(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 4 239 635 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
25.06.2025 Bulletin 2025/26

(21) Application number: **23187229.2**

(22) Date of filing: **04.11.2011**

(51) International Patent Classification (IPC):
G10L 19/005 *(2013.01)* ***G10L 25/21*** *(2013.01)*

(52) Cooperative Patent Classification (CPC):
G10L 19/005; G10L 25/21

(54) AUDIO ENCODING DEVICE AND METHOD

AUDIOCODIERUNGSVORRICHTUNG UND -VERFAHREN

DISPOSITIF ET PROCÉDÉ DE CODAGE AUDIO

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

(30) Priority: **22.11.2010 JP 2010260447**
18.02.2011 JP 2011033915

(43) Date of publication of application:
06.09.2023 Bulletin 2023/36

(60) Divisional application:
25164672.5 / 4 550 326

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
19161209.2 / 3 518 234
15184203.6 / 2 975 610
11842953.9 / 2 645 366

(73) Proprietor: **NTT DOCOMO, INC.**
Tokyo 100-6150 (JP)

(72) Inventors:
- **TSUTSUMI, Kimitaka**
Tokyo, 100-6150 (JP)
- **KIKUIRI, Kei**
Tokyo, 100-6150 (JP)

(74) Representative: **Viering, Jentschura & Partner mbB**
Patent- und Rechtsanwälte
Am Brauhaus 8
01099 Dresden (DE)

(56) References cited:
US-A1- 2005 154 584

- **GEISER BERND ET AL: "Joint pre-echo control and frame erasure concealment for VoIP audio codecs", 2009 17TH EUROPEAN SIGNAL PROCESSING CONFERENCE, IEEE, 24 August 2009 (2009-08-24), pages 1259 - 1263, XP032758754, ISBN: 978-1-61738-876-7, [retrieved on 20150401]**
- **BERND GEISER ET AL: "Candidate proposal for ITU-T super-wideband speech and audio coding", ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 2009. ICASSP 2009. IEEE INTERNATIONAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 19 April 2009 (2009-04-19), pages 4121 - 4124, XP031460181, ISBN: 978-1-4244-2353-8**
- **"G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729; G.729.1 (05/06)", ITU-T STANDARD, INTERNATIONAL TELECOMMUNICATION UNION, GENEVA ; CH, no. G.729.1 (05/06), 29 May 2006 (2006-05-29), pages 1 - 100, XP017466254**
- **MOHAMED CHIBANI ET AL: "Fast Recovery for a CELP-Like Speech Codec After a Frame Erasure", IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, USA, vol. 15, no. 8, 15 October 2007 (2007-10-15), pages 2485 - 2495, XP011192967, ISSN: 1558-7916, DOI: 10.1109/TASL.2007.907332**

EP 4 239 635 B1

- **KOMAKI N ET AL: "A PACKET LOSS CONCEALMENT TECHNIQUE FOR VOIP USING STEGANOGRAPHY", IEICE TRANSACTIONS ON FUNDAMENTALS OF ELECTRONICS,COMMUNICATIONS AND COMPUTER SCIENCES, ENGINEERING SCIENCES SOCIETY, TOKYO, JP, vol. E86-A, no. 8, 1 August 2003 (2003-08-01), pages 2069 - 2072, XP001177867, ISSN: 0916-8508**

Description

Technical Field

[0001] The present invention relates to error concealment in transmission of audio packets containing audio codes obtained by encoding an audio signal consisting of a plurality of frames, via an IP network or a mobile communication network and, more particularly, to an audio encoding device and audio encoding method to implement the error concealment.

Background Art

[0002] In transmitting an audio or acoustic signal (which will be generally referred to as an "audio signal") via an IP network or mobile communication, the audio signal is encoded to be expressed by a small bit count, the encoded data is divided into audio packets, and the audio packets are transmitted via the communication network. The audio packets received through the communication network are decoded by a receiver-side server, MCU, or terminal to obtain a decoded audio signal.

[0003] During the transmission of the audio packets via the communication network, a phenomenon can occur (so called packet losses) in which some audio packets are lost or errors are made in part of the information written in the audio packets. Such packet losses may occur because of a congestion condition of the communication network or the like. In such cases, the receiver side cannot correctly decode the audio packets and thus fails to obtain the desired decoded audio signal. Since the decoded audio signal corresponding to the audio packets subject to packet losses is perceived as noise, it significantly damages subjective quality for a human listener.

[0004] In order to overcome the inconvenience as described above, there are "concealment technologies on the receiver side" and "concealment technologies on the transmitter side," which may be known as packet loss concealment technologies to interpolate the audio or acoustic signal in the lost portions due to the packet losses.

[0005] The "concealment technologies on the receiver side" are, for example, like the technology of Non Patent Literature 1, to duplicate a decoded audio signal included in a packet normally received in the past, in pitch units, and multiply the duplication by a predetermined attenuation coefficient to generate an audio signal corresponding to a packet loss part. However, the "concealment technologies on the receiver side" are based on the premise that the property of audio of the packet loss part resembles that of audio immediately before the packet loss, and therefore these technologies cannot demonstrate a sufficient concealment effect if the packet loss part has a property different from that of the audio immediately before the loss, or if the power, or the energy of the audio, changes suddenly.

[0006] Furthermore, the "concealment technologies on the receiver side" also include the technology of Patent Literature 1 as a more advanced technology. This technology of Patent Literature 1 is different from the aforementioned technology of Non Patent Literature 1 in that, while the concealment signal is generated by duplicating the decoded audio contained in the packet normally received in the past, the duplication is multiplied by an attenuation coefficient that varies depending upon the property of the duplication source audio (shape of a power spectrum thereof), so as to implement high-quality shaping of the concealment signal with little abnormal sound.

[0007] On the other hand, the "concealment technologies on the transmitter side" can include the technology of Patent Literature 2 and the technology of Patent Literature 3.

[0008] The technology of Patent Literature 2 is to save audio signals contained in packets normally received in the past, in a buffer, and, with a packet loss, encode and transmit as auxiliary information, position information to indicate from which position in the buffer an audio signal should be duplicated. In addition to the position information, amplitude information to indicate whether the packet loss part is a silent interval is also contained in the auxiliary information, thereby preventing unwanted audio from being mixed in the case where the packet loss part is originally a silent interval.

[0009] In the technology of Patent Literature 3, a decoding device has a first concealment device to conceal a packet loss, a second concealment device to correct the first concealment signal output from the first concealment device, based on auxiliary information, and an auxiliary information decoding device to decode the auxiliary information. When the first concealment device fails to demonstrate a satisfactory concealment effect, the second concealment device corrects the first concealment signal, using the auxiliary information generated by the auxiliary information decoding device, to generate a second concealment signal. The auxiliary information to be used is a power spectrum envelope, or an encoded value of an error between an estimated value from a power spectrum envelope of an adjacent frame and an input power spectrum envelope. The second concealment device multiplies the first concealment signal by a gain in the frequency domain so as to provide the second concealment signal with the power spectrum envelope that can be used as the auxiliary information, to generate the second concealment signal with accuracy higher than the first concealment signal.

Citation List

Patent Literatures

[0010]

Patent Literature 1: Domestic re-publication of PCT publication WO2007/000988
Patent Literature 2: Japanese Patent Application Laid-open No. 2003-316670
Patent Literature 3: Japanese Patent Application Laid-open No. 2008-111991

Non Patent Literature

[0011]

Non Patent Literature 1: ITU-T G.711 Appendix I
Non Patent Literature 2: GEISER BERNDT ET AL: «Joint pre-echo control and frame erasure concealment of VOIP audio codecs", 2009 17th EUROPENA SIGNAL PROCESSING CONFERENCE, IEEE, 24 August 2009 ("009-08-24), pages 1259-1263, ISBN: 978-1-61738-876-7 [retrieved on 2015-04-01] discloses methods for concealment of lost audio packets in Voice-over-IP (VoIP) applications.
Non Patent Literature 3: BERND GEISER ET AL: «Candidate proposal for IUT-T super-wideband speech and audio coding", ACOUTICS, SPEECH AND SIGNAL PROCESSING, 2009. ICASSP 2009, NJ, USA, 19 April 2009 (2009-04-19), pages 4121-4124, ISBN: 978-1-4244-2353-8 discloses a speech and audio codec with increased encoded frequency range and maximum bit rate.

Summary of Invention

Technical Problem

[0012] Since the technology of Patent Literature 1 describes a technique to generate a concealment signal by prediction from the decoded signal normally received in the past, it is difficult to highly accurately generate the concealment signal with a power change of the audio signal that is significantly different than the prediction result, e.g., like generation of "clacks" of castanets as the concealment signal, from a past audio signal that does not include such "clacks."
[0013] The technology of Patent Literature 2 generates the amplitude information about the silent interval on the transmitter side so as to prevent the concealment signal from being generated in the case of the packet loss part being the silent interval, but fails to demonstrate a satisfactory concealment effect on sound with a sudden power change like the "clacks" of castanets as discussed above.
[0014] Since the technology of Patent Literature 3 is a method to perform the processing in the frequency domain after the time-frequency transform in frame units, the units of processing are the frame units and it is thus difficult to handle a sudden power change within a frame. Since the decoded audio of the packet loss part is recovered with high accuracy on the premise that there is a high correlation between the past signal and the packet loss signal, the correlation of signals becomes lower if the packet loss occurs in a part of the signal where the power changes suddenly. When the power changes suddenly, increases in a prediction error of the power spectrum envelope results, and it becomes difficult to encode the signal by a small bit count, and to generate the decoded audio with high accuracy.
[0015] As described above, the conventional technologies have the problem that they fail to show a satisfactory error concealment effect on a signal with a temporally quick power change (which will be referred to hereinafter as "transient signal") like hand claps and "clacks" of castanets. Namely, it is extremely difficult for the receiver side to accurately estimate at what timing the transient signal appears in the audio signal, based on the decoded signal obtained by decoding the audio packets normally received immediately before.
[0016] An object of the present invention is to provide an error concealment technology enabling high-accuracy concealment of a packet loss in a transient signal, the prediction of which from a preceding or following signal is difficult, while solving the above problem.
[0017] Non Patent Literature 2 discloses methods for concealment of lost audio packets in Voice-over-IP (VoIP) applications.
[0018] Non Patent Literature 3 discloses a speech and audio codec with increased encoded frequency range and maximum bit rate.

Solution to Problem

[0019] According to an aspect of the present invention there is provided an audio encoding device according to claim 1.
[0020] According to another aspect of the present invention there is provided an audio encoding method according to

claim 2.

Advantageous Effect of Invention

[0021] Since the present invention enables transmission of the information about a sudden power-changing part of a signal using the method described above, it realizes high-accuracy packet loss concealment of a signal upon occurrence of a sudden temporal change of power (transient signal), which by conventional technologies such packet loss concealment was difficult.

Brief Description of Drawings

[0022]

Fig. 1 is a drawing showing a system environment.
Fig. 2 is a configuration diagram of an encoding unit in the first, second, third, and sixth embodiments, which are informative examples.
Fig. 3 is a flowchart of processing by the encoding unit in Fig. 2.
Fig. 4 is a configuration diagram of an auxiliary information encoding unit in the first embodiment and others.
Fig. 5 is a drawing showing a temporal relation between signals as audio encoding targets and signals as auxiliary information encoding targets, and a configuration example of bitstreams.
Fig. 6 is a configuration diagram of a decoding unit in the first, second, third, fifth, and sixth embodiments, which are informative examples.
Fig. 7 is a flowchart of processing by the decoding unit in Fig. 6, which relates to informative examples.
Fig. 8 is a flowchart showing an example of processing by a concealment signal correction unit, which relates to an informative example.
Fig. 9 is a drawing showing an example of a configuration of the auxiliary information encoding unit.
Fig. 10 is a configuration diagram of the encoding unit in the fourth and fifth embodiments, which are informative examples.
Fig. 11 is a drawing showing an example of a configuration of a first concealment signal generation unit, which relates to an informative example.
Fig. 12 is a drawing showing an example of a configuration of the concealment signal correction unit, which relates to an informative example.
Fig. 13 is a configuration diagram of the decoding unit in the fourth embodiment, which is an informative example.
Fig. 14 is a drawing showing a temporal relation between signals as audio encoding targets and signals as auxiliary information encoding targets, and a configuration example of bitstreams in the sixth embodiment, which is an informative example.
Fig. 15 is a hardware configuration diagram of a computer.
Fig. 16 is an appearance diagram of the computer.
Fig. 17 is a drawing showing a configuration of an audio encoding program.
Fig. 18 is a drawing showing a configuration of an audio decoding program, which relates to an informative example.
Fig. 19 is a drawing showing another configuration example of the decoding unit, which relates to an informative example.
Fig. 20 is a configuration diagram of the auxiliary information encoding unit in the seventh embodiment, which is according to the present invention.
Fig. 21 is a flowchart of processing by the auxiliary information encoding unit in Fig. 20.
Fig. 22 is a configuration diagram of the auxiliary information decoding unit in the seventh and eleventh embodiments, the eleventh embodiment refers to an informative example.
Fig. 23 is a flowchart of processing by the auxiliary information decoding unit in Fig. 22.
Fig. 24 is a configuration diagram of the concealment signal correction unit in the seventh and eighth embodiments, the eighth embodiment refers to an informative example.
Fig. 25 is a flowchart of processing by the concealment signal correction unit in the seventh embodiment, which is according to the present invention.
Fig. 26 is a configuration diagram of the auxiliary information encoding unit in the eighth embodiment, which is an informative example.
Fig. 27 is a flowchart of processing by the auxiliary information encoding unit in Fig. 26.
Fig. 28 is a configuration diagram showing a modification example of the auxiliary information encoding unit in the eighth embodiment, which is an informative example.
Fig. 29 is a flowchart of processing by the auxiliary information encoding unit in Fig. 28.

Fig. 30 is a configuration diagram of the auxiliary information decoding unit in the eighth embodiment, which is an informative example.
Fig. 31 is a flowchart of processing by the auxiliary information decoding unit in Fig. 30, which relates to an informative example.
Fig. 32 is a flowchart of processing by the concealment signal correction unit in the eighth embodiment, which is an informative example.
Fig. 33 is a configuration diagram of the auxiliary information encoding unit in the tenth embodiment, which is an informative example.
Fig. 34 is a flowchart of processing by the auxiliary information encoding unit in Fig. 33.
Fig. 35 is a configuration diagram of the auxiliary information decoding unit in the tenth embodiment, which is an informative example.
Fig. 36 is a flowchart of processing by the auxiliary information decoding unit in Fig. 35.
Fig. 37 is a flowchart of processing by the concealment signal correction unit in the tenth embodiment, which is an informative example.
Fig. 38 is a configuration diagram of the auxiliary information encoding unit in the eleventh embodiment, which is an informative example.
Fig. 39 is a flowchart of processing by the auxiliary information encoding unit in Fig. 38.
Fig. 40 is a flowchart of processing by the auxiliary information decoding unit in the eleventh embodiment, which is an informative example.
Fig. 41 is a diagram showing an output content from a transient detection unit.
Fig. 42 is a drawing showing examples of scalar quantization methods for transient position information.
Fig. 43 is a configuration diagram of the auxiliary information encoding unit in the twelfth embodiment, which is an informative example.
Fig. 44 is a configuration diagram of the auxiliary information decoding unit in the twelfth embodiment, which is an informative example.
Fig. 45 is a configuration diagram of the auxiliary information encoding unit in the thirteenth embodiment.
Fig. 46 is a configuration diagram of the auxiliary information decoding unit in the thirteenth embodiment, which is an informative example.
Fig. 47 is a configuration diagram of the auxiliary information encoding unit in the fourteenth embodiment, which is an informative example.
Fig. 48 is a configuration diagram of the auxiliary information decoding unit in the fourteenth embodiment, which is an informative example.
Fig. 49 is a configuration diagram of the auxiliary information encoding unit in the fifteenth embodiment, which is an informative example.
Fig. 50 is a configuration diagram of the auxiliary information decoding unit in the fifteenth embodiment, which is an informative example.

Description of Embodiments and Informative Examples as such not covered by the claims but helpful for understanding the present invention

[0023] Various embodiments according to the present invention will be described below using the drawings.

[First Embodiment] - Informative Example

[0024] First, a system environment assumed by the present invention will be described using Fig. 1. As shown in Fig. 1, an audio signal acquired through a sensor such as a microphone is expressed in digital format and fed to an encoding unit 1.

[0025] The encoding unit 1 encodes digital signals in a buffer every time a predetermined amount of audio signals consisting of a predetermined number of samples are saved in a built-in buffer. The foregoing predetermined amount, i.e., the number of sample to be saved is called a frame length and an aggregate of digital signals saved in the buffer is called a frame. For example, in a case where audio is collected at the sampling frequency of 32 kHz and where the frame length is 20 ms, digital signals of 640 samples shall be saved in the buffer. The length of the buffer may be longer than one frame. For example, when the length of the buffer is set to that of two frames, encoding at the beginning is started only after digital signals of two frames have been saved in the buffer, whereby the digital signal of the next frame to the frame as an encoding target can be used for estimation of auxiliary information. The timing of execution of encoding may be determined so as to execute encoding in units of the frame length, or so as to execute encoding with an overlap of a certain length between frames. The encoding is performed using audio encoding such as 3GPP enhanced aacPlus and G.718. It should be noted that any method may be applicable as to the method of audio encoding. The auxiliary information is calculated using an audio or acoustic signal saved in the buffer for calculation of auxiliary information, and then is encoded and transmitted (auxiliary information code). The auxiliary information code may be transmitted in the same packet as an audio code, or may be transmitted in another packet different from a packet containing the audio code. The details of the operation of the encoding unit 1 will be described later.

[0026] A packet configuration unit 2 adds information necessary for communication such as an RTP header to the audio code acquired by the encoding unit 1, to generate an audio packet. The audio packet thus generated is sent through a network to a receiver.

[0027] A packet separation unit 3 separates the audio packet received through the network, into the packet header information and the other part (the audio code and auxiliary information code, which will be referred to hereinafter as "bitstream") and outputs the bitstream to a decoding unit 4.

[0028] The decoding unit 4 performs decoding of the audio code contained in the audio packet received normally, and, if it detects an abnormality (a packet error or a packet loss) in the received audio packet, it performs packet loss concealment. The detailed operation of the decoding unit 4 will be described in the below embodiment. The decoded audio output from the decoding unit 4 is sent to a buffer of audio or the like to be reproduced through a speaker or the like, or stored in a recording medium such as a memory or a hard disk.

[0029] Since the overall configuration in Fig. 1 described above is also applied similarly to the second to sixth embodiments described below, redundant description of the overall configuration will be omitted in the second to sixth embodiments.

[0030] Now, the encoding unit 1 and the decoding unit 4 will be described below in detail as characteristic portions of the first embodiment. The first embodiment will describe an example in which a parameter obtained by a functional approximation of powers of subframes shorter than one frame is used as auxiliary information about a temporal change of power.

(Configuration and Operation of Encoding Unit 1)

[0031] As shown in Fig. 2, the encoding unit 1 is provided with an audio encoding unit 11 to encode an audio signal, an auxiliary information encoding unit 12 to estimate and encode auxiliary information about a temporal change of power of the audio signal, which is used in packet loss concealment in decoding of the audio signal, and a code multiplexing unit 13 to multiplex an auxiliary information code obtained in encoding by the auxiliary information encoding unit 12 and an audio code obtained in encoding by the audio encoding unit 11, and output a bitstream of multiplex data.

[0032] The auxiliary information encoding unit 12 of these units, as shown in Fig. 4, is provided with a subframe power calculation unit 121, an attenuation coefficient estimation unit 122, and an attenuation coefficient quantization unit 123 which will be described later.

[0033] The operation of the encoding unit 1 will be described below using Fig. 3.

[0034] The audio encoding unit 11 saves audio signal for a predetermined period of time and encodes a signal of an encoding target out of the saved input audio (step S1101 in Fig. 3). The encoding may be performed, for example, using the audio encoding such as 3GPP enhanced aacPlus defined in Literature "3GPP TS26.401 'Enhanced aacPlus general audio codec General description'" and G.718 defined in Literature "Recommendation ITU-T G.718 'Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32kbit/s'", or using any other encoding method.

[0035] The subframe power calculation unit 121 in the auxiliary information encoding unit 12 saves the input audio for a predetermined period of time and later calculates a subframe power sequence for audio signals s(dT), s(1+dT), ..., s((d+1) T-1) out of the saved input audio. The calculation may occur later than encoding of target signals s(0), s(1), ..., s(T-1) by a predetermined number of frames (d frames in the present embodiment) (step S1211 in Fig. 3). The number of samples contained in one frame is defined as T herein. When a prediction target signal is defined by the following formula:

$$v(K \cdot l + k) = s(K \cdot l + k + dT) ,$$

a power P(1) of a subframe 1 (0 ≤ 1 ≤ L- 1) is obtained by the formula below. The letter k represents an index of a sample in each subframe (0 ≤ k ≤ K- 1). It is assumed herein that the number of samples in a digital signal in each subframe is K.

$$P(l) = 10\log_{10}\left(\frac{1}{K}\sum_{k=0}^{K-1} v^2(K \cdot l + k)\right)$$

[0036] Although it is assumed in this first embodiment that the length of subframes is K, it is also possible to use different lengths determined in advance for the respective subframes. The subframe power sequence may be calculated according to the following formula, where k^l_{start} represents an index of a start of the lth subframe and k^l_{end} represents an index of an end thereof.

$$P(l) = 10\log_{10}\left(\frac{1}{k^l_{end} - k^l_{start}} \sum_{k=k^l_{start}}^{k^l_{end}} v^2(k^{l-1}_{end} + k)\right)$$

[0037] The attenuation coefficient estimation unit 122 acquires from the subframe power sequence a slope γ_{opt} of a straight line representing a temporal change of power for example, by the least square method or the like (step S1221 in Fig. 3). More simply, the slope may be calculated from P(0) and P(L-1). Here, the letter L represents the number of subframes contained in one frame. In addition to the slope γ_{opt} of the straight line, an intercept P_{opt} may be calculated by a straight-line approximation of the subframe power sequence P(l).

[0038] The power of subframe m is expressed herein by the following formula.

$$\hat{P}(m) = \gamma_{opt} \cdot m + P_{opt}$$

At this time, the slope γ_{opt} and intercept P_{opt} of the straight line are acquired in accordance with the following formulas (the least square method).

$$\gamma_{opt} = \frac{L\sum_{m=0}^{L-1} m \cdot P(m) - \sum_{m=0}^{L-1} m \sum_{m=0}^{L-1} P(m)}{L\sum_{m=0}^{L-1} m^2 - \left(\sum_{m=0}^{L-1} m\right)^2}$$

$$P_{opt} = \frac{\sum_{m=0}^{L-1} m^2 \sum_{m=0}^{L-1} P(m) - \sum_{m=0}^{L-1} m \cdot P(m) \sum_{m=0}^{L-1} P(m)}{L\sum_{m=0}^{L-1} m^2 - \left(\sum_{m=0}^{L-1} m\right)^2}$$

[0039] The attenuation coefficient quantization unit 123 performs scalar quantization of the slope γ_{opt} of the straight line, then encodes the quantized data, and outputs the auxiliary information code (step S1231 in Fig. 3). It may use a scalar quantization codebook prepared in advance. In the case of the straight-line approximation of subframe powers P(l), the intercept P_{opt} may also be encoded in addition to the slope γ_{opt} of the straight line.

[0040] The code multiplexing unit 13 writes the audio code and the auxiliary information code in a predetermined order in a bitstream and outputs the bitstream (step S1301 in Fig. 3). Fig. 5 shows an example of the temporal relationship between signals as audio encoding targets and signals as auxiliary information encoding targets, and a configuration of bitstreams (in the case of d=1). For example, as shown in Fig. 5, the auxiliary information code of frame (N+1), for example, is added to the audio code of frame N to obtain a bitstream, which is output from the code multiplexing unit 13. Furthermore, the packet configuration unit 2 adds the packet header information to the bitstream to obtain an audio packet to be transmitted as the N-th packet.

[0041] The above processing of steps S1101 to S1301 is repeated to an end of the input audio (step S1401).

(Configuration and Operation of Decoding Unit 4)

[0042] As shown in Fig. 6, the decoding unit 4 is provided with an error/loss detection unit 41, a code separation unit 40, an audio decoding unit 42, an auxiliary information decoding unit 45, a first concealment signal generation unit 43, and a concealment signal correction unit 44. The first concealment signal generation unit 43 of these units, as shown in Fig. 11, is provided with a decoding coefficient storage unit 431 and a stored decoding coefficient repetition unit 432. The concealment signal correction unit 44, as shown in Fig. 12, is provided with an auxiliary information storage unit 441 and a subframe power correction unit 442.

[0043] The operation of the decoding unit 4 will be described below using Figs. 6 and 7.

[0044] The error/loss detection unit 41 detects an abnormality (a packet error or a packet loss) in a received audio packet and outputs an error flag indicative of the result of the detection (step S4101 in Fig. 7). The error flag is set off to indicate the normality of packet by default and, when the error/loss detection unit 41 detects an abnormality in the received audio packet, it sets the error flag on (to indicate the packet abnormality). For example, the error/loss detection unit 41 is provided with a counter that increases one for every reception of a new packet, and, when packets are assumed to be numbered in an order of transmission from the encoder, the error/loss detection unit 41 can compare a counter value with a number given to a packet to detect a packet loss if these values are different. It should be, however, noted that the packet loss detection method in the error/loss detection unit 41 described herein is just an example and the packet loss may be detected by any other method.

[0045] The operation will be described below in each of the case of the error flag being on (packet abnormality) and the case of the error flag being off (packet normality).

(Case of Error Flag Being Off (Case of NO in Step S4102 in Fig. 7))

[0046] The error/loss detection unit 41 sends the error flag to the audio decoding unit 42, the first concealment signal generation unit 43, the concealment signal correction unit 44, and the auxiliary information decoding unit 45 and sends the bitstream to the code separation unit 40.

[0047] The code separation unit 40 receives the bitstream from the error/loss detection unit 41, separates the bitstream into the audio code and the auxiliary information code, and sends the audio code to the audio decoding unit 42 and the auxiliary information code to the auxiliary information decoding unit 45 (step S4001 in Fig. 7).

[0048] The audio decoding unit 42 decodes the audio code to generate a decoded signal and outputs it as decoded audio. The decoding of audio code is performed using a decoding method corresponding to the aforementioned audio encoding unit 11. At this time, the audio decoding unit 42 also sends the decoded signal to the first concealment signal generation unit 43 (step S4311 in Fig. 7). At this time, the first concealment signal generation unit 43 stores the sent decoded signal into the decoding coefficient storage unit 431 shown in Fig. 11. The stored decoded signal in storage therein is denoted by b(k, l). The stored signal may be at least d or more past frames. The letter k herein represents an index of a sample in a subframe (provided that $0 \leq k \leq K - 1$) and the letter 1 an index of a subframe stored in the decoding coefficient storage unit 431 (provided that $0 \leq 1 \leq dL - 1$).

[0049] The auxiliary information decoding unit 45 decodes the auxiliary information code output from the code separation unit 40, to generate the auxiliary information, and then sends the auxiliary information to the concealment signal correction unit 44 (step S4202 in Fig. 7). At this time, the concealment signal correction unit 44 stores the auxiliary information into the auxiliary information storage unit 441 shown in Fig. 12. The auxiliary information stored at this time is preferably that of several past frames (that of at least d frames or more).

[0050] In above step S4202 the auxiliary information decoding unit 45 decodes the auxiliary information code output from the code separation unit 40, to generate an index, and obtains a slope γ_J of a straight line corresponding to the index from a codebook. Here, P(-1) represents a power of the last subframe in a signal received normally immediately before a frame loss.

$$\hat{P}(m) = \gamma_J \cdot m + P(-1)$$

In the case where an intercept of the straight line is simultaneously encoded by a straight-line approximation of powers of subframes, the subframe power is obtained by the following formula using the intercept P_J.

$$\hat{P}(m) = \gamma_J \cdot m + P_J$$

(Case of Error Flag Being On (Case of YES in Step S4102 in Fig. 7))

[0051] The error/loss detection unit 41 sends the error flag to the audio decoding unit 42, the first concealment signal generation unit 43, the concealment signal correction unit 44, and the auxiliary information decoding unit 45.

[0052] The stored decoding coefficient repetition unit 432 in the first concealment signal generation unit 43 obtains a first concealment signal z(k) using a stored decoding signal stored in the decoding coefficient storage unit 431 (step S4321 in Fig. 7). Specifically, it calculates the first concealment signal by repetition of the last subframe, for example, as expressed by the following formula.

$$z(K \cdot l + k) = b(k, dL - 1)$$

(provided that 0 ≤ 1 ≤ dL- 1 and 0 ≤ k ≤ K - 1)

[0053] It should be noted herein that the unit of repetition does not have to be limited to the last subframe but instead any part of b(k, l) may be extracted and repeated. Generation of the first concealment signal is not limited to the repetition as described above, and instead the first concealment signal may be calculated by extracting and repeating a waveform in a pitch unit from the decoding coefficient storage unit 431 or the first concealment signal may be generated by a prediction, for example, using the linear prediction. Alternatively, the first concealment signal may be generated in accordance with a model determined in advance, for example, as shown below.

$$[z(K \cdot (L-1)), \ldots, z(K \cdot L - 1)] = f(b(0,0), b(1,0) \ldots, b(K-1, dL-1))$$

[0054] The subframe power correction unit 442 corrects the first concealment signal for a value of power of the first concealment signal in each of the subframes in accordance with the formula below to acquire a concealment signal y(K·l+k). Specifically, it performs the correction according to the below formula (provided that 0 ≤ 1 ≤ L-1 and 0 ≤ k ≤ K- 1). In the formula, $P^{-d}(m)$ represents a power about a subframe contained in the auxiliary information code transmitted in the d-th packet before the packet (packet as a first concealment signal generation target) (step S4421 in Fig. 7).

$$\hat{P}(m) = P^{-d}(m)$$

$$z'(K \cdot l + k) = \frac{z(K \cdot l + k)}{\sqrt{\frac{1}{K}\sum_{k=0}^{K-1} z^2(K \cdot l + k)}}$$

$$y(K \cdot l + k) = 10^{\hat{P}(m)/20} \cdot z'(K \cdot l + k)$$

[0055] For example, the subframe power correction unit 442, as shown in Fig. 8, extracts the auxiliary information previously transmitted in the d-th packet, from the auxiliary information storage unit 441 (step S60 in Fig. 8), calculates a mean square amplitude value for each subframe as to the first concealment signal, and divides a value contained in each subframe, by the mean square amplitude value (step S61 in Fig. 8). This operation results in obtaining z'(K·l+k). Then it calculates a power of each subframe from the auxiliary information and multiplies the foregoing value of the subframe by a mean amplitude value obtained from the power (step S62 in Fig. 8). This multiplication results in obtaining the concealment signal y(K-l+k).

[0056] The above processing of steps S4101 to S4421 in Fig. 7 is repeated to the end of the input audio (step S4431 in Fig. 7).

[0057] As described above, the first embodiment can use the parameter obtained by the functional approximation of powers of subframes shorter than one frame, as the auxiliary information about the temporal change of power.

[Second Embodiment] - Informative Example

[0058] The auxiliary information may be auxiliary information obtained by encoding a subframe power sequence by vector quantization using preliminarily-learned or empirically-determined vectors $c_i(l)$. The second embodiment will describe an example of encoding or decoding, using as the auxiliary information, information about a vector obtained by vector quantization of powers of subframes, in the auxiliary information encoding unit 12 or in the auxiliary information decoding unit 45 in the first embodiment.

[0059] Since the second embodiment is different only in the auxiliary information encoding unit 12 and the auxiliary information decoding unit 45 from the first embodiment, these two elements will be described below.

[0060] The auxiliary information encoding unit 12, as shown in Fig. 9, is provided with the subframe power calculation unit 121 and a subframe power vector quantization unit 124. The function and operation of the subframe power calculation unit 121 is the same as in the first embodiment.

[0061] The subframe power vector quantization unit 124 performs vector quantization of powers P(l) of subframes l (provided that $0 \leq l \leq L - 1$), encodes the result, and outputs the auxiliary information code. The letter I represents the number of entries of straight lines or vectors in a codebook and the letter J represents an index of a straight line or a vector selected. $c_i(l)$ represents the lth element of the ith code vector in the codebook.

$$J = \underset{i=0,..,I-1}{\arg\min} \sum_{l=0}^{L-1} (c_i(l) - P(l))^2$$

Selected J is encoded by binary encoding to obtain the auxiliary information code.

[0062] On the other hand, the auxiliary information decoding unit 45 decodes the auxiliary information code output from the code separation unit 40, to generate the index J, obtains a vector $c_J(l)$ corresponding to the index J from the codebook, and outputs it.

$$\hat{P}(m) = c_J(l)$$

[0063] As described above, the second embodiment involves the encoding of the subframe power sequence by vector quantization using the preliminarily-learned or empirically-determined vectors, and uses the result as the auxiliary information.

[Third Embodiment] - Informative example

[0064] The calculation of the auxiliary information in above-described first and second embodiments used a signal that is later by d or more frames than the signal encoded by the audio encoding unit 11, whereas the below third embodiment will describe an example in which a signal that is earlier by d frames than the signal encoded by the audio encoding unit 11 is used in the calculation of the auxiliary information.

[0065] Since the following third embodiment is different from the first embodiment only in the subframe power calculation unit 121 included in the auxiliary information encoding unit 12, and the subframe power correction unit 442 included in the concealment signal correction unit 44, the subframe power calculation unit 121 and subframe power correction unit 442 will be described below.

[0066] The subframe power calculation unit 121 saves input audio for a predetermined period of time and the subframe power sequence for audio signals s(-dT), s(1-dT), ..., s(-1) is calculated earlier by a predetermined number of frames (d frames in the present embodiment) than the encoding of target signals s(0), s(1), ..., s(T-1) out of the saved input audio. It is assumed herein that the number of samples contained in one frame is T. When a prediction target signal is expressed by the following formula:

$$v(K \cdot l + k) = s(K \cdot l + k + dT),$$

the power P(l) of subframe l ($0 \leq l \leq L - 1$) is obtained by the formula below. The letter k represents an index of a sample in a subframe ($0 \leq k \leq K - 1$). It is assumed herein that the number of samples of digital signals contained in each subframe is K.

$$P(l) = 10\log_{10}\left(\frac{1}{K}\sum_{k=0}^{K-1} v^2(K \cdot l + k)\right)$$

[0067] On the other hand, the subframe power correction unit 442 corrects the first concealment signal for a value of power of the first concealment signal in each subframe in accordance with the formula below to obtain the concealment signal y(K·l+k). Specifically, it performs the correction in accordance with the below formula (provided that $0 \leq 1 \leq L - 1$ and $0 \leq k \leq K - 1$). $P^d(m)$ represents the power about the subframe contained in the auxiliary information code transmitted in the d-th packet after the pertinent packet (packet of a first concealment signal generation target).

$$\hat{P}(m) = P^d(m)$$

$$z'(K \cdot l + k) = \frac{z(K \cdot l + k)}{\sqrt{\frac{1}{K}\sum_{k=0}^{K-1} z^2(K \cdot l + k)}}$$

$$y(K \cdot l + k) = 10^{\hat{P}(m)/20} \cdot z'(K \cdot l + k)$$

[0068] As described above, the third embodiment allows use of the signal earlier by several frames than the signal encoded by the audio encoding unit for the calculation of the auxiliary information.

[Fourth Embodiment] - Informative Example

[0069] The fourth embodiment will describe an example in which the processing as executed in the first and second embodiments is applied to signals resulting from time-frequency transform.

[0070] The encoding unit 1 in the fourth embodiment has a configuration, as shown in Fig. 10, in which a time-frequency transform unit 10 is added to the input side of the audio encoding unit 11 and the auxiliary information encoding unit 12, in comparison to the encoding unit 1 (Fig. 2) in the first and second embodiments.

[0071] The time-frequency transform unit 10 performs a time-frequency transform of an audio signal using an analysis QMF. Specifically, it performs the time-frequency transform by the following formula.

$$V(k,l) = \sum_{n=-E \cdot K}^{2K-1} p_A(n) \cdot x(n) \cos\left[\frac{\pi}{K}(n + \frac{1}{2} - \frac{K}{2})(k + \frac{1}{2})\right]$$

In this formula, the letter E represents the number of subframes in the time direction and the letter K represents the number of frequency bins. The letter k represents an index of a frequency bin (provided that $0 \leq k \leq K-1$) and the letter l represents an index of a subframe (provided that $0 \leq l \leq L-1$). As an alternative to the analysis QMF, the time-frequency transform can also be executed by MDCT (Modified Discrete Cosine Transform) or the like.

[0072] The audio encoding unit 11 encodes the audio signal resulting from the time-frequency transform. For example, it may perform the encoding by an encoding method, for example, such as SBR (Spectral Band Replication), but the encoding may be executed by any encoding method.

[0073] The auxiliary information encoding unit 12, as shown in Fig. 4, is provided with the subframe power calculation unit 121, attenuation coefficient estimation unit 122, and attenuation coefficient quantization unit 123. Since only the subframe power calculation unit 121 of these constituent elements is different from that in the first and second embodiments, the subframe power calculation unit 121 will be described below. The attenuation coefficient quantization unit 123 may employ the vector quantization as described in the second embodiment.

[0074] The subframe power calculation unit 121 saves the audio signal for a predetermined period of time, and calculates the auxiliary information out of the saved audio signal as described below, using an audio signal V(k, l+d) obtained by transforming into the time-frequency domain an audio signal that is later by a predetermined number of frames (d frames) than the encoding of the target signal V(k, l). The power P(l+d) of subframe l+d is calculated by the following formula.

$$P(l + d) = 10\log_{10}(\frac{1}{K}\sum_{k=0}^{K-1} V^2(k, l + d))$$

The code multiplexing unit 13 writes the audio code and the auxiliary information code in a predetermined order, in the same manner as in the first and second embodiments, and outputs the resulting bitstream.

[0075] On the other hand, the decoding unit 4 in the fourth embodiment has a configuration, as shown in Fig. 13, in which an inverse transform unit 46 is added to the output side of the audio decoding unit 42 and the concealment signal correction unit 44, in comparison to the decoding unit 4 (Fig. 6) in the first and second embodiments.

[0076] In the decoding unit 4 in Fig. 13 as described above, the operations of the error/loss detection unit 41, code separation unit 40, and audio decoding unit 42 are the same as in the first and second embodiments, and thus the operations of the first concealment signal generation unit 43, auxiliary information decoding unit 45, concealment signal correction unit 44, and inverse transform unit 46 will be described below.

[0077] As shown in Fig. 11, the first concealment signal generation unit 43 is provided with the decoding coefficient storage unit 431 and the stored decoding coefficient repetition unit 432. The decoding coefficient storage unit 431 stores the decoded signal fed from the audio decoding unit 42. The stored decoded signal in storage is denoted by B(k, 1). The letter k herein represents an index of a sample in a subframe (provided that $0 \le k \le K-1$) and l represents an index of a subframe stored in the decoding coefficient storage unit 431 (provided that $0 \le 1 \le L-1$).

[0078] When the error flag is on (to indicate a packet abnormality), the stored decoding coefficient repetition unit 432 obtains the first concealment signal z(k, l) using the stored decoded signal stored in the decoding coefficient storage unit 431. Specifically, it calculates the first concealment signal, for example, by repetition of the last subframe in accordance with the following formula.

$$z(k,l) = B(k, L-1)$$

(provided that $0 \le l \le L-1$ and $0 \le k \le K-1$)

The unit of repetition does not have to be limited to the last subframe, and any part of B(k, l) may be extracted and repeated, or the first concealment signal may be generated, for example, by prediction using the linear prediction. Altermatively, the first concealment signal may be generated, for example, in accordance with a model determined in advance as described below.

$$[z(k,0)\ldots, z(k, L-1)] = f(B(0,0), B(1,0)\ldots, B(K-1, L-1))$$

[0079] The auxiliary information decoding unit 45 decodes the auxiliary information code output by the code separation unit 40 to generate an index, obtains a slope γ_J of a straight line corresponding to the index from the codebook, and outputs it. Here, P(-1) represents the power of the last subframe in the signal received normally immediately before the frame loss.

$$\hat{P}(m) = \gamma_J \cdot m + P(-1)$$

In the case where the intercept of the straight line is simultaneously encoded based on the straight-line approximation of powers of subframes, the subframe powers are obtained by the following formula using the intercept P_J.

$$\hat{P}(m) = \gamma_J \cdot m + P_J$$

[0080] In the case where the vector quantization is used in the attenuation coefficient quantization unit 123 included in the auxiliary information encoding unit 12 as in the second embodiment, the auxiliary information decoding unit 45 in the present embodiment calculates the powers of the subframes using the codebook, as does the auxiliary information decoding unit 45 in the second embodiment.

