value of, e.g., 0.75, to which the value is not necessarily limited. ω_i^C and β may be varied by an index to express a property of an encoding target frame, for example, as in the ISF concealment described in Non Patent Literature 1.

[0036] Furthermore, the LP coefficient calculator 122 also updates the internal states of the ISF residual parameters in accordance with the following equation.

[Mathematical Equation 31]

$$\dot{r}_i^{-1} = \dot{r}_i^0$$

[0037] The excitation vector synthesizer 126 updates the internal states by the excitation signal vectors in accordance with the below equation.

[Mathematical Equation 32]

$$\begin{aligned} u(n) &= u(n+L) && (0 \le n < N-L) \\ u(n+N-L+jL') &= e^j(n) && (0 \le n < L') \end{aligned}$$

[0038] Furthermore, the excitation vector synthesizer 126 updates the internal states of the gain parameters by the following equation.

[Mathematical Equation 33]

$$g_c^{(-M_{la}+j)} = g_c^j$$

[0039] The adaptive codebook calculator 123 updates the internal states of the parameters of the pitch lags by the following equation.

[Mathematical Equation 34]

$$T_p^{(-M_{la}+j)} = T_p^j$$

The range of j is defined as ($-2 \le j < M_{la}$) but different values may be selected as the range of j, depending upon the the design principle.

[0040] <Case of Packet Loss>

Fig. 6 shows an exemplary functional configuration of the concealment signal generator 13. As shown in this Fig. 6, the

concealment signal generator 13 has an LP coefficient interpolator 130, a pitch lag interpolator 131, a gain interpolator 132, a noise signal generator 133, a post-filter 134, a synthesis filter 135, an adaptive codebook calculator 136, and an excitation vector synthesizer 137. It should be noted, however, that the post-filter 134 is not an indispensable constitutive element.

[0041] The LP coefficient interpolator 130 calculates

[Mathematical Equation 35]

$$\dot{\omega}_i$$

by the following equation. In this respect, $\omega_i^{(-j)}$ represents the ISF parameters j frames prior, which are stored in the buffer.

[Mathematical Equation 36]

$$\dot{\omega}_i = \alpha\omega_i^{(-1)} + (1-\alpha)\vec{\omega}_i$$

In this equation,

[Mathematical Equation 37]

$$\vec{\omega}_i$$

represents the internal states of the ISF parameters calculated upon normal reception of a packet. α is also a constant and can be a value of, e.g., 0.9 to which the value is not necessarily limited. α may be varied by an index to express a property of an encoding target frame, for example, as in the ISF concealment described in Non Patent Literature 1.

[0042] The procedure of obtaining the LP coefficients from the ISF parameters is the same as performed in the case of normal reception of a

packet.

[0043] The pitch lag interpolator 131 uses the internal state parameters about the pitch lags

[Mathematical Equation 38]

$$T_p^{(-M_{la}+j)}$$

to calculate predictive values of the pitch lags

[Mathematical Equation 39]

$$\hat{T}_p \quad .$$

A specific processing procedure to be used can be the technique disclosed in Non Patent Literature 1.

[0044] In order to interpolate the fixed codebook gains, the gain interpolator 132 can use the technique according to the below equation as described in Non Patent Literature 1.

[Mathematical Equation 40]

$$g_s = 0.4 \cdot g_c^{-1} + 0.3 \cdot g_c^{-2} + 0.2 \cdot g_c^{-3} + 0.1 \cdot g_c^{-4}$$

[0045] The noise signal generator 133 generates white noise for the same length as the fixed codebook vectors and uses the resultant noise for the fixed codebook vectors.

[0046] The operations of the post-filter 134, the synthesis filter 135, the adaptive codebook calculator 136, and the excitation vector synthesizer 137 are the same as those in the aforementioned case of normal reception of a packet.

[0047] The internal state update is the same as performed in the case of

normal reception of a packet, except for an update of the ISF residual parameters. The updating of the ISF parameters is carried out in accordance with the following equation by the LP coefficient interpolator 130.

[Mathematical Equation 41]

$$\dot{r}_i^0 = \dot{\omega}_i^0 - mean_i - \frac{1}{3}\dot{r}_i^{-1}$$

Citation List

Patent Literatures

[0048] Patent Literature 1: International Publication WO 2002/035520

Patent Literature 2: International Publication WO 2008/108080

Non Patent Literature

[0049] Non Patent Literature 1: ITU-T Recommendation G.718, June 2008

Summary of Invention

Technical Problem

[0050] As described above, since the CELP encoding involves the internal states, a degradation of audio quality occurs because of a deviation between the parameters obtained by interpolations implemented upon a packet loss and the parameters that would have been used for decoding. Particularly, as to the ISF parameters, intraframe/interframe predictive encoding is carried out, and thus there is the problem that an influence by a packet loss continues even after recovery from the packet loss.

[0051] More specifically, a problem of a sudden increase of power is identified in the first frame after recovery from a packet loss occurring

in the vicinity of an audio start portion. This is caused for the following reason: That is, in the audio start portion where the power of the excitation signal becomes high, the impulse response of the LP coefficients calculated from the ISF coefficients obtained by the interpolation process upon a packet loss has a higher gain than the one that would have been originally expected for the decoder. This is perceived, according to the subjective quality standard, as an unpleasant discontinuity of audio .

[0052] The method described in Patent Literature 1 generates the interpolated ISF coefficients for a lost frame. However, since the ISF parameters are generated by a normal decoding process for the first frame after recovery from the loss, it fails to suppress the sudden increase of power.

[0053] On the other hand, the method described in Patent Literature 2 transmits a gain adjustment parameter (normalized prediction residual power) obtained on the encoding side and uses it for a power adjustment on the decoding side, thereby controlling the power of the excitation signal of a lost packet frame and enabling prevention of the sudden increase of power.

[0054] Fig. 7 shows an exemplary functional configuration of an audio decoder 1X implemented by the technology of Patent Literature 2, and Fig. 8 shows an exemplary functional configuration of a concealment signal generator 13X. In Patent Literature 2, an audio packet includes auxiliary information of at least a normalized prediction residual power in addition to the parameters described in the conventional technique.

[0055] A normalized prediction residual power decoder 15 provided in

the audio signal generator 1X decodes the auxiliary information of the normalized prediction residual power from a received audio packet to calculate a reference normalized prediction residual power, and outputs it to the concealment signal generator 13X.

[0056] Since the constitutive elements of the concealment signal generator 13X, other than normalized prediction residual adjuster 138, are the same as those in the aforementioned conventional technology, only the normalized prediction residual adjuster 138 will be described below.

[0057] The normalized prediction residual adjuster 138 calculates the normalized prediction residual power from the LP coefficients output by the LP coefficient interpolator 130. Next, the normalized prediction residual adjuster 138 calculates a synthesis filter gain adjustment coefficient, using the normalized prediction residual power and the reference normalized prediction residual power. Finally, the normalized prediction residual adjuster 138 multiplies the excitation signal by the synthesis filter gain adjustment coefficient and output the result to the synthesis filter 135.

[0058] The above-described technology of Patent Literature 2 can control the power of the concealment signal upon an occurrence of a packet loss in the same manner as performed in the normal reception. However, it is difficult to secure a bit rate necessary for transmission of the foregoing gain adjustment parameter in the process of low-bit-rate audio encoding. In addition, since it is the processing in the concealment signal generator, it is difficult to deal with a sudden change of power caused by a disagreement of the ISF parameters in a recovery

frame.

[0059] An object of the present invention is therefore to reduce a discontinuity of audio which can occur upon recovery from a packet loss at the audio start point, and thereby improve the subjective quality.

Solution to Problem

[0060] An audio signal processing device according to one embodiment of the present invention comprises: a discontinuity detector configured to determine an occurrence of a discontinuity occurring with a sudden increase of the amplitude of a decoded audio obtained by decoding a first audio packet which is received correctly after an occurrence of a packet loss; and a discontinuity corrector configured to correct the discontinuity of the decoded audio.

[0061] The discontinuity detector may determine an occurrence of a discontinuity of the decoded audio with the power of an excitation signal.

[0062] The discontinuity detector may detect an occurrence of a discontinuity of the decoded audio with quantized codebook gains used for calculation of an excitation signal.

[0063] The audio signal processing device may further comprise: an auxiliary information decoder configured to decode auxiliary information for determination on an occurrence of a discontinuity transmitted from an encoder, and the discontinuity detector may determine an occurrence of a discontinuity of the decoded audio, using the auxiliary information decoded and output as an auxiliary information code by the auxiliary information decoder .

[0064] The discontinuity corrector may correct ISF parameters or LSF

parameters (hereinafter referred to as "ISF/LSF parameters") according to a result of determination on an occurrence of a discontinuity.