[0081] As shown in Fig. 12, the concealment signal correction unit 44 is provided with the auxiliary information storage unit 441 and the subframe power correction unit 442. The auxiliary information storage unit 441 stores the auxiliary information fed from the auxiliary information decoding unit 45 when the error flag is off (to indicate packet normality). The auxiliary information to be stored is preferably that of several past frames. The subframe power correction unit 442 corrects the first concealment signal for a value of power of the first concealment signal in each subframe in accordance with the formula below to obtain the concealment signal Y(k, l). Specifically, it performs the correction in accordance with the below formula (provided that $0 \le l \le L-1$ and $0 \le k \le K-1$). $P^{-d}(m)$ represents the power about the subframe contained in the auxiliary information code transmitted in the d-th packet before the pertinent packet (packet of a first concealment signal generation target).

$$\hat{P}(m) = P^{-d}(m)$$

$$z'(k,l) = \frac{z(k,l)}{\sqrt{\frac{1}{K}\sum_{k=0}^{K-1} z^2(k,l)}}$$

$$Y(k,l) = 10^{\hat{P}(m)/20} \cdot z'(k,l)$$

[0082] The inverse transform unit 46 transforms the concealment signal or the decoded signal in the time-frequency domain into a signal in the time domain. For example, the transform is performed by the following formula indicating a synthesis QMF.

$$y(k,l) = \frac{1}{K}\sum_{k=0}^{K-1} p_S(n) \cdot Y(k,l) \cos\left[\frac{\pi}{K}(n + \frac{1}{2} - \frac{K}{2})(k + \frac{1}{2})\right]$$

In this formula, the letter l represents an index of a signal in the time domain, provided that $0 \le l \le K(2 + L)$.

[0083] As described above, the fourth embodiment allows the processing procedures as executed in the first and second embodiments to be applied to the signals resulting from the time-frequency transform.

[Fifth Embodiment] - Informative Example

[0084] The fifth embodiment will describe an example in which the technique described in the first embodiment is applied to each of subbands.

[0085] Since, in the encoding unit 1 in the fifth embodiment, the operation of the auxiliary information encoding unit 12 is different from that in the first embodiment, the operation of the auxiliary information encoding unit 12 will be described below. The auxiliary information encoding unit 12, as shown in Fig. 4, is provided with the subframe power calculation unit 121, attenuation coefficient estimation unit 122, and attenuation coefficient quantization unit 123.

[0086] The subframe power calculation unit 121 saves the input audio for the predetermined period of time, and calculates the subframe power sequence for the audio signal v(k, l+d) that is later by the predetermined number of frames (d frames in the present embodiment) than the encoding of the target signal v(k, l) out of the saved input audio. It is assumed herein that the number of samples contained in one frame is T. Supposing a prediction target signal is defined as v(k, l+d) = s(k, l+d), the power $P^i(l)$ of the ith subband in the subframe l ($0 \le l \le L-1$) is obtained by the following formula. The letter k represents an index of a sample in a subframe (provided that $0 \le k \le K-1$).

$$P^i(l+d) = 10\log_{10}\left(\frac{1}{K_{\max}^i - K_{\min}^i}\sum_{k=K_{\min}^i}^{K_{\max}^i} v^2(k,l+d)\right)$$

The subbands may be determined so that the widths of the subbands are unequal intervals, or they may be set to the width of the critical band, or the subband widths may be set to 1.

[0087] The attenuation coefficient estimation unit 122 obtains a slope γ^i_{opt} of a straight line indicative of a temporal change of power for each subframe from the subframe power sequence, for example, by the least square method or the like. More simply, the slope may be determined from $P^i(0)$ and $P^i(L-1)$. In addition to the slope γ^i_{opt} of the straight line, an intercept P^i_{opt} obtained by a straight-line approximation of the subframe power sequence $P^i(l)$ may be obtained. The power of subframe m is represented herein by the following formula. $\hat{P}^i(m) = \gamma^i_{opt} \cdot m + P^i_{opt}$ In this case, a slope γ_{opt} and an intercept P_J of a straight line are determined according to the following formulas (the least square method).

$$\gamma_{opt} = \frac{L\sum_{m=0}^{L-1} m \cdot P(m) - \sum_{m=0}^{L-1} m \sum_{m=0}^{L-1} P(m)}{L\sum_{m=0}^{L-1} m^2 - \left(\sum_{m=0}^{L-1} m\right)^2}$$

$$P_{opt} = \frac{\sum_{m=0}^{L-1} m^2 \sum_{m=0}^{L-1} P(m) - \sum_{m=0}^{L-1} m \cdot P(m) \sum_{m=0}^{L-1} P(m)}{L\sum_{m=0}^{L-1} m^2 - \left(\sum_{m=0}^{L-1} m\right)^2}$$

[0088] The attenuation coefficient quantization unit 123 performs scalar quantization of slopes γ^i_{opt} of straight lines, encodes the result, and outputs the auxiliary information code. The scalar quantization may be performed using a scalar quantization codebook prepared in advance. In the case of the straight-line approximation of the subframe powers $P^i(l)$, the intercept P^i_{opt} may be encoded in addition to the slope γ^i_{opt} of the straight line. The vector quantization and subsequent encoding may be applied to a vector obtained by arranging γ^i_{opt} of all the subbands, or the vector quantization and subsequent encoding may be applied to a vector obtained by arranging γ^i_{opt} and P^i_{opt}.

[0089] Since in the decoding unit 4 in the fifth embodiment the operations of the stored decoding coefficient repetition unit 432, auxiliary information decoding unit 45, and subframe power correction unit 442 are different from those in the first embodiment, the operations of these elements will be described below.

[0090] When the error flag is on (to indicate a packet abnormality), the stored decoding coefficient repetition unit 432 obtains the first concealment signal Z(k, l), using the stored decoded signal stored in the decoding coefficient storage unit 431. The stored decoded signal stored in the decoding coefficient storage unit 431 is denoted by B(k, l). The letter k herein represents an index of a sample in a subframe ($0 \le k \le K-1$) and the letter l represents an index of a subframe stored in the decoding coefficient storage unit 431 ($0 \le l \le L-1$).

[0091] Specifically, the stored decoding coefficient repetition unit 432 calculates the first concealment signal by repetition of the last subframe, as represented by the following formula.

$$Z(k,l) = B(k, dL-1)$$

(provided that $0 \le l \le L-1$ and $0 \le k \le K-1$)

The unit of repetition does not have to be limited to the last subframe, and any part of B(k, l) may be extracted and repeated. Without being limited to the generation of the first concealment signal by the repetition as described above, the first concealment signal may be generated, for example, by a prediction using the linear prediction. Alternatively, the first concealment signal may be generated, for example, in accordance with a model determined in advance as described below.

$$[Z(0,0),\ldots,Z(K-1,L-1)] = f(b(0,0),b(1,0)\ldots,b(K-1,dL-1))$$

[0092] The auxiliary information decoding unit 45 decodes the auxiliary information code output from the code separation unit 40, to generate indexes, and obtains a slope γ^i_J of a straight line corresponding to each of the indexes from the codebook. Here, $P^i(-1)$ represents the power of the last subframe in the signal received normally immediately before the packet loss.

$$\hat{P}^i(m) = \gamma^i_J \cdot m + P^i(-1)$$

[0093] In the case where the intercepts of the straight lines are simultaneously encoded based on the straight-line approximation of subframe powers, the subframe powers are obtained by the following formula using the intercepts P^i_J.

$$\hat{P}^i(m) = \gamma^i_J \cdot m + P^i_J$$

[0094] The auxiliary information storage unit 441 included in the concealment signal correction unit 44 stores the auxiliary information fed from the auxiliary information decoding unit 45 when the error flag indicates the value indicative of the normal packet. The auxiliary information to be stored is preferably that of several past frames (at least d frames or more).

[0095] In the concealment signal correction unit 44 as described above, the subframe power correction unit 442 corrects the first concealment signal for a value of power of the first concealment signal in each subframe in accordance with the formula below to obtain the concealment signal Y(k, l). Specifically, it performs the correction according to the below formula (provided that $0 \le l \le L-1$ and $0 \le k \le K-1$). $P^i_{-d}(m)$ represents the power of the ith subband about the subframe contained in the auxiliary information code transmitted in the d-th packet before the pertinent packet (packet of a first concealment signal generation target).

$$\hat{P}^i(m) = P^i_{-d}(m)$$

$$Z'(k,l) = \frac{Z(k,l)}{\sqrt{\frac{1}{K^i_{\max} - K^i_{\min}} \sum_{k=K^i_{\min}}^{K^i_{\max}} Z^2(k,l)}} \text{、} (K^i_{\min} \le k \le K^i_{\max}, \quad 0 \le i \le I-1)$$

$$Y(k,l) = 10^{\hat{P}^i(m)/20} \cdot Z'(k,l) \text{、} (K^i_{\min} \le k \le K^i_{\max}, \quad 0 \le i \le I-1)$$

The above fifth embodiment showed the example in which the auxiliary information was calculated and encoded for the frame "later by d frames" than the encoding of the target signal, but the auxiliary information may be calculated and encoded for the frame "earlier by d frames" than the encoding of the target signal, as in the third embodiment.

[0096] As described above, the fifth embodiment allows the technique described in the first embodiment to be applied to each of a plurality of subbands.

[Sixth Embodiment] - Informative Example

[0097] The sixth embodiment will describe an example in which the auxiliary information encoding unit obtains two or more pieces of auxiliary information, encodes them separately, and puts the encoded data into a bitstream. The differences from the first embodiment will be mainly described below.

[0098] The encoding unit 1 in the sixth embodiment, as shown in Fig. 2, is provided with the audio encoding unit 11, auxiliary information encoding unit 12, and code multiplexing unit 13. The audio encoding unit 11 is the same as in the first embodiment. The auxiliary information encoding unit 12, as shown in Fig. 4, is provided with the subframe power calculation unit 121, attenuation coefficient estimation unit 122, and attenuation coefficient quantization unit 123.

[0099] The subframe power calculation unit 121 saves the input audio for a predetermined period of time, and calculates a subframe power sequence $P_1(l)$ for audio signals s(dT), s(1+dT), ..., s((d+1)T-1) that are later by a predetermined number of frames (d frames in the present embodiment) than the encoding of the target signals s(0), s(1), ..., s(T-1) out of the saved input audio.

[0100] Furthermore, the subframe power calculation unit 121 calculates a subframe power sequence $P_2(l)$ for audio signals s((d+1)T), s(1+(d+1)T), ..., s((d+2)T-1) later by a predetermined number of frames ((d+1) frames in the present embodiment).

[0101] It is assumed herein that the number of samples contained in one frame is T. When a prediction target signal is expressed by the following formula:

$$v(K \cdot l + k) = s(K \cdot l + k + dT),$$

the powers $P_1(l)$, $P_2(l)$ of subframe l ($0 \le l \le L-1$) are obtained by the following formulas. The letter k represents an index of a sample in each subframe ($0 \le k \le K-1$).

$$P_1(l) = 10\log_{10}\left(\frac{1}{K}\sum_{k=0}^{K-1} v^2(K \cdot l + k)\right)$$

$$P_2(l) = 10\log_{10}\left(\frac{1}{K}\sum_{k=0}^{K-1} v^2(K \cdot l + k + T)\right)$$

[0102] The present embodiment defines K as the length of each subframe, but different lengths may be used for the respective subframes, which are determined in advance for the respective subframes. The subframe power sequence may also be calculated in accordance with the following formula where k^l_{start} represents an index of a start of the lth subframe and k^l_{end} represents an index of an end thereof.

$$P(l)=10\log_{10}\left(\frac{1}{k_{end}^{l}-k_{start}^{l}}\sum_{k=k_{start}^{l}}^{k_{end}^{l}} v^{2}(k_{end}^{l-1}+k)\right)$$

[0103] The attenuation coefficient estimation unit 122 calculates slopes γ^1_{opt}, γ^2_{opt} of straight lines indicative of respective temporal changes of power from the subframe power sequences $P_1(l)$, $P_2(l)$, for example, by the least square method or the like. The calculation method is the same as that performed by the attenuation coefficient estimation unit 122 in the first embodiment.

[0104] The attenuation coefficient quantization unit 123 performs the scalar quantization of each of the slopes γ^1_{opt}, γ^2_{opt} of the straight lines, encodes the results of the scalar quantization, and outputs auxiliary information codes C^1, C^2. It may use the scalar quantization codebook prepared in advance. In the case of the straight-line approximation of subframe power P(l), intercepts P^1_{opt}, P^2_{opt} may also be encoded in addition to the slopes γ^1_{opt}, γ^2_{opt} of the straight lines.

[0105] The code multiplexing unit 13 writes the audio code and the auxiliary information codes C^1, C^2 in a predetermined order and outputs a bitstream. Fig. 14 shows an example of temporal relationship between signals as audio encoding targets and signals as auxiliary information encoding targets, and a configuration of bitstreams. As shown in Fig. 14, for example, the auxiliary information code of frame (N+1) and the auxiliary information code of frame (N+2) are added to the audio code of frame N to obtain a bitstream, which is output from the code multiplexing unit 13. Furthermore, the packet configuration unit 2 in Fig. 1 adds the packet header information to the bitstream to obtain an audio packet to be transmitted as the N-th packet. Although the present embodiment shows the generation of the two pieces of auxiliary information, the auxiliary information to be generated may be three or more pieces of auxiliary information. The auxiliary information may be calculated for a target of an audio signal that is earlier by one or more frames than the audio signal encoded by the audio encoding unit.

[0106] The decoding unit 4 in the sixth embodiment, as shown in Fig. 6, is provided with the error/loss detection unit 41, code separation unit 40, audio decoding unit 42, auxiliary information decoding unit 45, first concealment signal generation unit 43, and concealment signal correction unit 44. Since the operations of the error/loss detection unit 41, audio decoding unit 42, and first concealment signal generation unit 43 are the same as those in the first embodiment, redundant description is omitted herein.

[0107] The code separation unit 40 reads the audio code and auxiliary information codes C^1, C^2 from the bitstream, and sends the audio code to the audio decoding unit 42 and the auxiliary information codes C^1, C^2 to the auxiliary information decoding unit 45.

[0108] The auxiliary information decoding unit 45 decodes the auxiliary information codes C^1, C^2, calculates the auxiliary information, and sends the result to the concealment signal correction unit 44. For example, the auxiliary information decoding unit 45 decodes the auxiliary information codes C^1, C^2 output from the code separation unit 40, to generate indexes, and obtains slopes γ_J of straight lines corresponding to the respective indexes from the codebook. Here, P(-1) represents the power of the last subframe in the signal received normally immediately before the frame loss.

$$\hat{P}(m)=\gamma_J\cdot m+P(-1)$$

When the intercepts of the straight lines are simultaneously encoded based on the straight-line approximation of subframe powers, the subframe powers are obtained according to the following formula using the intercepts P_J.

$$\hat{P}(m)=\gamma_J\cdot m+P_J$$

[0109] The concealment signal correction unit 44, as shown in Fig. 12, is provided with the auxiliary information storage unit 441 and the subframe power correction unit 442.

[0110] The auxiliary information storage unit 441 stores the auxiliary information fed from the auxiliary information decoding unit 45 when the error flag indicates the value indicative of the normal packet. The auxiliary information to be stored is preferably that of several past frames (at least d frames or more). In the present embodiment, the auxiliary information of two frames is acquired per packet.

[0111] The subframe power correction unit 442 corrects the first concealment signal for a value of power of the first concealment signal in each subframe in accordance with the formula below to obtain the concealment signal Y(K·l+k). Specifically, it performs the correction according to the below formula (provided that $0 \le l \le L-1$ and $0 \le k \le K-1$). $P^{-d}(m)$ represents the power about the subframe contained in the auxiliary information code C^1 transmitted in the d-th packet before the pertinent packet (packet of a first concealment signal generation target).

$$\hat{P}(m) = P^{-d}(m)$$

$$z'(K \cdot l + k) = \frac{z(K \cdot l + k)}{\sqrt{\frac{1}{K}\sum_{k=0}^{K-1} z^2(K \cdot l + k)}}$$

$$Y(K \cdot l + k) = 10^{\hat{P}(m)/20} \cdot z'(K \cdot l + k)$$

[0112] For example, the subframe power correction unit 442, as shown in Fig. 8, earlier extracts the auxiliary information transmitted in the d-th packet, from the auxiliary information storage unit 441 (step S60 in Fig. 8), calculates the mean square amplitude value for each subframe as to the first concealment signal, and divides the value contained in the subframe, by the mean square amplitude value (step S61). This calculation results in obtaining z'(K·l+k). Then powers of respective subframes are calculated from the auxiliary information and the value of the subframe is multiplied by a mean amplitude value obtained from the powers (step S62). This multiplication results in obtaining the concealment signal Y(K·l+k). The above processing of steps S4101 to S4421 (Fig. 7) is repeated to the end of the input audio (step S4431).

[0113] When a consecutive packet loss further occurs, the packet loss can also be concealed in the case of occurrence of the consecutive packet loss by carrying out the same processing, using the power about the subframe contained in the auxiliary information code C^2 transmitted in the d-th packet before the pertinent packet (packet of a first concealment signal generation target).

[0114] As described above, the sixth embodiment allows the auxiliary information encoding unit to obtain two or more pieces of auxiliary information, encode them separately, and put them into the bitstream.

[0115] Incidentally, Fig. 19 shows a configuration diagram of a modification example of the decoding unit 4. The decoding unit 4 in Fig. 13 in the fourth embodiment described above was configured to feed the error flag to the audio decoding unit 42, the first concealment signal generation unit 43, the concealment signal correction unit 44, and the auxiliary information decoding unit 45, whereas the configuration in Fig. 19 omits these inputs. Even in the configuration with omission of these inputs, there is no input to the audio decoding unit 42 and the auxiliary information decoding unit 45 with the error flag being on and therefore the error flag can be determined to be on by the absence of the input. Namely, the state of the error flag can be determined, depending upon the presence/absence of the input to the audio decoding unit 42 and the auxiliary information decoding unit 45. The first concealment signal generation unit 43 and the concealment signal correction unit 44 can also determine the state of the error flag in the same manner. The decoding unit 4 in Fig. 13 is configured so that an audio parameter storage unit 47 shown in Fig. 19 is included in the first concealment signal generation unit 43, but the audio parameter storage unit 47 may be configured as a constituent element independent of the first concealment signal generation unit 43, as shown in Fig. 19. The function of the decoding unit 4 of the configuration in Fig. 19 is substantially the same as that of the decoding unit 4 in Fig. 13. The decoding unit 4 in the first, second, third, fifth, and sixth embodiments shown in Fig. 6 may also be configured so that the input of the error flag to the audio decoding unit 42, the first concealment signal generation unit 43, the concealment signal correction unit 44, and the auxiliary information decoding unit 45 is omitted and/or so that the audio parameter storage unit is a constituent element independent of the first concealment signal generation unit 43, as described above.

[Seventh Embodiment]

[0116] The seventh embodiment will describe an example in which the auxiliary information about a sudden change of power (which will be referred to hereinafter as "transient") to be used herein is a position of the transient in a frame as an auxiliary information encoding target, and a power of a subframe at the position of the transient.

(Configuration and Operation of Encoding Unit 1)

[0117] In the seventh embodiment the overall configuration of the encoding unit 1 is also as shown in Fig. 2 and the overall configuration of the decoding unit 4 is as shown in Fig. 6. In the seventh embodiment as well, the description about the overall configuration is omitted as in the second to sixth embodiments.

[0118] The auxiliary information encoding unit 12 will be described below in detail as a characteristic portion of the encoding unit 1 in the seventh embodiment. The auxiliary information encoding unit 12, as shown in Fig. 20, is provided with a transient detection unit 124A, a transient position quantization unit 125, a transient power scalar quantization unit 126, and a parameter encoding unit 127.

[0119] The operation of the auxiliary information encoding unit 12 of this configuration will be described based on Fig. 21. The transient detection unit 124A saves the input audio for a predetermined period of time, and detects a transient using audio signals s(dT), s(1+dT), ..., s((d+1)T-1) that is later by a predetermined number of frames (d frames in the present embodiment) than the encoding of the target signals s(0), s(1), ..., s(T-1) out of the saved input audio (step S7401 in Fig. 21). The auxiliary information encoding target frame may be a frame that is later by one or more frames than an audio encoding target frame or may be a frame that is earlier by one or more frames than an audio encoding target frame. The auxiliary information codes may be calculated from two or more frames selected from frames that are earlier or later by one or more frames than the audio encoding target frame.

[0120] A method for detection of the transient can be, for example, the method described in Section 7.1 in "ITU-T Recommendation G.719." The transient may also be detected using one of other standard technologies and non-standard technologies. In the above method described in Section 7.1, the power is calculated in each subframe and then a temporal change of each subframe is compared with a threshold to determine whether or not there is a transient. Calculated as a result of the transient detection are: a transient flag F_{tran} indicative of whether a transient is contained in the auxiliary information encoding target frame, a position l_{tran} of the transient, and a subframe power sequence P(l). When a power of a subframe at the position l_{tran} of the transient is represented by $P(l_{tran})$ as shown in Fig. 41, the transient detection unit 124A outputs the position l_{tran} of the transient through line 1L45, outputs the power $P(l_{tran})$ of the subframe at the position l_{tran} of the transient through line 1L46, and outputs the transient flag F_{tran} through line 1L47. The transient detection unit 124A may be configured to output the position l_{tran} of the transient and the subframe power sequence P(l) through line 1L46.

[0121] For example, when the transient detection is carried out by the method described in Section 7.2 in "ITU-T Recommendation G.719," the transient detection unit 124A is supposed to calculate the same parameter as the subframe power sequence calculated by the subframe power calculation unit 121 in Fig. 4. When the transient detection is carried out by other methods, the transient detection unit 124A also calculates and outputs the same parameter as the subframe power sequence calculated by the subframe power calculation unit 121 in Fig. 4.

[0122] When the transient flag F_{tran} does not indicate a value for inclusion of a transient in a frame, a value indicative of a normal frame is entered in F_{tran}. In this case, the parameter encoding unit 127 encodes only the transient flag and outputs the encoded data as an auxiliary information code (step S7702 in Fig. 21).

[0123] On the other hand, when the transient flag F_{tran} indicates a value for inclusion of a transient in a frame, the transient position quantization unit 125 performs the scalar quantization of the position l_{tran} of the transient by a predetermined bit count and outputs quantized position information (step S7501 in Fig. 21). The scalar quantization may be performed by a method of binary coding with l_{tran} being regarded as a binary number, or by a method of providing predetermined positions with indexes, and performing binary encoding of an index at the closest position to l_{tran}, or by entropy coding such as Huffman coding, or by any other quantization method. Fig. 42(a) shows a schematic diagram of an example of transient position information encoding by the binary coding, and Fig. 42(b) a schematic diagram of an example of transient position information encoding by the scalar quantization. As a modification example, another available method is as follows: two or more subframe indexes are selected as "information indicative of a change of power," in addition to the position of the transient, and the two or more subframe indexes thus selected are encoded and transmitted. There are no particular restrictions on the method of encoding herein.

[0124] When the value for inclusion of a transient in a frame is set in the transient flag F_{tran}, the transient power scalar quantization unit 126 performs the scalar quantization of the power of the subframe corresponding to the position l_{tran} of the transient and outputs the quantized transient power (step S7601 in Fig. 21). For example, in a case where the quantization is performed between 0 dB and 96 dB with use of a 6-bit linear encoder, the quantization is carried out according to the below formula. In this formula, C can be the value of 1.55 and ε can be the value of 0.001 or the like, but these constants may be changed according to the quantization bit count or the like.

$$I_E = \left\lfloor \frac{10\log(P(l_{tran}) + \varepsilon)}{C} \right\rfloor$$

According to the above formula, the power of the transient is quantized into an index ranging from 0 to 63. The quantization may be carried out using a codebook determined in advance by learning or the like, or any other quantization means may be applied. When the transient flag F_{tran} does not indicate the value for inclusion of a transient in a frame, the value indicative of a normal frame is entered in I_E in the above formula.

[0125] The parameter encoding unit 127 combines the transient flag, the quantized position information, and the quantized transient power together and outputs the auxiliary information code (step S7701 in Fig. 21). It is also possible to adopt a method in which the transient flag, the quantized position information, and the quantized transient power are regarded together as a vector and then the vector is encoded by vector quantization or by any other encoding method.

There are no particular restrictions on the method of encoding.

(Configuration and Operation of Decoding Unit 4)

[0126] The overall configuration of the decoding unit 4 is as shown in Fig. 6 described in the first embodiment. The following will describe the configurations and operations of the auxiliary information decoding unit 45 and the concealment signal correction unit 44 which are characteristic configurations in the seventh embodiment. The first concealment signal generation unit 43 may generate the first concealment signal by an existing standard technique, for example, as described in Section 5.2 in TS26.402, in addition to the techniques described in the first to sixth embodiments, or may generate the first concealment signal by another concealment signal generation technique which is not a standard.

[0127] The auxiliary information decoding unit 45, as shown in Fig. 22, is provided with a transient flag decoding unit 129, a transient position decoding unit 1212, and a transient power decoding unit 1213.

[0128] The operation of the auxiliary information decoding unit 45 of this configuration will be described based on Fig. 23. The auxiliary information decoding unit 45 decodes the auxiliary information code and determines whether the obtained transient flag F_{tran} is on (indicative of a frame including a transient) or off (indicative of a frame including no transient) (step S7901 in Fig. 23).

[0129] When the transient flag F_{tran} indicates a frame containing no transient, only the value of the transient flag F_{tran} is output as auxiliary information (step S7142 in Fig. 23).

[0130] On the other hand, when the transient flag F_{tran} indicates a frame including a transient, the auxiliary information decoding unit reads the quantized position information I_{tran} out of the auxiliary information code, decodes it, and outputs the quantized position information (step S7121 in Fig. 23). Furthermore, the unit reads and decodes the quantized transient power I_E from the auxiliary information code and outputs the decoded transient power (step S7131 in Fig. 23). For example, where the linear quantization as described above is used, the decoded transient power is obtained from the quantized transient power in accordance with the following formula.

$$\hat{P}_{tran} = 10^{C \cdot I_E / 20}$$

[0131] Then the auxiliary information decoding unit 45 outputs the calculated transient flag F_{tran}, quantized position information, and decoded transient power as auxiliary information (step S7141 in Fig. 23).

[0132] Next, the concealment signal correction unit 44 will be described. As shown in Fig. 24, the concealment signal correction unit 44 is provided with the auxiliary information storage unit 441 and the subframe power correction unit 442. The first to sixth embodiments showed the configuration in which the error flag was fed to the subframe power correction unit 442, whereas the concealment signal correction unit 44 in Fig. 24 is configured not to feed the error flag to the subframe power correction unit 442 and is further configured to determine the state of the error flag by the presence/absence of input of the first concealment signal from the first concealment signal generation unit 43. Namely, the error flag is determined to be off, with input of the first concealment signal from the first concealment signal generation unit 43; the error flag is determined to be on, without input of the first concealment signal from the first concealment signal generation unit 43. It is a matter of course that the concealment signal correction unit may be configured to perform the determination on the error flag by supplying the error flag to the auxiliary information storage unit 441 and the subframe power correction unit 442.

[0133] The operation of the concealment signal correction unit 44 is as shown in the flowchart of Fig. 25. First, the state of the error flag is determined by the presence/absence of input of the first concealment signal from the first concealment signal generation unit 43 as described above (step S7800 in Fig. 25). When the error flag is off herein (to indicate no packet loss), the auxiliary information decoding unit 45 decodes the auxiliary information code and outputs the transient flag, the transient position information, and the decoded transient power through line 6L001 in Fig. 24 (step S7101 in Fig. 25). Then the auxiliary information storage unit 441 stores the transient flag, the transient position information, and the decoded transient power (step S7111 in Fig. 25).

[0134] On the other hand, when the error flag is on (to indicate a packet loss), the subframe power correction unit 442 reads the transient flag, quantized position information, and decoded transient power from the auxiliary information storage unit 441, and corrects the first concealment signal for a value of power of the first concealment signal $z(K \cdot l+k)$ in each subframe to obtain a concealment signal $y(K \cdot l+k)$ (provided that $0 \leq 1 \leq L-1$ and $0 \leq k \leq K-1$) (step S7901 in Fig. 25). Specifically, the subframe power correction unit 442 corrects the value of the power of the first concealment signal $z(K \cdot l+k)$ in accordance with the following procedure. First, the first concealment signal output from the first concealment signal generation unit 43 is fed through line 6L002 in Fig. 24 to the subframe power correction unit 442. Next, the subframe power correction unit 442 reads the transient flag F_{tran}, the transient position information I_{tran}, and the decoded transient power represented by

$$\hat{P}_{tran},$$

from the auxiliary information storage unit 441.

[0135] Next, the subframe power correction unit 442 calculates a corrected power of each subframe from the transient position information l_{tran} and the decoded transient power represented by

$$\hat{P}_{tran},$$

which are read from the auxiliary information storage unit 441 (step S7121 in Fig. 25). Specifically, the calculation is carried out according to the following procedure. First, the power of each subframe is calculated according to the following formula.

$$P(m) = \left\{ 10\log 10\left(\frac{1}{K}\sum_{k=0}^{K-1} z^2(K \cdot m + k)\right) \right.$$

[0136] Next, the subframe power correction unit calculates a difference between the power of the first concealment signal at the position of the transient and the decoded transient power (differential transient power).

$$\dot{P}_{tran} = P(l_{tran}) - \hat{P}_{tran}$$

[0137] Then the subframe power correction unit corrects the power of the first concealment signal corresponding to each subframe after the position of the transient, using the foregoing differential transient power, to obtain a corrected concealment signal subframe power.

$$\hat{P}(m) = \begin{cases} P(m)(0 \le m < l_{tran}) \\ P(m) + \dot{P}_{tran}(l_{tran} \le m < L-1) \end{cases}$$

[0138] Next, after calculating the power of each subframe for the first concealment signal, the subframe power correction unit 442 normalizes each of the resulting powers (step S7801 in Fig. 25). The lengths of the respective subframes may be set to be unequal as in the second to sixth embodiments. The present embodiment will detail the case where the lengths of the respective subframes are equal.

$$z'(K \cdot l + k) = \frac{z(K \cdot l + k)}{\sqrt{\frac{1}{K}\sum_{k=0}^{K-1} z^2(K \cdot l + k)}}$$

[0139] Finally, the subframe power correction unit multiplies the normalized first concealment signal by the corrected concealment signal subframe power to calculate a concealment signal (step S7131 in Fig. 25).

$$y(K \cdot l + k) = 10^{\hat{P}(m)/20} \cdot z'(K \cdot l + k)$$

[0140] As a modification example of step S7121 in Fig. 25, the method of calculating from the subframe power P(m) and the decoded transient power:

$$\hat{P}_{tran}\ ,$$

the corrected concealment signal subframe power:

$$\hat{P}(m)$$

may be a method as represented by the following formula.

$$\dot{P}_{tran} = P(l_{tran}) - \hat{P}_{tran}$$

$$P'(m) = \begin{cases} P(m)(0 \le m < l_{tran}) \\ P(m) + \dot{P}_{tran}(l_{tran} \le m < L-1) \end{cases}$$

[0141] Finally, a corrected concealment signal power is calculated using a predetermined prediction coefficient a_p. The prediction coefficient may be switched to another, depending upon properties of subframe power sequences.

$$\hat{P}(m) = \sum_{p=0}^{P} a_p \cdot P'(m-p)$$

[0142] Alternatively, smoothing may be carried out using a model determined in advance.

$$\hat{P}(m) = f(P'(0), \cdots, P'(L-1))$$

[0143] The function f to be used herein may be, for example, a sigmoid function, a spline function, or the like and there are no particular restrictions thereon as long as smoothing can be implemented.

[0144] The seventh embodiment as described above can realize the high-accuracy packet loss concealment for the transient signal, using the indication information indicative of the presence/absence of a sudden change of power, the position of the transient in the frame as an auxiliary information encoding target, and the power of the subframe at the position of the transient, as the auxiliary information about the sudden change of power (transient).

[Eighth Embodiment] - Informative Example

(Configuration and Operation of Encoding Unit 1)

[0145] The auxiliary information encoding unit 12 in the eighth embodiment, as shown in Fig. 26, is provided with the transient detection unit 124A, the transient position quantization unit 125, the transient power scalar quantization unit 126, a transient power vector quantization unit 128, and the parameter encoding unit 127. The eighth embodiment is different in the provision of the transient power vector quantization unit 128, in addition to the transient power scalar quantization unit 126 in the seventh embodiment, and in the configuration and operation of the auxiliary information decoding unit 45, from the seventh embodiment.

[0146] The operation of the auxiliary information encoding unit 12 in the eighth embodiment is shown in Fig. 27. First, the transient detection unit 124A detects a transient in an auxiliary information encoding target frame (step S7401 in Fig. 27). A detection method of the transient is the same as in step S7401 in Fig. 21 in the seventh embodiment. The auxiliary information encoding target frame may be a frame later by one or more frames than the audio encoding target frame or a frame earlier by one or more frames than it. Furthermore, two or more frames may be selected from frames earlier or later by one or more frames than the audio encoding target frame, and the auxiliary information codes are calculated therefrom and used herein.

[0147] When a transient is detected, the following procedure is carried out. First, the transient position quantization unit

125 quantizes the transient position information (step S7501 in Fig. 27). A method of the quantization is the same as in step S7501 in Fig. 21 in the seventh embodiment.

[0148] Next, the transient power scalar quantization unit 126 performs the scalar quantization of the power of the subframe corresponding to the transient position and outputs the quantized transient power. The operation of the transient power scalar quantization unit 126 is the same as in the seventh embodiment (step S7601 in Fig. 27).

[0149] Next, the transient power vector quantization unit 128 normalizes the subframe power sequence, using the power of the subframe indicated by the quantized position information, and then performs vector quantization (step S8701 in Fig. 27).

$$\overline{P}(m) = \frac{P(m)}{P(l_{tran})}$$

[0150] The vector quantization is carried out according to the following formula.

$$J = \underset{i=0,..,I-1}{\arg\min} \sum_{l=0}^{L-1} (c_i(l) - \overline{P}(l + l_{tran} - L + 1))^2$$

[0151] The letter I represents the number of entries of straight lines or vectors in a codebook and the letter J represents an index of a selected straight line or vector (which will be referred to hereinafter as "code vector index"). $c_i(l)$ indicates the lth element of the ith code vector in the codebook.

[0152] The present embodiment showed the example of the vector quantization after the normalization of the subframe power sequence, whereas a modification example may adopt a configuration to perform the vector quantization without execution of the normalization as shown in Fig. 28. The operation of the auxiliary information encoding unit 12 in Fig. 28 is as shown in Fig. 29, and the vector quantization is carried out according to the following formula (step S8901 in Fig. 29), instead of S8701 in Fig. 27. The other is the same as in Fig. 27.

$$J = \underset{i=0,..,I-1}{\arg\min} \sum_{l=0}^{L-1} (c_i(l) - P(l + l_{tran} - L + 1))^2$$

[0153] Returning to Fig. 27, the parameter encoding unit 127 then outputs the transient flag, the quantized position information, the quantized transient power, and the code vector index as auxiliary information code (step S8801 in Fig. 27). The transient flag, the quantized position information, and the quantized transient power may be encoded by vector quantization or by another encoding method. There are no particular restrictions on the method of encoding. The auxiliary information may be encoded by variable length coding to encode the auxiliary information by a value of 2 or more bits only if the value of the transient flag indicates the existence of the transient, and to use only one bit indicative of the transient flag as auxiliary information if the value of the transient flag indicates the absence of the transient.

(Configuration and Operation of Decoding Unit 4)

[0154] The eighth embodiment is different from the seventh embodiment, in the configuration and operation of the auxiliary information decoding unit 45 in Fig. 30 and in the operations of the auxiliary information storage unit 441 and the subframe power correction unit 442 in the concealment signal correction unit 44. As shown in Fig. 30, the auxiliary information decoding unit 45 is provided with the transient flag decoding unit 129, the transient position decoding unit 1212, the transient power decoding unit 1213, and a transient power vector decoding unit 1214.

[0155] The operation of the auxiliary information decoding unit 45 is shown in Fig. 31. The auxiliary information decoding unit 45 reads the transient flag F_{tran}, the quantized position information I_{tran}, the quantized transient power I_E, and the code vector index J from the auxiliary information code and determines the state of the transient flag F_{tran} (step S901 in Fig. 31). When the value of the transient flag F_{tran} indicates no transient, only the value of the transient flag F_{tran} is output as auxiliary information (step S906 in Fig. 31), as in the seventh embodiment.

[0156] On the other hand, when the value of the transient flag F_{tran} indicates a transient, the quantized position information I_{tran} is decoded by the same method as in step S7121 in Fig. 23 in the seventh embodiment and the decoded position information is output (step S902 in Fig. 31).

[0157] Next, the decoded transient power is calculated from the quantized transient power by the same method as in step S7131 in Fig. 23 in the seventh embodiment (step S903 in Fig. 31).

[0158] A code vector $c_J(m)$ corresponding to the code vector index J is output (step S904 in Fig. 31).

[0159] Finally, the transient flag, decoded position information, decoded transient power, and code vector are output (step S905 in Fig. 31).

[0160] Next, the operation of the concealment signal correction unit 44 shown in Fig. 32 will be described with reference to the configuration of the concealment signal correction unit 44 shown in Fig. 24.

[0161] First, the state of the error flag is determined (step S1500 in Fig. 32). For the determination on the state of the error flag, the value of the error flag entered from the outside may be read or it may be determined whether the first concealment signal from the first concealment signal generation unit 43 is fed to the subframe power correction unit 442. Specifically, the value of the error flag may be determined to indicate no packet loss (which is off), with input of the first concealment signal to the subframe power correction unit 442; the value of the error flag may be determined to indicate a packet loss (which is on), without input of the first concealment signal to the subframe power correction unit 442.

[0162] When the value of the error flag indicates no packet loss (off), the auxiliary information storage unit 441 stores the transient flag, decoded position information, decoded transient power, and code vector (step S1501 in Fig. 32).

[0163] On the other hand, when the value of the error flag indicates a packet loss (on), the subframe power correction unit 442 corrects the first concealment signal $z(K \cdot l+k)$ for a value of power of the first concealment signal in each subframe in accordance with the below-described formula to obtain the concealment signal $y(K \cdot l+k)$ (provided that $0 \leq 1 \leq L-1$ and $0 \leq k \leq K-1$). Specifically, the value of power of the first concealment signal is corrected in each subframe in accordance with the following procedure.

[0164] First, the correction unit reads the transient flag, decoded position information, decoded transient power, and code vector from the auxiliary information storage unit (step S1502 in Fig. 32).

[0165] Next, the power of each subframe is calculated using the auxiliary information (step S1503 in Fig. 32). In this step, first, the subframe power is calculated.

$$P(m) = \left\{ 10 \log 10 \left(\frac{1}{K} \sum_{k=0}^{K-1} z^2 (K \cdot m + k) \right) \right.$$

[0166] Next, the correction unit calculates the differential transient power which is the difference between the subframe power corresponding to the transient position and the decoded transient power.

$$\dot{P}_{tran} = P(l_{tran}) - P_{tran}$$

[0167] Next, the corrected concealment signal subframe power is calculated using the differential transient power and the code vector.

$$\hat{P}(m) = \begin{cases} P_{tran} \cdot c_J (L - l_{tran} - 1 + m)(0 \leq m < l_{tran}) \\ P(m) + \dot{P}_{tran}(l_{tran} \leq m < L-1) \end{cases}$$

[0168] The present embodiment shows the example of the vector quantization after the normalization of the values of the subframe power sequence on the encoder side, but it is also possible to adopt a method in which the vector quantization of the subframe power sequence is carried out without execution of the normalization. In the case without execution of the normalization, the corrected concealment signal subframe power is calculated as follows.

$$\hat{P}(m) = \begin{cases} c_J (L - l_{tran} - 1 + m)(0 \leq m < l_{tran}) \\ P(m) + \dot{P}_{tran}(l_{tran} \leq m < L-1) \end{cases}$$

[0169] Next, the first concealment signal is normalized in each subframe (step S1504 in Fig. 32).

$$z'(K \cdot l+k)=\frac{z(K \cdot l+k)}{\sqrt{\frac{1}{K}\sum_{k=0}^{K-1} z^2(K \cdot l+k)}}$$

[0170] Finally, the normalized first concealment signal is multiplied by the corrected subframe power and the concealment signal is output (step S1505 in Fig. 32).

$$y(K \cdot l+k)=10^{\hat{P}(m)/20} \cdot z'(K \cdot l+k)$$

[0171] The eighth embodiment as described above can realize the high-accuracy packet loss concealment for the transient signal, further using the information obtained by the vector quantization of the transient power change, as the auxiliary information about the sudden change of power (transient).