[0065] More specifically, the discontinuity corrector may change a distance between elements of the ISF/LSF parameters given for ensuring stability of a synthesis filter, according to a result of determination on an occurrence of a discontinuity.

[0066] At this time, the discontinuity corrector may extend the distance between the elements of the ISF/LSF parameters given for ensuring the stability of the synthesis filter to become larger than an ordinary distance given for ensuring stability.

[0067] For the distance between the elements of the ISF/LSF parameters given for ensuring the stability of the synthesis filter, the discontinuity corrector may use a distance, which is obtained by equally dividing the ISF/LSF parameters into those of a predetermined length.

[0068] Furthermore, the discontinuity corrector may replace a part of or all of the ISF/LSF parameters with predetermined vectors.

[0069] An audio signal processing device according to one embodiment of the present invention comprises: an ISF/LSF quantizer configured to quantize ISF/LSF parameters; an ISF/LSF concealer configured to generate concealment ISF/LSF parameters, which are concealment information for the ISF/LSF parameters; a discontinuity detector configured to determine an occurrence of a discontinuity occurring in a first audio packet which is received correctly after an occurrence of a packet loss, using distances between the quantized ISF/LSF parameters obtained in the quantization process by the ISF/LSF quantizer and the concealment ISF/LSF parameters generated by the ISF/LSF concealer;

and an auxiliary information encoder configured to encode auxiliary information for determination on an occurrence of a discontinuity.

[0070] An audio signal processing device according to one embodiment of the present invention comprises: a discontinuity detector configured to determine an occurrence of a discontinuity occurring in a first audio packet which is received correctly after an occurrence of a packet loss; an auxiliary information encoder configured to encode auxiliary information for determination on an occurrence of a discontinuity; and an ISF/LSF quantizer configured to use past quantized ISF/LSF residual parameters for ISF/LSF quantization in a given frame when the discontinuity detector does not determine an occurrence of a discontinuity, and avoid using the past quantized ISF/LSF residual parameters for ISF/LSF quantization in the given frame when the discontinuity detector determines an occurrence of a discontinuity.

[0071] An audio signal processing device according to one embodiment of the present invention comprises: an auxiliary information decoder configured to decode and output auxiliary information for determination on an occurrence of a discontinuity occurring in a first audio packet which is received correctly after an occurrence of a packet loss; a discontinuity corrector configured to correct the discontinuity of a decoded audio; and an ISF/LSF decoder configured to use past quantized ISF/LSF residual parameters for ISF/LSF calculation in a pertinent frame when the auxiliary information from the auxiliary information decoder does not indicate an occurrence of a discontinuity, and avoid using the past quantized ISF/LSF residual parameters for the ISF/LSF calculation in the pertinent frame when the auxiliary

information from the auxiliary information decoder indicates an occurrence of a discontinuity.

[0072] The audio signal processing device may adopt a configuration in which the audio signal processing device further comprises: a reception state determiner configured to determine packet reception states of a predetermined number of past frames; the discontinuity corrector corrects a discontinuity on the basis of a determination result of the packet reception states as well, in addition to a result of determination on an occurrence of a discontinuity.

[0073] Now, the audio signal processing device according to one embodiment of the present invention may be taken as an invention associated with an audio signal processing method, and as an invention associated with an audio signal processing program, and can be described as below.

[0074] An audio signal processing method according to one embodiment of the present invention is an audio signal processing method to be executed by an audio signal processing device, comprising: a step of determining an occurrence of a discontinuity of decoded audio occurring with a sudden increase of the amplitude of a decoded audio obtained by decoding a first audio packet which is received correctly after an occurrence of a packet loss; and a step of correcting the discontinuity of the decoded audio.

[0075] An audio signal processing method according to one embodiment of the present invention is an audio signal processing method to be executed by an audio signal processing device, comprising: a step of quantizing ISF/LSF parameters; a step of

generating concealment ISF/LSF parameters which are concealment information for the ISF/LSF parameters; a step of determining an occurrence of a discontinuity occurring in a first audio packet which is received correctly after an occurrence of a packet loss, using distances between quantized ISF/LSF parameters obtained in a quantization process of the ISF/LSF quantizer and the generated concealment ISF/LSF parameters; and a step of encoding auxiliary information for determination on an occurrence of a discontinuity.

[0076] An audio signal processing method according to one embodiment of the present invention is an audio signal processing method to be executed by an audio signal processing device, comprising: a step of determining an occurrence of a discontinuity occurring in a first audio packet which is received correctly after an occurrence of a packet loss; a step of encoding auxiliary information for determination on an occurrence of a discontinuity; and a step of using past quantized ISF/LSF residual parameters for ISF/LSF quantization in a given frame when an occurrence of a discontinuity is not determined, and avoiding using the past quantized ISF/LSF residual parameters for the ISF/LSF quantization in the pertinent frame when an occurrence of a discontinuity is determined.

[0077] An audio signal processing method according to one embodiment of the present invention is an audio signal processing method to be executed by an audio signal processing device, comprising: a step of decoding and outputting auxiliary information for determination on an occurrence of a discontinuity of decoded audio occurring in a first audio packet which is received correctly after an

occurrence of a packet loss; a step of correcting the discontinuity of decoded audio; and a step of using past quantized ISF/LSF residual parameters for ISF/LSF calculation in a given frame when the auxiliary information does not indicate an occurrence of a discontinuity, and avoiding using the past quantized ISF/LSF residual parameters for the ISF/LSF calculation in the given frame when the auxiliary information indicates an occurrence of a discontinuity.

[0078] An audio signal processing program according to one embodiment of the present invention is an audio signal processing program that programs a computer to operate as: a discontinuity detector operable to determine an occurrence of a discontinuity of decoded audio occurring with a sudden increase of amplitude of a decoded audio obtained by decoding a first audio packet which is received correctly after an occurrence of a packet loss; and a discontinuity corrector operable to correct the discontinuity of the decoded audio.

[0079] An audio signal processing program according to one embodiment of the present invention is an audio signal processing program that programs a computer to operate as: an ISF/LSF quantizer operable to quantize ISF/LSF parameters; an ISF/LSF concealer operable to generate concealment ISF/LSF parameters which are concealment information for the ISF/LSF parameters; a discontinuity detector operable to determine an occurrence of a discontinuity occurring in a first audio packet which is received correctly after an occurrence of a packet loss, using distances between quantized ISF/LSF parameters obtained in a quantization process of the ISF/LSF quantizer

and the concealment ISF/LSF parameters generated by the ISF/LSF concealer; and an auxiliary information encoder operable to encode auxiliary information for determination on an occurrence of a discontinuity.

[0080] An audio signal processing program according to one embodiment of the present invention is an audio signal processing program that programs a computer to operate as: a discontinuity detector operable to determine an occurrence of a discontinuity occurring in a first audio packet which is received correctly after an occurrence of a packet loss; an auxiliary information encoder operable to encode auxiliary information for determination on an occurrence of a discontinuity; and an ISF/LSF quantizer operable to use past quantized ISF/LSF residual parameters for ISF/LSF quantization in a pertinent frame when the discontinuity detector does not determine an occurrence of a discontinuity, and avoid using the past quantized ISF/LSF residual parameters for the ISF/LSF quantization in the pertinent frame when the discontinuity detector determines an occurrence of a discontinuity.

[0081] An audio signal processing program according to one embodiment of the present invention is an audio signal processing program that programs a computer to operate as: an auxiliary information decoder operable to decode and output auxiliary information for determination on an occurrence of a discontinuity of decoded audio occurring in a first audio packet which is received correctly after an occurrence of a packet loss; a discontinuity corrector operable to correct the discontinuity of decoded audio; and an ISF/LSF decoder operable to use past quantized ISF/LSF residual parameters for

ISF/LSF calculation in a given frame when the auxiliary information from the auxiliary information decoder does not indicate an occurrence of a discontinuity, and avoid using the past quantized ISF/LSF residual parameters for the ISF/LSF calculation in the pertinent frame when the auxiliary information from the auxiliary information decoder indicates an occurrence of a discontinuity.

Advantageous Effect of Invention

[0082] The present invention as described above can reduce a discontinuity of audio possibly occurring subsequent to recovery from a packet loss at the audio start point and thus improve the subjective quality.

Brief Description of Drawings

[0083] Fig. 1 is a configuration diagram of the audio decoder.

Fig. 2 is a processing flow of the audio decoder.

Fig. 3 is a functional configuration diagram of the audio code decoder.

Fig. 4 is a functional configuration diagram of the LP coefficient calculator.