[Ninth Embodiment] - Informative Example

[0172] The ninth embodiment will describe an example in which the processing as executed in the seventh and eighth embodiments is applied to signals resulting from a time-frequency transform. The auxiliary information encoding target frame may be a frame later by one or more frames than the audio encoding target frame or a frame earlier by one or more frames than it. The auxiliary information codes may be calculated from two or more frames selected from frames that are earlier or later by one or more frames than the audio encoding target frame, and used herein.

(Configuration and Operation of Encoding Unit 1)

[0173] The encoding unit 1 in the ninth embodiment has the same configuration as in Fig. 2 described in the first embodiment, and thus the detailed description of the entire unit will be omitted herein. The time-frequency transform is as described in the fourth embodiment and the signals after the transform into the frequency domain are denoted by V(k, l). The letter k herein is an index of a frequency bin (provided that $0 \leq k \leq K-1$) and l an index of a subframe (provided that $0 \leq 1 \leq L-1$).

[0174] The auxiliary information encoding unit will be described below in detail as a characteristic portion of the ninth embodiment. The auxiliary information encoding unit, as shown in Fig. 20, is provided with the transient detection unit 124A, transient position quantization unit 125, transient power scalar quantization unit 126, and parameter encoding unit 127. The ninth embodiment will describe an example using a position of a transient in a frame as an auxiliary information encoding target, and a power of at least one subband out of subbands resulting from division of the entire band into the subbands, out of powers in a subframe at the position of the transient, as auxiliary information about a sudden change of power (transient). In the encoding of the auxiliary information, the auxiliary information may be encoded by the vector quantization as executed in the eighth embodiment. The number of subbands to be encoded is not limited to one, but the same processing may be carried out for two or more subbands.

[0175] The transient detection unit 124A detects a transient, using the signals obtained by the transform into the frequency domain. The detection of transient may be carried out using the means used in the seventh embodiment, or using TS26.404 or the like which is the standard technology of transient detection for signals in the frequency domain, or using another transient detection technology for frequency-domain signals. The subband power sequence is calculated herein about values in a range ($K_s \leq k < K_e$) in the frequency domain preliminarily determined in the transient detection. The signals in the frequency band to be used in the detection of transient may be signals in the entire band or only at least one specific subband may be used.

$$P(l)=10\log_{10}\left(\frac{1}{K_e-K_s}\sum_{k=K_s}^{K_e-1} V^2(k,l)\right)$$

[0176] Concerning the method of encoding the transient position information, and, the value of the subband power corresponding to the transient position or the quantized value of the subband power corresponding to the transient position, the same method as in the seventh embodiment and the eighth embodiment can be applied to the subband power sequence calculated as described above. The subband power sequence to be encoded as auxiliary information may be calculated using the entire band or using only at least one specific subband. The subband power sequence to be encoded as auxiliary information may be a subband power sequence calculated for subbands used in the transient detection, or a subband power sequence calculated for subbands not used in the transient detection.

(Configuration and Operation of Decoding Unit 4)

[0177] The overall configuration of the decoding unit 4 is the same as in Fig. 6 described in the first embodiment. The below will describe the configurations and operations of the auxiliary information decoding unit 45 and the concealment signal correction unit 44 which are characteristic configurations in the eighth embodiment. The first concealment signal generation unit 43 may generate the first concealment signal, for example, by the existing standard technology as described in Section 5.2 in TS26.402, in addition to the means described in the first to sixth embodiments, or by another concealment signal generation technology which is not a standard.

[0178] When the error flag indicates a normal frame, the auxiliary information decoding unit 45 reads the transient flag F_{tran}, quantized position information I_{tran}, and quantized transient power I_E from the auxiliary information code. In the case of the transient flag, quantized position information, and quantized transient power being encoded, the auxiliary information decoding unit 45 decodes the auxiliary information code by corresponding decoding means to obtain these parameters. For example, in the case using the linear quantization as described above, the decoded transient power is obtained from the quantized transient power in accordance with the following formula.

$$\hat{P}_{tran} = 10^{C \cdot I_E / 20}$$

[0179] Next, the operation of the concealment signal correction unit will be described. When the error flag indicates a packet loss, the subframe power correction unit 442 reads the auxiliary information from the auxiliary information storage unit 441 and corrects the first concealment signal Z(l, k) for a value of power of the first concealment signal in each subframe in accordance with the below formula to obtain the concealment signal Y(l, k). Specifically, it performs the correction in accordance with the below formula (provided that $0 \le 1 \le L- 1$ and $0 \le k \le K - 1$).

[0180] First, it reads the transient flag from the auxiliary information storage unit and determines the state of the transient. With indication of a transient, a power is obtained in each subframe as to the first concealment signal. The lengths of the respective subframes may be set to be unequal as in the second to sixth embodiments. The present embodiment will detail the case where the lengths of the respective subframes are equal.

$$P(m) = \begin{cases} 10\log 10\left(\frac{1}{K_e - K_s} \sum_{k=K_s}^{K_e - 1} Z^2(m,k)\right) \end{cases}$$

[0181] Furthermore, the correction unit calculates the difference between the power of the first concealment signal at the position of the transient and the decoded transient power (differential transient power).

$$\dot{P}_{tran} = P(l_{tran}) - \hat{P}_{tran}$$

[0182] Furthermore, it corrects the power of the first concealment signal corresponding to each subframe after the position of the transient, using the aforementioned differential transient power, to obtain the corrected concealment signal subframe power.

$$\hat{P}(m) = \begin{cases} P(m)(0 \le m < l_{tran}) \\ P(m) + \dot{P}_{tran}(l_{tran} \le m < L-1) \end{cases}$$

[0183] Next, the first concealment signal is normalized in each subframe.

$$Z'(l,k)=\frac{Z(l,k)}{\sqrt{\frac{1}{K_e-K_s}\sum_{k=K_s}^{K_e-1}Z^2(l,k)}},(K_s\le k<K_e)$$

[0184] Finally, the normalized first concealment signal is multiplied by the corrected concealment signal subband power to calculate the concealment signal.

$$Y(l,k)=10^{\hat{P}(l)/20}\cdot Z'(l,k),(K_s\le k<K_e)$$

[0185] The smoothing as described in the seventh embodiment may be applied or the vector quantization as described in the eighth embodiment may be combined.

[0186] The concealment signal obtained finally is transformed into a signal in the time domain by the inverse transform unit 46 and the resulting concealment signal is output.

[0187] The ninth embodiment as described above allows the processing as executed in the seventh and eighth embodiments to be applied to the signals obtained by the time-frequency transform.

[Tenth Embodiment] - Informative Example

[0188] In the tenth embodiment, the encoder side outputs the auxiliary information code by the means in the seventh or eighth embodiment with the input signal being the transient signal, and conceals a packet loss signal with higher quality by the means in the first to third embodiments as to the part other than the transient signal. For the input signal expressed in the frequency domain, the method in the ninth embodiment may be used in the case of the transient and the methods in the fourth to sixth embodiments may be used in the case other than the transient.

(Operation and Configuration of Encoding Unit 1)

[0189] As shown in Fig. 33, the auxiliary information encoding unit 12 is provided with the attenuation coefficient estimation unit 122, attenuation coefficient quantization unit 123, transient detection unit 124A, transient position quantization unit 125, transient power scalar quantization unit 126, and parameter encoding unit 127. The operations of the individual constituent elements are the same as those described in the first, second, seventh, and eighth embodiments. The overall operation of the auxiliary information encoding unit 12 will be described below. The operation of the auxiliary information encoding unit 12 is shown in the flowchart of Fig. 34.

[0190] First, the transient detection unit 124A determines whether there is a transient in the input signal. The operation of the transient detection unit 124A is the same as in the seventh embodiment (step S1701 in Fig. 34). When there is no transient in the signal as an auxiliary information encoding target, the attenuation coefficient estimation unit 122 estimates the attenuation coefficient from the subframe power sequence by the same operation as in the first embodiment (step S1702 in Fig. 34).

[0191] Next, the attenuation coefficient quantization unit 123 quantizes the attenuation coefficient by the same operation as in the first embodiment, and outputs the quantized attenuation coefficient (step S1703 in Fig. 34).

[0192] Next, the parameter encoding unit 127 outputs the quantized attenuation coefficient as an auxiliary information code (step S1704 in Fig. 34).

[0193] The operations of the transient position quantization unit 125 and the transient power scalar quantization unit 126 with the signal as an auxiliary information encoding target containing a transient are the same as in the seventh embodiment (steps S1705-S1706 in Fig. 34).

[0194] Next, when the transient flag indicates the value for inclusion of a transient in the auxiliary information encoding target frame, the parameter encoding unit 127 encodes the transient flag, transient position information, and quantized transient power and outputs the auxiliary information code (step S1707 in Fig. 34).

(Operation and Configuration of Decoding Unit 4)

[0195] The overall configuration of the tenth embodiment is also the same as in the first embodiment to the ninth embodiment and therefore the operations of the auxiliary information decoding unit 45 and the concealment signal correction unit 44 being the major differences will be described below.

[0196] The auxiliary information decoding unit 45, as shown in Fig. 35, is provided with the transient flag decoding unit 129, attenuation coefficient decoding unit 1210, transient position decoding unit 1212, and transient power decoding unit 1213. The operation of the auxiliary information decoding unit 45 will be described below. The flowchart to show the flow of operation is as shown in Fig. 36.

[0197] The transient flag decoding unit 129 reads the transient flag from the auxiliary information code and determines whether the auxiliary information code corresponds to a transient signal (step S1901 in Fig. 36).

[0198] When the transient flag indicates that the auxiliary information code does not correspond to a transient, the attenuation coefficient decoding unit 1210 reads the quantized attenuation coefficient code from the auxiliary information code, decodes the quantized attenuation coefficient code, and outputs the resulting decoded attenuation coefficient and transient flag as auxiliary information (steps S1902-S1903 in Fig. 36). The basic operation of the attenuation coefficient decoding unit 1210 is the same as the calculation of the attenuation coefficient in the auxiliary information decoding unit in the first embodiment.

[0199] On the other hand, when the transient flag indicates that the auxiliary information code corresponds to a transient, the transient position decoding unit 1212 decodes the quantized transient position information and outputs the resulting transient position information (which will be referred to hereinafter as "decoded position information") (step S1904 in Fig. 36), and the transient power decoding unit 1213 decodes the encoded quantized power and outputs the resulting decoded transient power (step S1905 in Fig. 36), thereby outputting the transient flag, the decoded position information, and the decoded transient power as auxiliary information (step S1906 in Fig. 36). The operations of the transient position decoding unit 1212 and the transient power decoding unit 1213 are the same as in the seventh embodiment.

[0200] The flowchart to show the flow of the operation by the concealment signal correction unit 44 in Fig. 24 is as shown in Fig. 37. The operation of the concealment signal correction unit 44 will be described below.

[0201] With reference to the error flag, the unit determines whether the packet contains an error (step S2001 in Fig. 37). When the error flag indicates a normal frame, the auxiliary information storage unit 441 refers to the value of the transient flag (step S2002 in Fig. 37) and, in the case of a transient, it stores the transient flag, decoded position information, and decoded transient power (step S2003 in Fig. 37). On the other hand, when there is no transient, it stores the transient flag and decoded attenuation coefficient (step S2004 in Fig. 37).

[0202] On the other hand, when the error flag indicates a packet loss, the subframe power correction unit 442 normalizes the first concealment signal (step S2005 in Fig. 37). The method of normalization is the same as the normalization of the first concealment signal in the seventh embodiment.

[0203] Next, the subframe power correction unit 442 reads the transient flag from the auxiliary information storage unit 441 and determines the value of the transient flag (step S2006 in Fig. 37). When the transient flag shows the value indicative of a transient, the subframe power correction unit 442 reads the decoded position information and decoded transient power from the auxiliary information storage unit 441, calculates powers of respective subframes from these decoded position information and decoded transient power, and multiplies the value of the subframe obtained in step S2005, by a mean amplitude value calculated from the foregoing powers, to obtain the concealment signal (step S2007 in Fig. 37).

[0204] On the other hand, when the transient flag shows no transient, the subframe power correction unit 442 reads the decoded attenuation coefficient from the auxiliary information storage unit 441 and calculates the subframe power sequence from the decoded attenuation coefficient by the same method as the method described in the first embodiment. Next, the subframe power correction unit 442 calculates a gain from the calculated subframe power sequence and multiplies the normalized first concealment signal by the obtained gain to obtain the concealment signal (step S2008 in Fig. 37).

[0205] The technique of the tenth embodiment described above may be applied to the input signal resulting from the transform into the frequency domain. In applying the technique to the input signal resulting from the transform into the frequency domain, the calculation and encoding of auxiliary information may be carried out for at least one subband.

[0206] In the tenth embodiment as described above, the encoder side can output the auxiliary information code by the means in the seventh or eighth embodiment with the input signal being a transient signal, and conceal a packet loss signal with higher quality with the use of the means in the first to third embodiments for the part other than the transient signal as well.

[Eleventh Embodiment] - Informative Example

[0207] As shown in Fig. 38, a code length selection unit 128A is added to the auxiliary information encoding unit 12, whereby the auxiliary information is encoded by a value of 2 or more bits only if the value of the transient flag is the value indicating the existence of a transient and whereby the auxiliary information is encoded by only one bit indicative of the transient flag if the value of the transient flag is the value indicative of the absence of a transient. The auxiliary information may be encoded by the variable length coding as described above, or may be always encoded by the same bit count so as to fill zeros as many as the same bit count as the transient position information and the quantized transient power in the absence of a transient as well, or any other information may be encoded instead to form the auxiliary information code.

[0208] It is a matter of course that the configuration wherein the auxiliary information encoding unit is provided with the code length selection unit to make the code length of auxiliary information variable as in the present embodiment can be applied to all of the first embodiment to the tenth embodiment.

[0209] The below will describe the configuration and operation in the case where the code length selection unit is added to the configuration of the seventh embodiment to allow the variable code length. The auxiliary information encoding unit 12, as shown in Fig. 38, is provided with the transient detection unit 124A, transient position quantization unit 125, transient power scalar quantization unit 126, parameter encoding unit 127, and code length selection unit 128A.

[0210] The operation of the auxiliary information encoding unit 12 will be described based on Fig. 39. The transient detection unit 124A performs the detection of transient by the same operation as in the seventh embodiment (step S2201 in Fig. 39).

[0211] When the transient flag F_{tran} indicates the value for inclusion of a transient in a frame, the code length selection unit 128A outputs a predetermined bit count larger than one bit (step S2204 in Fig. 39).

[0212] The transient position quantization unit 125 scalar-quantizes the position l_{tran} of the transient by the predetermined bit count and outputs the quantized position information (step S2205 in Fig. 39). The operation of the transient position quantization unit 125 is the same as in the seventh embodiment.

[0213] Next, the transient power scalar quantization unit 126 performs the scalar quantization of the power of the subframe corresponding to the position l_{tran} of the transient and outputs the quantized transient power (step S2206 in Fig. 39). The operation of the transient power scalar quantization unit 126 is the same as in the seventh embodiment.

[0214] The parameter encoding unit 127 outputs the transient flag, quantized position information, and quantized transient power together as an auxiliary information code (step S2207 in Fig. 39). At this time, the total length of the auxiliary information code is the value determined in step S2204 in Fig. 39.

[0215] On the other hand, when it is determined in step S2201 that the transient flag F_{tran} does not show the value for inclusion of a transient in a frame, the code length selection unit 128A determines the code length to be one bit (step S2202 in Fig. 39). Next, the parameter encoding unit 127 encodes only the transient flag by one bit and outputs it (step S2203 in Fig. 39).

(Configuration and Operation of Decoding Unit 4)

[0216] The auxiliary information decoding unit 45, as shown in Fig. 22, is provided with the transient flag decoding unit 129, transient position decoding unit 1212, and transient power decoding unit 1213, as in the seventh embodiment.

[0217] The operation of the auxiliary information decoding unit 45 of this configuration will be described based on Fig. 40. The auxiliary information decoding unit 45 decodes the auxiliary information code and determines whether the resulting transient flag F_{tran} is on (to indicate a frame containing a transient) or off (to indicate a frame containing no transient) (step S2401 in Fig. 40).

[0218] When the transient flag F_{tran} shows a frame containing a transient, the transient flag decoding unit 129 further reads the quantized position information from the auxiliary information code and outputs the information to the transient position decoding unit 1212, and it further reads the quantized transient power I_E from the auxiliary information code and outputs the power to the transient power decoding unit 1213 (step S2402 in Fig. 40).

[0219] Next, the transient position decoding unit 1212 decodes the quantized position information and outputs the resulting decoded position information l_{tran} (step S2403 in Fig. 40). Furthermore, the transient power decoding unit 1213 decodes the quantized transient power I_E and outputs the resulting decoded transient power $P(l_{tran})$ (step S2404 in Fig. 40).

[0220] This operation results in outputting the transient flag F_{tran}, decoded position information l_{tran}, and decoded transient power $P(l_{tran})$ as auxiliary information (step S2405 in Fig. 40). The steps S2403 to S2405 in Fig. 40 are the same as in the seventh embodiment.

[0221] On the other hand, when the transient flag F_{tran} shows a frame containing no transient, only the transient flag F_{tran} is output as auxiliary information (step S2406 in Fig. 40).

[0222] The operation of the concealment signal correction unit 44 (Fig. 24) is the same as in the seventh embodiment.

[0223] The eleventh embodiment as described above allows the code length of the auxiliary information to be made variable.

[Twelfth Embodiment] - Informative Example

[0224] The twelfth embodiment will describe a modification example of the seventh embodiment. The present embodiment will describe an example in which only the quantized transient power is transmitted as auxiliary information.

(Configuration and Operation of Encoding Unit 1)

[0225] The configuration of the encoding unit 1 is the same as in the first embodiment. The below will describe the configuration and operation of the auxiliary information encoding unit 12 which is a characteristic configuration in the present embodiment. The configuration of the auxiliary information encoding unit 12, as shown in Fig. 43, is provided with the transient detection unit 124A, transient power scalar quantization unit 126, and parameter encoding unit 127.

[0226] The transient detection unit 124A outputs the subframe power sequence by the same processing as in the seventh embodiment. The position of the transient may be determined to be a position where the subframe power exceeds a predetermined threshold, or a position where a ratio of subframe power to power of an immediately-preceding subframe becomes maximum. It may also be such a position that a dispersion of subframe powers for a fixed period of time stored in a buffer is calculated and the resulting dispersion becomes maximum at the position.

[0227] Next, the transient power scalar quantization unit 126 quantizes the subframe power at the transient position by the same method as in the seventh embodiment and outputs the quantized transient power to the parameter encoding unit 127.

[0228] Then the parameter encoding unit 127 encodes only the quantized transient power to generate the auxiliary information code.

(Configuration and Operation of Decoding Unit 4)

[0229] The overall configuration of the decoding unit 4 is the same as in the first embodiment (as shown in Fig. 6). The below will describe the configuration and operation of the auxiliary information decoding unit 45 which is a characteristic configuration in the present embodiment. The first concealment signal generation unit 43 generates the first concealment signal by the same method as in the seventh embodiment.

[0230] The configuration of the auxiliary information decoding unit 45 in the present embodiment is as shown in Fig. 44. In the present embodiment, the auxiliary information code transmitted from the encoding unit 1 does not contain the transient flag and the quantized position information. Then, in the present embodiment the transient flag is always set to the value of on and a predetermined value I_{const} is always set as the transient position information. The transient power decoding unit 1213 decodes the auxiliary information code (quantized power code) containing only the quantized transient power by the same processing as in the seventh embodiment and outputs the decoded transient power.

[0231] The concealment signal correction unit 44 in Fig. 6 processes the foregoing transient flag, transient position information, and output decoded transient power as auxiliary information.

[0232] As described above, it is feasible to realize the embodiment to transmit only the quantized transient power as the auxiliary information, while achieving the same effect as the seventh embodiment.

[Thirteenth Embodiment]

[0233] The thirteenth embodiment will describe another modification example of the seventh embodiment. The present embodiment will describe an example in which only the transient flag and the quantized transient power are transmitted as auxiliary information.

(Configuration and Operation of Encoding Unit 1)

[0234] The below will describe the configuration and operation of the auxiliary information encoding unit 12 which is a characteristic configuration in the present embodiment. The configuration of the auxiliary information encoding unit 12, as shown in Fig. 45, is provided with the transient detection unit 124A, transient power scalar quantization unit 126, and parameter encoding unit 127.

[0235] The operations of the transient detection unit 124A and the transient power scalar quantization unit 126 are the same as in the seventh embodiment.

[0236] The parameter encoding unit 127 encodes the transient flag and the quantized transient power together to generate the auxiliary information code. When the value of the transient flag is off, the parameter encoding unit 127 does not enter the quantized transient power in the auxiliary information code, as in the seventh embodiment.

(Configuration and Operation of Decoding Unit 4)

[0237] The overall configuration of the decoding unit 4 is the same as in the first embodiment (as shown in Fig. 6). The below will describe the configuration and operation of the auxiliary information decoding unit 45 which is a characteristic configuration in the present embodiment. The configuration of the auxiliary information decoding unit 45 in the present embodiment is as shown in Fig. 46.

[0238] The operation of the transient flag decoding unit 129 and the operation of the transient power decoding unit 1213 are the same as in the seventh embodiment. In the present embodiment, the predetermined value l_{const} is always set in the transient position information, as in the twelfth embodiment.

[0239] As described above, it is feasible to realize the embodiment to transmit only the transient flag and the quantized transient power as the auxiliary information, while achieving the same effect as the seventh embodiment.

[Fourteenth Embodiment] - Informative Example

[0240] In the fourteenth embodiment, the subframe at the transient position is divided into subbands and a power of at least one subband is quantized as auxiliary information. In the quantization of the power of at least one subband, at least one subband among one or more subbands is defined as "core subband." Next, for a subband except for the core subband, a difference between a power of the subband (the subband except for the core subband) and a power of the core subband is calculated and the power of the core subband and the foregoing difference are quantized as auxiliary information. The power of the core subband may be contained in the auxiliary information or, may not be contained in the auxiliary information while a value contained in the audio code itself may be used instead.

(Configuration and Operation of Encoding Unit 1)

[0241] The encoding unit 1 in the present embodiment has the same configuration as in Fig. 10 described in the first embodiment, and the detailed description of the entire unit is omitted herein. The time-frequency transform is as described in the fourth embodiment. The signal after the transform into the frequency domain is denoted by V(k, l). The letter k herein represents an index of a frequency bin (provided that $0 \le k \le K-1$) and l an index of a subframe (provided that $0 \le l \le L-1$). The time-frequency transform unit 10 supplies both of the signal V(k, l) after the transform into the frequency domain and the audio signal before the time-frequency transform to the auxiliary information encoding unit 12.

[0242] The configuration of the auxiliary information encoding unit 12 in the present embodiment is shown in Fig. 47. The auxiliary information encoding unit 12 is provided with the transient detection unit 124A, a subband power calculation unit 128B, a core subband power quantization unit 129A, a difference quantization unit 1210A, and the parameter encoding unit 127. Furthermore, it may be configured including the transient position quantization unit 125, but the below will describe the configuration without the transient position quantization unit 125.

[0243] The operation of the transient detection unit 124A is the same as in the seventh embodiment.

[0244] The subband power calculation unit 128B calculates subband powers of the subframe corresponding to the transient position, in accordance with the formula below. $P^{(i)}(l_{tran})$ represents the power of the ith subband at the transient position. Furthermore, $K_s^{(i)}$ and $K_e^{(i)}$ represent an index of the first frequency bin of the ith subband and an index of the last frequency bin of the ith subband, respectively.

$$P^{(i)}(l_{tran}) = 10\log_{10}\left(\frac{1}{K_e^{(i)} - K_s^{(i)}} \sum_{k=K_s^{(i)}}^{K_e^{(i)}-1} V^2(k, l_{tran})\right)$$

[0245] The core subband power quantization unit 129A defines a predetermined i_{core}-th subband as a core subband, quantizes the power of the core subband defined as follows:

$$P^{(i_{core})}(l_{tran}) ,$$

and outputs a core subband power code. The quantization may be quantization using a predetermined quantization codebook or quantization by entropy coding using the Huffman coding or the like. In another method, J subbands of not less than one subband preliminarily determined as follows:

$$(i_{core}^{(1)} \cdots i_{core}^{J})$$

are defined as core subbands, and an average of powers of the J subbands is defined as a power of the core subbands. It is also possible to adopt a maximum, a minimum, or the median of the J subbands as a power of the core subbands. Furthermore, the core subband power quantization unit 129A decodes the core subband power code and outputs the decoded core subband power denoted as follows.

$$\hat{P}^{(i_{core})}(l_{tran})$$

[0246] The difference quantization unit 1210A calculates a differential subband power sequence expressed as follows:

$$\dot{P}^{(i)}(l_{tran})$$

in accordance with the formula below, quantizes the sequence, and outputs the differential subband power code. The quantization may be quantization using a predetermined quantization codebook, quantization by entropy coding using the Huffman coding or the like, or quantization by the vector quantization if the differential subband power sequence has two or more subbands.

$$\dot{P}^{(i)}(l_{tran}) = P^{(i)}(l_{tran}) - \hat{P}^{(i_{core})}(l_{tran})$$

[0247] The parameter encoding unit 127 encodes the transient flag, core subband power code, and differential subband power code together and outputs the auxiliary information code. However, if the value of the transient flag is off, the core subband power code and the differential subband power code are not contained in the auxiliary information code.

(Configuration and Operation of Decoding Unit 4)

[0248] The configuration of the auxiliary information decoding unit 45 in the present embodiment is shown in Fig. 48. The auxiliary information decoding unit 45 is provided with the transient flag decoding unit 129, a core subband power decoding unit 1214A, and a difference decoding unit 1215. Furthermore, it may have a configuration including the transient position decoding unit 1212, but the below will describe the configuration without the transient position decoding unit 1212.

[0249] The operation of the transient flag decoding unit 129 is the same as in the seventh embodiment.

[0250] The core subband power decoding unit 1214A decodes the quantized core subband power and outputs the decoded core subband power expressed as follows.

$$\hat{P}^{(i_{core})}(l_{tran})$$

[0251] The difference decoding unit 1215 decodes the differential subband power code and outputs the decoded differential subband power sequence expressed as follows.

$$\tilde{P}^{(i)}(l_{tran})$$

[0252] Furthermore, the difference decoding unit 1215 adds the decoded differential subband power sequence and the decoded core subband power in accordance with the formula

$$\hat{P}^{(i)}(l_{tran}) = \tilde{P}^{(i)}(l_{tran}) + \hat{P}^{(i_{core})}(l_{tran})$$

to calculate a transient power spectrum expressed as follows.

$$\hat{P}^{(i)}(l_{tran})$$

[0253] Next, the operation of the subframe power correction unit 442 (Fig. 24) in the present embodiment will be described. The auxiliary information storage unit 441 stores the transient flag and the transient power spectrum obtained by the forgoing auxiliary information decoding unit 45, as auxiliary information, and the subframe power correction unit 442 reads the transient flag and the transient power spectrum from the auxiliary information storage unit 441, and corrects the first concealment signal $z(K\cdot l+k)$ for a value of power thereof in each subframe to obtain the concealment signal $y(K\cdot l+k)$. Specifically, it performs the correction in accordance with the following procedure (provided that $0 \le 1 \le L-1$ and $0 \le k \le K-1$).

[0254] First, the first concealment signal output from the first concealment signal generation unit 43 is fed to the subframe power correction unit 442. Furthermore, the transient flag and the transient power spectrum stored in the auxiliary information storage unit 441 are fed to the subframe power correction unit 442.

[0255] Next, the subframe power correction unit 442 sets a predetermined value in the transient position information l_{tran}.

[0256] Next, the subframe power correction unit 442 calculates the subband power sequence in accordance with the formula below.

$$\widehat{P}^{(i)}(l_{tran}) = 10\log_{10}\left(\frac{1}{K_e^{(i)} - K_s^{(i)}}\sum_{k=K_s^{(i)}}^{K_e^{(i)}-1} Z^2(k, l_{tran})\right)$$

[0257] Next, the subframe power correction unit 442 calculates a difference between the subband power sequence of the first concealment signal at the position of the transient and the transient power spectrum (differential transient power) in accordance with the formula below.

$$\overline{P}^{(i)}(l) = \hat{P}^{(i)}(l) - \widehat{P}^{(i)}(l_{tran})$$

[0258] Next, the subframe power correction unit 442 corrects the power of the first concealment signal corresponding to each subframe after the position of the transient, using the differential transient power, to obtain a corrected concealment signal subframe power.

[0259] Finally, the subframe power correction unit 442 multiplies the first concealment signal by the corrected concealment signal subframe power in accordance with the formula below for all the subbands i, to calculate the concealment signal. However, $K_s^{(i)} \le k < K_e^{(i)}$ and $l \ge l_{tran}$.

$$y(k,l) = 10^{\overline{P}^{(i)}(l)/20} \cdot z(k,l)$$

[0260] By making use of the difference between the power of the core subband and the power of each subband except for the core subband as auxiliary information, as described above, it is feasible to realize the high-accuracy packet loss concealment for the transient signal.

[0261] The present embodiment described the configurations without the transient position quantization unit 125 in the auxiliary information encoding unit 12 in Fig. 47 and without the transient position decoding unit 1212 in the auxiliary information decoding unit 45 in Fig. 48, but it is also possible to adopt the configurations including them.

[Fifteenth Embodiment] - Informative Example

[0262] The fifteenth embodiment will describe a case without the core subband power quantization unit 129A in Fig. 47 and without the core subband power decoding unit 1214A in Fig. 48 in the fourteenth embodiment.

(Configuration and Operation of Encoding Unit 1)

[0263] The encoding unit 1 in the present embodiment has the same configuration as in Fig. 10 described in the first embodiment and thus the detailed description of the entire unit is omitted herein. The time-frequency transform is the same as in the fourteenth embodiment.

[0264] The audio encoding unit 11 is configured to perform calculation and quantization of power of the audio signal to calculate the core subband power code, and enter it in the audio code. In output of the core subband power code, a power of a frame or at least one subframe obtained in the time domain may be quantized, a power of a frame or at least one subframe obtained in the frequency domain may be quantized, or a power of at least one subsample of a signal resulting from transform into QMF domain may be quantized. In the quantization in the frequency domain and in the QMF domain, a power calculated for at least one subband may be quantized.

[0265] The configuration of the auxiliary information encoding unit 12 in the present embodiment is shown in Fig. 49. The auxiliary information encoding unit 12 is provided with the transient detection unit 124A, subband power calculation unit 128B, difference quantization unit 1210A, and parameter encoding unit 127. Furthermore, it may have a configuration including the transient position quantization unit 125, but the below will describe the configuration without the transient position quantization unit 125.

[0266] The operation of the transient detection unit 124A is the same as in the seventh embodiment and the subband power calculation unit 128B is the same as in the fourteenth embodiment.

[0267] The audio encoding unit 11 feeds the decoded core subband power P_{core} obtained by decoding the code about the power included in the audio code, to the difference quantization unit 1210A.

[0268] The difference quantization unit 1210A calculates the differential subband power sequence expressed as follows:

$$\dot{P}^{(i)}(l_{tran})$$

in accordance with the formula below, quantizes the sequence, and outputs the resulting differential subband power code. The quantization may be quantization using a predetermined quantization codebook, quantization by entropy coding using the Huffman coding or the like, or quantization by vector quantization if the differential subband power sequence has two or more subbands.

$$\dot{P}^{(i)}(l_{tran}) = P^{(i)}(l_{tran}) - P_{core}$$

[0269] The parameter encoding unit 127 is the same as in the fourteenth embodiment.

(Configuration and Operation of Decoding Unit 4)

[0270] The configuration of the auxiliary information decoding unit 45 in the present embodiment is shown in Fig. 50. The auxiliary information decoding unit 45 is provided with the transient flag decoding unit 129 and the difference decoding unit 1215. Furthermore, it may have a configuration including the transient position decoding unit 1212, but the below will describe the configuration without the transient position decoding unit 1212.

[0271] The operation of the transient flag decoding unit 129 is the same as in the seventh embodiment.

[0272] The audio decoding unit 42 decodes the code about the power included in the audio code and feeds the resulting decoded core subband power P_{core} to the difference decoding unit 1215. If P_{core} is a value obtained in a domain different from the signal V(k, l) after the transform into the frequency domain, e.g., a value in the time domain, an offset is added to express P_{core} in the same unit, and then P_{core} is fed to the difference decoding unit 1215.

[0273] The difference decoding unit 1215 decodes the differential subband power code and outputs the decoded differential subband power sequence expressed as follows.

$$\widetilde{P}^{(i)}(l_{tran})$$

[0274] Furthermore, the difference decoding unit 1215 adds the decoded differential subband power sequence and the decoded core subband power to calculate the transient power spectrum expressed as follows:

$$\hat{P}^{(i)}(l_{tran}),$$

in accordance with the formula below.

$$\hat{P}^{(i)}(l_{tran}) = \widetilde{P}^{(i)}(l_{tran}) + P_{core}$$

[0275] The operation of the subframe power correction unit 442 in Fig. 24 is the same as in the fourteenth embodiment.

[0276] As described above, it is feasible to realize the embodiment without the core subband power quantization unit 129A in Fig. 47 and without the core subband power decoding unit 1214A in Fig. 48 in the fourteenth embodiment, while achieving the same effect as the fourteenth embodiment.

[0277] The present embodiment described the configurations without the transient position quantization unit 125 in the auxiliary information encoding unit 12 in Fig. 49 and without the transient position decoding unit 1212 in the auxiliary information decoding unit 45 in Fig. 50, but it is also possible to adopt the configurations including them.

[Audio Encoding Program and Audio Decoding Program]

[0278] First, an audio encoding program for letting a computer operate as the audio encoding device according to the present invention will be described.

[0279] Fig. 17 is a drawing showing a configuration of an audio encoding program according to an embodiment. Fig. 15 is a hardware configuration diagram of a computer according to an embodiment. Fig. 16 is an appearance diagram of the computer according to an embodiment. The audio encoding program P1 shown in Fig. 17 can make the computer C10 shown in Fig. 15 and Fig. 16, operate as the encoding unit 1. It is noted that the program described in the present specification can make any information processing device such as a cell phone, a portable information terminal, or a portable personal computer, without having to be limited to the computer as shown in Figs. 15 and 16, operate in accordance with the program.

[0280] The audio encoding program P1 can be provided as stored in a recording medium M. The recording medium M can be, for example, a recording medium such as a flexible disk, CD-ROM, DVD, or ROM, or a semiconductor memory or the like.

[0281] As shown in Fig. 15, the computer C10 is provided with a reading device C12 such as a flexible disk drive unit, CD-ROM drive unit, or DVD drive unit, a working memory (RAM) C14, a memory C16 to store the program stored in the recording medium M, a display C18, a mouse C20 and a keyboard C22 as input devices, a communication device C24 to perform transmission/reception of data or the like, and a central processing unit (CPU) C26 to control execution of the program.

[0282] When the recording medium M is set in the reading device C12, the computer C10 becomes accessible to the audio encoding program P1 stored in the recording medium M, through the reading device C12 and can operate as the audio encoding device according to the present invention, based on the audio encoding program P1.

[0283] As shown in Fig. 16, the audio encoding program P1 may be a program provided as a computer data signal W superimposed on a carrier wave, through a network. In this case, the computer C10 stores the audio encoding program P1 received by the communication device C24, into the memory C16 and then can execute the audio encoding program P1.

[0284] As shown in Fig. 17, the audio encoding program P1 is provided with an audio encoding module P11 and an auxiliary information encoding module P12. These audio encoding module P11 and auxiliary information encoding module P12 make the computer C10 execute the same functions as the aforementioned audio encoding unit 11 and auxiliary information encoding unit 12. According to this audio encoding program P1, the computer C10 can operate as the audio encoding device according to the present invention.

[0285] Next, an audio decoding program for letting a computer operate as the audio decoding device will be described. Fig. 18 is a drawing showing a configuration of an audio decoding program according to an embodiment.

[0286] The audio decoding program P4 shown in Fig. 18 can be used in the computer shown in Figs. 15 and 16. The audio decoding program P4 can be provided in the same manner as the audio encoding program P1.

[0287] As shown in Fig. 18, the audio decoding program P4 is provided with an error/loss detection module P41, an audio decoding module P42, an auxiliary information decoding module P45, a first concealment signal generation module P43, and a concealment signal correction module P44. These error/loss detection module P41, audio decoding module P42, auxiliary information decoding module P45, first concealment signal generation module P43, and concealment signal correction module P44 make the computer C10 execute the same functions as the aforementioned error/loss detection unit 41, audio decoding unit 42, auxiliary information decoding unit 45, first concealment signal generation unit 43, and concealment signal correction unit 44, respectively. According to this audio decoding program P4, the computer C10 can operate as the audio decoding device.

[0288] The various embodiments described above allow the effective auxiliary information about the part where power changes suddenly, to be sent from the encoder side to the decoder side, and realize the high-accuracy packet loss concealment for the signal with the sudden temporal change of power (transient signal), for which the packet loss concealment was difficult by the conventional technologies, so as to reduce degradation of subjective quality with occurrence of a packet loss.

Reference Signs List

[0289] 1: encoding unit; 2: packet configuration unit; 3: packet separation unit; 4: decoding unit; 10: time-frequency transform unit; 11: audio encoding unit; 12: auxiliary information encoding unit; 13: code multiplexing unit; 40: code separation unit; 41: error/loss detection unit; 42: audio decoding unit; 43: first concealment signal generation unit; 44:

concealment signal correction unit; 45: auxiliary information decoding unit; 46: inverse transform unit; 47: audio parameter storage unit; 121: subframe power calculation unit; 122: attenuation coefficient estimation unit; 123: attenuation coefficient quantization unit; 124: subframe power vector quantization unit; 124A: transient detection unit; 125: transient position quantization unit; 126: transient power scalar quantization unit; 127: parameter encoding unit; 128: transient power vector quantization unit; 128A: code length selection unit; 128B: subband power calculation unit; 129: transient flag decoding unit; 129A: core subband power quantization unit; 1210: attenuation coefficient decoding unit; 1210A: difference quantization unit; 1212: transient position decoding unit; 1213: transient power decoding unit; 1214: transient power vector decoding unit; 1214A: core subband power decoding unit; 1215: difference decoding unit; 431: decoding coefficient storage unit; 432: stored decoding coefficient repetition unit; 441: auxiliary information storage unit; 442: subframe power correction unit; C10: computer; C12: reading device; C14: working memory; C16: memory; C18: display; C20: mouse; C22: keyboard; C24: communication device; C26: CPU; M: recording medium; W: computer data signal; P1: audio encoding program; P11: audio encoding module; P12: auxiliary information encoding module; P4: audio decoding program; P41: error/loss detection module; P42: audio decoding module; P43: first concealment signal generation module; P44: concealment signal correction module; P45: auxiliary information decoding module.

Claims

1. An audio encoding device (1) for encoding an audio signal consisting of a plurality of frames, the audio encoding device (1) comprising:

an audio encoding unit (11) for encoding an audio encoding target frame of the plurality of frames of the audio signal; and
an auxiliary information encoding unit (12) for estimating and encoding auxiliary information about a temporal change of power of the audio signal, the auxiliary information being used in packet loss concealment in decoding of the audio signal,
wherein the auxiliary information encoding unit (12) is configured:

- to estimate and to encode a flag of sudden change of power, as the auxiliary information,
- when the flag indicates a predetermined mode, to estimate and to encode quantized transient power as the auxiliary information, wherein the auxiliary information contains only the flag and the quantized transient power, and
- when the flag does not indicate the predetermined mode, to not include the quantized transient power in the auxiliary information, and the auxiliary information contains only the flag;

wherein the audio signal comprises a frame having a plurality of sub frames, and the quantized transient power is estimated from the sub frame corresponding to a position of a sudden change of power; and
wherein the auxiliary information encoding unit (12) is further configured:

- to use the sudden change of power as a position of a transient in the frame,
- to perform a quantization of the transient power corresponding to the position of the transient in the frame to obtain the quantized transient power, and
- to calculate the auxiliary information from two or more frames selected from frames that are earlier or later by one or more frames than the audio encoding target frame.