Fig. 5 is a processing flow of calculating the LP coefficients.

Fig. 6 is a functional configuration diagram of the concealment signal generator.

Fig. 7 is a configuration diagram of the audio decoder of Patent Literature 2.

Fig. 8 is a functional configuration diagram of the concealment signal generator of Patent Literature 2.

Fig. 9 is a functional configuration diagram of the audio code

decoder in a first embodiment.

Fig. 10 is a processing flow of the LP coefficient calculator in the first embodiment.

Fig. 11 is a functional configuration diagram of the audio code decoder in the first embodiment.

Fig. 12 is a processing flow of a second stability processor in modification example 1 of the first embodiment.

Fig. 13 is a functional configuration diagram of the audio code decoder in a second embodiment.

Fig. 14 is a functional configuration diagram of the LP coefficient calculator in the second embodiment.

Fig. 15 is a processing flow of calculation of the LP coefficients in the second embodiment.

Fig. 16 is a configuration diagram of an audio encoder in fourth embodiment.

Fig. 17 is a configuration diagram of the audio encoder in the fourth embodiment.

Fig. 18 is a configuration diagram of an LP analyzer/encoder in the fourth embodiment.

Fig. 19 is a processing flow of the LP analyzer/encoder in the fourth embodiment.

Fig. 20 is a functional configuration diagram of the audio code decoder in the fourth embodiment.

Fig. 21 is a processing flow of the LP coefficient calculator in the fourth embodiment.

Fig. 22 is a configuration diagram of the LP analyzer/encoder in

the fifth embodiment.

Fig. 23 is a processing flow of the LP analyzer/encoder in the fifth embodiment.

Fig. 24 is a functional configuration diagram of the audio code decoder in the fourth embodiment.

Fig. 25 is a processing flow of the LP coefficient calculator in the fifth embodiment.

Fig. 26 is a configuration diagram of the audio decoder in the seventh embodiment.

Fig. 27 is a processing flow of the audio decoder in the seventh embodiment.

Fig. 28 is a functional configuration diagram of the audio code decoder in the seventh embodiment.

Fig. 29 is a processing flow of calculation of the LP coefficients in the seventh embodiment.

Fig. 30 is a drawing showing a hardware configuration example of a computer.

Fig. 31 is an appearance diagram of the computer.

Figs. 32 (a), (b), (c), and (d) are drawings showing various examples of audio signal processing programs.

Description of Embodiments

[0084] Preferred embodiments of an audio signal processing device, an audio signal processing method, and an audio signal processing program according to the present invention will be described below in detail using the drawings. The same elements will be denoted by similar reference signs in the description of the drawings to avoid duplicate

descriptions.

[0085] [First Embodiment]

The audio signal processing device in the first embodiment has the same configuration as the aforementioned audio decoder 1 shown in Fig. 1 and has a novel feature in the audio code decoder, and thus the audio code decoder will be described below.

[0086] Fig. 9 is a diagram showing a functional configuration of an audio code decoder 12A in the first embodiment, and Fig. 10 shows a flowchart of the LP coefficient calculation process. The audio code decoder 12A shown in Fig. 9 is configured by adding a discontinuity detector 129 to the aforementioned configuration of Fig. 3. Since the present embodiment differs from the conventional technology only in the LP coefficient calculation process, the operations of respective parts associated with the LP coefficient calculation process will be described below.

[0087] A discontinuity detector 129 refers to a fixed codebook gain g_c^0 acquired by decoding and a fixed codebook gain g_c^{-1} included in the internal states and compares a change of the gain with a threshold in accordance with the following equation (step S11 in Fig. 10).

[Mathematical Equation 42]

$$\log(g_c^0) - \log(g_c^{-1}) > Thres$$

[0088] When the gain change exceeds the threshold, the detector detects an occurrence of a discontinuity (also referred to hereinafter simply as "detects a discontinuity") and outputs a control signal indicating a detection result of a discontinuity occurrence to the stability

processor 121.

[0089] The following equation may be used for the comparison between the gain change and the threshold.

[Mathematical Equation 43]

$$g_c^0 - g_c^{-1} > Thres$$

[0090] Furthermore, the comparison between the gain change and the threshold may be made by the following equation, where a $g_c^{(c)}$ represents the maximum among the fixed codebook gains of the first to fourth subframes included in the current frame and a $g_c^{(p)}$ represents the minimum among the fixed codebook gains included in the internal states.

[Mathematical Equation 44]

$$\log(g_c^{(c)}) - \log(g_c^{(p)}) > Thres$$

[0091] The flowing equation can also be used.

[Mathematical Equation 45]

$$g_c^{(c)} - g_c^{(p)} > Thres$$

[0092] The above example of the first embodiment shows an example in which a discontinuity detection is conducted using the fixed codebook gain g_c^{-1} of the fourth subframe of the immediately preceding frame (lost frame) and the fixed codebook gain g_c^0 of the first subframe of the current frame. However, comparison between the gain change and the threshold may be made using averages calculated from the fixed codebook gains included in the internal states and the fixed codebook gains included in the current frame.

[0093] The ISF decoder 120 performs the same operation as in the conventional technology (step S12 in Fig. 10).

[0094] The stability processor 121 corrects the ISF parameters by the following process when the discontinuity detector 129 detects a discontinuity (step S13 in Fig. 10).

[0095] First, the stability processor 121 subjects the ISF parameters

[Mathematical Equation 46]

$$\dot{\omega}_i^{-1}$$

stored in the internal state buffer 14 to a process of expanding a distance between two adjacent element to become M_{-1} times wider than the ordinary distance. The process of placing a very wide distance than the ordinary distance provides an effect to suppress excessive peaks and dips in the spectrum envelope. Here, min_dist represents the minimum ISF distance, and isf_min represents the minimum of ISF necessary for securing the distance of min_dist. isf_min is successively updated by adding the distance of min_dist to a value of neighboring ISF. On the other hand, isf_max is the maximum of ISF necessary for securing the distance of min_dist. isf_max is successively updated by subtracting the distance of min_dist from a value of neighboring ISF.

[Mathematical Equation 47]

```
isf_min = min_dist = 50M_1
for i=0 to 14
        if ω_i^-1 <isf_min then  ω_i^-1 =isf_min
        isf_min =ω_i^-1 + min_dist
isf_max = 6400 - min_dist
if ω_14^-1 >isf_max
        for i=14 down to 1
              if ω_i^-1 > isf_max then ω_i^-1 =isf_max
              isf_max = ω_i^-1 - min_dist
```

[0096] Next, a stability processor 121 subjects the ISF parameters of the current frame to a process of expanding a distance between two adjacent element to become M_0 times wider than the ordinary distance. $1 < M_0 < M_{-1}$ is assumed herein, but it is also possible to set one of r M_{-1} and M_0 to 1 and the other to a value larger than 1.

[Mathematical Equation 48]

isf_min = min_dist = $50M_0$

for i=0 to 14

if $\dot{\omega}_i^0$<isf_min then $\dot{\omega}_i^0$=isf_min

isf_min = $\dot{\omega}_i^0$ + min_dist

isf_max = 6400 - min_dist

if $\dot{\omega}_{14}^0$>isf_max

for i=14 down to 1

if $\dot{\omega}_i^0$ > isf_max then $\dot{\omega}_i^0$=isf_max

isf_max = $\dot{\omega}_i^0$ - min_dist

[0097] Furthermore, the stability processor 121 performs the following process in the same manner as carried out in the ordinary decoding process, when the discontinuity detector detects no discontinuity.

[Mathematical Equation 49]

```
isf_min = min_dist = 50
for i=0 to 14
        if ω_i^0<isf_min then ω_i^0=isf_min
        isf_min = ω_i^0 + min_dist
isf_max = 6400 - min_dist
if ω_14^0>isf_max
        for i=14 down to 1
                if ω_i^0 > isf_max then ω_i^0=isf_max
                isf_max = ω_i^0 - min_dist
```

[0098] The minimum distance placed between elements when a discontinuity is detected may be varied depending upon the frequency of ISF. The minimum distance placed between elements when a discontinuity is detected needs only to be different from the minimum distance placed between elements in the ordinary decoding process.

[0099] The ISF-ISP converter 122A in the LP coefficient calculator 122 converts the ISF parameters

[Mathematical Equation 50]

$\dot{\omega}_i$、$\dot{\omega}_i^{-1}$

into the ISP parameters

[Mathematical Equation 51]

$\dot{\mathbf{q}}_i$、$\dot{\mathbf{q}}_i^{-1}$,

respectively, in accordance with the following equation (step S14 in Fig.

10). Here, C is a constant determined in advance.

[Mathematical Equation 52]

$$\dot{\mathbf{q}}_i = \cos(C \cdot \dot{\omega}_i)$$

[0100] The ISP interpolator 122B calculates the ISP parameters for the respective subframes from the past ISP parameters

[Mathematical Equation 53]

$$\dot{\mathbf{q}}_i^{-1}$$

and the foregoing ISP parameters

[Mathematical Equation 54]

$$\dot{\mathbf{q}}_i$$

in accordance with the following equation (step S15 in Fig. 10). Other coefficients may be used for the interpolation.

[Mathematical Equation 55]

$$\mathbf{q}_i^{(1)} = 0.75 \cdot \dot{\mathbf{q}}_i^{-1} + 0.25 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(2)} = 0.5 \cdot \dot{\mathbf{q}}_i^{-1} + 0.5 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(3)} = 0.25 \cdot \dot{\mathbf{q}}_i^{-1} + 0.75 \cdot \dot{\mathbf{q}}_i$$
$$\mathbf{q}_i^{(4)} = \dot{\mathbf{q}}_i$$

[0101] The ISP-LPC converter 122C converts the ISP parameters for the respective subframes into the LP coefficients

[Mathematical Equation 56]

$$\dot{a}_i^j \quad (0 < i \leq P,\ 0 \leq j < 4)$$

(step S16 in Fig. 10). Here, the number of subframes included in a look-ahead signal was assumed to be 4, but the number of subframes may differ depending upon the design principle. A specific conversion procedure to be used can be the processing procedure described in Non

Patent Literature 1.

[0102] Furthermore, the ISF-ISP converter 122A updates the ISF parameters stored in the internal state buffer 14

[Mathematical Equation 57]

$$\dot{\omega}_i^{-1}$$

in accordance with the following equation.

[Mathematical Equation 58]

$$\dot{\omega}_i^{-1} = \dot{\omega}_i^{0}$$

At this time, even when a discontinuity is detected, the ISF-ISP converter 122A may carry out the below procedure to update the ISF parameters

[Mathematical Equation 59]

$$\dot{\omega}_i^{-1}$$

stored in the internal state buffer, using the calculation result of the ISF parameters.

[Mathematical Equation 60]

isf_min = min_dist = 50

for i=0 to 14

if $\dot{\omega}_i^0$<isf_min then $\dot{\omega}_i^0$=isf_min

isf_min = $\dot{\omega}_i^0$ + min_dist

isf_max = 6400 - min_dist

if $\dot{\omega}_{14}^0$>isf_max

for i=14 down to 1

if $\dot{\omega}_i^0$ > isf_max then $\dot{\omega}_i^0$=isf_max

isf_max = $\dot{\omega}_i^0$ - min_dist

[0103] As in the above first embodiment, a discontinuity of decoded audio can be determined with the quantized codebook gains used in the calculation of the excitation signal and the ISF/LSF parameters (e.g., the distance between elements of the ISF/LSF parameters given for ensuring stability of the synthesis filter) can be corrected according to a result of the determination for a discontinuity. This reduces the discontinuity of audio which can occur upon recovery from a packet loss at the audio start point, and thereby improves the subjective quality.

[0104] [Modification Example of First Embodiment]

Fig. 11 is a diagram showing a functional configuration of an audio code decoder 12S according to a modification example of the first embodiment. Since it differs from the configuration of the conventional technology shown in Fig. 3 only in the discontinuity detector 129 and the second stability processor 121S, the operations of

these will be described. The second stability processor 121S has a gain adjustor 121X and a gain multiplier 121Y, and a processing flow of the second stability processor 121S is shown in Fig. 12.

[0105] The discontinuity detector 129 refers to the fixed codebook gain g_c^0 obtained by decoding and the fixed codebook gain g_c^{-1} included in the internal states and compares the gain change with a threshold, in the same manner as performed by the discontinuity detector 129 in the first embodiment. Then, the discontinuity detector 129 sends to the gain adjustor 121X, a control signal including information about whether the gain change exceeds the threshold.

[0106] The gain adjustor 121X reads from the control signal the information about whether the gain change exceeds the threshold, and, when the gain change exceeds the threshold, it outputs a predetermined gain g_{on} to the gain multiplier 121Y. On the other hand, when the gain change does not exceed the threshold, the gain adjustor 121X outputs a predetermined gain g_{off} to the gain multiplier 121Y. This operation of the gain adjustor 121X corresponds to step S18 in Fig. 12.

[0107] The gain multiplier 121Y multiplies the synthesized signal output from the synthesis filter 128 by the foregoing gain g_{on} or gain g_{off} (step S19 in Fig. 12) and outputs the resultant decoded signal.

[0108] Here, the audio code decoder may be configured such that the LP coefficient calculator 122 outputs the LP coefficients or the ISF parameters to feed them to the second stability processor 121S (as indicated by a dotted line from the LP coefficient calculator 122 to the gain adjustor 121X in Fig. 11). In this case, the gains to be multiplied are determined using the LP coefficients or the ISF parameters

calculated by the LP coefficient calculator 122.

[0109] By adding the second stability processor 121S to the audio code decoder 12S and adjusting the gain, depending upon whether the gain change exceeds the threshold as described in the above modification example, an appropriate decoded signal can be obtained.

[0110] The second stability processor 121S may be configured to multiply the excitation signal by the foregoing calculated gain and output the result to the synthesis filter 128.

[0111] [Second Embodiment]

An audio signal processing device according to the second embodiment has the same configuration as that of the aforementioned audio decoder 1 in Fig. 1 and has a novel feature in an audio code decoder, and thus the audio code decoder will be described below. Fig. 13 shows an exemplary functional configuration of the audio code decoder 12B, Fig. 14 shows an exemplary functional configuration associated with the calculation process of the LP coefficients, and Fig. 15 shows a flow of the calculation process of the LP coefficients. The audio code decoder 12B in Fig. 13 is configured by adding the discontinuity detector 129 to the aforementioned configuration shown in Fig. 3.

[0112] The ISF decoder 120 calculates the ISF parameters in the same manner as performed in the conventional technology (step S21 in Fig. 15).

[0113] The stability processor 121 performs the process of placing a distance of not less than 50 Hz between elements of the ISF parameters

[Mathematical Equation 61]

$\dot{\omega}_i$

in order to secure the stability of the filter in the same manner as performed in the conventional technology (step S22 in Fig. 15).

[0114] The ISF-ISP converter 122A converts the ISF parameters output by the stability processor 121 into the ISP parameters in the same manner as performed in the first embodiment (step S23 in Fig. 15).

[0115] The ISP interpolator 122B, in the same manner as performed in the first embodiment (step S24 in Fig. 15), calculates the ISP parameters for the respective subframes from the past ISP parameters

[Mathematical Equation 62]

$\dot{\mathbf{q}}_i^{-1}$

and the ISP parameters

[Mathematical Equation 63]

$\dot{\mathbf{q}}_i$

obtained by the conversion by the ISF-ISP converter 122A.

[0116] The ISP-LPC converter 122C, in the same manner as performed in the first embodiment (step S25 in Fig. 15), converts the ISP parameters for the respective subframes into the LP coefficients

[Mathematical Equation 64]

$\dot{a}_i^j \quad (0 < i \leq P \text{、} \ 0 \leq j < 4)$

. Here, the number of subframes included in the look-ahead signal is assumed to be 4, but the number of subframes may differ depending upon the design principle.

[0117] The internal state buffer 14 updates the ISF parameters stored in

the past with the new ISF parameters.

[0118] The discontinuity detector 129 reads the LP coefficients of the fourth subframe in the lost packet frame from the internal state buffer 14 and calculates the power of the impulse response of the LP coefficients of the fourth subframe in the lost packet frame. The LP coefficients of the fourth subframe in the lost packet frame to be used can be the coefficients output by the LP coefficient interpolator 130 included in the concealment signal generator 13 shown in Fig. 6 and accumulated in the internal state buffer 14 upon the packet loss.

[Mathematical Equation 65]

$$E_{-1} = 10\log\left(\sum_{n=0}^{L'-1} h_{-1}^2(n)\right)$$

$$h_{-1}(n) = \delta(n) - \sum_{i=1}^{P} \dot{a}_i^{(-1)} \cdot h_{-1}(n-i)$$

[0119] Then, the discontinuity detector 129 detects a discontinuity, for example, by the below equation (step S26 in Fig. 15).