2. An audio encoding method executed by an audio encoding device (1) for encoding an audio signal consisting of a plurality of frames, the audio encoding method comprising:

an audio encoding step of encoding an audio encoding target frame of the plurality of frames of the audio signal; and
an auxiliary information encoding step of estimating and encoding auxiliary information about a temporal change of power of the audio signal, the auxiliary information being used in packet loss concealment in decoding of the audio signal,
wherein in the auxiliary information encoding step, the audio encoding device (1):

- estimates and encodes a flag of sudden change of power, as the auxiliary information,
- when the flag indicates a predetermined mode, further estimates and encodes quantized transient power as the auxiliary information, wherein the auxiliary information contains only the flag and the quantized transient

power, and
- when the flag does not indicate the predetermined mode, not includes the quantized transient power in the auxiliary information, and the auxiliary information contains only the flag,

wherein the audio signal comprises a frame having a plurality of sub frames, and the quantized transient power is estimated from the sub frame corresponding to a position of a sudden change of power,
wherein in the auxiliary encoding step, the audio encoding method further comprising:

- using the sudden change of power as a position of a transient in the frame,
- performing a quantization of the transient power corresponding to the position of the transient in the frame to obtain the quantized transient power, frames and
- calculating the auxiliary information from two or more frames selected from frames that are earlier or later by one or more frames than the audio encoding target frame.

Patentansprüche

1. Eine Audiokodierungsvorrichtung (1) zum Kodieren eines Audiosignals, das aus einer Mehrzahl von Datenübertragungsblöcken besteht, wobei die Audiokodierungsvorrichtung (1) folgendes aufweist:

eine Audiokodierungseinheit (11) zum Kodieren eines Audiokodierungszieldatenübertragungsblocks aus der Mehrzahl von Datenübertragungsblöcken des Audiosignals; und
eine Hilfsinformationskodierungseinheit (12) zum Schätzen und Kodieren von Hilfsinformation über eine zeitliche Leistungsänderung des Audiosignals, wobei die Hilfsinformation bei der Paketverlustverbergung bei dem Dekodieren des Audiosignals verwendet wird,
wobei die Hilfsinformationskodierungseinheit (12) eingerichtet ist, um:

- ein Kennzeichen einer plötzlichen Leistungsänderung als die Hilfsinformation zu schätzen und zu kodieren,
- wenn das Kennzeichen einen vorbestimmten Modus angibt, quantisierte Übergangsleistung als die Hilfsinformation zu schätzen und zu kodieren, wobei die Hilfsinformation nur das Kennzeichen und die quantisierte Übergangsleistung enthält, und
- wenn das Kennzeichen nicht den vorbestimmten Modus angibt, die quantisierte Übergangsleistung nicht in die Hilfsinformation aufzunehmen, und wobei die Hilfsinformation nur das Kennzeichen enthält;

wobei das Audiosignal einen Datenübertragungsblock mit einer Mehrzahl von Unterdatenübertragungsblöcken aufweist, und die quantisierte Übergangsleistung aus dem Unterdatenübertragungsblock geschätzt ist, der einer Position einer plötzlichen Leistungsänderung entspricht; und
wobei die Hilfsinformationskodierungseinheit (12) weiter eingerichtet ist, um:

- die plötzliche Leistungsänderung als eine Position eines Übergangs in dem Datenübertragungsblock zu verwenden,
- eine Quantisierung der Übergangsleistung entsprechend der Position des Übergangs in dem Datenübertragungsblock durchzuführen, um die quantisierte Übergangsleistung zu erhalten, und
- die Hilfsinformation aus zwei oder mehr Datenübertragungsblöcken zu berechnen, die aus Datenübertragungsblöcken ausgewählt sind, die um einen oder mehrere Datenübertragungsblöcke vor oder nach dem Audiokodierungsziel-Datenübertragungsblock liegen.

2. Ein Audiokodierungsverfahren, das mittels einer Audiokodierungsvorrichtung (1) zum Kodieren eines aus einer Mehrzahl von Datenübertragungsblöcken bestehenden Audiosignals ausgeführt wird, wobei das Audiokodierungsverfahren aufweist:

einen Audiokodierungsschritt der Kodierung eines Audiokodierungszieldatenübertragungsblocks der Mehrzahl von Datenübertragungsblöcken des Audiosignals; und
einen Hilfsinformationskodierungsschritt zum Schätzen und Kodieren von Hilfsinformation über eine zeitliche Änderung der Leistung des Audiosignals, wobei die Hilfsinformation bei der Paketverlustverbergung bei der Dekodierung des Audiosignals verwendet wird,
wobei die Audiokodierungsvorrichtung (1) in dem Hilfsinformationskodierungsschritt:

- ein Kennzeichen einer plötzlichen Leistungsänderung als Hilfsinformation schätzt und kodiert,
- wenn das Kennzeichen einen vorbestimmten Modus angibt, weiter quantisierte Übergangsleistung als die Hilfsinformation schätzt und kodiert, wobei die Hilfsinformation nur das Kennzeichen und die quantisierte Übergangsleistung enthält, und
- wenn das Kennzeichen nicht den vorbestimmten Modus angibt, die quantisierte Übergangsleistung nicht in die Hilfsinformation aufnimmt, und wobei die Hilfsinformation nur das Kennzeichen enthält;

wobei das Audiosignal einen Datenübertragungsblock mit einer Mehrzahl von Unterdatenübertragungsblöcken aufweist, und die quantisierte Übergangsleistung aus dem Unterdatenübertragungsblock geschätzt ist, der einer Position einer plötzlichen Leistungsänderung entspricht; und
wobei in dem Hilfsinformationskodierungsschritt das Audiokodierungsverfahren weiter aufweist:

- Verwenden der plötzlichen Leistungsänderung als eine Position eines Übergangs in dem Datenübertragungsblock,
- Durchführen einer Quantisierung der Übergangsleistung entsprechend der Position des Übergangs in dem Datenübertragungsblock, um die quantisierte Übergangsleistung zu erhalten, und
- Berechnen der Hilfsinformation aus zwei oder mehr Datenübertragungsblöcken, die aus Datenübertragungsblöcken ausgewählt sind, die um einen oder mehrere Datenübertragungsblöcke vor oder nach dem Audiokodierungsziel-Datenübertragungsblock liegen.

Revendications

1. Dispositif de codage audio (1) pour coder un signal audio constitué d'une pluralité de trames, le dispositif de codage audio (1) comprenant :

une unité de codage audio (11) pour coder une trame cible de codage audio de la pluralité de trames du signal audio ; et
une unité de codage d'informations auxiliaires (12) pour estimer et coder des informations auxiliaires concernant une variation temporelle de puissance du signal audio, les informations auxiliaires étant utilisées dans la dissimulation de perte de paquets lors du décodage du signal audio,
dans lequel l'unité de codage d'informations auxiliaires (12) est configurée :

- pour estimer et coder un indicateur de variation soudaine de puissance, en tant qu'informations auxiliaires,
- lorsque l'indicateur indique un mode prédéterminé, pour estimer et coder une puissance transitoire quantifiée en tant qu'informations auxiliaires, dans lequel les informations auxiliaires ne contiennent que l'indicateur et la puissance transitoire quantifiée, et
- lorsque l'indicateur n'indique pas le mode prédéterminé, pour ne pas inclure la puissance transitoire quantifiée dans les informations auxiliaires, et les informations auxiliaires ne contiennent que l'indicateur ;

dans lequel le signal audio comprend une trame ayant une pluralité de sous-trames, et la puissance transitoire quantifiée est estimée à partir de la sous-trame correspondant à une position d'une variation soudaine de puissance ; et
dans lequel l'unité de codage d'informations auxiliaires (12) est en outre configurée :

- pour utiliser la variation soudaine de puissance comme position d'un transitoire dans la trame,
- pour effectuer une quantification de la puissance transitoire correspondant à la position du transitoire dans la trame afin d'obtenir la puissance transitoire quantifiée, et
- pour calculer les informations auxiliaires à partir de deux ou plusieurs trames sélectionnées parmi des trames qui sont antérieures ou postérieures d'une ou plusieurs trames à la trame cible de codage audio.

2. Procédé de codage audio exécuté par un dispositif de codage audio (1) pour coder un signal audio constitué d'une pluralité de trames, le procédé de codage audio comprenant :

une étape de codage audio consistant à coder une trame cible de codage audio de la pluralité de trames du signal audio ; et
une étape de codage d'informations auxiliaires consistant à estimer et à coder des informations auxiliaires concernant une variation temporelle de puissance du signal audio, les informations auxiliaires étant utilisées

dans la dissimulation de perte de paquets lors du décodage du signal audio,
dans lequel, dans l'étape de codage d'informations auxiliaires, le dispositif de codage audio (1) :

- estime et code un indicateur de variation soudaine de puissance, en tant qu'informations auxiliaires,
- lorsque l'indicateur indique un mode prédéterminé, estime et code en outre une puissance transitoire quantifiée en tant qu'informations auxiliaires, dans lequel les informations auxiliaires ne contiennent que l'indicateur et la puissance transitoire quantifiée, et
- lorsque l'indicateur n'indique pas le mode prédéterminé, n'inclut pas la puissance transitoire quantifiée dans les informations auxiliaires, et les informations auxiliaires ne contiennent que l'indicateur,

dans lequel le signal audio comprend une trame ayant une pluralité de sous-trames, et la puissance transitoire quantifiée est estimée à partir de la sous-trame correspondant à une position d'une variation soudaine de puissance,
dans lequel, dans l'étape de codage auxiliaire, le procédé de codage audio comprend en outre :

- l'utilisation de la variation soudaine de puissance comme position d'un transitoire dans la trame,
- la réalisation d'une quantification de la puissance transitoire correspondant à la position du transitoire dans la trame afin d'obtenir la puissance transitoire quantifiée, et
- le calcul des informations auxiliaires à partir de deux ou plusieurs trames sélectionnées parmi des trames qui sont antérieures ou postérieures d'une ou plusieurs trames à la trame cible de codage audio.

Fig.1

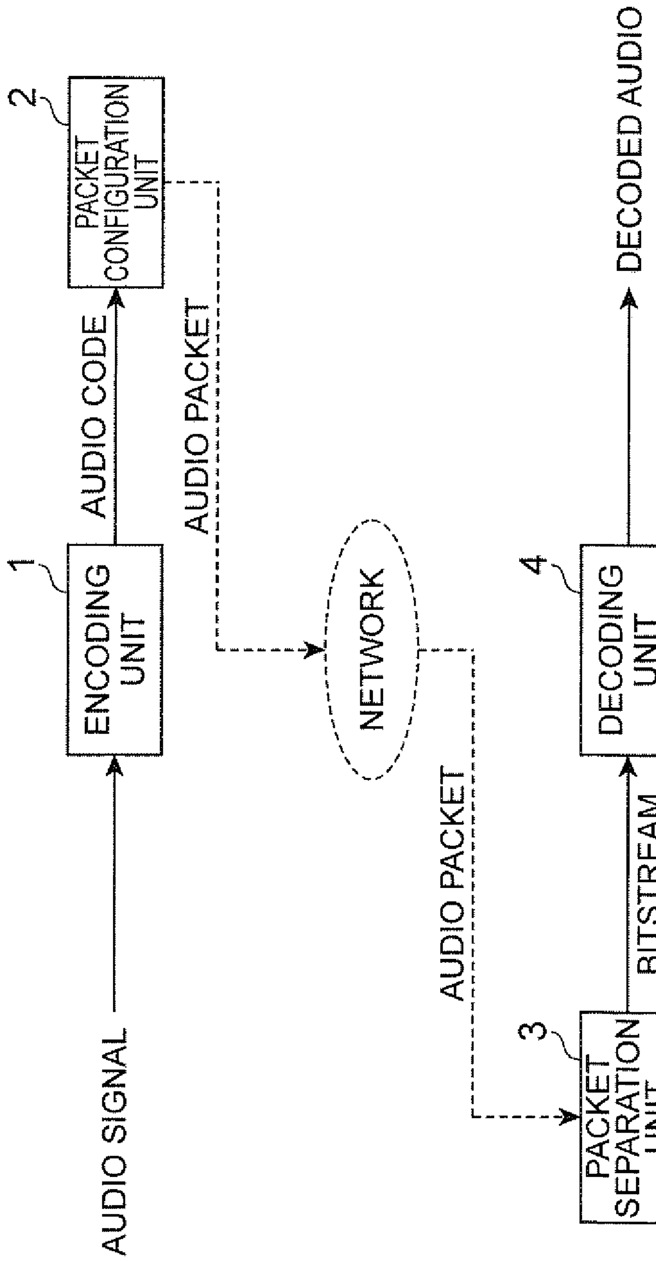

AUDIO SIGNAL
1
ENCODING UNIT
AUDIO CODE
2
PACKET CONFIGURATION UNIT
AUDIO PACKET
NETWORK
AUDIO PACKET
3
PACKET SEPARATION UNIT
BITSTREAM
4
DECODING UNIT
DECODED AUDIO

Fig.2

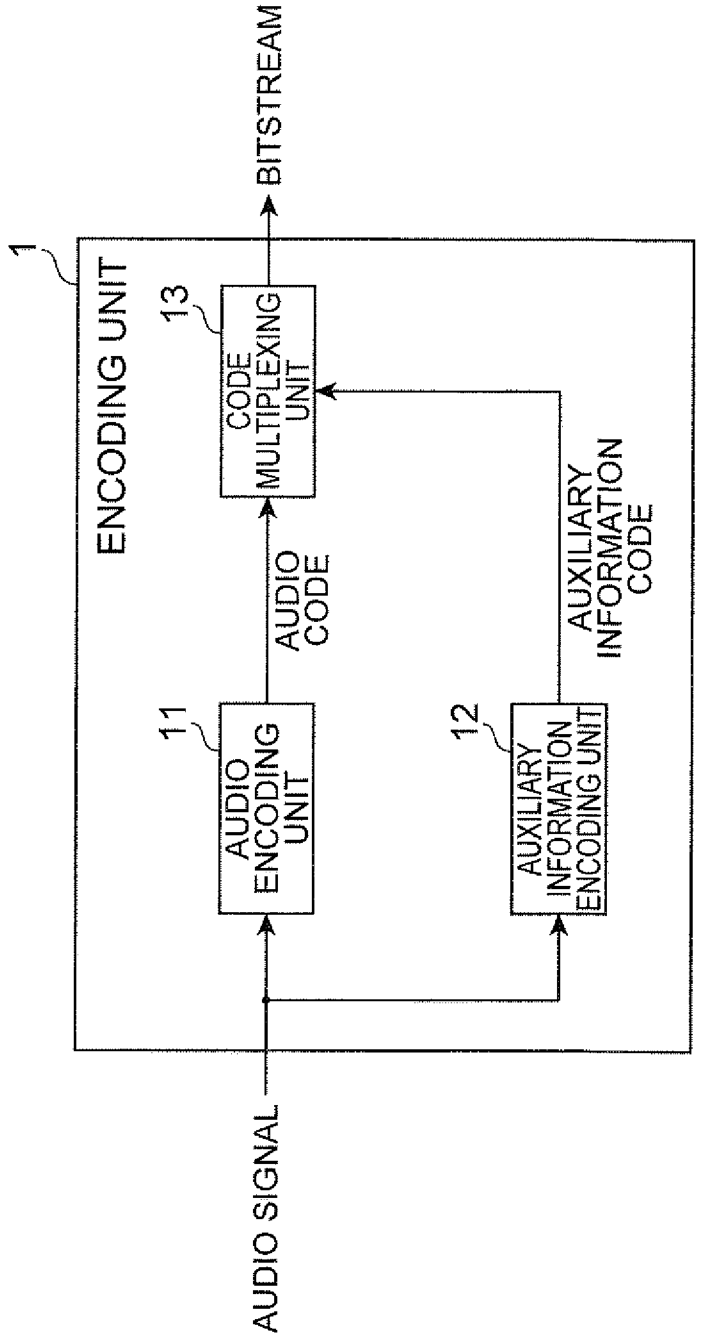

ENCODING UNIT
1
AUDIO SIGNAL
AUDIO ENCODING UNIT
11
AUXILIARY INFORMATION ENCODING UNIT
12
AUDIO CODE
AUXILIARY INFORMATION CODE
CODE MULTIPLEXING UNIT
13
BITSTREAM

Fig.3

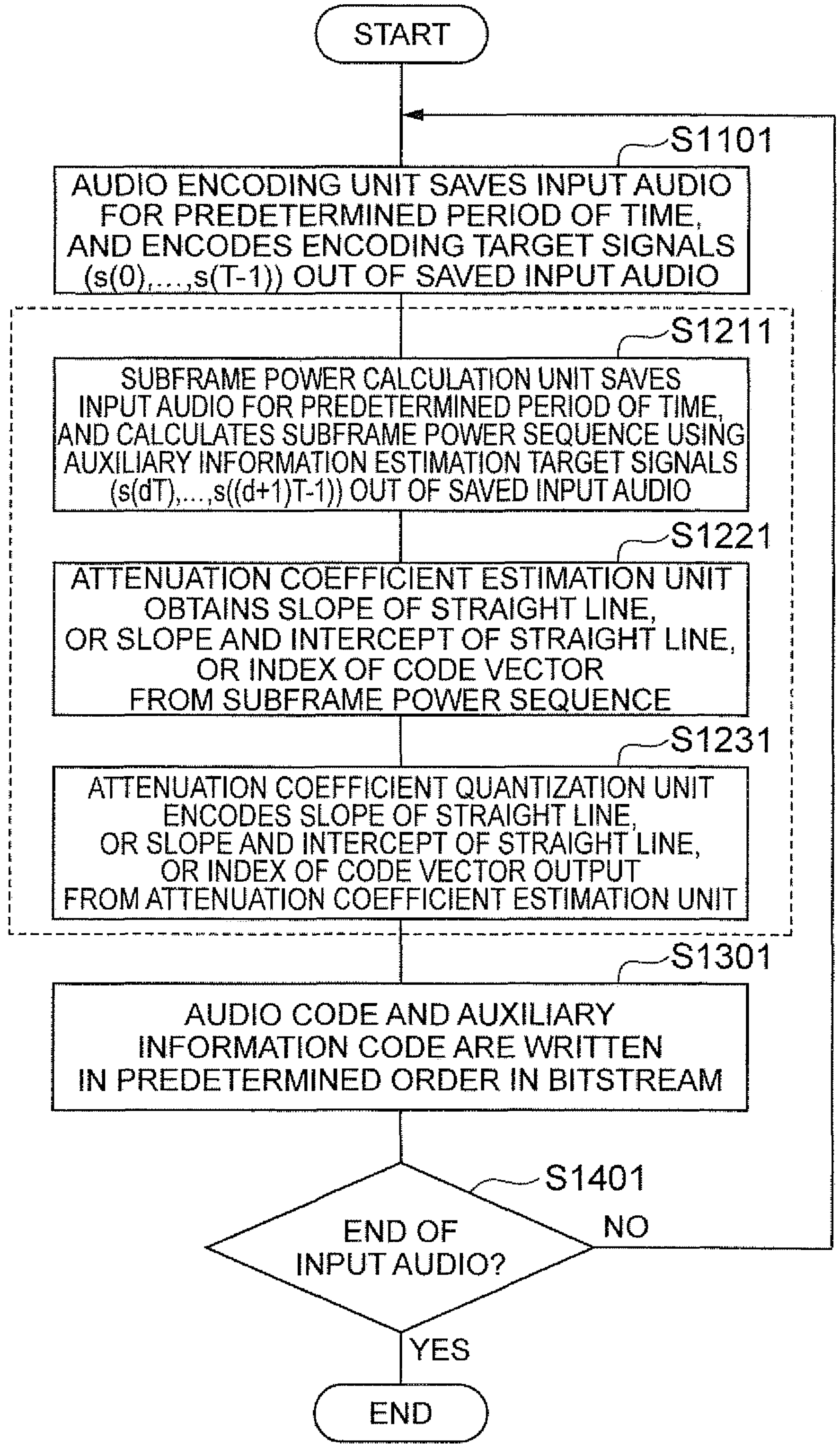

START
S1101
AUDIO ENCODING UNIT SAVES INPUT AUDIO FOR PREDETERMINED PERIOD OF TIME, AND ENCODES ENCODING TARGET SIGNALS (s(0),...,s(T-1)) OUT OF SAVED INPUT AUDIO
S1211
SUBFRAME POWER CALCULATION UNIT SAVES INPUT AUDIO FOR PREDETERMINED PERIOD OF TIME, AND CALCULATES SUBFRAME POWER SEQUENCE USING AUXILIARY INFORMATION ESTIMATION TARGET SIGNALS (s(dT),...,s((d+1)T-1)) OUT OF SAVED INPUT AUDIO
S1221
ATTENUATION COEFFICIENT ESTIMATION UNIT OBTAINS SLOPE OF STRAIGHT LINE, OR SLOPE AND INTERCEPT OF STRAIGHT LINE, OR INDEX OF CODE VECTOR FROM SUBFRAME POWER SEQUENCE
S1231
ATTENUATION COEFFICIENT QUANTIZATION UNIT ENCODES SLOPE OF STRAIGHT LINE, OR SLOPE AND INTERCEPT OF STRAIGHT LINE, OR INDEX OF CODE VECTOR OUTPUT FROM ATTENUATION COEFFICIENT ESTIMATION UNIT
S1301
AUDIO CODE AND AUXILIARY INFORMATION CODE ARE WRITTEN IN PREDETERMINED ORDER IN BITSTREAM
S1401
END OF INPUT AUDIO?
NO
YES
END

Fig.4

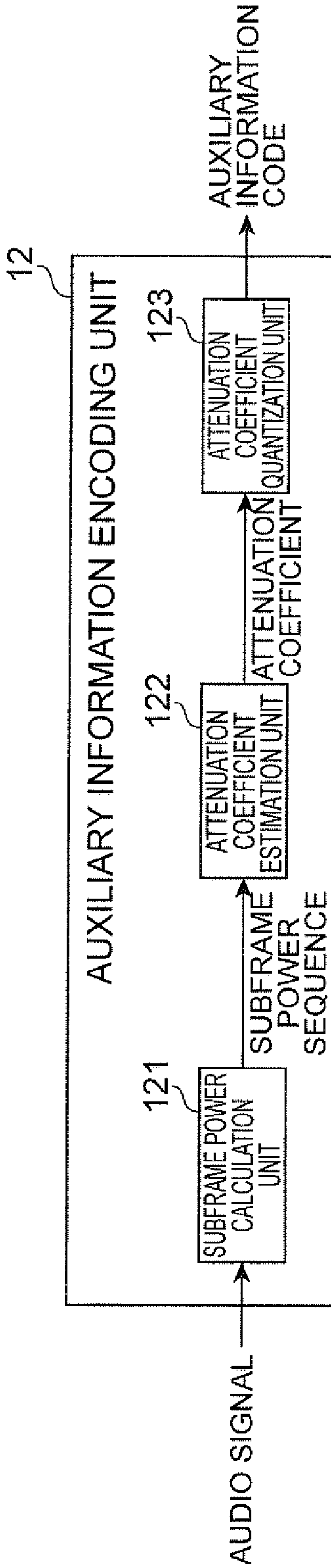

AUXILIARY INFORMATION ENCODING UNIT
12
AUDIO SIGNAL
SUBFRAME POWER CALCULATION UNIT
121
SUBFRAME POWER SEQUENCE
ATTENUATION COEFFICIENT ESTIMATION UNIT
122
ATTENUATION COEFFICIENT
ATTENUATION COEFFICIENT QUANTIZATION UNIT
123
AUXILIARY INFORMATION CODE

Fig.5

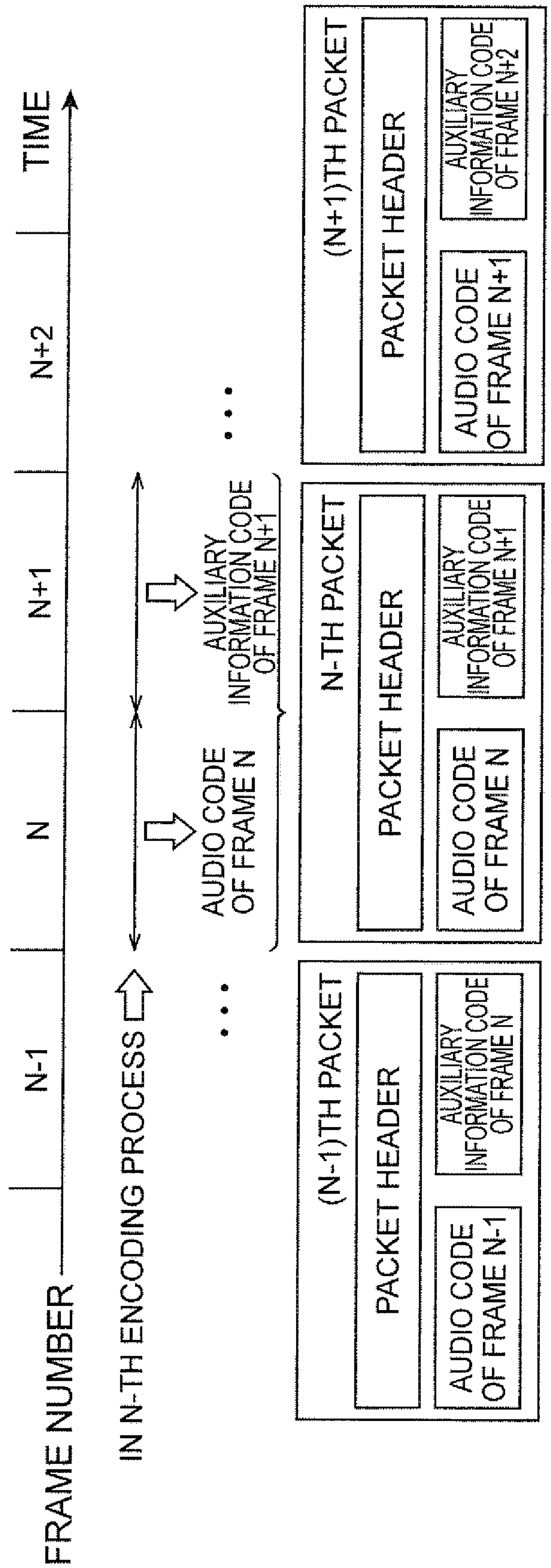

FRAME NUMBER
N-1
N
N+1
N+2
TIME
IN N-TH ENCODING PROCESS
AUDIO CODE OF FRAME N
AUXILIARY INFORMATION CODE OF FRAME N+1
(N-1)TH PACKET
PACKET HEADER
AUDIO CODE OF FRAME N-1
AUXILIARY INFORMATION CODE OF FRAME N
N-TH PACKET
PACKET HEADER
AUDIO CODE OF FRAME N
AUXILIARY INFORMATION CODE OF FRAME N+1
(N+1)TH PACKET
PACKET HEADER
AUDIO CODE OF FRAME N+1
AUXILIARY INFORMATION CODE OF FRAME N+2

Fig.6

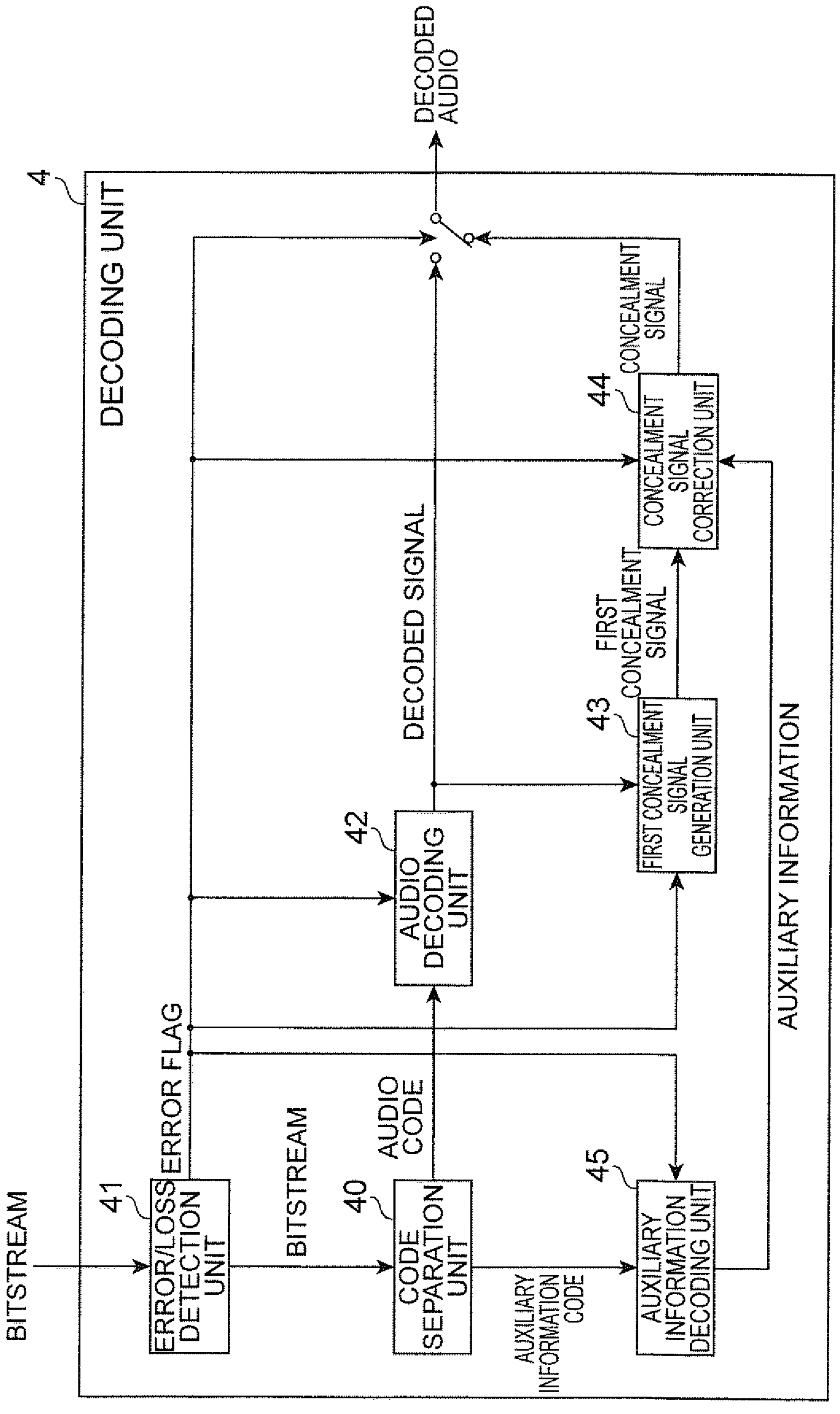

4
DECODING UNIT
BITSTREAM
41
ERROR/LOSS DETECTION UNIT
ERROR FLAG
BITSTREAM
40
CODE SEPARATION UNIT
AUDIO CODE
AUXILIARY INFORMATION CODE
42
AUDIO DECODING UNIT
DECODED SIGNAL
45
AUXILIARY INFORMATION DECODING UNIT
AUXILIARY INFORMATION
43
FIRST CONCEALMENT SIGNAL GENERATION UNIT
FIRST CONCEALMENT SIGNAL
44
CONCEALMENT SIGNAL CORRECTION UNIT
CONCEALMENT SIGNAL
DECODED AUDIO

Fig.7

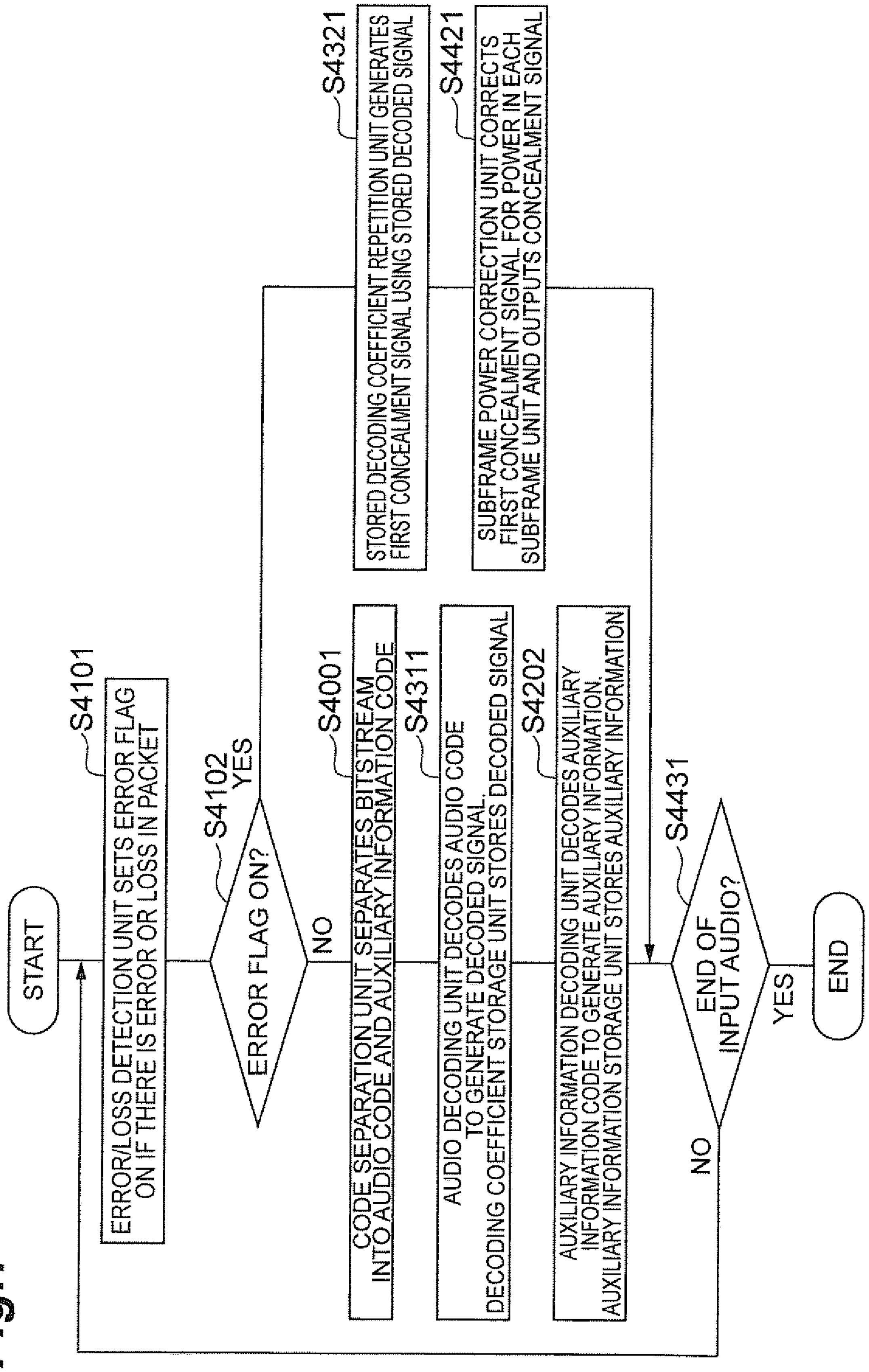

START
S4101
ERROR/LOSS DETECTION UNIT SETS ERROR FLAG ON IF THERE IS ERROR OR LOSS IN PACKET
S4102
ERROR FLAG ON?
YES
NO
S4001
CODE SEPARATION UNIT SEPARATES BITSTREAM INTO AUDIO CODE AND AUXILIARY INFORMATION CODE
S4311
AUDIO DECODING UNIT DECODES AUDIO CODE TO GENERATE DECODED SIGNAL. DECODING COEFFICIENT STORAGE UNIT STORES DECODED SIGNAL
S4202
AUXILIARY INFORMATION DECODING UNIT DECODES AUXILIARY INFORMATION CODE TO GENERATE AUXILIARY INFORMATION. AUXILIARY INFORMATION STORAGE UNIT STORES AUXILIARY INFORMATION
S4321
STORED DECODING COEFFICIENT REPETITION UNIT GENERATES FIRST CONCEALMENT SIGNAL USING STORED DECODED SIGNAL
S4421
SUBFRAME POWER CORRECTION UNIT CORRECTS FIRST CONCEALMENT SIGNAL FOR POWER IN EACH SUBFRAME UNIT AND OUTPUTS CONCEALMENT SIGNAL
S4431
END OF INPUT AUDIO?
YES
NO
END

Fig.8

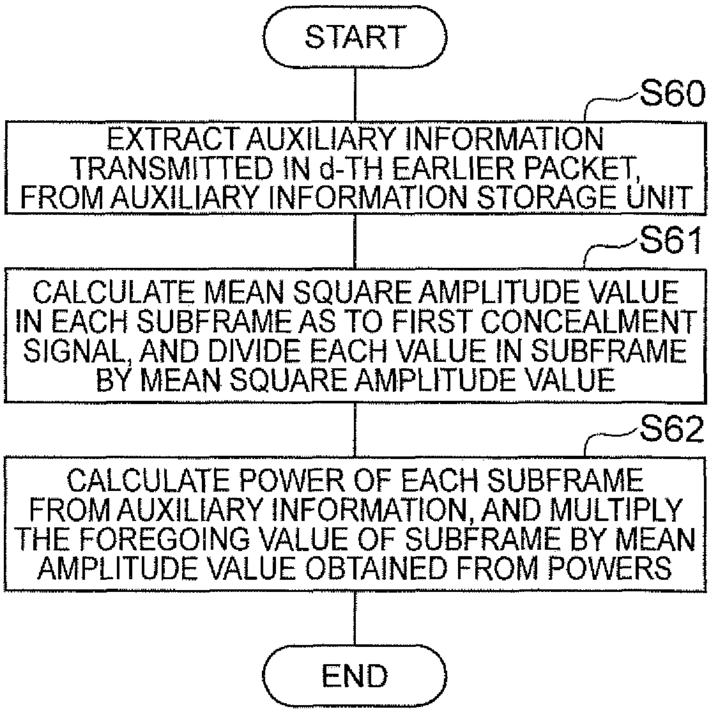
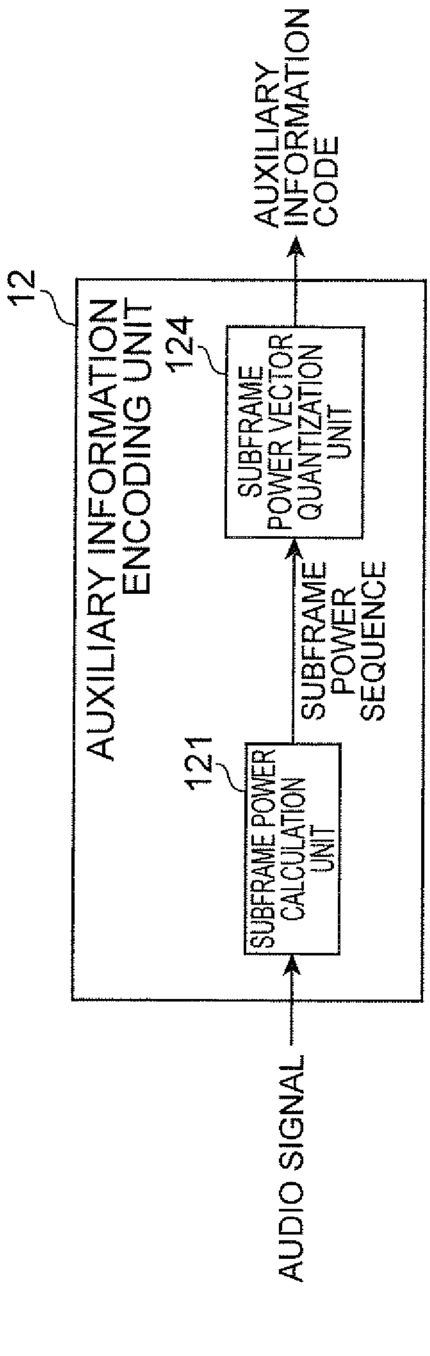

START
S60
EXTRACT AUXILIARY INFORMATION TRANSMITTED IN d-TH EARLIER PACKET, FROM AUXILIARY INFORMATION STORAGE UNIT
S61
CALCULATE MEAN SQUARE AMPLITUDE VALUE IN EACH SUBFRAME AS TO FIRST CONCEALMENT SIGNAL, AND DIVIDE EACH VALUE IN SUBFRAME BY MEAN SQUARE AMPLITUDE VALUE
S62
CALCULATE POWER OF EACH SUBFRAME FROM AUXILIARY INFORMATION, AND MULTIPLY THE FOREGOING VALUE OF SUBFRAME BY MEAN AMPLITUDE VALUE OBTAINED FROM POWERS
END

Fig.9


12
AUXILIARY INFORMATION ENCODING UNIT
121
SUBFRAME POWER CALCULATION UNIT
124
SUBFRAME POWER VECTOR QUANTIZATION UNIT
SUBFRAME POWER SEQUENCE
AUDIO SIGNAL
AUXILIARY INFORMATION CODE

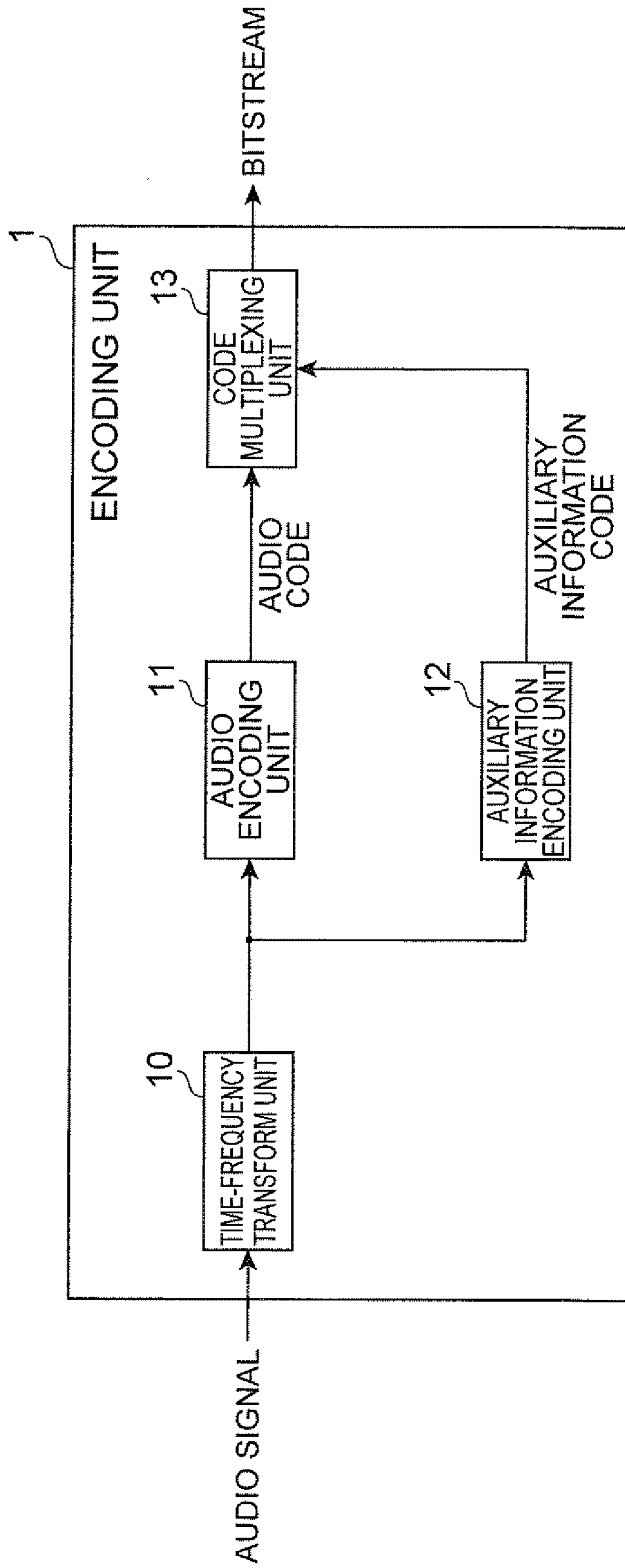

ENCODING UNIT
1
AUDIO SIGNAL
TIME-FREQUENCY TRANSFORM UNIT
10
AUDIO ENCODING UNIT
11
AUXILIARY INFORMATION ENCODING UNIT
12
AUDIO CODE
AUXILIARY INFORMATION CODE
CODE MULTIPLEXING UNIT
13
BITSTREAM

Fig.10

Fig.11

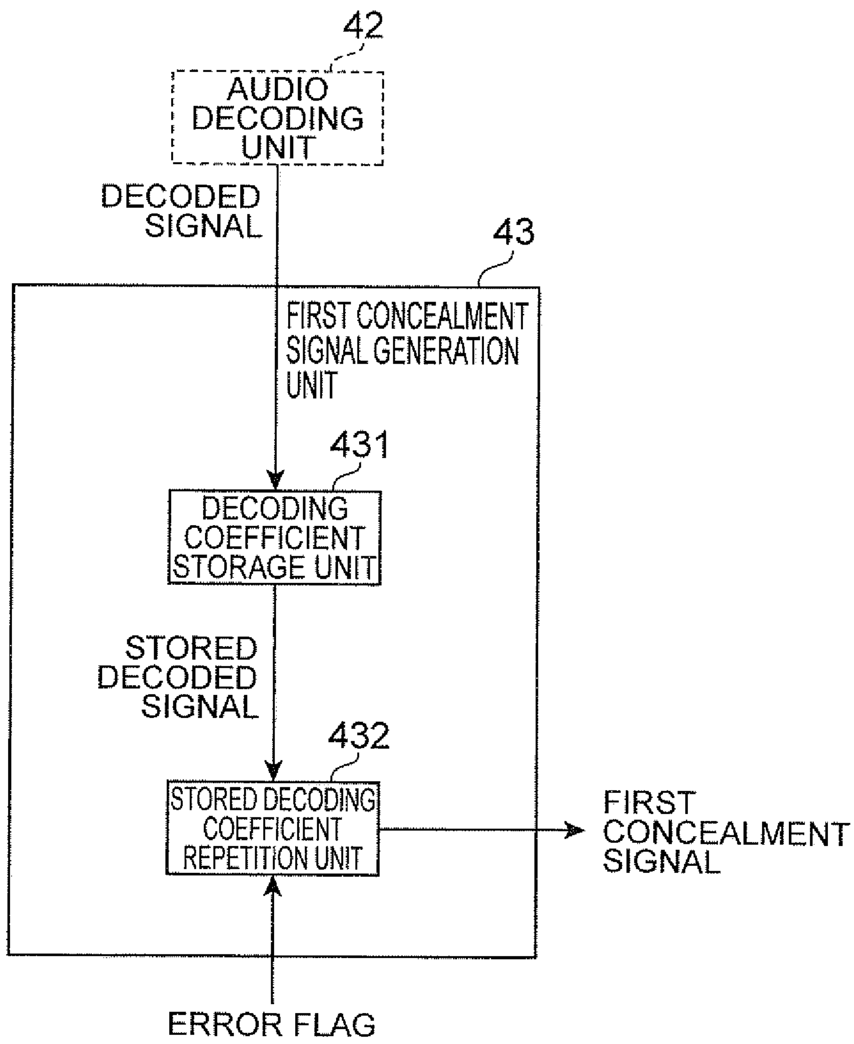
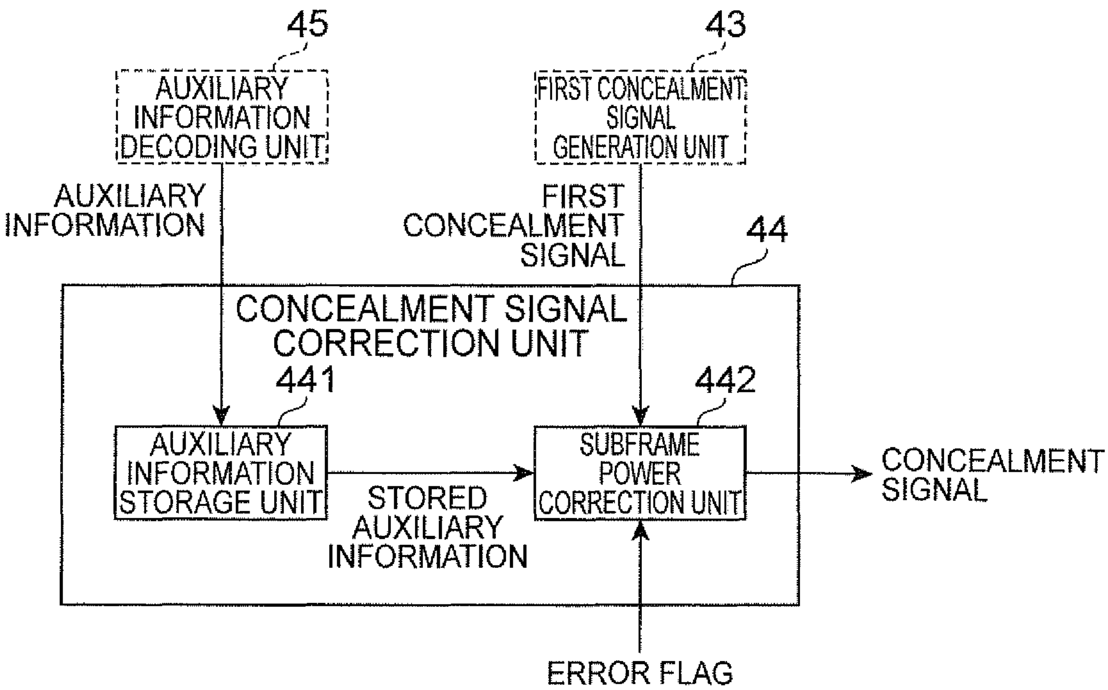

42
AUDIO DECODING UNIT
DECODED SIGNAL
43
FIRST CONCEALMENT SIGNAL GENERATION UNIT
431
DECODING COEFFICIENT STORAGE UNIT
STORED DECODED SIGNAL
432
STORED DECODING COEFFICIENT REPETITION UNIT
FIRST CONCEALMENT SIGNAL
ERROR FLAG

Fig.12


45
AUXILIARY INFORMATION DECODING UNIT
43
FIRST CONCEALMENT SIGNAL GENERATION UNIT
AUXILIARY INFORMATION
FIRST CONCEALMENT SIGNAL
44
CONCEALMENT SIGNAL CORRECTION UNIT
441
AUXILIARY INFORMATION STORAGE UNIT
442
SUBFRAME POWER CORRECTION UNIT
STORED AUXILIARY INFORMATION
CONCEALMENT SIGNAL
ERROR FLAG

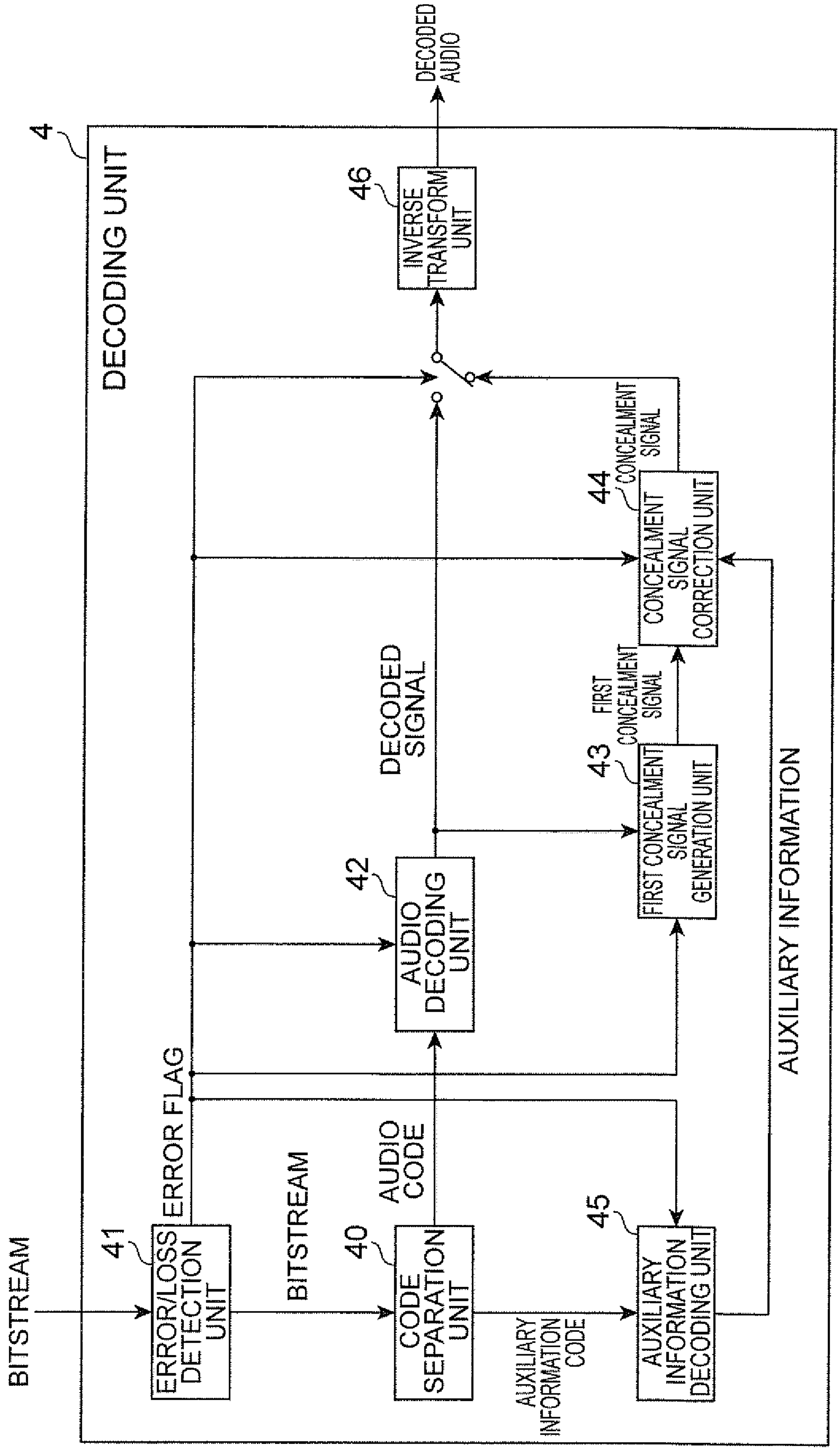

4
DECODING UNIT
BITSTREAM
41
ERROR/LOSS DETECTION UNIT
ERROR FLAG
BITSTREAM
40
CODE SEPARATION UNIT
AUDIO CODE
AUXILIARY INFORMATION CODE
45
AUXILIARY INFORMATION DECODING UNIT
AUXILIARY INFORMATION
42
AUDIO DECODING UNIT
DECODED SIGNAL
43
FIRST CONCEALMENT SIGNAL GENERATION UNIT
FIRST CONCEALMENT SIGNAL
44
CONCEALMENT SIGNAL CORRECTION UNIT
CONCEALMENT SIGNAL
46
INVERSE TRANSFORM UNIT
DECODED AUDIO

Fig.13

Fig.14

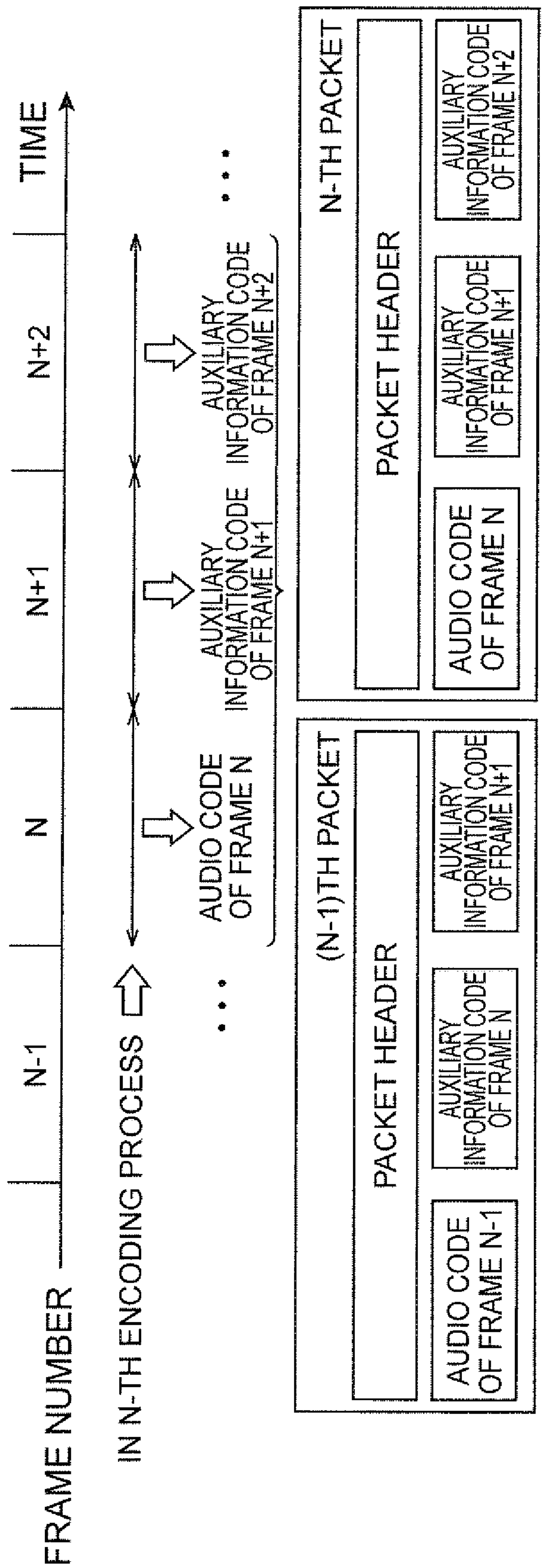

FRAME NUMBER
N-1
N
N+1
N+2
TIME
IN N-TH ENCODING PROCESS
AUDIO CODE OF FRAME N
AUXILIARY INFORMATION CODE OF FRAME N+1
AUXILIARY INFORMATION CODE OF FRAME N+2
(N-1)TH PACKET
PACKET HEADER
AUDIO CODE OF FRAME N-1
AUXILIARY INFORMATION CODE OF FRAME N
AUXILIARY INFORMATION CODE OF FRAME N+1
N-TH PACKET
PACKET HEADER
AUDIO CODE OF FRAME N
AUXILIARY INFORMATION CODE OF FRAME N+1
AUXILIARY INFORMATION CODE OF FRAME N+2

Fig.15

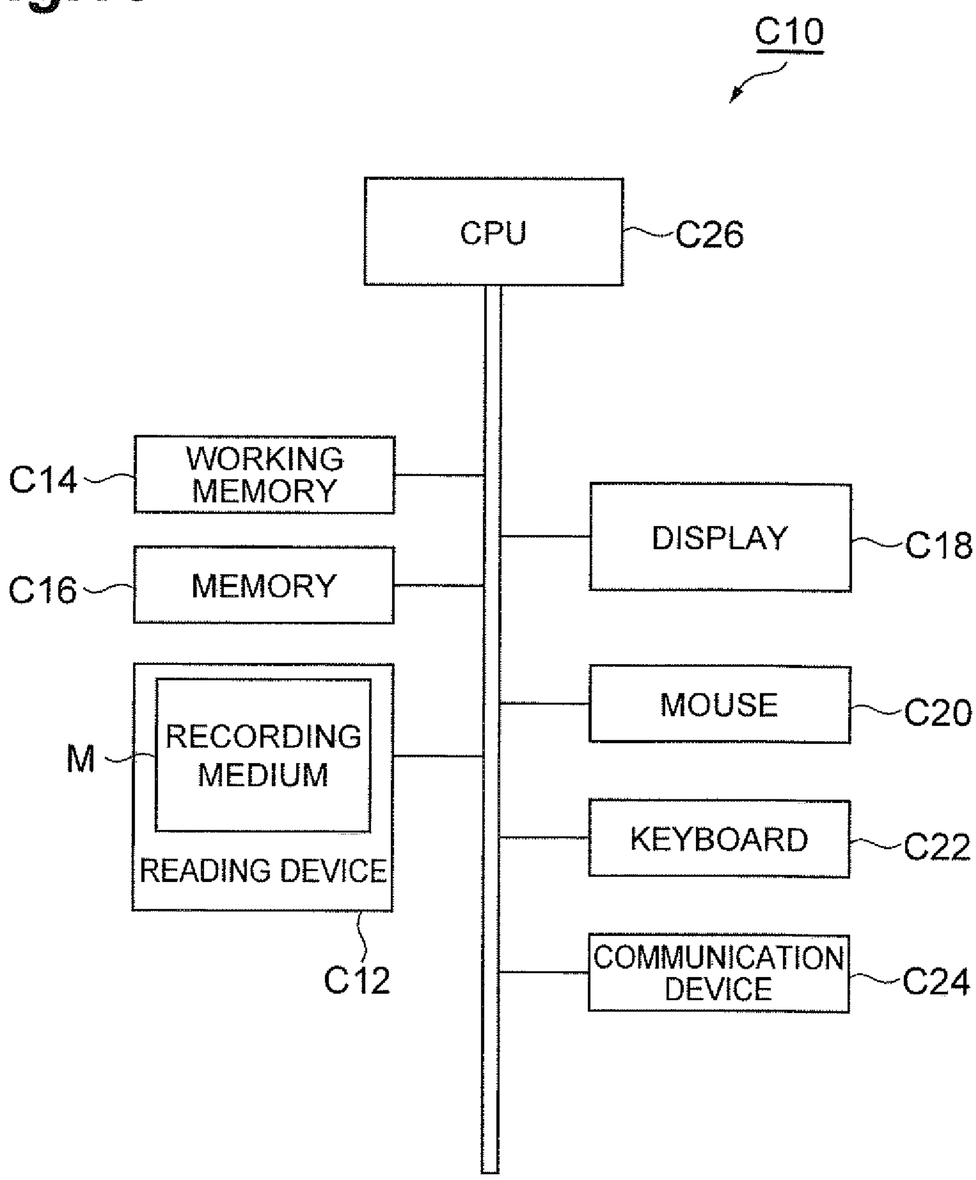

C10
CPU
C26
C14
WORKING MEMORY
C16
MEMORY
M
RECORDING MEDIUM
READING DEVICE
C12
DISPLAY
C18
MOUSE
C20
KEYBOARD
C22
COMMUNICATION DEVICE
C24

Fig.16

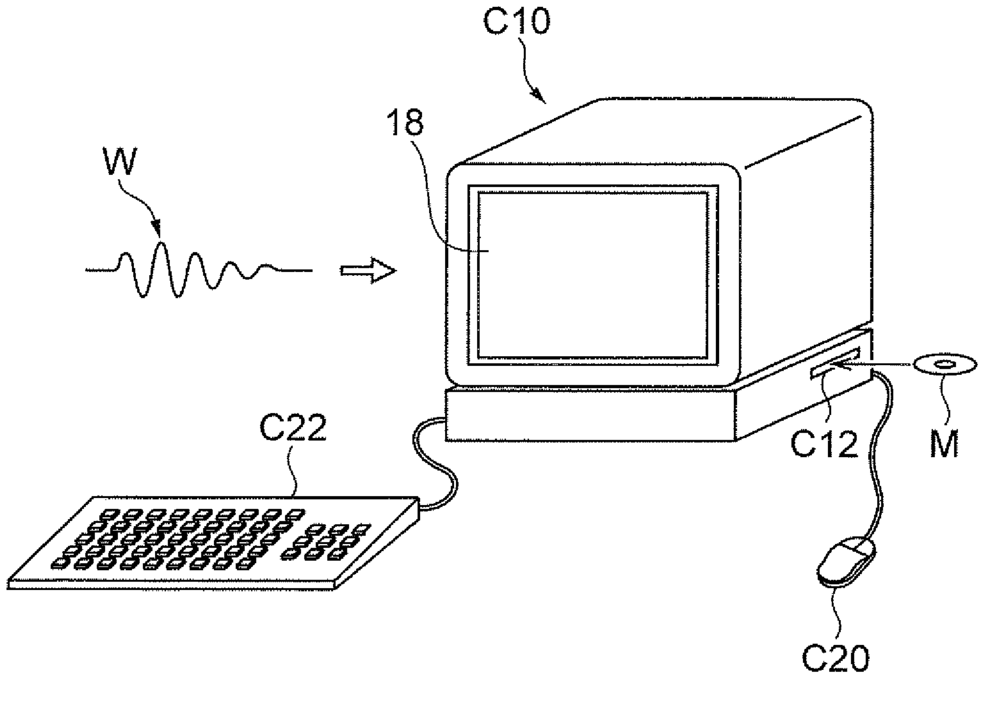


Fig.17

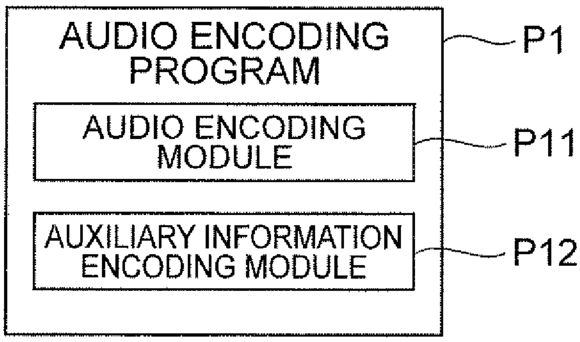

AUDIO ENCODING PROGRAM
P1
AUDIO ENCODING MODULE
P11
AUXILIARY INFORMATION ENCODING MODULE
P12

Fig.18

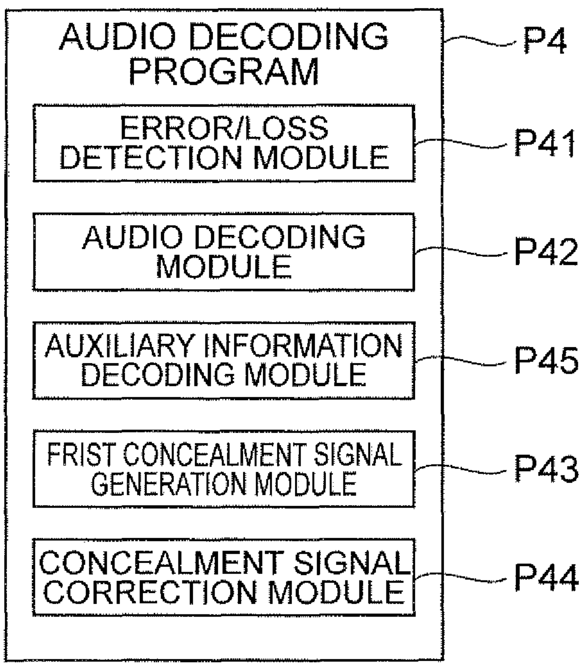

AUDIO DECODING PROGRAM
P4
ERROR/LOSS DETECTION MODULE
P41
AUDIO DECODING MODULE
P42
AUXILIARY INFORMATION DECODING MODULE
P45
FRIST CONCEALMENT SIGNAL GENERATION MODULE
P43
CONCEALMENT SIGNAL CORRECTION MODULE
P44

Fig.19

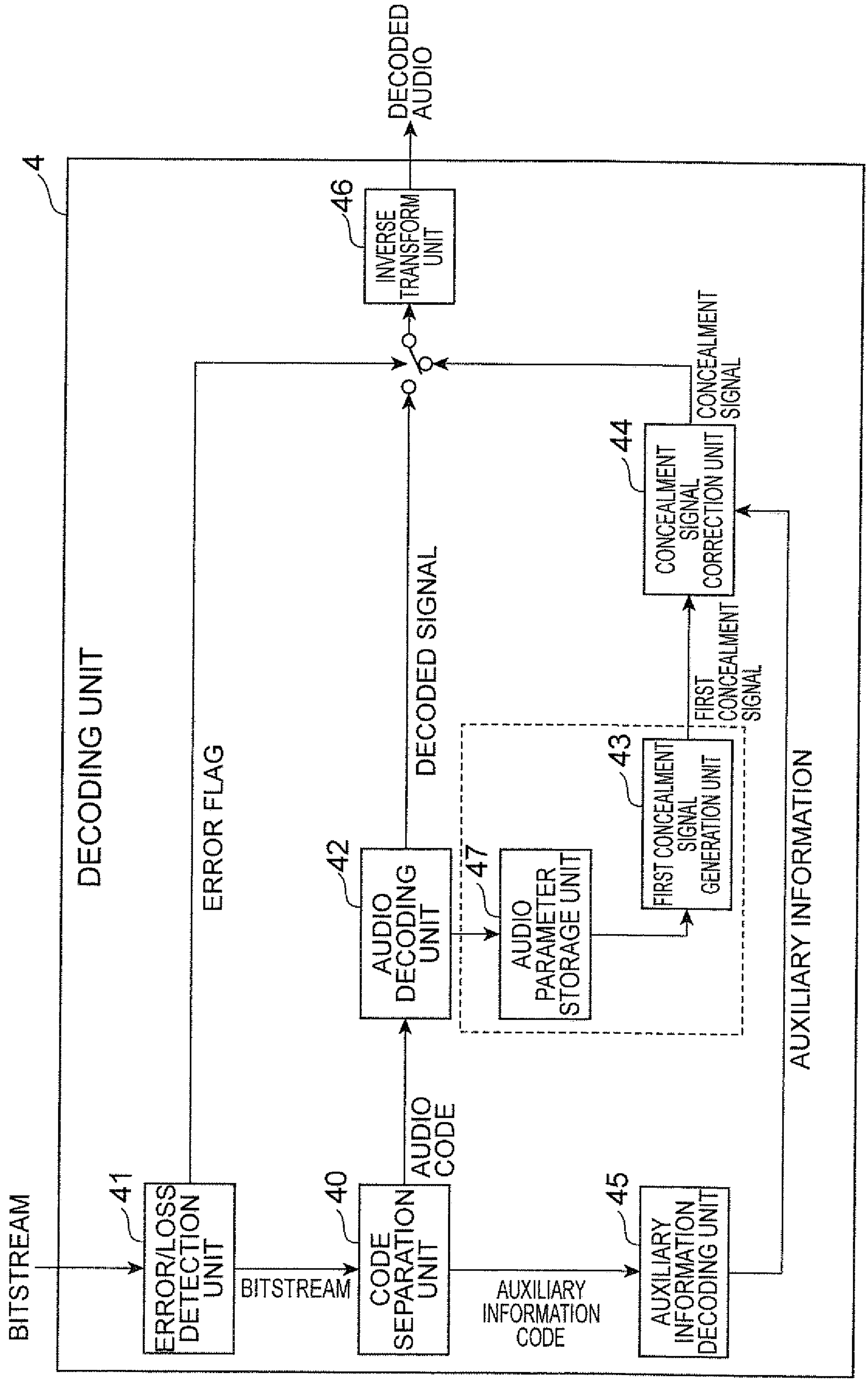

BITSTREAM
4
DECODING UNIT
41
ERROR/LOSS DETECTION UNIT
ERROR FLAG
BITSTREAM
40
CODE SEPARATION UNIT
AUDIO CODE
42
AUDIO DECODING UNIT
DECODED SIGNAL
47
AUDIO PARAMETER STORAGE UNIT
43
FIRST CONCEALMENT SIGNAL GENERATION UNIT
FIRST CONCEALMENT SIGNAL
44
CONCEALMENT SIGNAL CORRECTION UNIT
CONCEALMENT SIGNAL
46
INVERSE TRANSFORM UNIT
DECODED AUDIO
AUXILIARY INFORMATION CODE
45
AUXILIARY INFORMATION DECODING UNIT
AUXILIARY INFORMATION

Fig.20

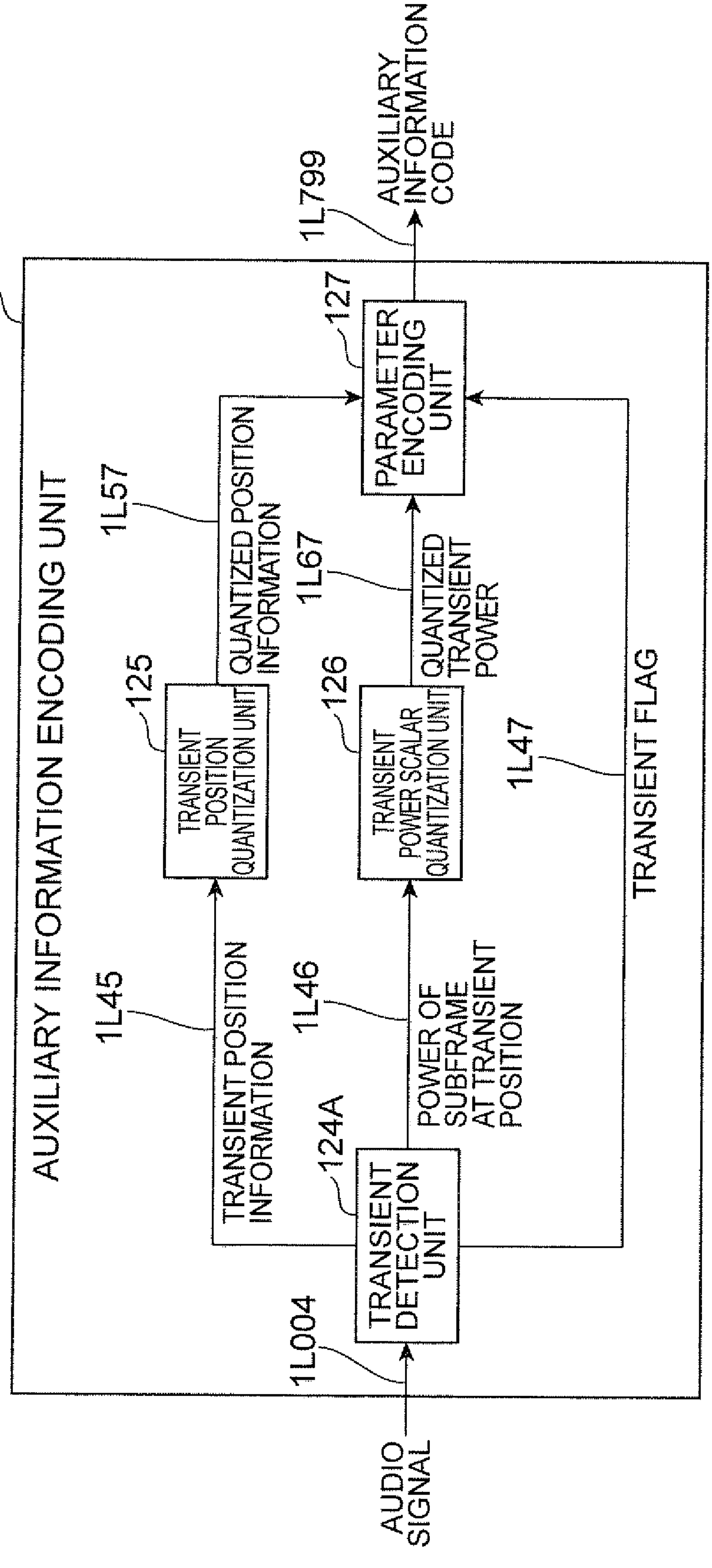

12
AUXILIARY INFORMATION ENCODING UNIT
AUDIO SIGNAL
1L004
TRANSIENT DETECTION UNIT
124A
TRANSIENT POSITION INFORMATION
1L45
POWER OF SUBFRAME AT TRANSIENT POSITION
1L46
TRANSIENT POSITION QUANTIZATION UNIT
125
TRANSIENT POWER SCALAR QUANTIZATION UNIT
126
QUANTIZED POSITION INFORMATION
1L57
QUANTIZED TRANSIENT POWER
1L67
1L47
TRANSIENT FLAG
PARAMETER ENCODING UNIT
127
1L799
AUXILIARY INFORMATION CODE

Fig.21

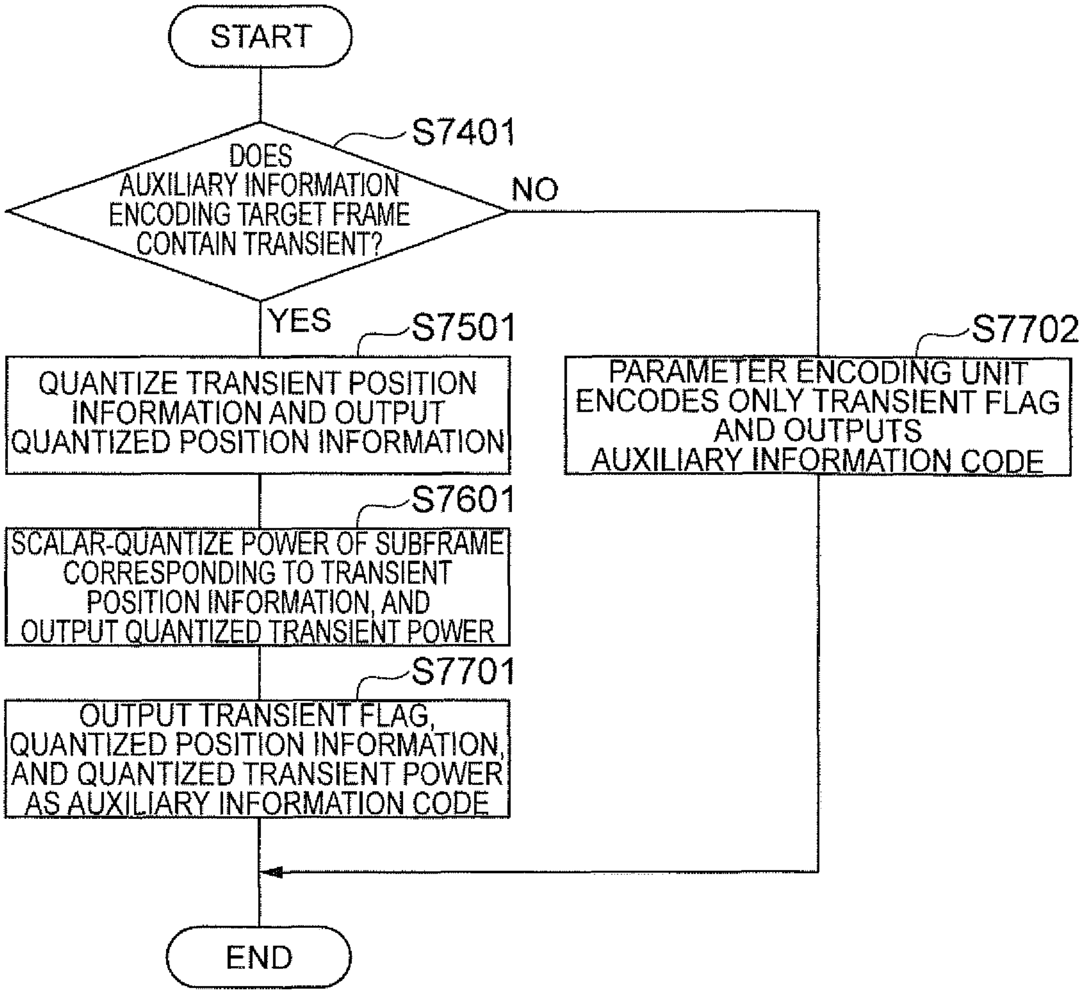

START
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
S7401
NO
YES
S7501
QUANTIZE TRANSIENT POSITION INFORMATION AND OUTPUT QUANTIZED POSITION INFORMATION
S7601
SCALAR-QUANTIZE POWER OF SUBFRAME CORRESPONDING TO TRANSIENT POSITION INFORMATION, AND OUTPUT QUANTIZED TRANSIENT POWER
S7701
OUTPUT TRANSIENT FLAG, QUANTIZED POSITION INFORMATION, AND QUANTIZED TRANSIENT POWER AS AUXILIARY INFORMATION CODE
S7702
PARAMETER ENCODING UNIT ENCODES ONLY TRANSIENT FLAG AND OUTPUTS AUXILIARY INFORMATION CODE
END