[Mathematical Equation 66]

$$E_0 - E_{-1} > Thres$$

[0120] When the gain change does not exceed the threshold (NO in step S27 of Fig. 15), the discontinuity detector 129 does not detect an occurrence of a discontinuity, and the ISP-LPC converter 122C outputs the LP coefficients and ends the processing. On the other hand, when the gain change exceeds the threshold (YES in step S27 of Fig. 15), the discontinuity detector 129 detects an occurrence of a discontinuity and sends a control signal indicative of a result of the detection for an occurrence of a discontinuity to the stability processor 121. When

receiving the control signal, the stability processor 121 corrects the ISP parameters in the same manner as performed in the first embodiment (step S28 in Fig. 15). The subsequent operations of the ISF-ISP converter 122A, ISP interpolator 122B, and ISP-LPC converter 122C (steps S29, S2A, and S2B in Fig. 15) are the same as above.

[0121] As discussed in the above second embodiment, a discontinuity of decoded audio can be determined by the power of the excitation signal, and the discontinuous audio is reduced to improve the subjective quality in the same manner as performed in the first embodiment.

[0122] [Third Embodiment]

Upon a detection of discontinuity, the ISF parameters may be corrected by another method. The third embodiment differs from the first embodiment only in the stability processor 121, and thus only the operation of the stability processor 121 will be described.

[0123] When the discontinuity detector 129 detects a discontinuity, the stability processor 121 performs the following process to correct the ISF parameters.

[0124] With respect to the ISF parameters stored in the internal state buffer 14,

[Mathematical Equation 67]

$$\dot{\omega}_i^{-1}$$

the stability processor 121 replaces the ISF parameters up to a low-order P' dimension ($0 < P' \leq P$) in accordance with the below equation. Here, the following definition is adopted.

[Mathematical Equation 68]

$$\delta^{-1} = \dot{\omega}^{-1}_{P'-1} / P'$$

[Mathematical Equation 69]

$$\dot{\omega}^{-1}_{i} = \dot{\omega}^{-1}_{i-1} + \delta^{-1}$$
$$\dot{\omega}^{-1}_{0} = \delta^{-1} \qquad (0 \le i < P')$$

[0125] The stability processor 121 may overwrite the ISF parameters of the low-order P' dimensions with P'-dimension vectors obtained in advance by learning as follows.

[Mathematical Equation 70]

$$\dot{\omega}^{-1}_{i} = \omega^{0}_{i} \qquad (0 \le i < P')$$

[0126] Next, as to the ISF parameters of the current frame, the stability processor 121 may, as performed in the first embodiment, perform the process of expanding the distance between elements to becomeM_0 times wider than the ordinary distance or may determine them in accordance with the below equation. Here, the following definition is adopted.

[Mathematical Equation 71]

$$\delta^{0} = \dot{\omega}^{0}_{P'-1} / P'$$

[Mathematical Equation 72]

$$\dot{\omega}^{0}_{i} = \dot{\omega}^{0}_{i-1} + \delta^{0}$$
$$\dot{\omega}^{0}_{0} = \delta^{0}$$

[0127] The stability processor 121 may overwrite them with P'-dimensional vectors learned in advance.

[Mathematical Equation 73]

$$\dot{\omega}_i^0 = \boldsymbol{\omega}_i^0 \qquad (0 \le i < P')$$

[0128] Furthermore, the foregoing P'-dimensional vectors may be learned in the decoding process or may be defined, for example, as follows.

[Mathematical Equation 74]

$$\boldsymbol{\omega}_i^0 = (1-\lambda)\boldsymbol{\omega}_i^{-1} + \lambda\dot{\omega}_i^{-1}$$

In a frame at the start of decoding, however, ω_i^{-1} may be defined as predetermined P'-dimensional vector ω_i^{init}.

[0129] The internal state buffer 14 updates the ISF parameters stored in the past with the new ISF parameters.

[0130] As discussed in the above third embodiment, the distance obtained by equally dividing the ISF/LSF parameters into those of a predetermined dimension can be used as the distance between elements of the ISF/LSF parameters given for ensuring the stability of the synthesis filter, whereby the discontinuous audio is reduced to improve the subjective quality as performed in the first and second embodiments.

[0131] [Fourth Embodiment]

A fourth embodiment will be described in which the encoding side detects an occurrence of a discontinuity and transmits a discontinuity determination code (indicative of a detection result) as included in audio codes to the decoding side and also in which the decoding side determines the operation of the stability process, based on the discontinuity determination code included in the audio codes.

[0132] (Regarding Encoding Side)

Fig. 16 shows an exemplary functional configuration of the encoder 2, and Fig. 17 is a flowchart showing the processes performed in the encoder 2. As shown in Fig. 16, the encoder 2 has an LP analyzer/encoder 21, a residual encoder 22, and a code multiplexer 23.

[0133] An exemplary functional configuration of the LP analyzer/encoder 21 among them is shown in Fig. 18, and a flowchart showing the processes performed in the LP analyzer/encoder 21 is shown in Fig. 19. As shown in Fig. 18, the LP analyzer/encoder 21 has an LP analyzer 210, an LP-ISF converter 211, an ISF encoder 212, a discontinuity determiner 213, an ISF concealer 214, an ISF-LP converter 215, and an ISF buffer 216.

[0134] In the LP analyzer/encoder 21, the LP analyzer 210 performs a linear prediction analysis on an input signal to obtain linear prediction coefficients (step T41 in Fig. 17 and step U41 in Fig. 18). For the calculation of linear prediction coefficients, an autocorrelation function is first calculated from the audio signal, and then the Levinson-Durbin algorithm or the like can be applied.

[0135] The LP-ISF converter 211 converts the calculated linear prediction coefficients into the ISP parameters in the same manner as performed in the first embodiment (steps T42, U42). The conversion from linear prediction coefficients into ISF parameters may be implemented by use of the method described in the Non Patent Literature.

[0136] The ISF encoder 212 encodes the ISF parameters using a predetermined method to calculate ISF codes (steps T43, U43) and

outputs quantized ISF parameters obtained in the process of encoding to the discontinuity determiner 213, the ISF concealer 214, and the ISF-LP converter 215 (step U47). Here, the quantized ISF parameters are equal to the ISF parameters obtained by an inverse quantization of the ISF codes. A method of encoding may be vector-encoding, or encoding by a vector quantization or the like of error vectors from ISFs of the immediately preceding frame and mean vectors determined in advance by learning.

[0137] The discontinuity determiner 213 encodes a discontinuity determination flag stored in an internal buffer (not shown) built in the discontinuity determiner 213 and outputs a resultant discontinuity determination code (step U47). In addition, the discontinuity determiner 213 uses concealment ISF parameters

[Mathematical Equation 75]

$$\tilde{\omega}_i$$

read from the ISF buffer 216 and the quantized ISF parameters

[Mathematical Equation 76]

$$\dot{\omega}_i$$

to make a determination on a discontinuity in accordance with the below equation (steps T44, U46). Here, Thres_{ω} represents a threshold determined in advance, and P' an integer satisfying the following equation ($0 < P' \leq P$).

[Mathematical Equation 77]

$$\sum_{i=0}^{P'-1}(\dot{\omega}_i - \tilde{\omega}_i)^2 > Thres_{\omega}$$

[0138] The example is described above in which the discontinuity determination is made using the Euclidean distances between the ISF parameters. However, the discontinuity determination may be made by other methods.

[0139] The ISF concealer 214 calculates the concealment ISF parameters from the quantized ISF parameters by the same process as performed by the decoder-side ISF concealer and outputs the resultant concealment ISF parameters to the ISF buffer 216 (steps U44, U45). The operation of the ISF concealment process may be performed by any method as long as it is the same process as that of the decoder-side packet loss concealer.

[0140] The ISF-LP converter 215 calculates quantized linear prediction coefficients by converting the foregoing quantized ISF parameters and outputs a resultant quantized linear prediction coefficients to the residual encoder 22 (step T45). A method used for converting the ISF parameters into the quantized linear prediction coefficients may be the method described in the Non Patent Literature.

[0141] The residual encoder 22 filters the audio signal by use of the quantized liner prediction coefficients to calculate residual signals (step T46).

[0142] Next, the residual encoder 22 encodes the residual signals by encoding means using CELP or TCX (Transform Coded Excitation) or by encoding means switchably using CELP and TCX and outputs resultant residual codes (step T47). Since the operation of the residual

encoder 22 is less relevant to the present invention, description thereof is omitted herein.

[0143] The code multiplexer 23 assembles the ISF codes, the discontinuity determination code and the residual codes in a predetermined order and outputs resultant audio codes (step T48).

[0144] (Regarding Decoding Side)

An audio signal processing device according to the fourth embodiment has the same configuration as that of the aforementioned audio decoder 1 in Fig. 1 and has a novel feature in the audio code decoder, and thus the audio code decoder will be described below. Fig. 20 shows an exemplary functional configuration of an audio code decoder 12D, and Fig. 21 is a flowchart showing the process of calculating the LP coefficients. The audio code decoder 12D shown in Fig. 20 is configured by adding the discontinuity detector 129 to the aforementioned configuration shown in Fig. 3.

[0145] The ISF decoder 120 decodes the ISF codes and outputs resultant codes to the stability processor 121 and the internal state buffer 14 (step S41 in Fig. 21).

[0146] The discontinuity detector 129 decodes the discontinuity determination code and outputs a resultant discontinuity detection result to the stability processor 121 (step S42 in Fig. 21).

[0147] The stability processor 121 performs the stability process according to the discontinuity detection result (step S43 in Fig. 21). The processing procedure of the stability processor to be used can be the same method as executed in the first embodiment and the third embodiment.

[0148] The stability processor 121 may perform the stability process as described below, on the basis of other parameters included in the audio codes, in addition to the discontinuity detection result acquired from the discontinuity determination code. For example, the stability processor 121 may be configured to perform the stability process in such a manner that an ISF stability stab is calculated in accordance with the below equation and that when the ISF stability exceeds a threshold, even if the discontinuity determination code shows a detection of a discontinuity, the process is performed as if no discontinuity is detected. Here, C is a constant determined in advance.

[Mathematical Equation 78]

$$stab = 1.25 - \sum_{i=0}^{P'-1} (\dot{\omega}_i^0 - \dot{\omega}_i^{-1})^2 \Big/ C$$

[0149] The ISF-ISP converter 122A in the LP coefficient calculator 122 converts the ISF parameters into the ISP parameters by the same processing procedure as performed in the first embodiment (step S44 in Fig. 21).

[0150] The ISP interpolator 122B calculates the ISP parameters for the respective subframes by the same processing procedure as performed in the first embodiment (step S45 in Fig. 21).

[0151] The ISP-LPC converter 122C converts the ISP parameters calculated for the respective subframes into the LPC parameters by the same processing procedure as performed in the first embodiment (step S46 in Fig. 21).

[0152] In the fourth embodiment as described above, the encoding side performs the discontinuity determination (the discontinuity

determination using the Euclidian distances between concealment ISF parameters and quantized ISF parameters, as an example) encodes auxiliary information about a result of the determination and outputs encoded information to the decoding side, and the decoding side determine a discontinuity using the auxiliary information obtained by decoding. In this manner, the appropriate processing can be executed according to the discontinuity determination result made by the encoding side while the encoding side and the decoding side work in concert with each other.

[0153] [Fifth Embodiment]

(Regarding Encoding Side)

The functional configuration of the encoder is the same as that of the fourth embodiment shown in Fig. 16, and the processing flow of the encoder is the same as the processing flow of the fourth embodiment shown in Fig. 17. The below will describe the LP analyzer/encoder according to the fifth embodiment which is different from that in the fourth embodiment.

[0154] Fig. 22 shows an exemplary functional configuration of the LP analyzer/encoder, and Fig. 23 shows a flow of the processes performed by the LP analyzer/encoder. As shown in Fig. 22, the LP analyzer/encoder 21S has the LP analyzer 210, the LP-ISF converter 211, the ISF encoder 212, the discontinuity determiner 213, the ISF concealer 214, the ISF-LP converter 215, and the ISF buffer 216.

[0155] In this LP analyzer/encoder 21S, the LP analyzer 210 performs the linear prediction analysis on the input signal by the same process as performed in the fourth embodiment to obtain the linear prediction

coefficients (step U51 in Fig. 23).

[0156] The LP-ISF converter 211 converts the calculated linear prediction coefficients into the ISF parameters by the same process as performed in the fourth embodiment (step U52 in Fig. 23). The method described in the Non Patent Literature may be used for the conversion from the linear prediction coefficients into the ISF parameters.

[0157] The ISF encoder 212 reads the discontinuity determination flag stored in the internal buffer (not shown) of the discontinuity determiner 213 (step U53 in Fig. 23).

[0158] <Case Where Discontinuity Determination Flag Indicates Detection of Discontinuity>

The ISF encoder 212 calculates the ISF codes by vector-quantization of ISF residual parameters r_i calculated by the below equation (step U54 in Fig. 23). Here, the ISF parameters calculated by the LP-ISF converter are denoted by ω_i and mean vectors, which are $mean_i$, obtained in advance by learning.

[Mathematical Equation 79]

$$r_i = \omega_i - mean_i$$

[0159] Next, the ISF encoder 212 uses the quantized ISF residual parameters

[Mathematical Equation 80]

$$\hat{r}_i$$

obtained by quantization of the ISF residual parameters r_i to update the ISF residual parameter buffer in accordance with the following equation

(step U55 in Fig. 23).

[Mathematical Equation 81]

$$\dot{r}_i^{-1} = \hat{r}_i$$

[0160] <Case Where Discontinuity Determination Flag does not Indicate Detection of Discontinuity>

The ISF encoder 212 calculates the ISF codes by vector-quantization of the ISF residual parameters r_i calculated by the below equation (step U54 in Fig. 23). Here, the ISF residual parameters obtained by decoding in the immediately preceding frame are denoted as follows.

[Mathematical Equation 82]

$$\dot{r}_i^{-1}$$

[Mathematical Equation 83]

$$r_i = \omega_i - mean_i - \frac{1}{3}\dot{r}_i^{-1}$$

[0161] Next, the ISF encoder 212 uses the quantized ISF residual parameters

[Mathematical Equation 84]

$$\hat{r}_i$$

obtained by quantization of the ISF residual parameters r_i to update the ISF residual parameter buffer in accordance with the following equation (step U55 in Fig. 23).

[Mathematical Equation 85]

$$\dot{r}_i^{-1} = \hat{r}_i$$

[0162] By the above procedure, the ISF encoder 212 calculates the ISF codes and outputs quantized ISF parameters obtained in the process of encoding to the discontinuity determiner 213, the ISF concealer 214, and the ISF-LP converter 215.

[0163] The ISF concealer 214 calculates the concealment ISF parameters from the quantized ISF parameters by the same process as performed by the decoder-side ISF concealer in the same manner as executed in the fourth embodiment and outputs them to the ISF buffer 216 (steps U56, U58 in Fig. 23). The operation of the ISF concealment process may be performed by any method as long as it is the same process as that of the decoder-side packet loss concealer.

[0164] The discontinuity determiner 213 performs a determination of a discontinuity by the same process as performed in the fourth embodiment and stores a determination result in the internal buffer (not shown) of the discontinuity determiner 213 (step U57 in Fig. 23).

[0165] The ISF-LP converter 215 converts the quantized ISF parameters, in the same manner as performed in the fourth embodiment, to calculate the quantized linear prediction coefficients and outputs them to the residual encoder 22 (Fig. 16) (step U58 in Fig. 23).

[0166] (Regarding Decoding Side)

An audio signal processing device according to the fifth embodiment has the same configuration as that of the aforementioned audio decoder 1 in Fig. 1 and has a novel feature in the audio code decoder, and thus the audio code decoder will be described below. Fig.

24 shows an exemplary functional configuration of the audio code decoder 12E, and Fig. 25 shows a flow of the calculation process performed by the LP coefficients. The audio code decoder 12E shown in Fig. 24 is configured by adding the discontinuity detector 129 to the aforementioned configuration shown in Fig. 3.

[0167] The discontinuity detector 129 decodes the discontinuity determination code and outputs the resultant discontinuity determination flag to the ISF decoder 120 (step S51 in Fig. 25).

[0168] The ISF decoder 120 calculates the ISF parameters as follows, depending upon the value of the discontinuity determination flag, and outputs the ISF parameters to the stability processor 121 and the internal state buffer 14 (step S52 in Fig. 25).

[0169] <Case Where Discontinuity Determination Flag Indicates Detection of Discontinuity>

The ISF decoder 120 uses the quantized ISF residual parameters

[Mathematical Equation 86]

$$\dot{r}_i$$

obtained by decoding of the ISF codes, and the mean vectors $mean_i$ obtained in advance by learning to obtain the quantized ISF parameters

[Mathematical Equation 87]

$$\dot{\omega}_i$$

in accordance with the following equation.

[Mathematical Equation 88]

$$\dot{\omega}_i = mean_i + \dot{r}_i$$

[0170] Next, the ISF decoder 120 updates the ISF residual parameters stored in the internal state buffer 14 in accordance with the following equation.