Fig.22

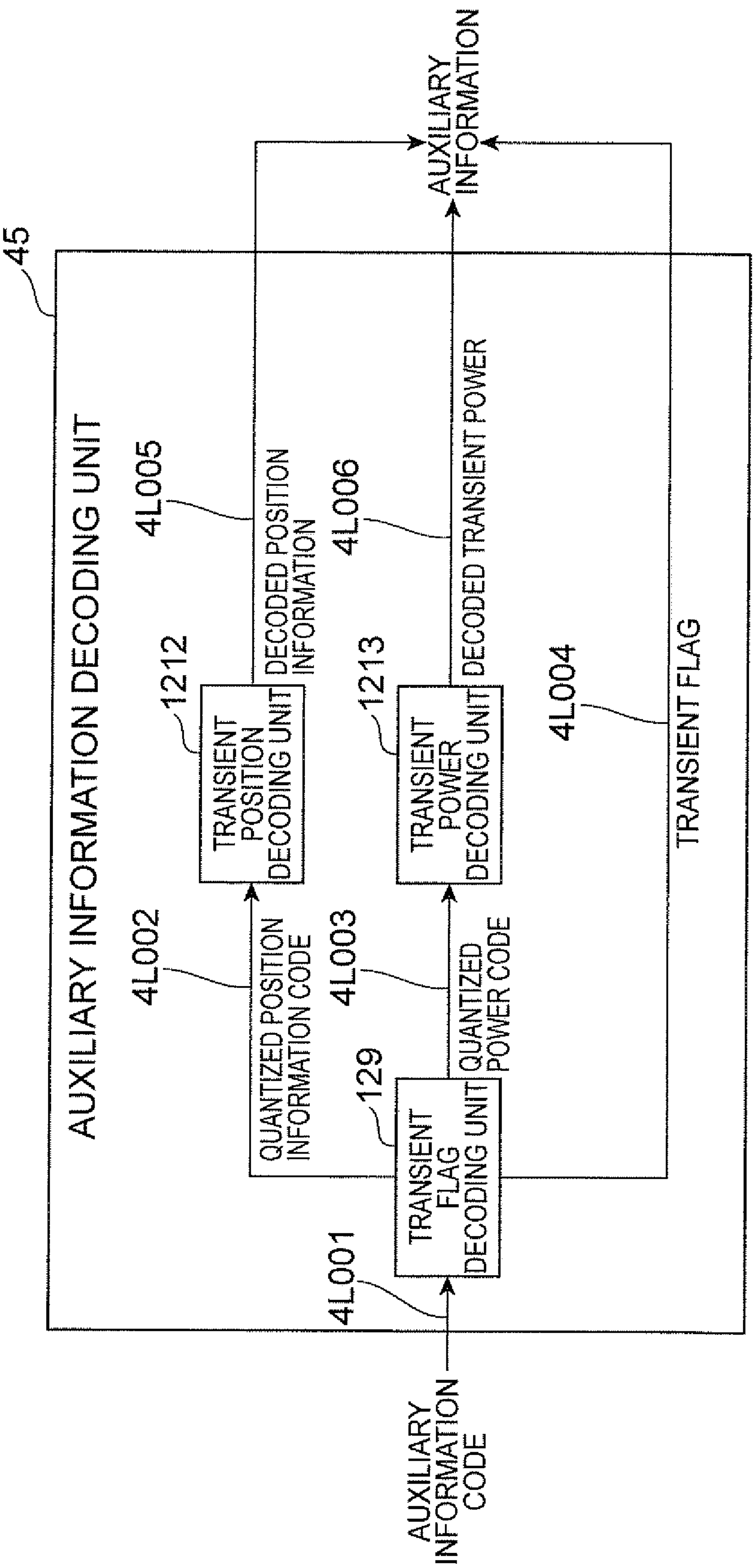

45
AUXILIARY INFORMATION DECODING UNIT
AUXILIARY INFORMATION CODE
4L001
TRANSIENT FLAG DECODING UNIT
129
QUANTIZED POSITION INFORMATION CODE
4L002
QUANTIZED POWER CODE
4L003
TRANSIENT POSITION DECODING UNIT
1212
TRANSIENT POWER DECODING UNIT
1213
DECODED POSITION INFORMATION
4L005
DECODED TRANSIENT POWER
4L006
TRANSIENT FLAG
4L004
AUXILIARY INFORMATION

Fig.23

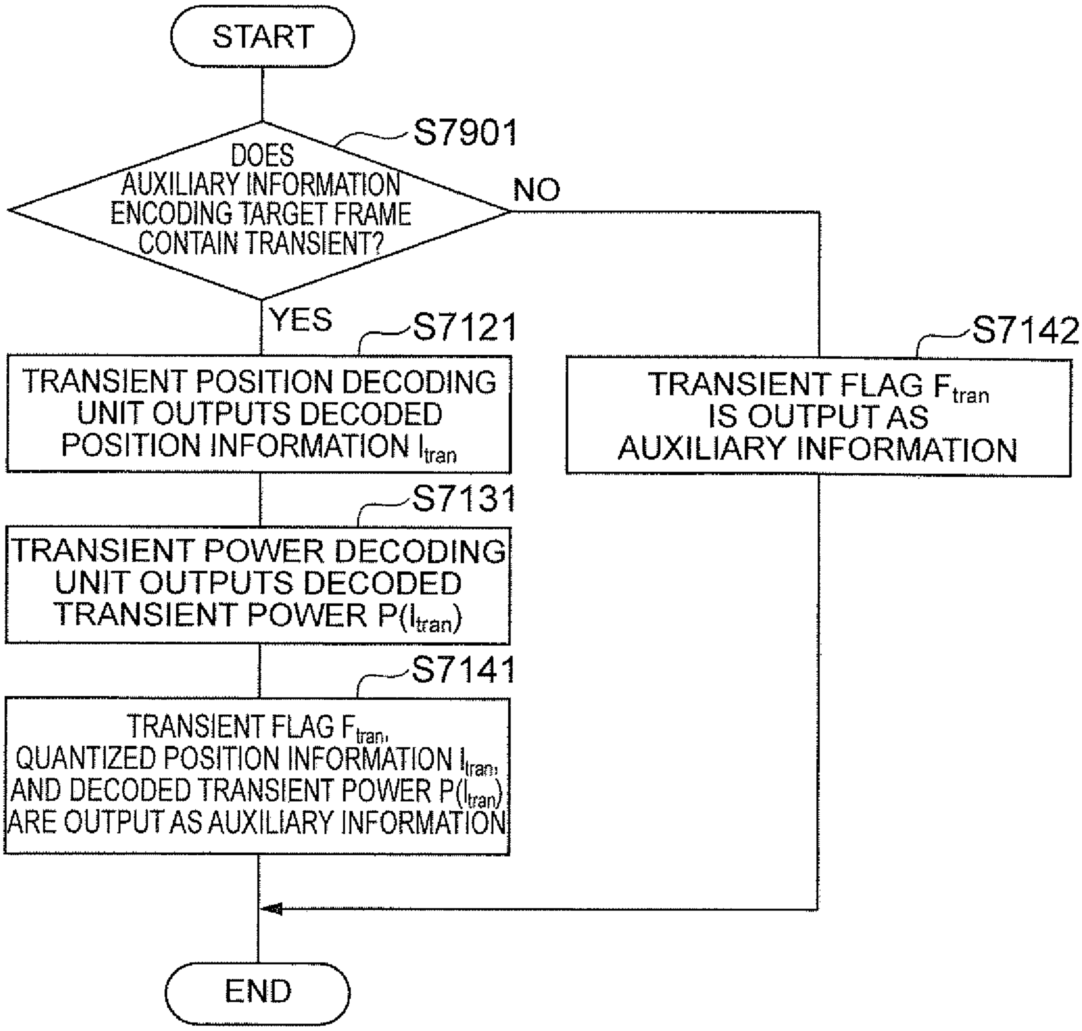

START
S7901
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S7121
TRANSIENT POSITION DECODING UNIT OUTPUTS DECODED POSITION INFORMATION I_{tran}
S7131
TRANSIENT POWER DECODING UNIT OUTPUTS DECODED TRANSIENT POWER $P(I_{tran})$
S7141
TRANSIENT FLAG F_{tran}, QUANTIZED POSITION INFORMATION I_{tran}, AND DECODED TRANSIENT POWER $P(I_{tran})$ ARE OUTPUT AS AUXILIARY INFORMATION
S7142
TRANSIENT FLAG F_{tran} IS OUTPUT AS AUXILIARY INFORMATION
END

Fig.24

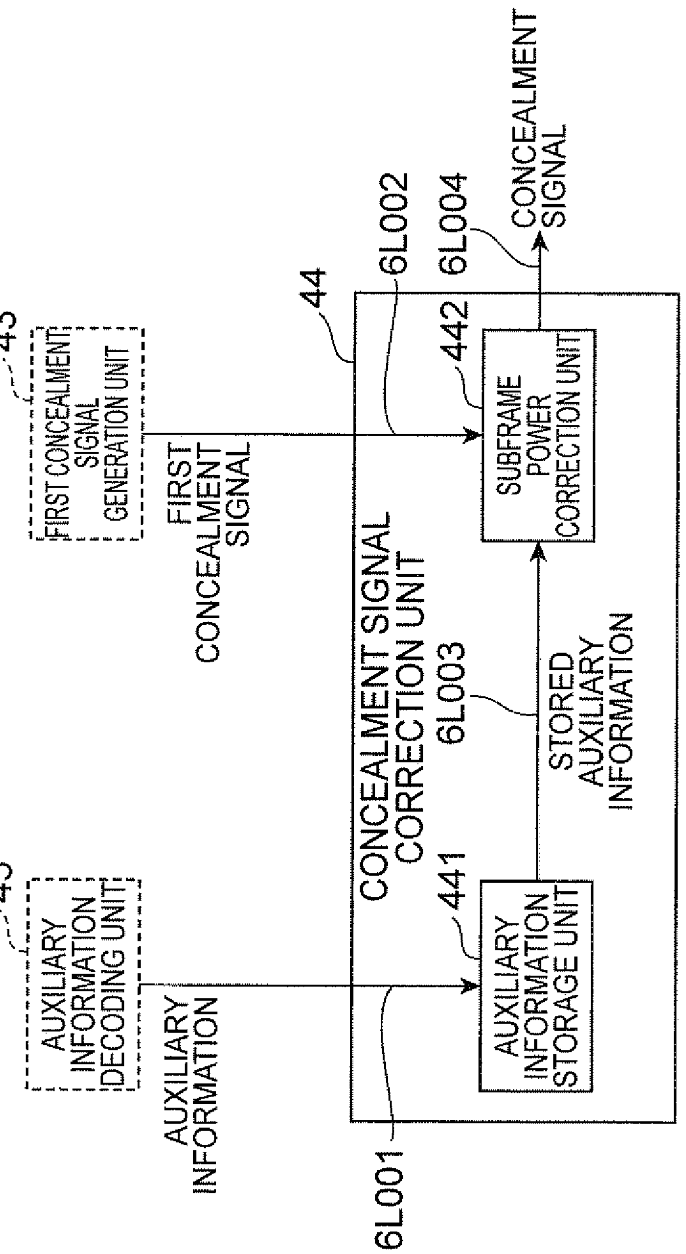

43
FIRST CONCEALMENT SIGNAL GENERATION UNIT
FIRST CONCEALMENT SIGNAL
45
AUXILIARY INFORMATION DECODING UNIT
AUXILIARY INFORMATION
6L001
44
CONCEALMENT SIGNAL CORRECTION UNIT
441
AUXILIARY INFORMATION STORAGE UNIT
6L003
STORED AUXILIARY INFORMATION
442
SUBFRAME POWER CORRECTION UNIT
6L002
6L004
CONCEALMENT SIGNAL

Fig.25

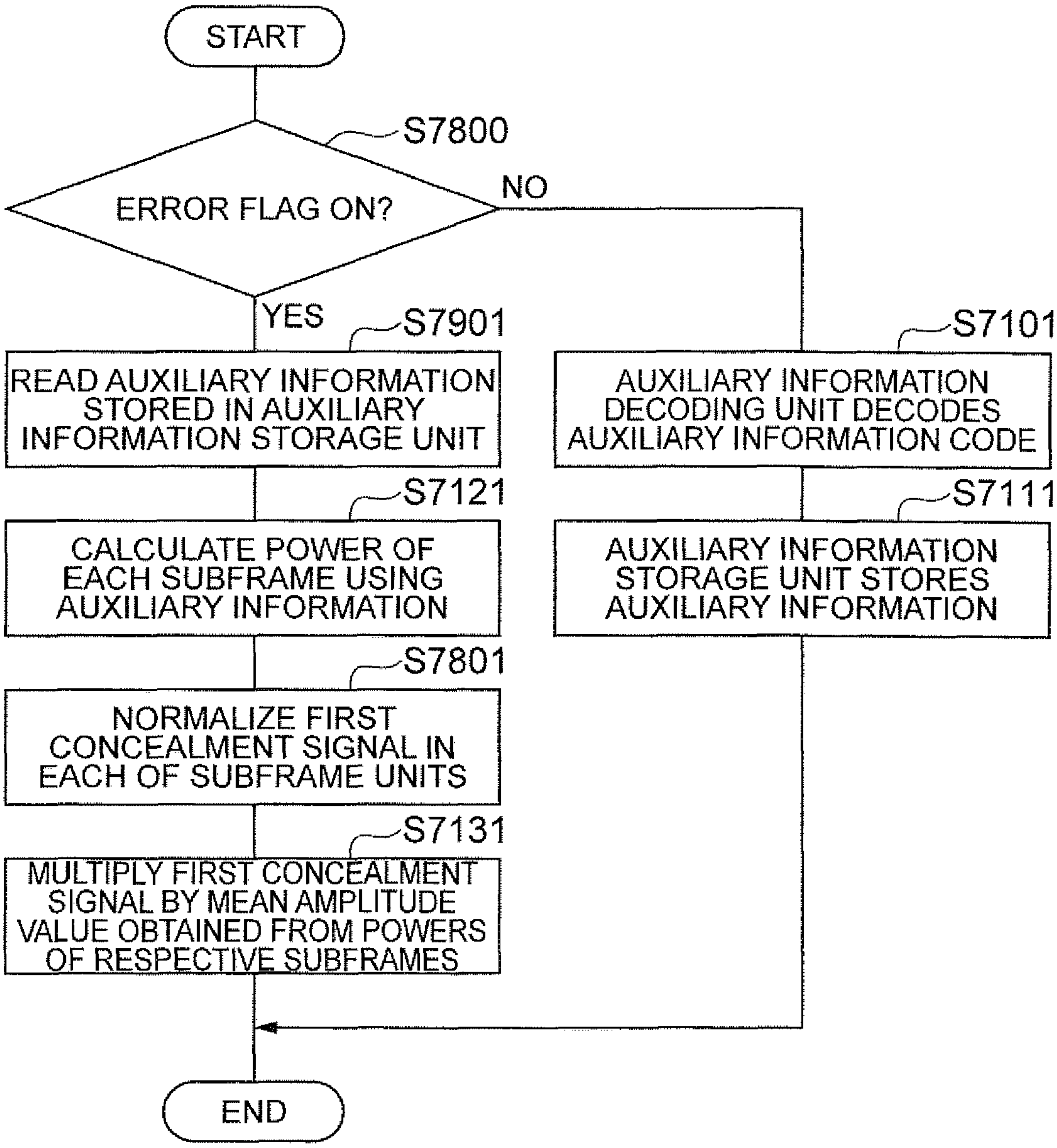

START
S7800
ERROR FLAG ON?
NO
YES
S7901
READ AUXILIARY INFORMATION STORED IN AUXILIARY INFORMATION STORAGE UNIT
S7121
CALCULATE POWER OF EACH SUBFRAME USING AUXILIARY INFORMATION
S7801
NORMALIZE FIRST CONCEALMENT SIGNAL IN EACH OF SUBFRAME UNITS
S7131
MULTIPLY FIRST CONCEALMENT SIGNAL BY MEAN AMPLITUDE VALUE OBTAINED FROM POWERS OF RESPECTIVE SUBFRAMES
S7101
AUXILIARY INFORMATION DECODING UNIT DECODES AUXILIARY INFORMATION CODE
S7111
AUXILIARY INFORMATION STORAGE UNIT STORES AUXILIARY INFORMATION
END

Fig.26

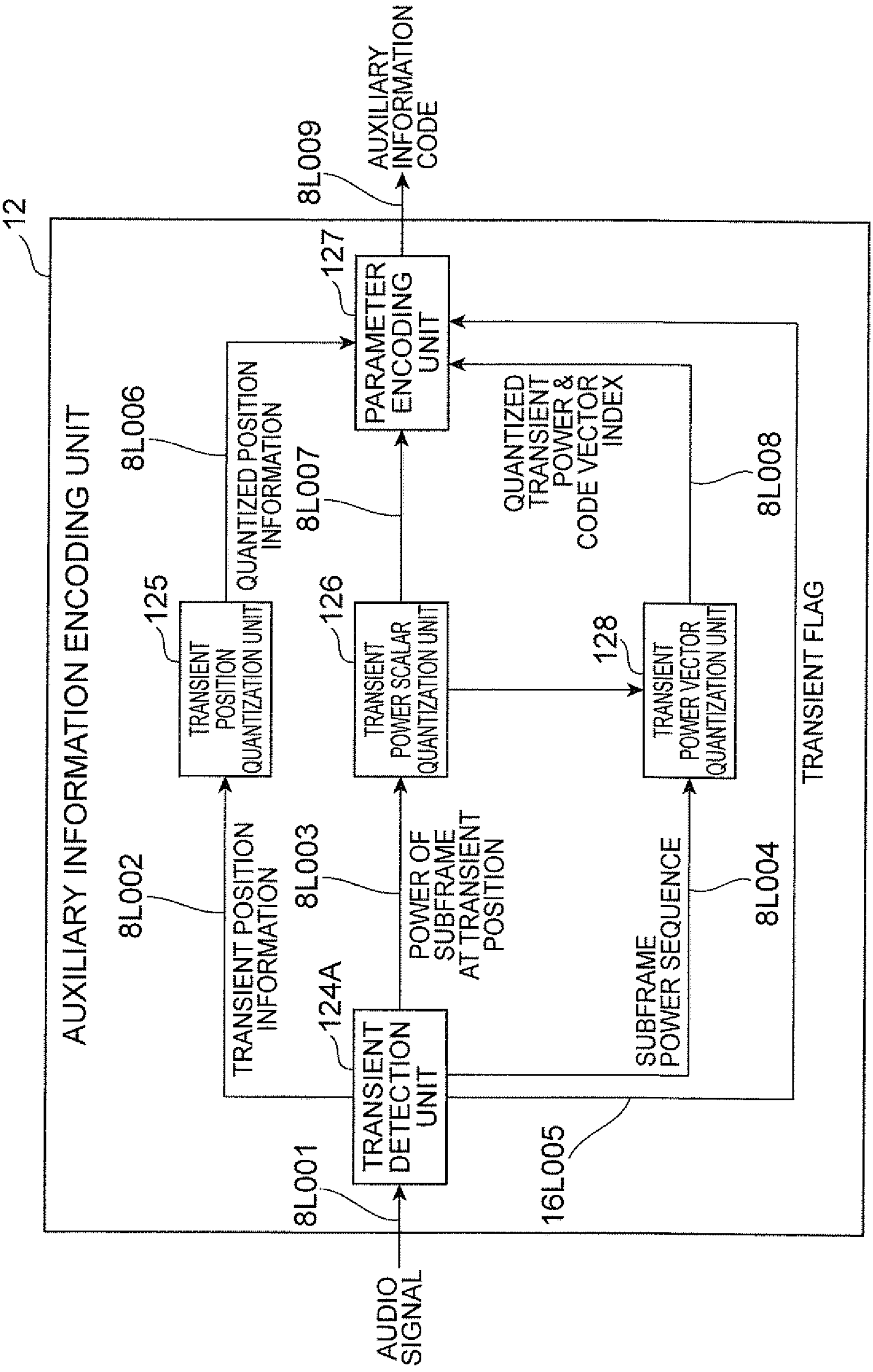

12
AUXILIARY INFORMATION ENCODING UNIT
AUDIO SIGNAL
8L001
TRANSIENT DETECTION UNIT
124A
8L002
TRANSIENT POSITION INFORMATION
8L003
POWER OF SUBFRAME AT TRANSIENT POSITION
SUBFRAME POWER SEQUENCE
8L004
16L005
TRANSIENT FLAG
125
TRANSIENT POSITION QUANTIZATION UNIT
126
TRANSIENT POWER SCALAR QUANTIZATION UNIT
128
TRANSIENT POWER VECTOR QUANTIZATION UNIT
8L006
QUANTIZED POSITION INFORMATION
8L007
QUANTIZED TRANSIENT POWER & CODE VECTOR INDEX
8L008
127
PARAMETER ENCODING UNIT
8L009
AUXILIARY INFORMATION CODE

Fig.27

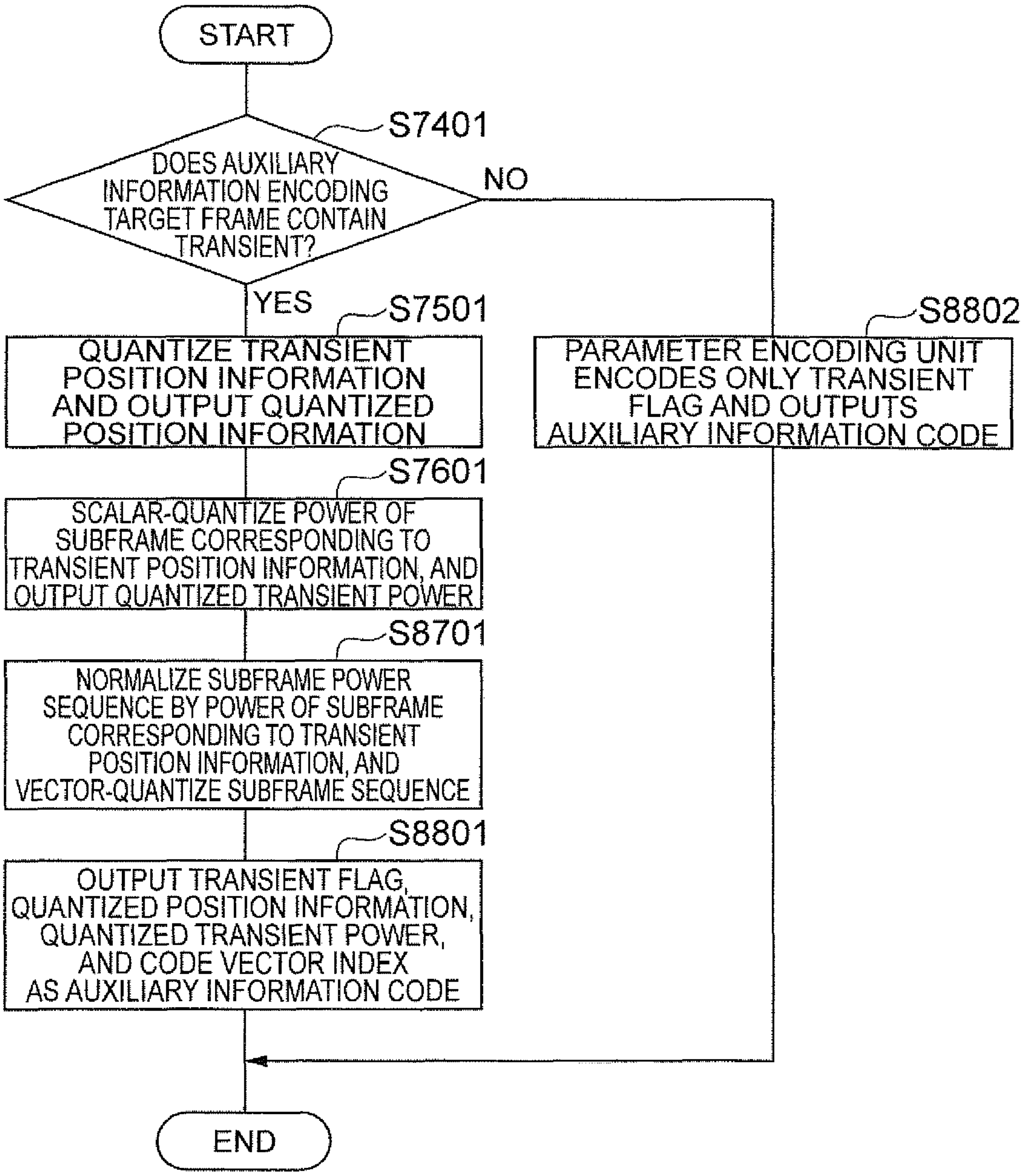

START
S7401
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S7501
QUANTIZE TRANSIENT POSITION INFORMATION AND OUTPUT QUANTIZED POSITION INFORMATION
S8802
PARAMETER ENCODING UNIT ENCODES ONLY TRANSIENT FLAG AND OUTPUTS AUXILIARY INFORMATION CODE
S7601
SCALAR-QUANTIZE POWER OF SUBFRAME CORRESPONDING TO TRANSIENT POSITION INFORMATION, AND OUTPUT QUANTIZED TRANSIENT POWER
S8701
NORMALIZE SUBFRAME POWER SEQUENCE BY POWER OF SUBFRAME CORRESPONDING TO TRANSIENT POSITION INFORMATION, AND VECTOR-QUANTIZE SUBFRAME SEQUENCE
S8801
OUTPUT TRANSIENT FLAG, QUANTIZED POSITION INFORMATION, QUANTIZED TRANSIENT POWER, AND CODE VECTOR INDEX AS AUXILIARY INFORMATION CODE
END

Fig.28

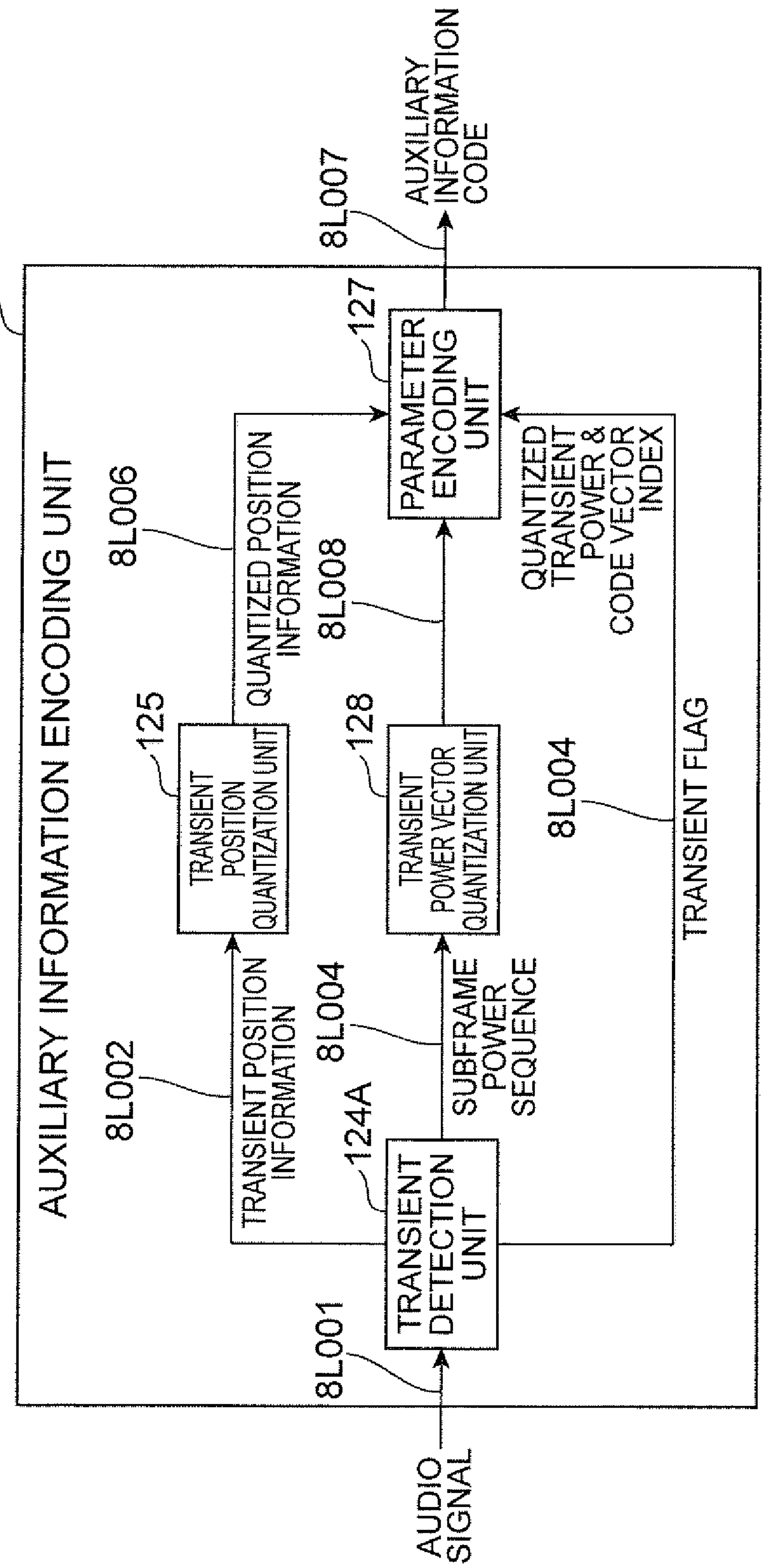

12
AUXILIARY INFORMATION ENCODING UNIT
AUDIO SIGNAL
8L001
TRANSIENT DETECTION UNIT
124A
8L002
TRANSIENT POSITION INFORMATION
8L004
SUBFRAME POWER SEQUENCE
TRANSIENT FLAG
TRANSIENT POSITION QUANTIZATION UNIT
125
TRANSIENT POWER VECTOR QUANTIZATION UNIT
128
8L004
8L006
QUANTIZED POSITION INFORMATION
8L008
QUANTIZED TRANSIENT POWER & CODE VECTOR INDEX
PARAMETER ENCODING UNIT
127
8L007
AUXILIARY INFORMATION CODE

Fig.29

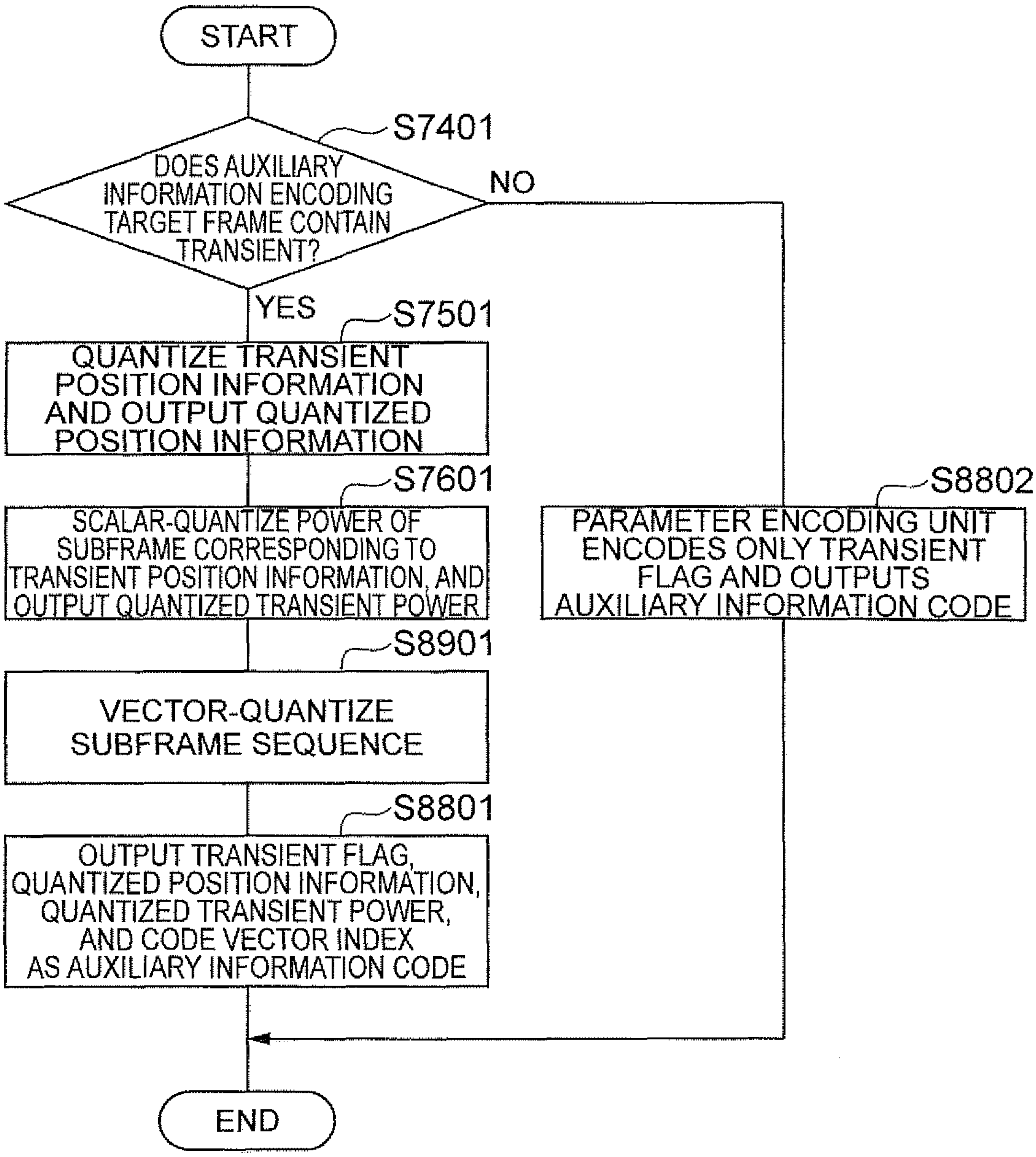

START
S7401
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S7501
QUANTIZE TRANSIENT POSITION INFORMATION AND OUTPUT QUANTIZED POSITION INFORMATION
S7601
SCALAR-QUANTIZE POWER OF SUBFRAME CORRESPONDING TO TRANSIENT POSITION INFORMATION, AND OUTPUT QUANTIZED TRANSIENT POWER
S8901
VECTOR-QUANTIZE SUBFRAME SEQUENCE
S8801
OUTPUT TRANSIENT FLAG, QUANTIZED POSITION INFORMATION, QUANTIZED TRANSIENT POWER, AND CODE VECTOR INDEX AS AUXILIARY INFORMATION CODE
S8802
PARAMETER ENCODING UNIT ENCODES ONLY TRANSIENT FLAG AND OUTPUTS AUXILIARY INFORMATION CODE
END

Fig.30

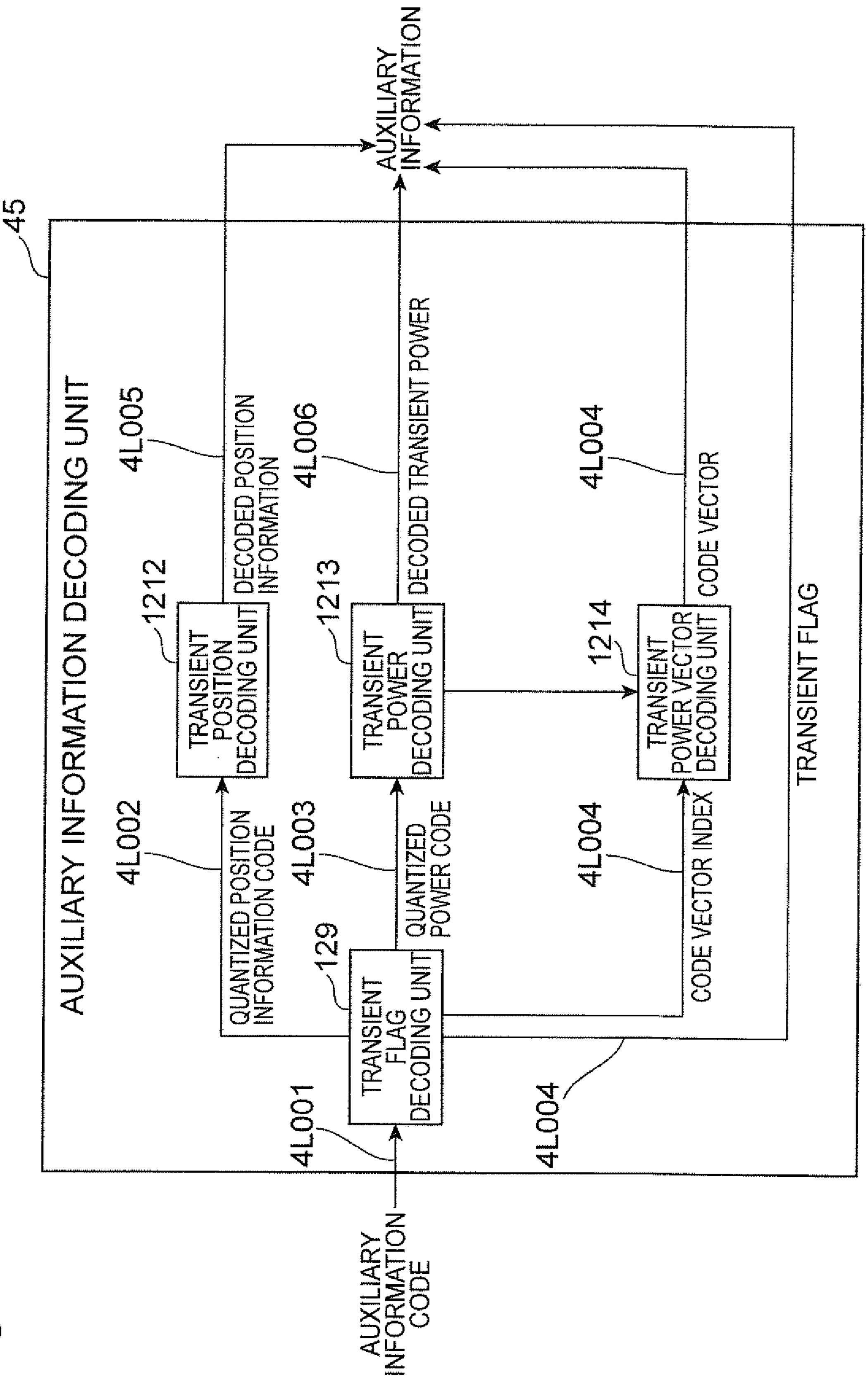

45
AUXILIARY INFORMATION DECODING UNIT
AUXILIARY INFORMATION CODE
4L001
TRANSIENT FLAG DECODING UNIT
129
QUANTIZED POSITION INFORMATION CODE
4L002
QUANTIZED POWER CODE
4L003
CODE VECTOR INDEX
4L004
4L004
TRANSIENT POSITION DECODING UNIT
1212
TRANSIENT POWER DECODING UNIT
1213
TRANSIENT POWER VECTOR DECODING UNIT
1214
DECODED POSITION INFORMATION
4L005
DECODED TRANSIENT POWER
4L006
CODE VECTOR
4L004
TRANSIENT FLAG
AUXILIARY INFORMATION