[Mathematical Equation 89]

$$\dot{r}_i^{-1} = \dot{r}_i$$

[0171] <Case Where Discontinuity Determination Flag Does Not Indicate Detection of Discontinuity>

The ISF decoder 120 reads, from the internal state buffer 14, the ISF residual parameters

[Mathematical Equation 90]

$$\dot{r}_i^{-1}$$

obtained by decoding of the immediately preceding frame and uses the resultant ISF residual parameters

[Mathematical Equation 91]

$$\dot{r}_i^{-1}$$

,

the mean vectors $mean_i$ obtained in advance by learning and the quantized ISF residual parameters

[Mathematical Equation 92]

$$\dot{r}_i$$

obtained by decoding of the ISF codes to calculate the quantized ISF parameters

[Mathematical Equation 93]

$$\dot{\omega}_i$$

in accordance with the following equation.

[Mathematical Equation 94]

$$\dot{\omega}_i = mean_i + \dot{r}_i + \frac{1}{3}\dot{r}_i^{-1}$$

[0172] Next, the ISF decoder 120 updates the ISF residual parameters stored in the internal state buffer 14 in accordance with the following equation.

[Mathematical Equation 95]

$$\dot{r}_i^{-1} = \dot{r}_i$$

[0173] The stability processor 121 performs the same process as performed in the first embodiment (step S53 in Fig. 25) when a discontinuity is not detected.

[0174] The ISF-ISP converter 122A in the LP coefficient calculator 122 converts the ISF parameters into the ISP parameters by the same processing procedure as described in the first embodiment (step S54 in Fig. 25).

[0175] The ISP interpolator 122B calculates the ISP parameters for the respective subframes by the same processing procedure as performed in the first embodiment (step S55 in Fig. 25).

[0176] The ISP-LPC converter 122C, by the same processing procedure as performed in the first embodiment (step S56 in Fig. 25), converts the ISP parameters calculated for the respective subframes into the LPC parameters.

[0177] In the fifth embodiment as described above, the encoding side is configured as follows: When the discontinuity determination flag does not indicate a detection of a discontinuity, the vector quantization of the ISF residual parameters is carried out using the ISF residual parameters obtained by decoding of the immediately preceding frame. On the other hand, when the discontinuity determination flag indicates a detection of a discontinuity, the encoder avoids using the ISF residual parameters obtained by decoding of the immediately preceding frame. Similarly, the decoding side is configured as follows: When the discontinuity determination flag does not indicate a detection of a discontinuity, the quantized ISF parameters are calculated using the ISF residual parameters obtained by decoding of the immediately preceding frame. On the other hand, when the discontinuity determination flag indicates a detection of discontinuity, the decoder avoids using the ISF residual parameters obtained by decoding of the immediately preceding frame. In this manner, the appropriate processing according to a discontinuity determination result can be executed while the encoding side and the decoding side work in concert with each other.

[0178] [Sixth Embodiment]

The above first to fifth embodiments may be applied in combination. For example, as described in the fourth embodiment, the decoding side decodes the discontinuity determination code included in the audio codes from the encoding side to detect a discontinuity. When a discontinuity is detected, it may carry out the subsequent operation as follows.

[0179] For the ISF parameters

[Mathematical Equation 96]

$$\dot{\omega}_i^{-1}$$

stored in the internal state buffer, the ISF parameters up to the low-degree P' dimension ($0 < P' \leq P$) are replaces in accordance with the following equation as described in the third embodiment.

[Mathematical Equation 97]

$$\dot{\omega}_i^{-1} = \omega_i^0 \qquad (0 \leq i < P')$$

[0180] On the other hand, the ISF parameters of the current frame are calculated in accordance with the following equation as described in the fifth embodiment.

[Mathematical Equation 98]

$$\dot{\omega}_i = mean_i + \dot{r}_i$$

[0181] Thereafter, using the ISF parameters obtained as described above, the LP coefficients are obtained by the processes of the ISF-ISP converter 122A, the ISP interpolator 122B, and the ISP-LPC converter 122C as performed in the first embodiment.

[0182] It is also effective to adopt optional combinations of the first to fifth embodiments as described above.

[0183] [Seventh Embodiment]

It may be considered in the decoding operation according to the above first to sixth embodiments and their modifications, how the frame is lost (e.g., whether a single frame is lost or consecutive frames are lost). In the seventh embodiment, it suffices that a discontinuity

detection is made using, for example, the result of decoding of the discontinuity determination code included in the audio codes, and the method of how it should be performed is not limited to the above.

[0184] An audio signal processing device according to the seventh embodiment has the same configuration as that of the aforementioned audio decoder 1 in Fig. 1 and has a novel feature in the audio code decoder, and thus the audio code decoder will be described below.

[0185] Fig. 26 shows an exemplary configuration of the audio decoder 1S according to the seventh embodiment, and Fig. 27 shows a flowchart of the processes performed in the audio decoder. As shown in Fig. 26, in addition to the aforementioned audio code decoder 12G, the concealment signal generator 13 and the internal state buffer 14, the audio decoder 1S has a reception state determiner 16 that determines packet reception states in some past frames and stores a packet loss history.

[0186] The reception state determiner 16 determines a packet reception state and updates the packet loss history information, based on a determination result (step S50 in Fig. 27).

[0187] When a packet loss is detected (NO in step S100), the reception state determiner 16 outputs a packet loss detection result of the pertinent frame to the concealment signal generator 13, and the concealment signal generator 13 generates the concealment signal as described above and updates the internal states (steps S300, S400). The concealment signal generator 13 may also utilize the packet loss history information for interpolation of parameters or the like.

[0188] On the other hand, when no packet loss is detected (YES in step

S100), the reception state determiner 16 outputs the packet loss history information including a packet loss detection result of the pertinent frame and the audio codes included in the received packet to the audio code decoder 12, and the audio code decoder 12 decodes the audio codes as described before and updates the internal states (steps S200, S400).

[0189] Thereafter, the processes of steps S50 to S400 are repeated until the communication ends (or until step S500 results in a determination of YES).

[0190] Fig. 28 shows an exemplary functional configuration of the audio code decoder 12G, and Fig. 29 shows a flowchart of the calculation processes performed by the LP coefficients. An example will be described below using the packet loss history information only for the LP coefficient calculator 122, but the audio code decoder may be configured to use the packet loss history information for other constitutive elements.

[0191] Since the audio code decoder 12G has the same configuration as described in the first embodiment, except for the configuration associated with the calculation process of LP coefficients, the below will describe the configuration and its operation associated with the calculation process of LP coefficients.

[0192] The ISF decoder 120 decodes the ISF codes in the same manner as performed in the first embodiment and outputs the ISF parameters to the stability processor 121 (step S71 in Fig. 29).

[0193] The discontinuity detector 129 refers to the packet loss history information to determine the reception state (step S72). The

discontinuity detector 129 may be designed, for example, as follows: It stores a specific reception pattern which indicates, for example, a packet loss occurred three frames prior, a normal reception occurred two frames prior, and a packet loss occurred one frame prior. When the reception pattern is recognized which has been looked for, it sets a reception state flag to off and, otherwise, it sets the reception state flag to on.

[0194] Furthermore, the discontinuity detector 129 detects a discontinuity in the same manner as described in one of the first to sixth embodiments.

[0195] Then, the stability processor 121 performs the stability process according to the reception state flag and a result of the discontinuity detection, for example, as described below (step S73).

[0196] When the reception state flag is off, the stability processor 121 performs the same process as performed when a discontinuity is not detected, regardless of a result of the discontinuity detection.

[0197] On the other hand, when the reception flag is on and when the result of the discontinuity detection indicates that a discontinuity is not detected, the stability processor 121 performs the same process as performed when a discontinuity is not detected.

[0198] Furthermore, when the reception flag is on and when the result of the discontinuity detection is detection of discontinuity, the stability processor 121 performs the same process as performed when a discontinuity is detected.

[0199] Thereafter, the operations (steps S74 to S76) of the ISF-ISP converter 122A, the ISP interpolator 122B, and the ISP-LPC converter

122C in the LP coefficient calculator 122 are performed in the same manners as performed in the first embodiment.

[0200] In the seventh embodiment as described above, the stability process is carried out depending upon a result of the discontinuity detection and the state of the reception state flag, whereby more accurate processing can be executed while it is considered how the frame is lost (e.g., whether a single frame is lost or consecutive frames are lost).

[0201] [Regarding Audio Signal Processing Programs]

The below will describe audio signal processing programs that program a computer to operate as an audio signal processing device according to the present invention.

[0202] Fig. 32 is a drawing showing various exemplary configurations of the audio signal processing programs. Fig. 30 is an exemplary hardware configuration of the computer, and Fig. 31 shows a schematic view of a computer. Audio signal processing programs P1-P4 (which will be referred to hereinafter generally as "audio signal processing program P") shown in Fig. 32 (a) to (d), respectively, can program the computer C10 shown in Figs. 31 and 32 to operate as an audio signal processing device. It should be noted that the audio signal processing program P described in the present specification can be implemented not only on the computer as shown in Figs. 31 and 32 but also on any information processing device such as a cell phone, a personal digital assistance, or a portable personal computer.