Fig.31

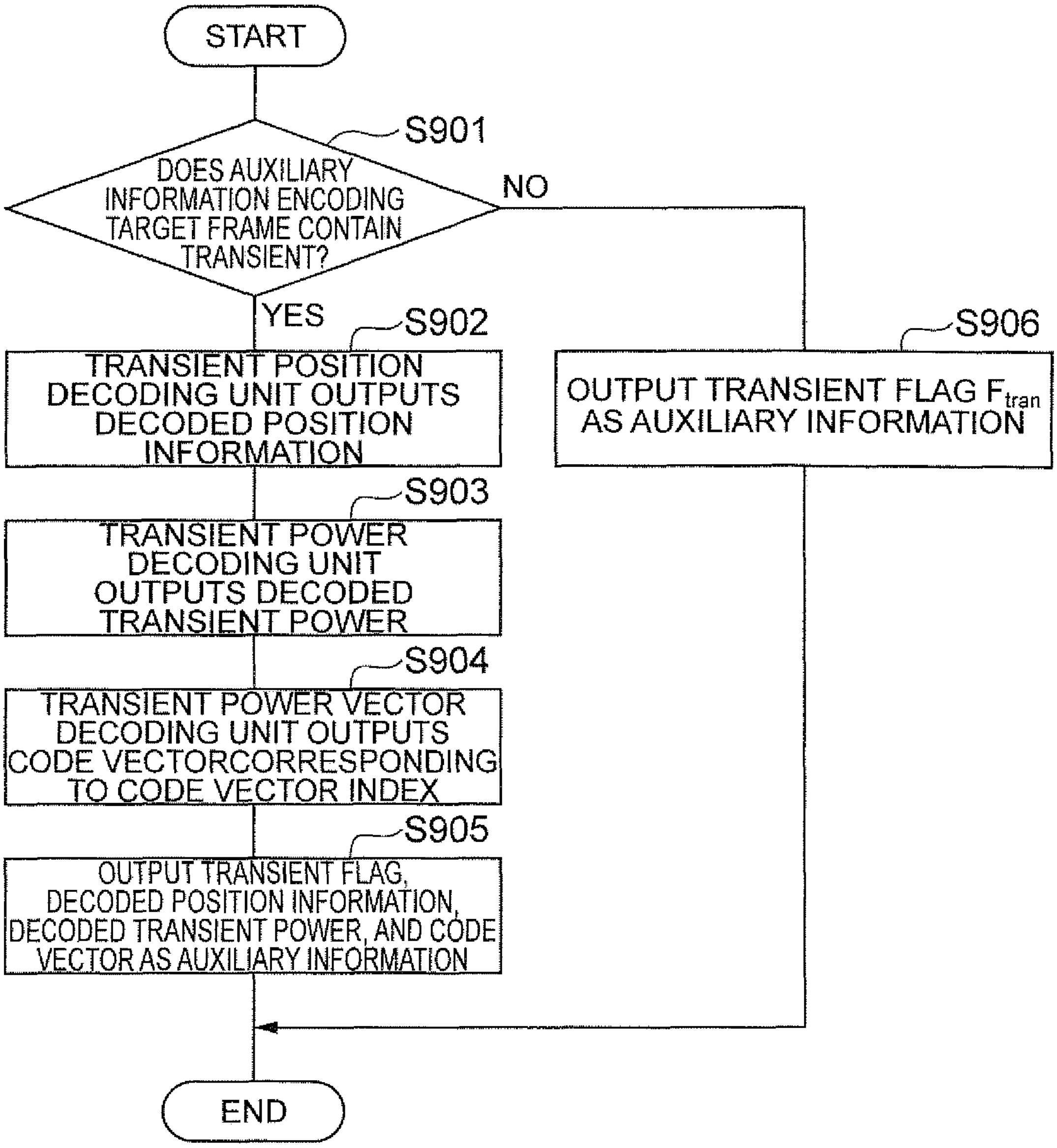

START
S901
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S902
TRANSIENT POSITION DECODING UNIT OUTPUTS DECODED POSITION INFORMATION
S903
TRANSIENT POWER DECODING UNIT OUTPUTS DECODED TRANSIENT POWER
S904
TRANSIENT POWER VECTOR DECODING UNIT OUTPUTS CODE VECTORCORRESPONDING TO CODE VECTOR INDEX
S905
OUTPUT TRANSIENT FLAG, DECODED POSITION INFORMATION, DECODED TRANSIENT POWER, AND CODE VECTOR AS AUXILIARY INFORMATION
S906
OUTPUT TRANSIENT FLAG F_{tran} AS AUXILIARY INFORMATION
END

Fig.32

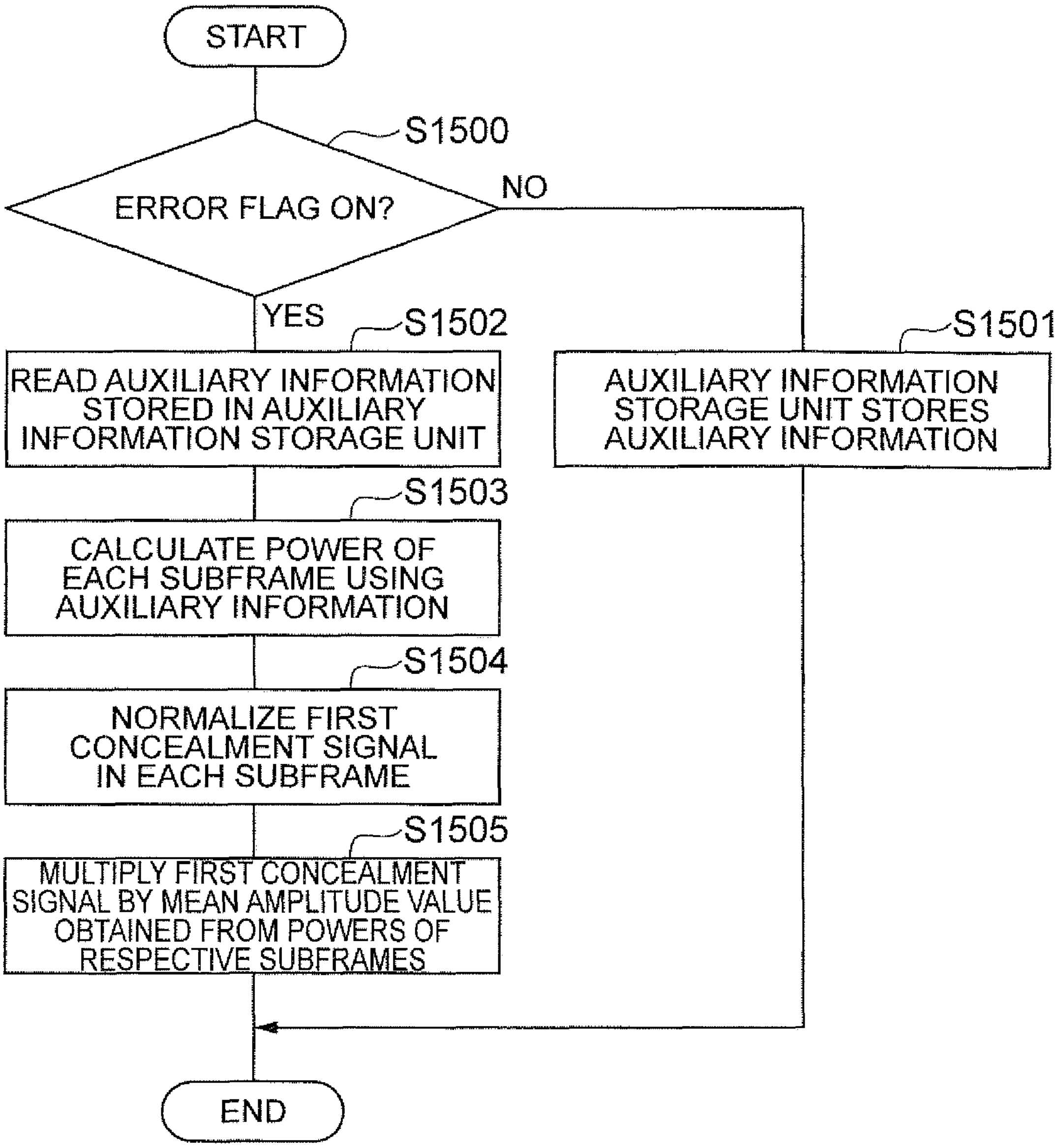

START
S1500
ERROR FLAG ON?
NO
YES
S1502
READ AUXILIARY INFORMATION STORED IN AUXILIARY INFORMATION STORAGE UNIT
S1501
AUXILIARY INFORMATION STORAGE UNIT STORES AUXILIARY INFORMATION
S1503
CALCULATE POWER OF EACH SUBFRAME USING AUXILIARY INFORMATION
S1504
NORMALIZE FIRST CONCEALMENT SIGNAL IN EACH SUBFRAME
S1505
MULTIPLY FIRST CONCEALMENT SIGNAL BY MEAN AMPLITUDE VALUE OBTAINED FROM POWERS OF RESPECTIVE SUBFRAMES
END

Fig.33

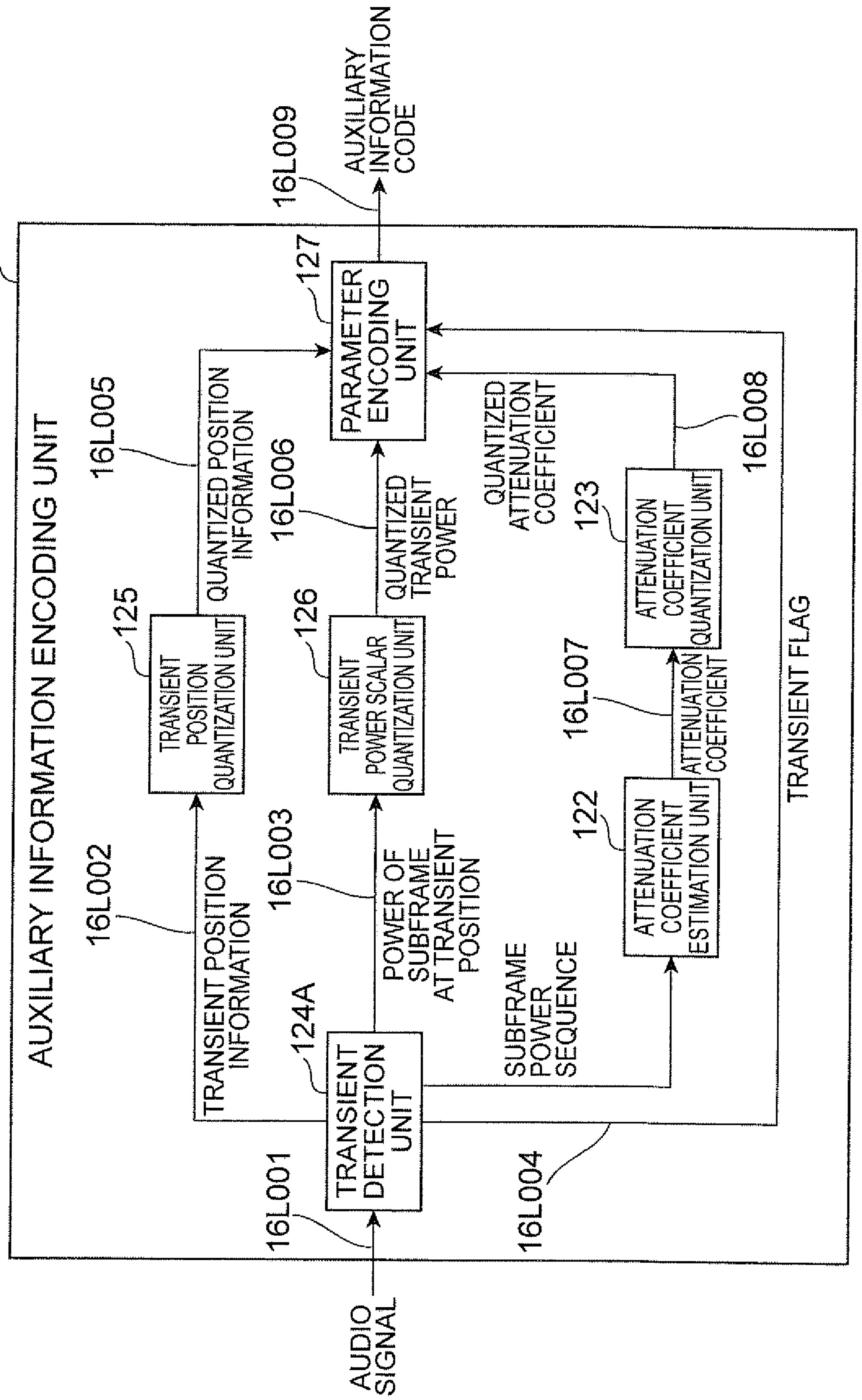

AUXILIARY INFORMATION ENCODING UNIT
12
AUDIO SIGNAL
16L001
TRANSIENT DETECTION UNIT
124A
16L002
TRANSIENT POSITION INFORMATION
TRANSIENT POSITION QUANTIZATION UNIT
125
16L005
QUANTIZED POSITION INFORMATION
16L003
POWER OF SUBFRAME AT TRANSIENT POSITION
TRANSIENT POWER SCALAR QUANTIZATION UNIT
126
16L006
QUANTIZED TRANSIENT POWER
SUBFRAME POWER SEQUENCE
ATTENUATION COEFFICIENT ESTIMATION UNIT
122
16L007
ATTENUATION COEFFICIENT
ATTENUATION COEFFICIENT QUANTIZATION UNIT
123
16L008
QUANTIZED ATTENUATION COEFFICIENT
16L004
TRANSIENT FLAG
PARAMETER ENCODING UNIT
127
16L009
AUXILIARY INFORMATION CODE

Fig.34

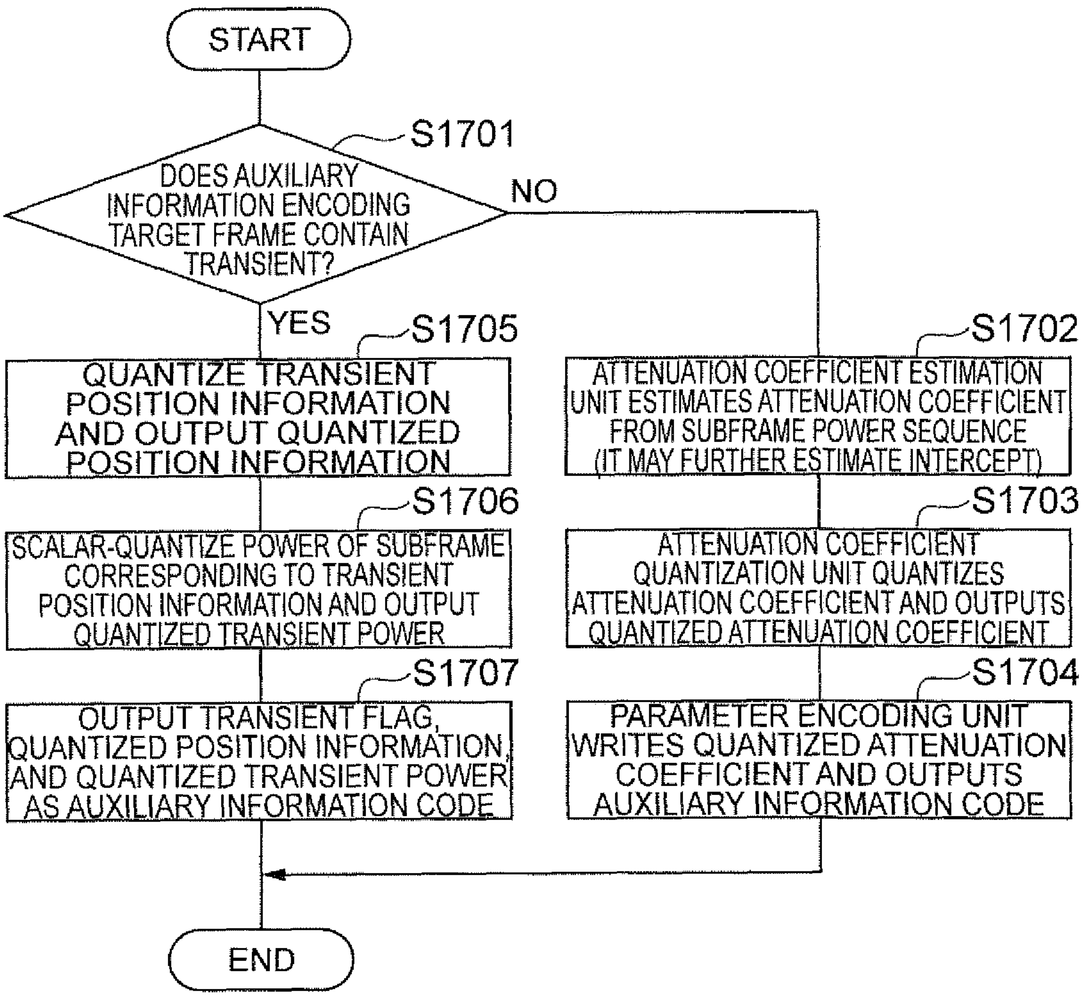

START
S1701
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S1705
QUANTIZE TRANSIENT POSITION INFORMATION AND OUTPUT QUANTIZED POSITION INFORMATION
S1706
SCALAR-QUANTIZE POWER OF SUBFRAME CORRESPONDING TO TRANSIENT POSITION INFORMATION AND OUTPUT QUANTIZED TRANSIENT POWER
S1707
OUTPUT TRANSIENT FLAG, QUANTIZED POSITION INFORMATION, AND QUANTIZED TRANSIENT POWER AS AUXILIARY INFORMATION CODE
S1702
ATTENUATION COEFFICIENT ESTIMATION UNIT ESTIMATES ATTENUATION COEFFICIENT FROM SUBFRAME POWER SEQUENCE (IT MAY FURTHER ESTIMATE INTERCEPT)
S1703
ATTENUATION COEFFICIENT QUANTIZATION UNIT QUANTIZES ATTENUATION COEFFICIENT AND OUTPUTS QUANTIZED ATTENUATION COEFFICIENT
S1704
PARAMETER ENCODING UNIT WRITES QUANTIZED ATTENUATION COEFFICIENT AND OUTPUTS AUXILIARY INFORMATION CODE
END

Fig.35

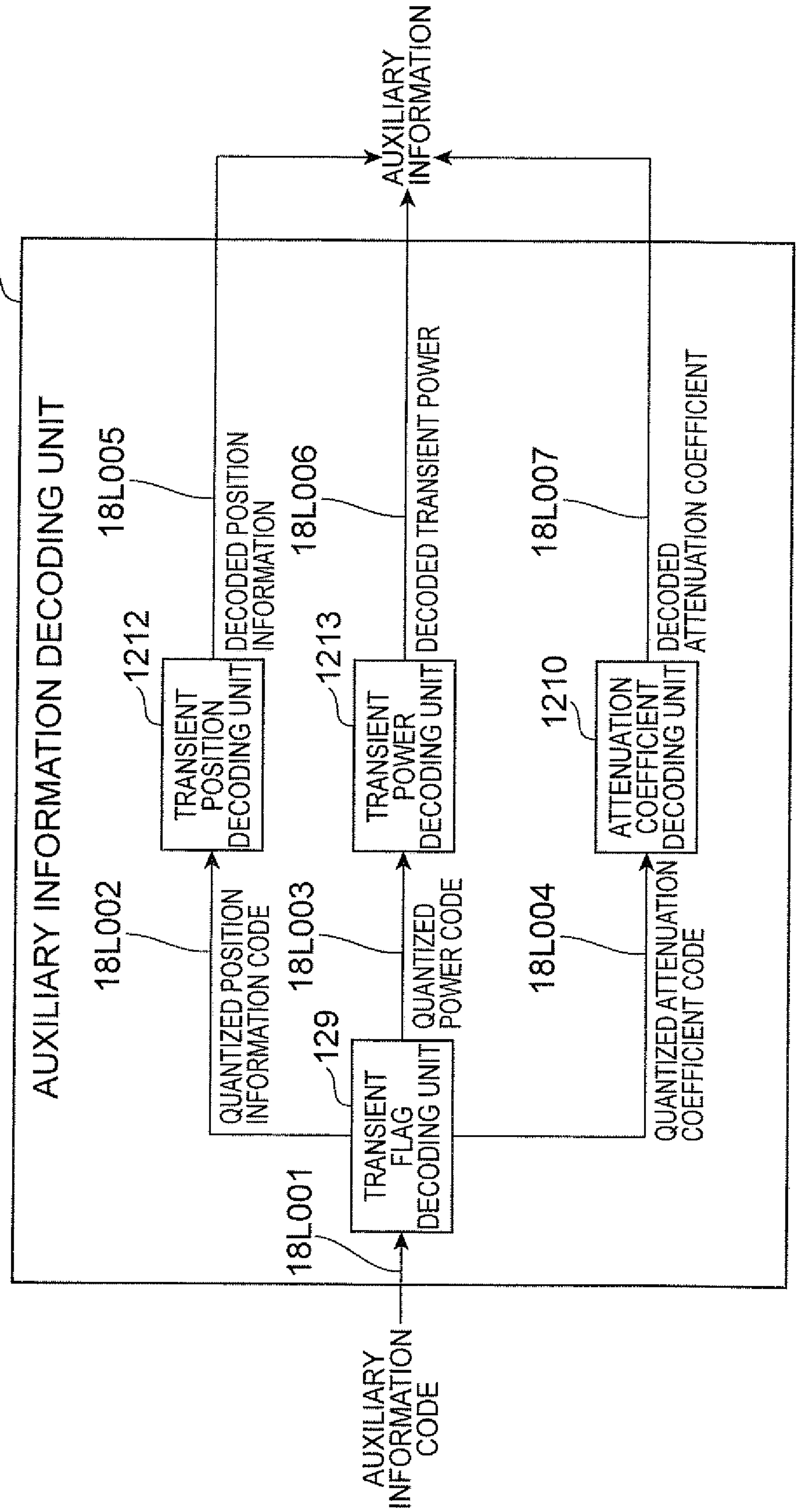

45
AUXILIARY INFORMATION DECODING UNIT
AUXILIARY INFORMATION CODE
18L001
TRANSIENT FLAG DECODING UNIT
129
QUANTIZED POSITION INFORMATION CODE
18L002
QUANTIZED POWER CODE
18L003
QUANTIZED ATTENUATION COEFFICIENT CODE
18L004
TRANSIENT POSITION DECODING UNIT
1212
TRANSIENT POWER DECODING UNIT
1213
ATTENUATION COEFFICIENT DECODING UNIT
1210
DECODED POSITION INFORMATION
18L005
DECODED TRANSIENT POWER
18L006
DECODED ATTENUATION COEFFICIENT
18L007
AUXILIARY INFORMATION

Fig.36

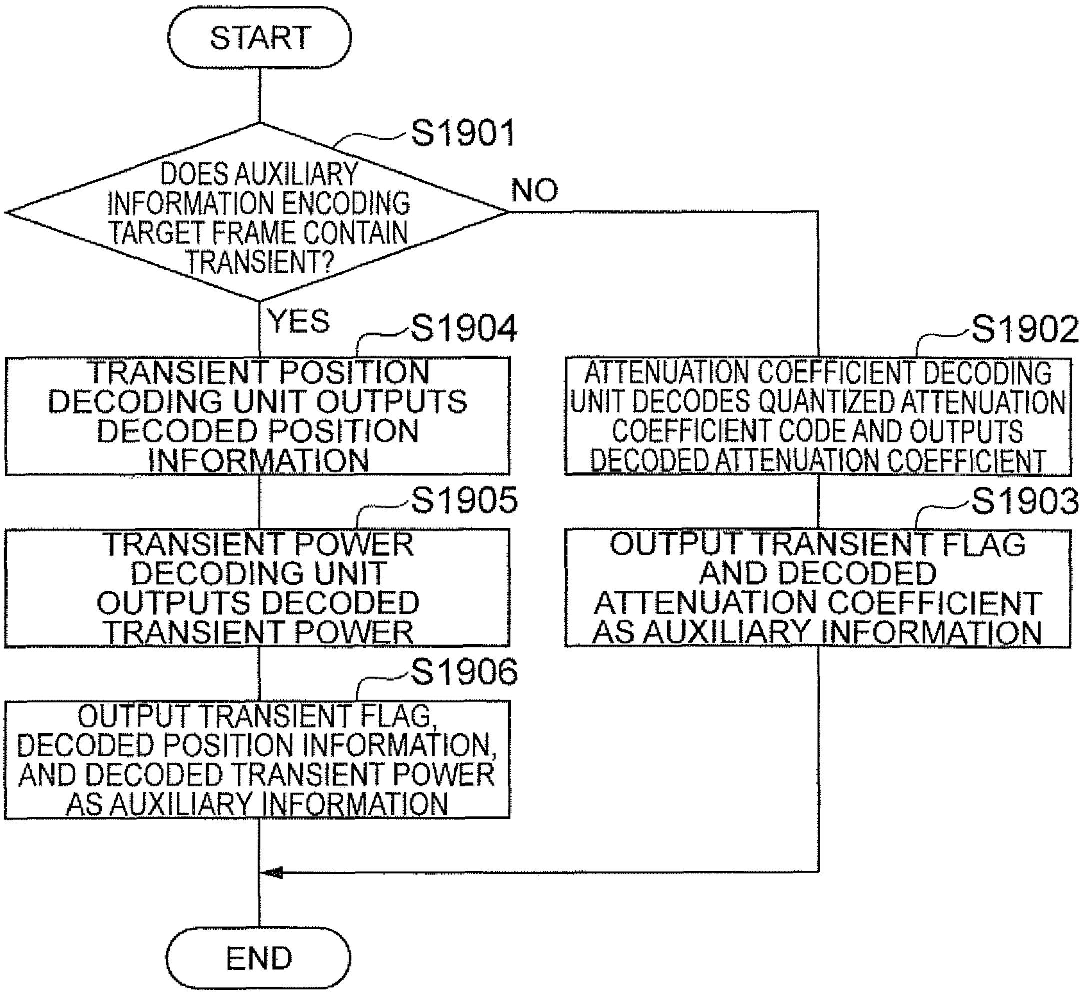

START
S1901
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S1904
TRANSIENT POSITION DECODING UNIT OUTPUTS DECODED POSITION INFORMATION
S1905
TRANSIENT POWER DECODING UNIT OUTPUTS DECODED TRANSIENT POWER
S1906
OUTPUT TRANSIENT FLAG, DECODED POSITION INFORMATION, AND DECODED TRANSIENT POWER AS AUXILIARY INFORMATION
S1902
ATTENUATION COEFFICIENT DECODING UNIT DECODES QUANTIZED ATTENUATION COEFFICIENT CODE AND OUTPUTS DECODED ATTENUATION COEFFICIENT
S1903
OUTPUT TRANSIENT FLAG AND DECODED ATTENUATION COEFFICIENT AS AUXILIARY INFORMATION
END

Fig.37

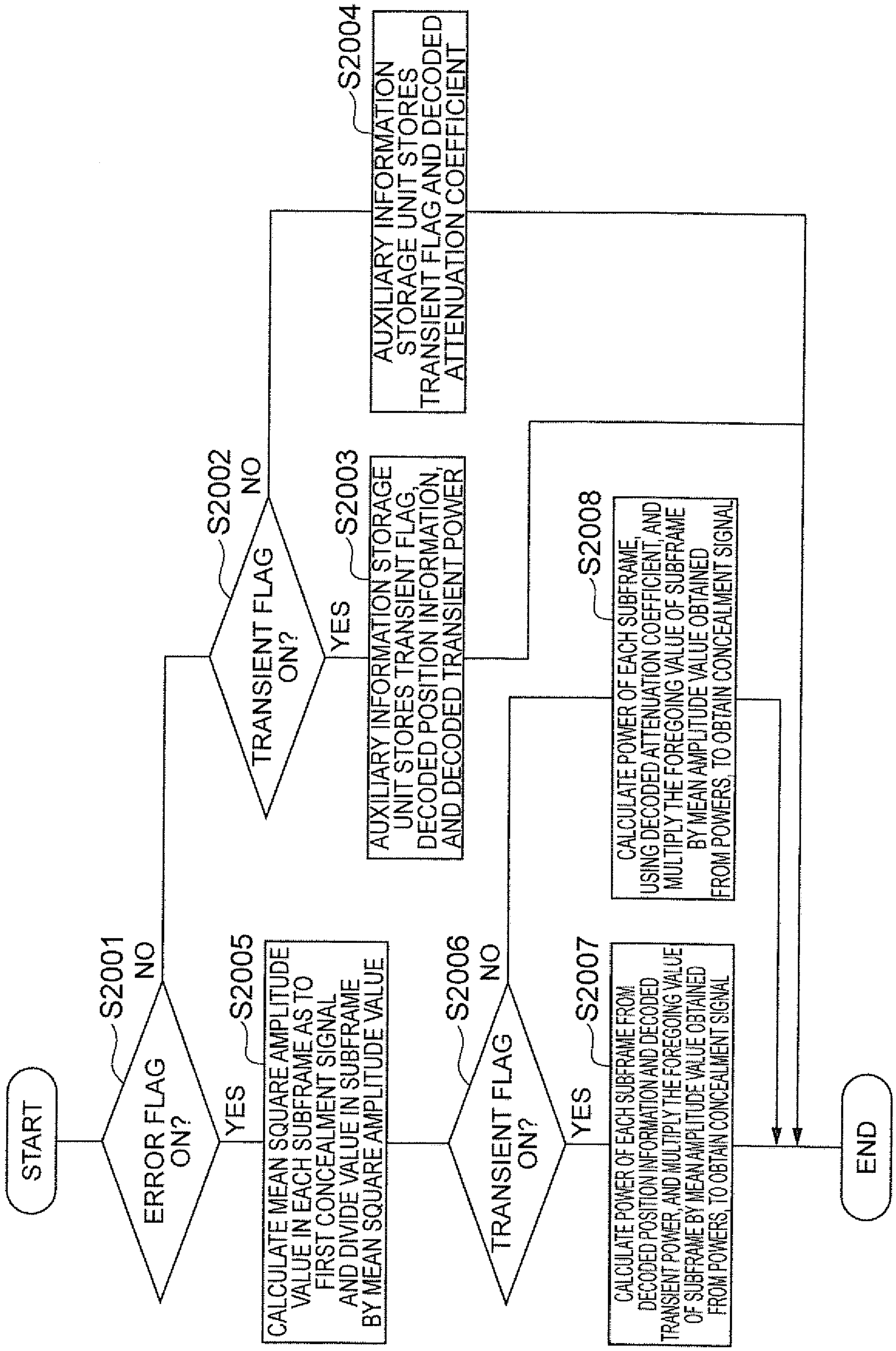

START
S2001
ERROR FLAG ON?
YES
NO
S2002
TRANSIENT FLAG ON?
YES
NO
S2003
AUXILIARY INFORMATION STORAGE UNIT STORES TRANSIENT FLAG, DECODED POSITION INFORMATION, AND DECODED TRANSIENT POWER
S2004
AUXILIARY INFORMATION STORAGE UNIT STORES TRANSIENT FLAG AND DECODED ATTENUATION COEFFICIENT
S2005
CALCULATE MEAN SQUARE AMPLITUDE VALUE IN EACH SUBFRAME AS TO FIRST CONCEALMENT SIGNAL AND DIVIDE VALUE IN SUBFRAME BY MEAN SQUARE AMPLITUDE VALUE
S2006
TRANSIENT FLAG ON?
YES
NO
S2007
CALCULATE POWER OF EACH SUBFRAME FROM DECODED POSITION INFORMATION AND DECODED TRANSIENT POWER, AND MULTIPLY THE FOREGOING VALUE OF SUBFRAME BY MEAN AMPLITUDE VALUE OBTAINED FROM POWERS, TO OBTAIN CONCEALMENT SIGNAL
S2008
CALCULATE POWER OF EACH SUBFRAME, USING DECODED ATTENUATION COEFFICIENT, AND MULTIPLY THE FOREGOING VALUE OF SUBFRAME BY MEAN AMPLITUDE VALUE OBTAINED FROM POWERS, TO OBTAIN CONCEALMENT SIGNAL
END

Fig.38

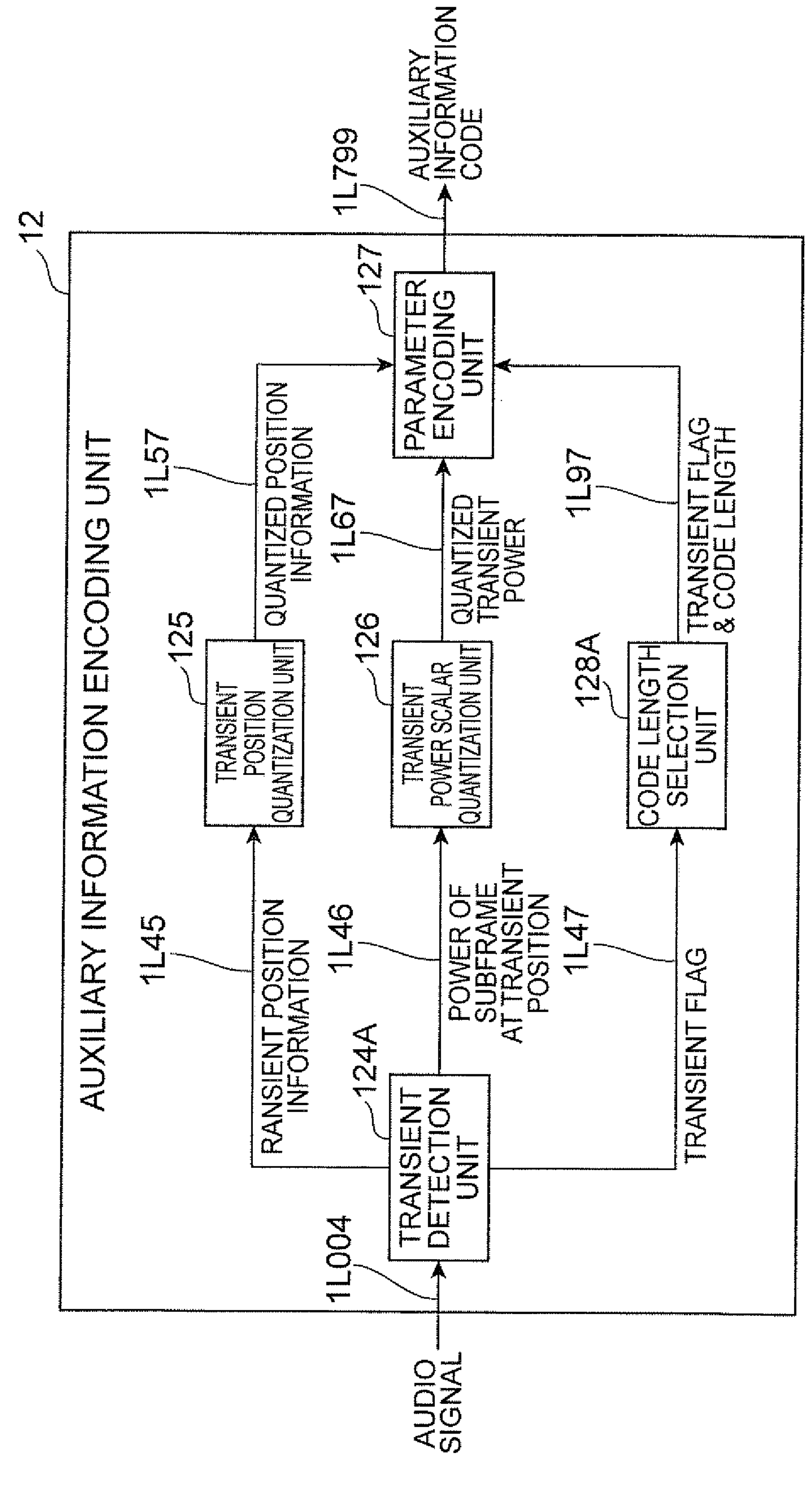

12
AUXILIARY INFORMATION ENCODING UNIT
AUDIO SIGNAL
1L004
TRANSIENT DETECTION UNIT
124A
RANSIENT POSITION INFORMATION
1L45
POWER OF SUBFRAME AT TRANSIENT POSITION
1L46
TRANSIENT FLAG
1L47
TRANSIENT POSITION QUANTIZATION UNIT
125
TRANSIENT POWER SCALAR QUANTIZATION UNIT
126
CODE LENGTH SELECTION UNIT
128A
QUANTIZED POSITION INFORMATION
1L57
QUANTIZED TRANSIENT POWER
1L67
TRANSIENT FLAG & CODE LENGTH
1L97
PARAMETER ENCODING UNIT
127
1L799
AUXILIARY INFORMATION CODE

Fig.39

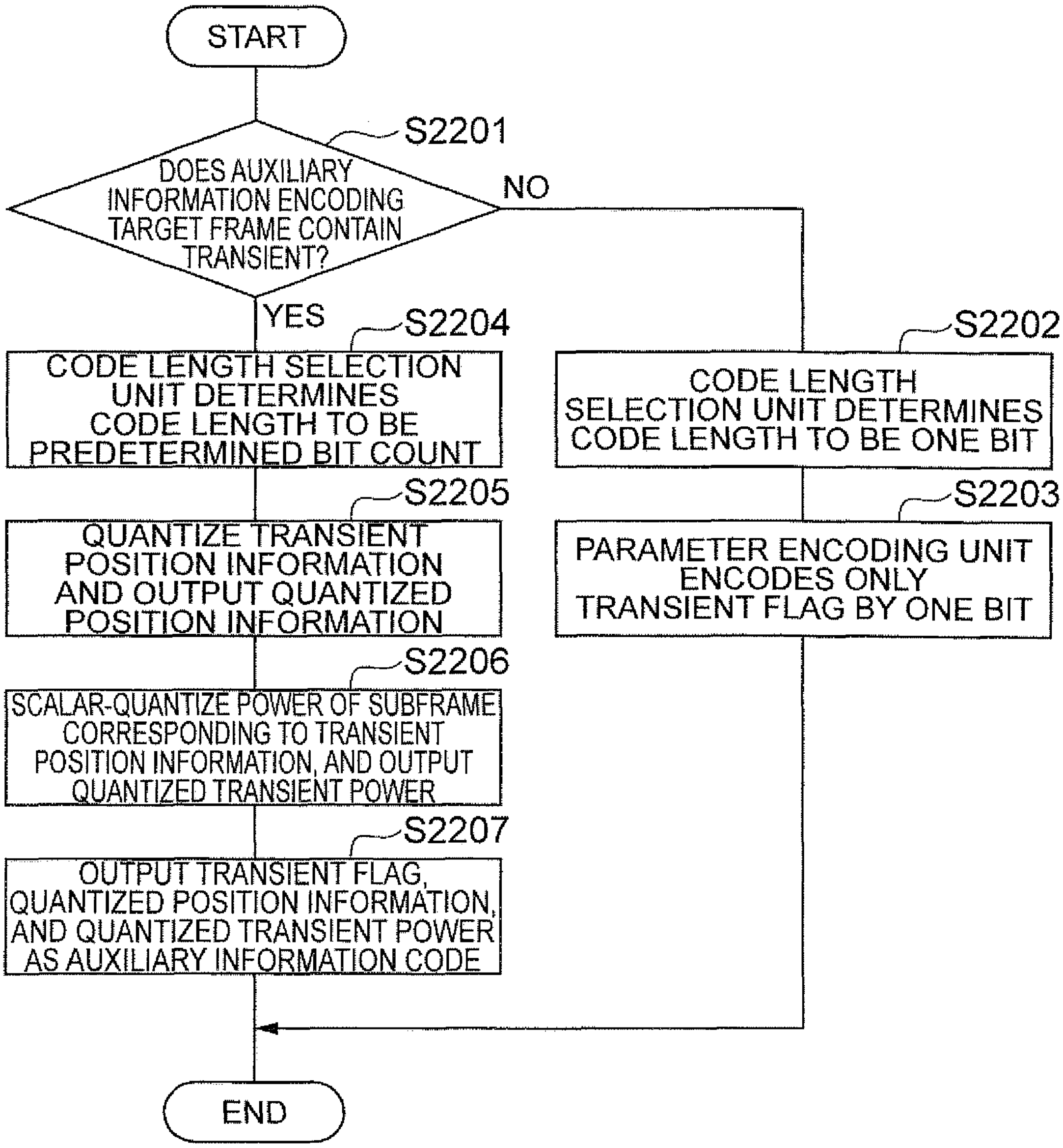

START
S2201
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S2204
CODE LENGTH SELECTION UNIT DETERMINES CODE LENGTH TO BE PREDETERMINED BIT COUNT
S2205
QUANTIZE TRANSIENT POSITION INFORMATION AND OUTPUT QUANTIZED POSITION INFORMATION
S2206
SCALAR-QUANTIZE POWER OF SUBFRAME CORRESPONDING TO TRANSIENT POSITION INFORMATION, AND OUTPUT QUANTIZED TRANSIENT POWER
S2207
OUTPUT TRANSIENT FLAG, QUANTIZED POSITION INFORMATION, AND QUANTIZED TRANSIENT POWER AS AUXILIARY INFORMATION CODE
S2202
CODE LENGTH SELECTION UNIT DETERMINES CODE LENGTH TO BE ONE BIT
S2203
PARAMETER ENCODING UNIT ENCODES ONLY TRANSIENT FLAG BY ONE BIT
END