[0203] The audio signal processing program P can be provided in a form stored in a recording medium M. Examples of the recording

medium M include recording media such as flexible disc, CD-ROM, DVD, or ROM, semiconductor memories, and so on.

[0204] As shown in Fig. 30, the computer C10 has a reading device C12 such as a flexible disc drive unit, a CD-ROM drive unit, or a DVD drive unit, a working memory (RAM) C14, a memory C16 for storing a program stored in the recording medium M, a display C16, a mouse C20 and a keyboard C22 as input devices, a communication device C24 for executing transmission/reception of data or the like, and a central processing unit (CPU) C26 for controlling execution of the program.

[0205] When the recording medium M is put into the reading device C12, the computer C10 becomes accessible to the audio signal processing program P stored in the recording medium M through the reading device C12 and becomes able to operate as an audio signal processing device programmed by the audio signal processing program P.

[0206] The audio signal processing program P may be one provided as computer data signal W superimposed on a carrier wave, as shown in Fig. 31, transmitted through a network. In this case, the computer C10 stores the audio signal processing program P received by the communication device C24 into the memory C16 and then can execute the audio signal processing program P.

[0207] The audio signal processing program P can be configured by adopting the various configurations shown in Fig. 32 (a) to (d). These correspond to the configurations recited in claims 18 to 21 associated with the audio signal processing programs as set forth in the scope of claims. For example, the audio signal processing program P1 shown

in Fig. 32 (a) has a discontinuity detection module P11 and a discontinuity correction module P12. The audio signal processing program P2 shown in Fig. 32 (b) has an ISF/LSF quantization module P21, an ISF/LSF concealment module P22, a discontinuity detection module P23, and an auxiliary information encoding module P24. The audio signal processing program P3 shown in Fig. 32 (c) has a discontinuity detection module P31, an auxiliary information encoding module P32, and an ISF/LSF quantization module P33. The audio signal processing program P4 shown in Fig. 32 (d) has an auxiliary information decoding module P41, a discontinuity correction module P42, and an ISF/LSF decoding module P43.

[0208] By implementing the various embodiments described above, the subjective quality can be improved while reducing a discontinuous audio which can occur in the recovery from a packet loss at the audio start point.

[0209] The stability processor, which is the first feature of the invention, is configured so that when a discontinuity is detected in the first packet which is received correctly after a packet loss occurs, for example, a distance between elements of the ISF parameters is set wider than normal, whereby it can prevent the gain of the LP coefficients from becoming too large. Since it can prevent both the gain of the LP coefficient and the power of the excitation signal from increasing, a discontinuity of the synthesized signal is reduced, whereby a degradation of the subjective quality can be suppressed. Furthermore, the stability processor may reduce a discontinuity of the synthesized signal by multiplying the synthesized signal by the gain calculated by

using the LP coefficients or the like.

[0210] The discontinuity detector, which is the second feature of the invention, monitors the gain of the excitation signal included in the first packet which is received correctly after a packet loss occurs, and determines a discontinuity for a packet whose gain of the excitation signal increased more than a certain level.

Reference Signs List

[0211] 1, 1S, 1X audio decoder; 11 packet loss detector; 12, 12A, 12B, 12D, 12E, 12G, 12S audio code decoder; 13, 13X concealment signal generator; 14 internal state buffer; 15 normalized prediction residual power decoder; 16 reception state determiner; 21, 21S analyzer/encoder; 22 residual encoder; 23 code multiplexer; 120 ISF decoder; 121, 121S stability processor; 121X gain adjustor; 121Y gain multiplier; 122 LP coefficient calculator; 122A ISF-ISP converter; 122B ISP interpolator; 122C ISP-LPC converter; 123 adaptive codebook calculator; 124 fixed codebook decoder; 125 gain decoder; 126 excitation vector synthesizer; 127 post-filter; 128 synthesis filter; 129 discontinuity detector; 130 LP coefficient interpolator; 131 pitch lag interpolator; 132 gain interpolator; 133 noise signal generator; 134 post-filter; 135 synthesis filter; 136 adaptive codebook calculator; 137 excitation vector synthesizer; 138 normalized prediction residual adjustor; 210 LP analyzer; 211 LP-ISF converter; 212 ISF encoder; 213 discontinuity determiner; 214 ISF concealer; 215 ISF-LP converter; 216 ISF buffer; C10 computer; C12 reading device; C14 working memory; C16 memory; C18 display; C20 mouse; C22 keyboard; C24 communication device; C26 CPU; M recording medium; P1 to P4 audio signal processing programs; P11

discontinuity detection module; P12 discontinuity correction module; P21 ISF/LSF quantization module; P22 ISF/LSF concealment module; P23 discontinuity detection module; P24 auxiliary information encoding module; P31 discontinuity detection module; P32 auxiliary information encoding module; P33 ISF/LSF quantization module; P41 auxiliary information decoding module; P42 discontinuity correction module; P43 ISF/LSF decoding module; W computer data signal.

Claims

1. An audio signal processing method to be executed by an audio signal processing device, comprising:

determining an occurrence of a discontinuity of decoded audio, the discontinuity occurring with a sudden increase of an amplitude of the decoded audio, the decoded audio being obtained by decoding a first audio packet which is received correctly after an occurrence of a packet loss; and

correcting the discontinuity of the decoded audio,

wherein the step of correcting changes, according to a result of determination on an occurrence of a discontinuity, a distance between elements of ISF/LSF parameters.

2. An audio signal processing device comprising:

a discontinuity detector configured to determine an occurrence of a discontinuity of decoded audio, the discontinuity occurring with a sudden increase of amplitude of the decoded audio, the decoded audio being obtained by decoding a first audio packet that is received correctly after an occurrence of a packet loss; and

a discontinuity corrector configured to correct the discontinuity of the decoded audio,

wherein the discontinuity corrector changes, according to a result of determination on an occurrence of a discontinuity, a distance between elements of ISF/LSF parameters,

3. An audio signal processing method to be executed by an audio signal processing device, comprising:

quantizing ISF/LSF parameters;

generating concealment ISF/LSF parameters being representative of concealment information about the ISF/LSF parameters;

determining an occurrence of a discontinuity which occurs in a first audio packet which is received correctly after an occurrence of a packet loss, determining an occurrence of a discontinuity comprising using distances between quantized ISF/LSF parameters obtained in a quantization process of the ISF/LSF parameters and the generated concealment ISF/LSF parameters; and

encoding auxiliary information for determination on an occurrence of a discontinuity

4. An audio signal processing device comprising:

an ISF/LSF quantizer configured to quantize ISF/LSF parameters;

an ISF/LSF concealer configured to generate concealment ISF/LSF parameters representative of concealment information about the ISF/LSF parameters;

a discontinuity detector configured to determine an occurrence of a discontinuity which occurs in a first audio packet which is received correctly after an occurrence of a packet loss, the discontinuity detector being configured to use distances between quantized ISF/LSF parameters obtained in a quantization process of the ISF/LSF quantizer and the concealment ISF/LSF parameters generated by the ISF/LSF concealer to determine an occurrence of a discontinuity; and

an auxiliary information encoder configured to encode auxiliary information for determination on an occurrence of a discontinuity.

5. An audio signal processing program that programs a computer to operate as:

a discontinuity detector operable to determine an occurrence of a discontinuity of a decoded audio, the discontinuity occurring with a sudden increase of an amplitude of the decoded audio, the decoded audio being obtained by decoding a first audio packet which is received correctly after an occurrence of a packet loss; and

a discontinuity corrector operable to correct the discontinuity of the decoded audio,

wherein the discontinuity corrector changes, according to a result of determination on an occurrence of a discontinuity, a distance between elements of the ISF/LSF parameters.

6. An audio signal processing program that programs a computer to operate as:

an ISF/LSF quantizer operable to quantize ISF/LSF parameters;

an ISF/LSF concealer operable to generate concealment ISF/LSF parameters being representative of concealment information about the

ISF/LSF parameters;

a discontinuity detector operable to determine an occurrence of a discontinuity which occurs in a first audio packet which is received correctly after an occurrence of a packet loss, the discontinuity detector being operable to use distances between quantized ISF/LSF parameters obtained in a quantization process of the ISF/LSF quantizer and the concealment ISF/LSF parameters generated by the ISF/LSF concealer in order to determine an occurrence of a discontinuity; and

an auxiliary information encoder operable to encode auxiliary information for determination on an occurrence of a discontinuity.