Fig.40

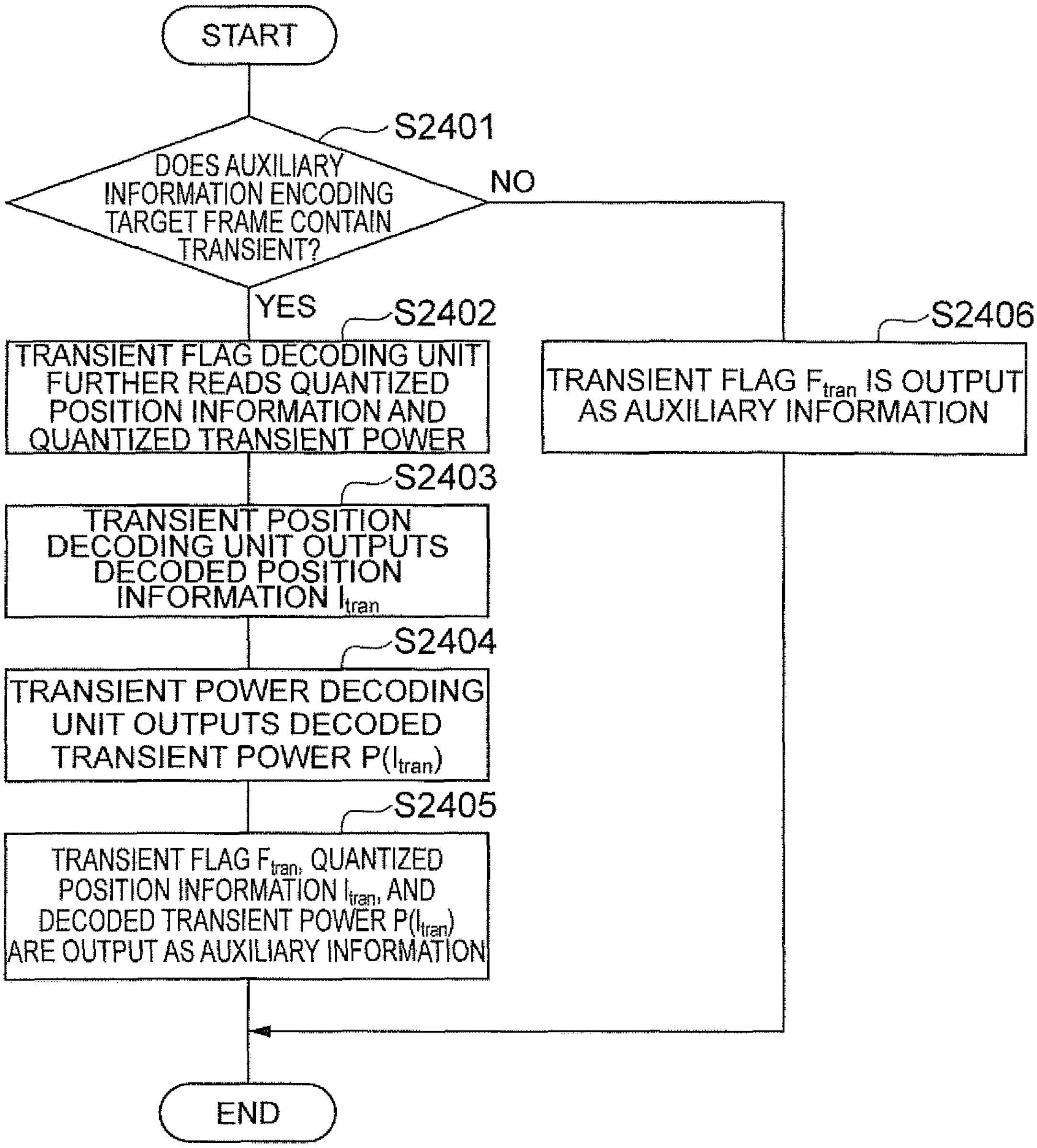

START
S2401
DOES AUXILIARY INFORMATION ENCODING TARGET FRAME CONTAIN TRANSIENT?
NO
YES
S2402
TRANSIENT FLAG DECODING UNIT FURTHER READS QUANTIZED POSITION INFORMATION AND QUANTIZED TRANSIENT POWER
S2403
TRANSIENT POSITION DECODING UNIT OUTPUTS DECODED POSITION INFORMATION I_{tran}
S2404
TRANSIENT POWER DECODING UNIT OUTPUTS DECODED TRANSIENT POWER $P(I_{tran})$
S2405
TRANSIENT FLAG F_{tran}, QUANTIZED POSITION INFORMATION I_{tran}, AND DECODED TRANSIENT POWER $P(I_{tran})$ ARE OUTPUT AS AUXILIARY INFORMATION
S2406
TRANSIENT FLAG F_{tran} IS OUTPUT AS AUXILIARY INFORMATION
END

Fig.41

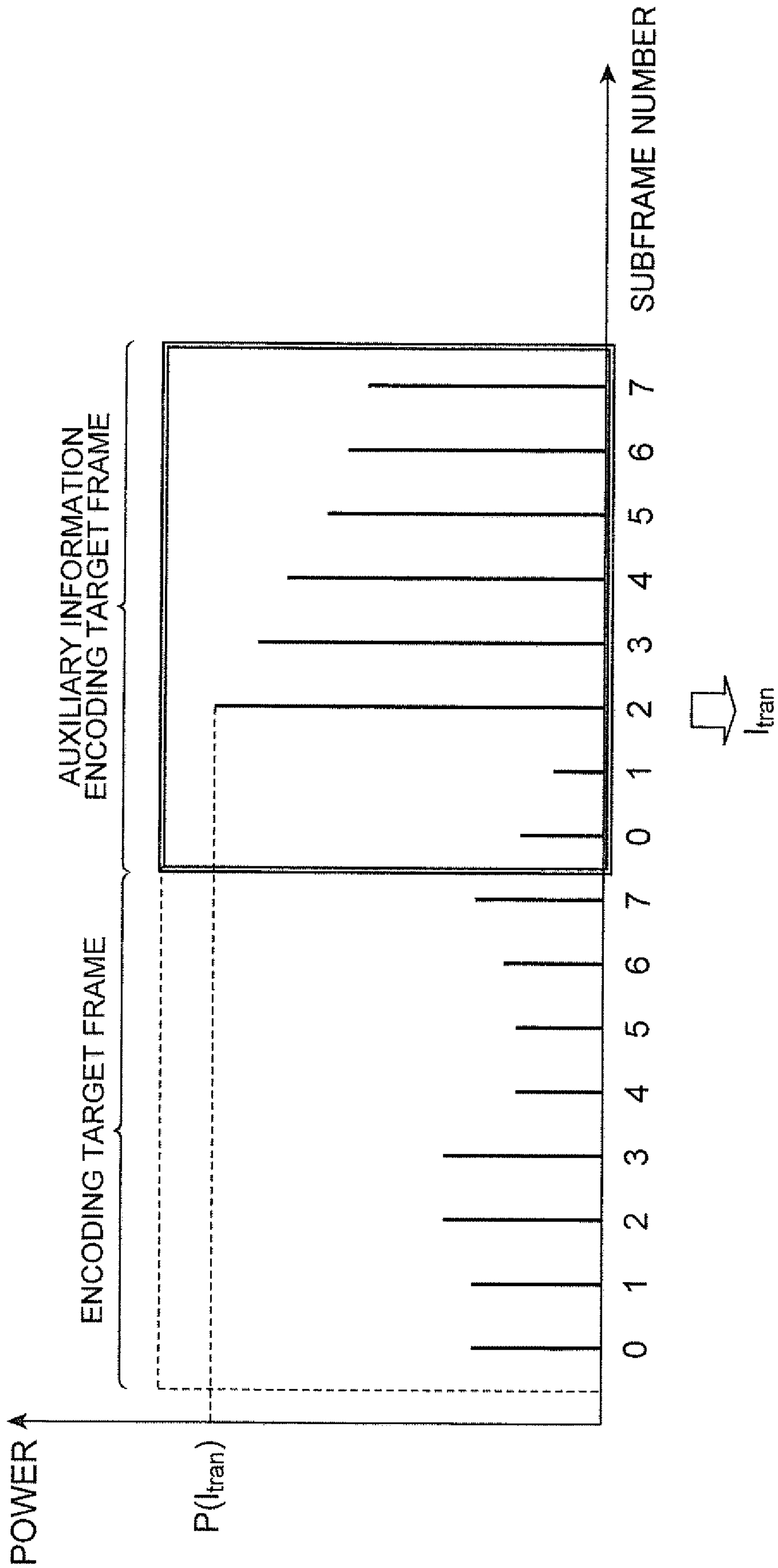

POWER
P(I_tran)
ENCODING TARGET FRAME
AUXILIARY INFORMATION ENCODING TARGET FRAME
0
1
2
3
4
5
6
7
0
1
2
3
4
5
6
7
I_tran
SUBFRAME NUMBER

Fig.42

(a)

BINARY CODING

I_{tran}	QUANTIZED POSITION INFORMATION
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

(b)

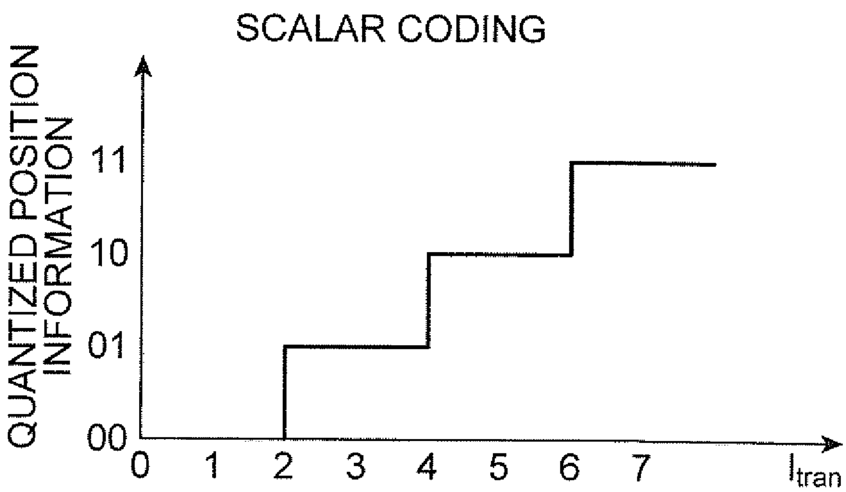


Fig.43

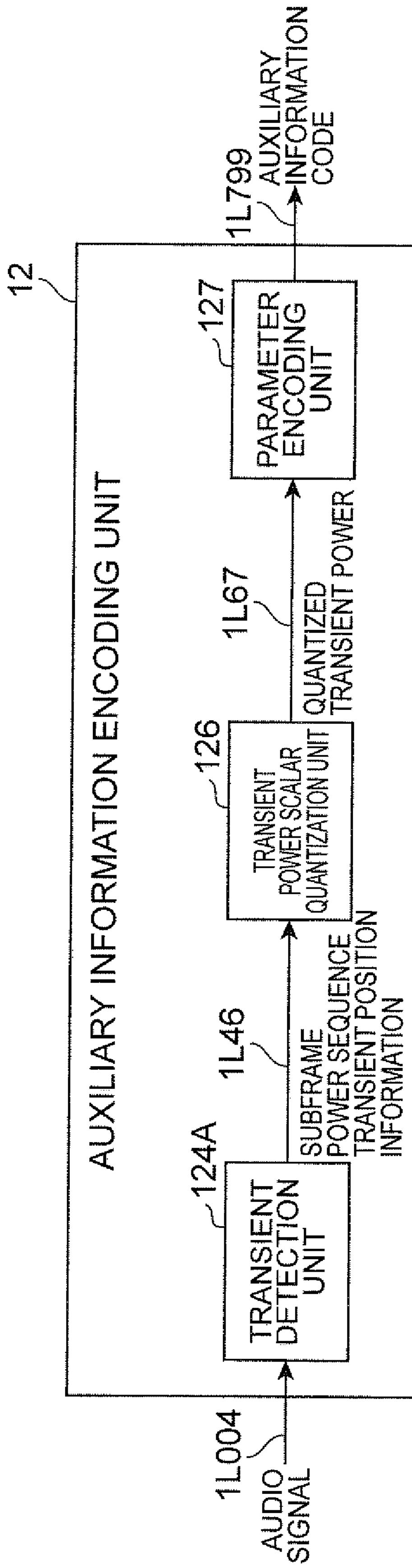

AUXILIARY INFORMATION ENCODING UNIT
12
AUDIO SIGNAL
1L004
TRANSIENT DETECTION UNIT
124A
1L46
SUBFRAME POWER SEQUENCE TRANSIENT POSITION INFORMATION
TRANSIENT POWER SCALAR QUANTIZATION UNIT
126
1L67
QUANTIZED TRANSIENT POWER
PARAMETER ENCODING UNIT
127
1L799
AUXILIARY INFORMATION CODE

Fig.44

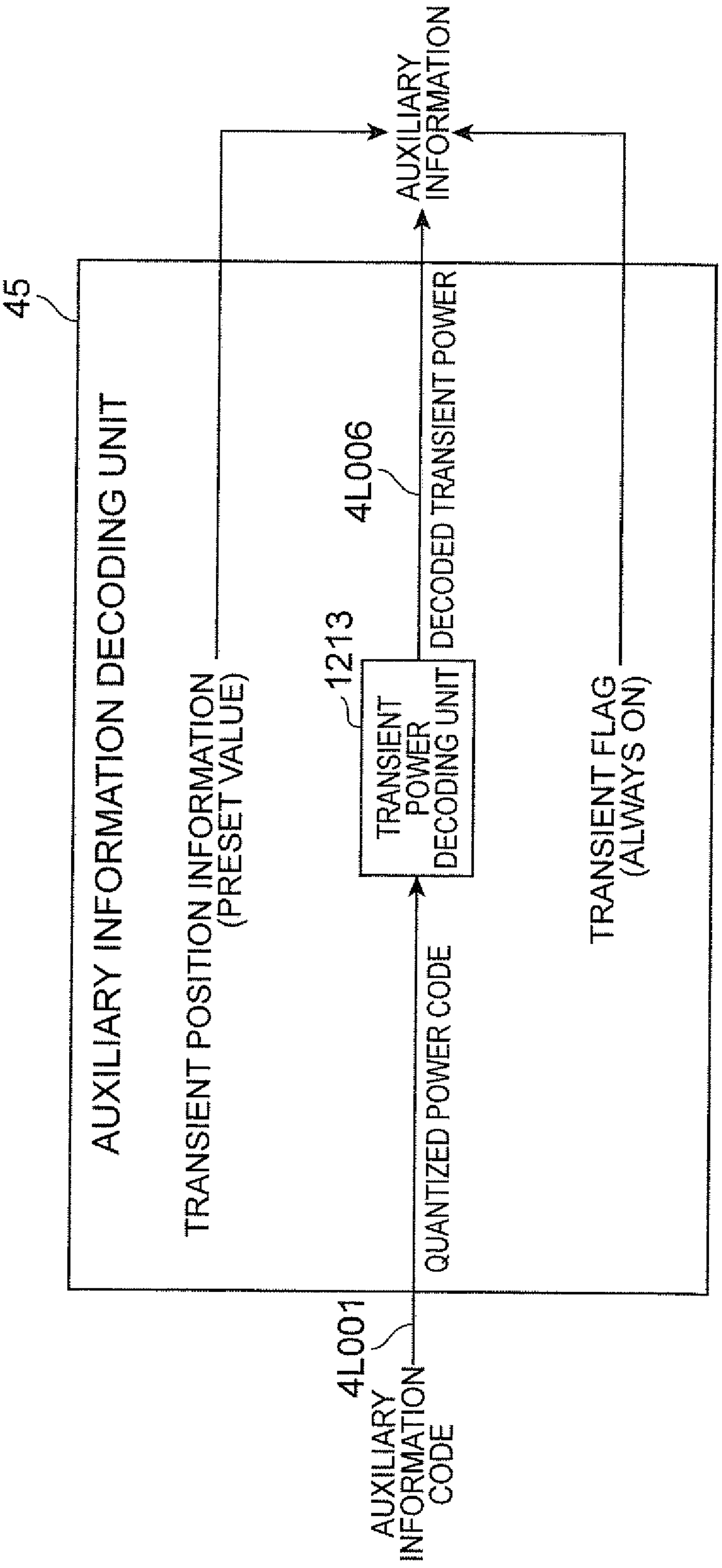

45
AUXILIARY INFORMATION DECODING UNIT
TRANSIENT POSITION INFORMATION
(PRESET VALUE)
1213
TRANSIENT
POWER
DECODING UNIT
4L006
DECODED TRANSIENT POWER
TRANSIENT FLAG
(ALWAYS ON)
4L001
AUXILIARY
INFORMATION
CODE
QUANTIZED POWER CODE
AUXILIARY
INFORMATION

Fig.45

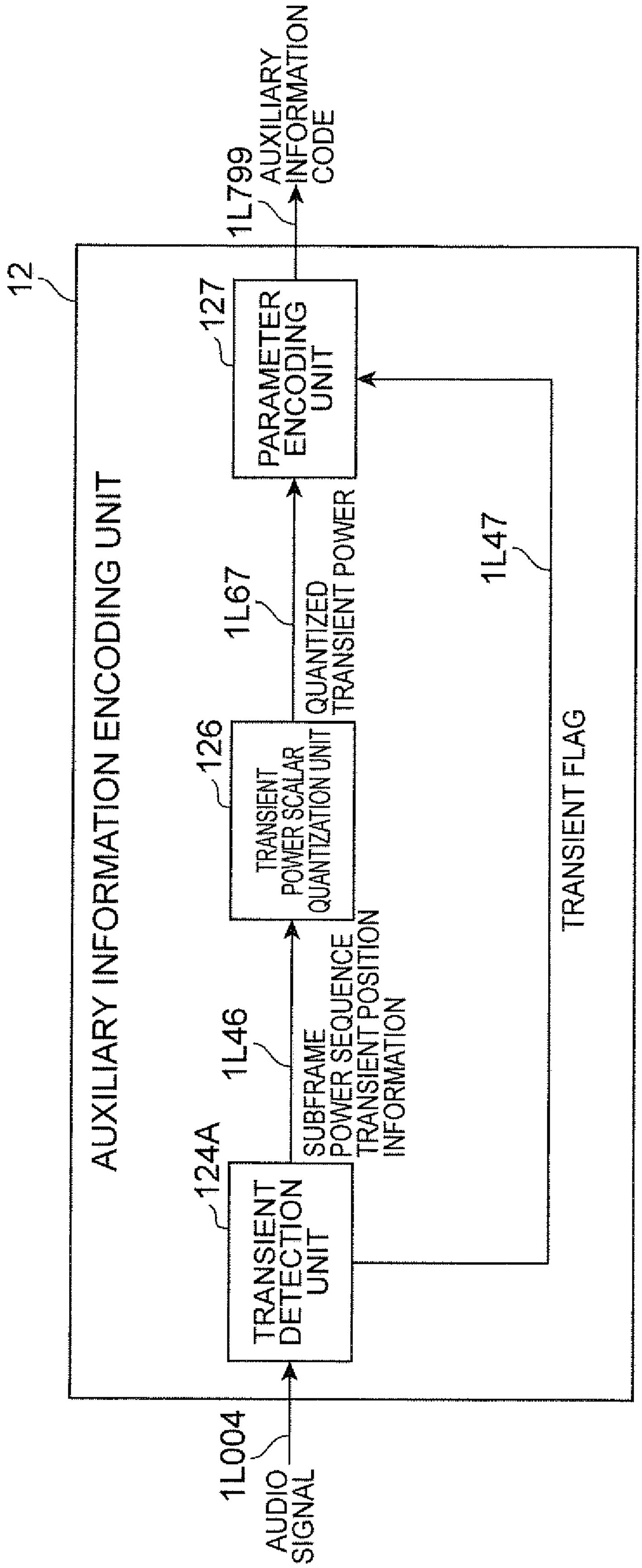

12
AUXILIARY INFORMATION ENCODING UNIT
AUDIO SIGNAL
1L004
TRANSIENT DETECTION UNIT
124A
1L46
SUBFRAME POWER SEQUENCE
TRANSIENT POSITION INFORMATION
TRANSIENT POWER SCALAR QUANTIZATION UNIT
126
1L67
QUANTIZED TRANSIENT POWER
PARAMETER ENCODING UNIT
127
1L799
AUXILIARY INFORMATION CODE
TRANSIENT FLAG
1L47

Fig.46

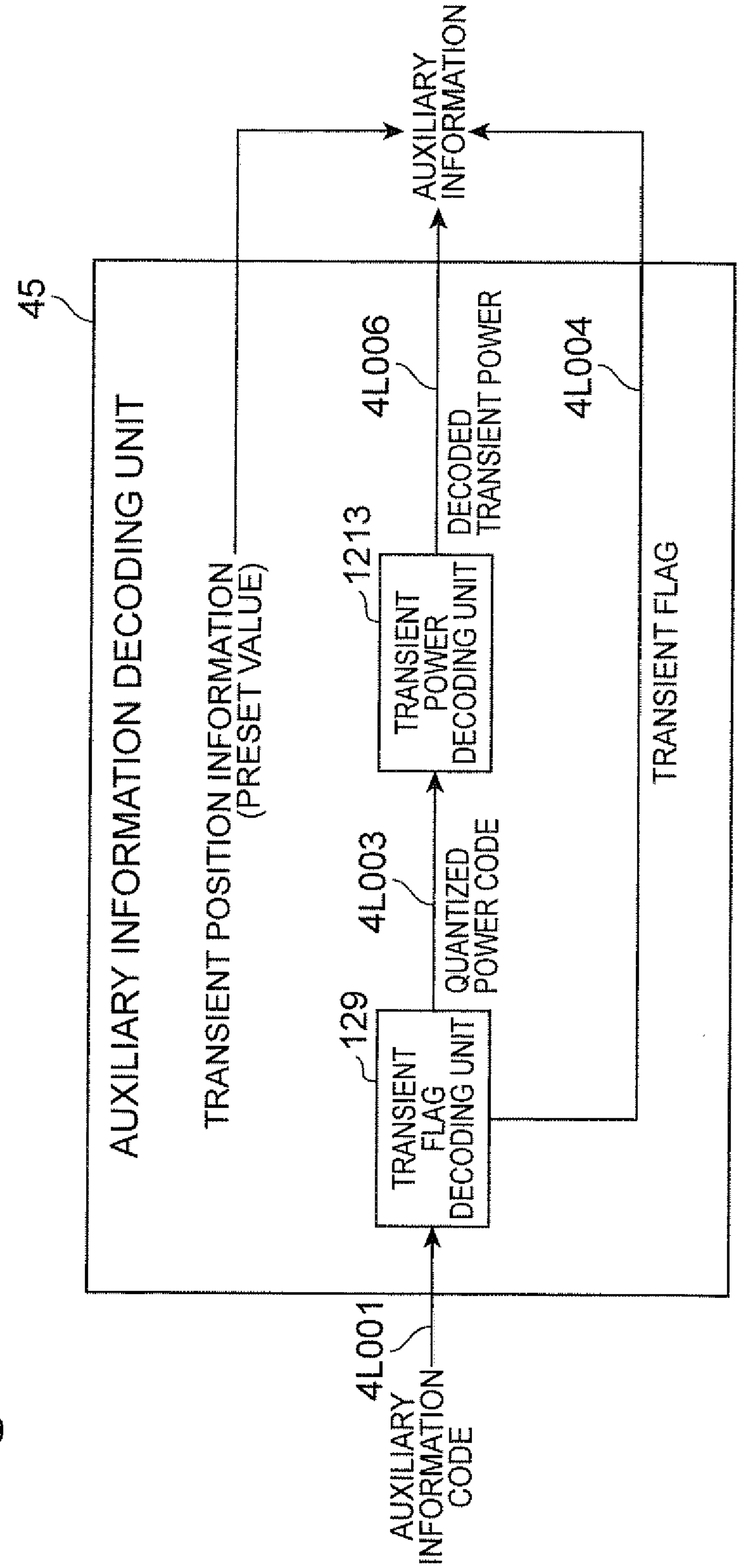

45
AUXILIARY INFORMATION DECODING UNIT
TRANSIENT POSITION INFORMATION (PRESET VALUE)
AUXILIARY INFORMATION CODE
4L001
129
TRANSIENT FLAG DECODING UNIT
4L003
QUANTIZED POWER CODE
1213
TRANSIENT POWER DECODING UNIT
DECODED TRANSIENT POWER
4L006
4L004
TRANSIENT FLAG
AUXILIARY INFORMATION

Fig.47

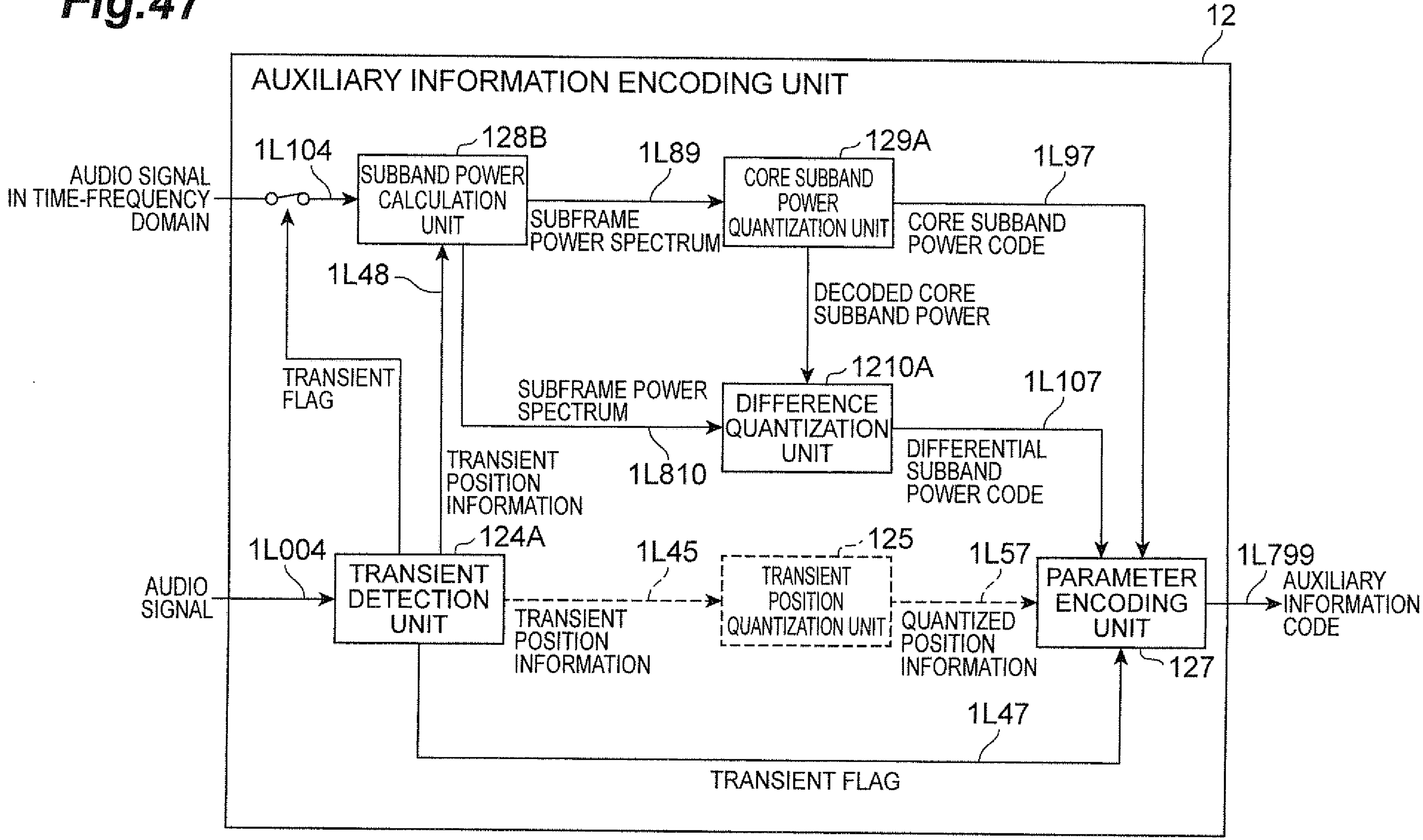

12
AUXILIARY INFORMATION ENCODING UNIT
AUDIO SIGNAL IN TIME-FREQUENCY DOMAIN
1L104
128B
SUBBAND POWER CALCULATION UNIT
1L89
SUBFRAME POWER SPECTRUM
129A
CORE SUBBAND POWER QUANTIZATION UNIT
1L97
CORE SUBBAND POWER CODE
1L48
DECODED CORE SUBBAND POWER
TRANSIENT FLAG
SUBFRAME POWER SPECTRUM
1210A
DIFFERENCE QUANTIZATION UNIT
1L107
1L810
DIFFERENTIAL SUBBAND POWER CODE
TRANSIENT POSITION INFORMATION
1L004
124A
AUDIO SIGNAL
TRANSIENT DETECTION UNIT
1L45
TRANSIENT POSITION INFORMATION
125
TRANSIENT POSITION QUANTIZATION UNIT
1L57
QUANTIZED POSITION INFORMATION
PARAMETER ENCODING UNIT
127
1L799
AUXILIARY INFORMATION CODE
1L47
TRANSIENT FLAG

Fig.48

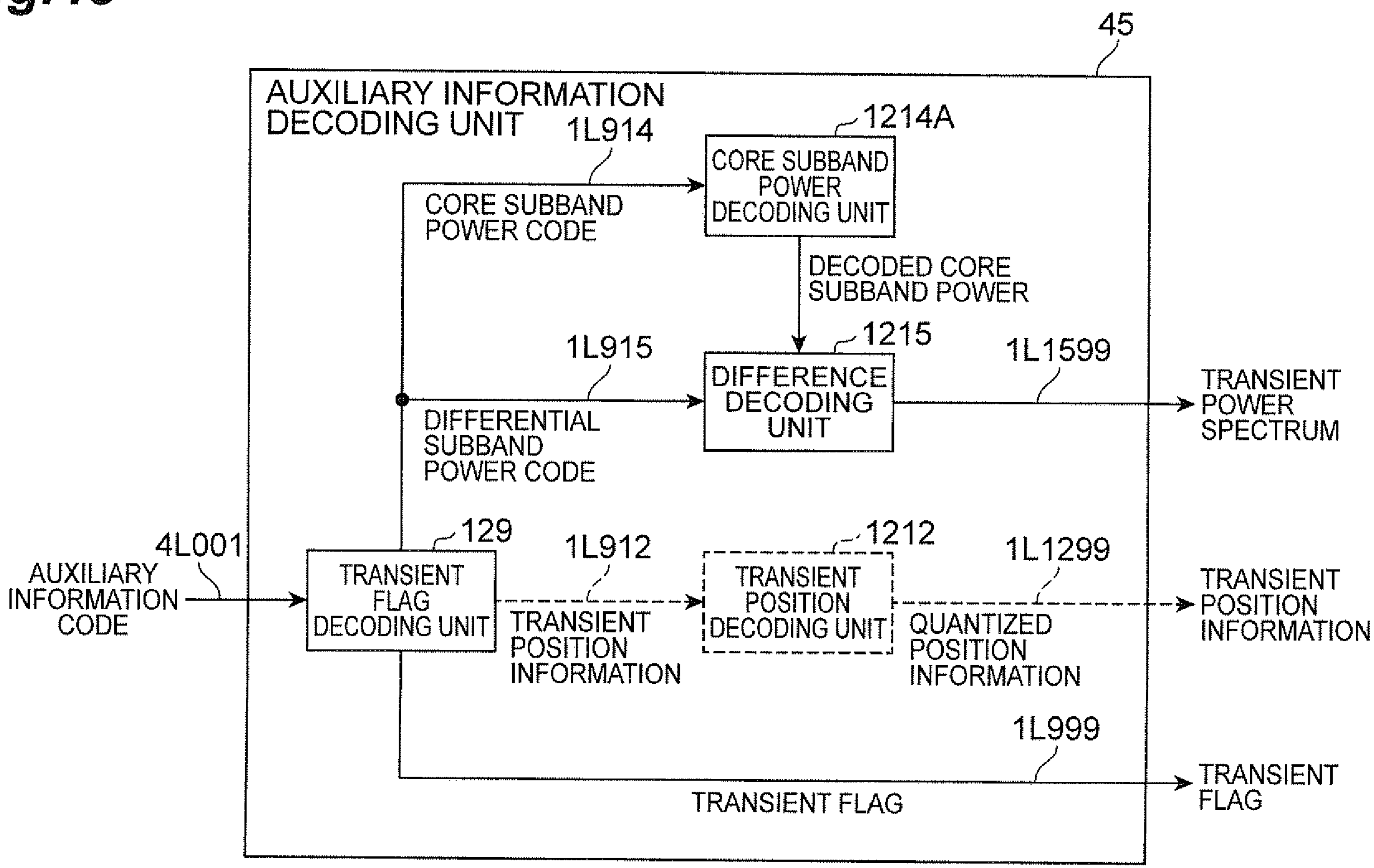
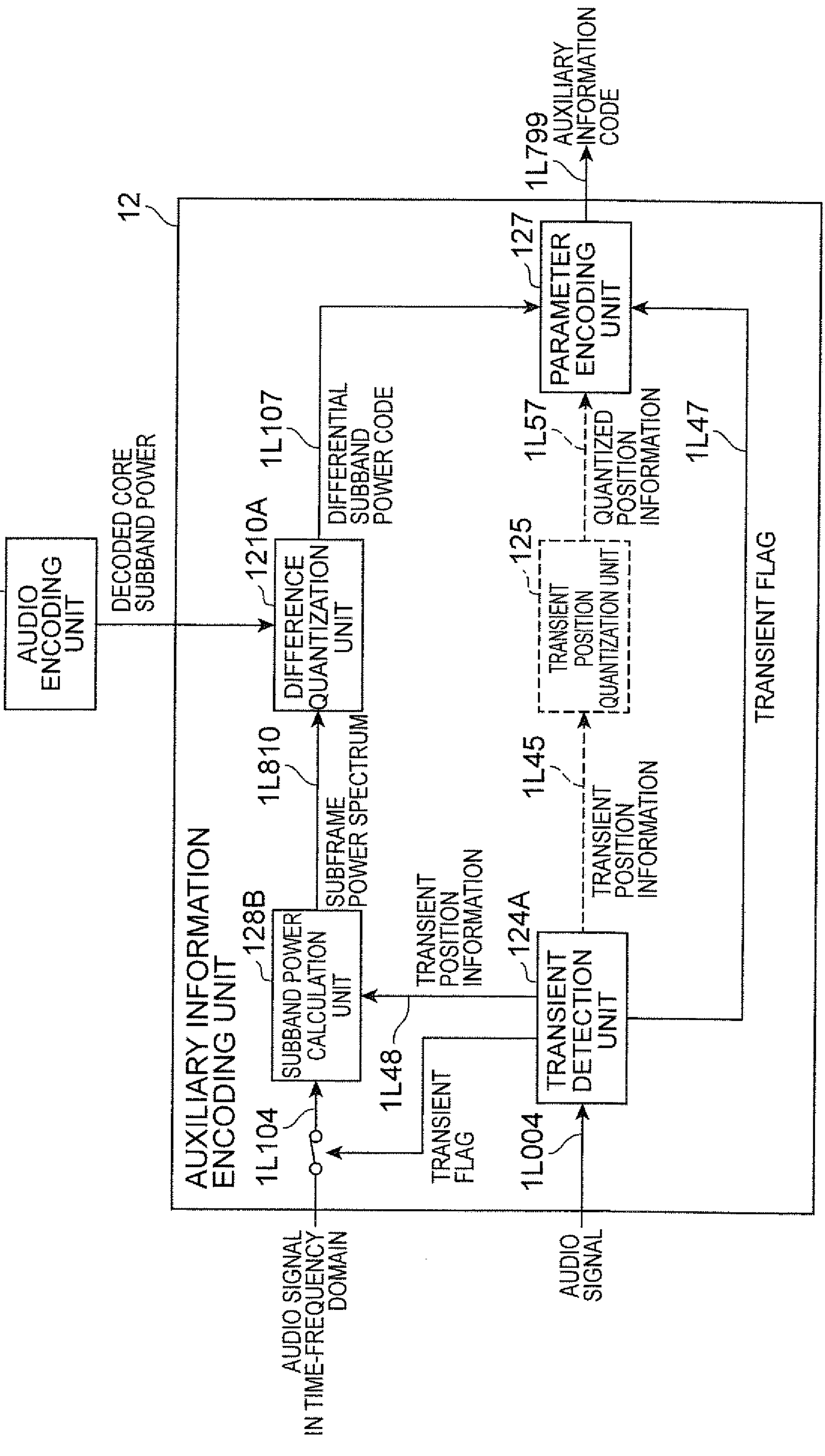
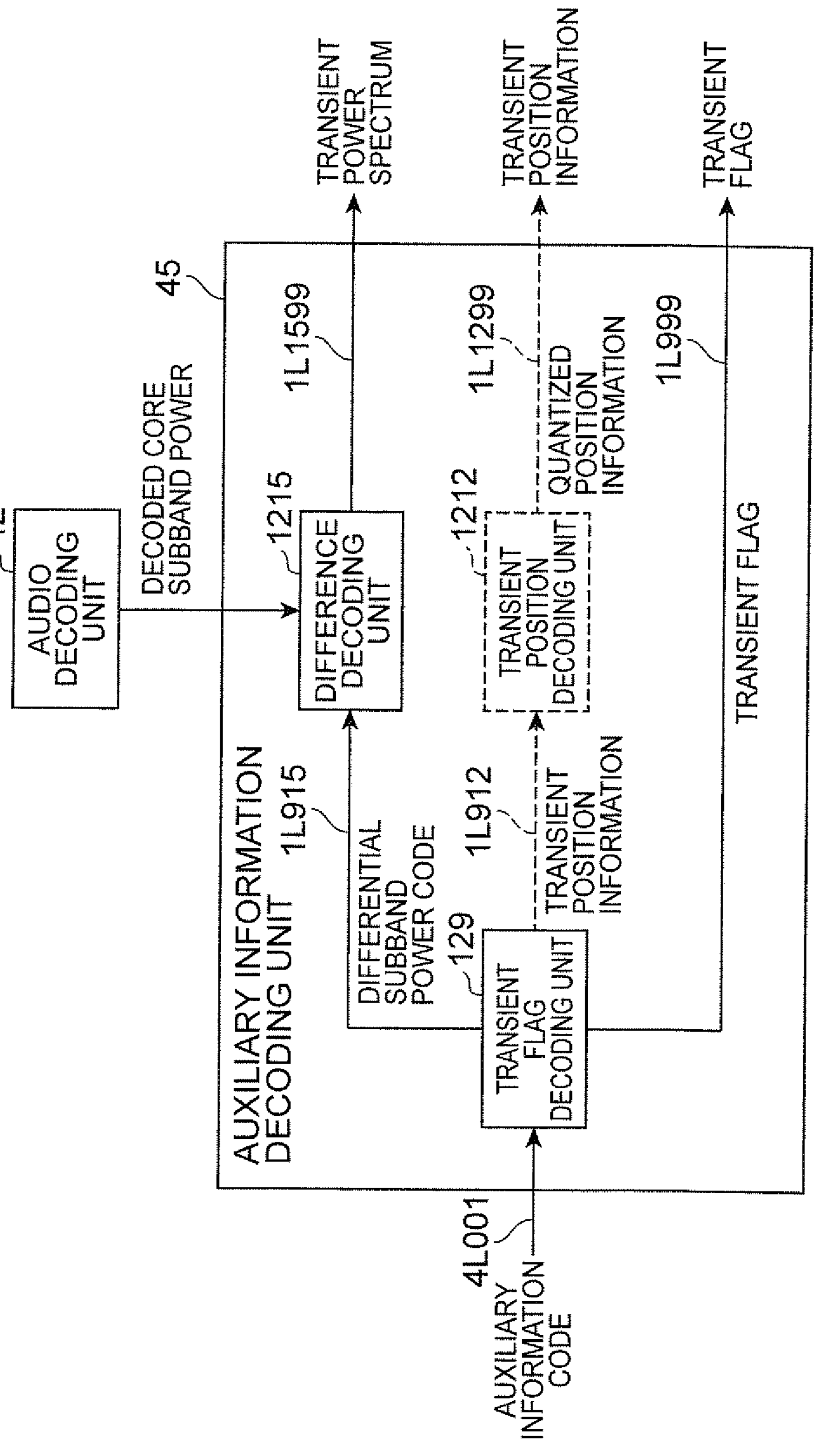

45
AUXILIARY INFORMATION DECODING UNIT
1L914
1214A
CORE SUBBAND POWER DECODING UNIT
CORE SUBBAND POWER CODE
DECODED CORE SUBBAND POWER
1L915
1215
1L1599
DIFFERENCE DECODING UNIT
TRANSIENT POWER SPECTRUM
DIFFERENTIAL SUBBAND POWER CODE
4L001
129
1L912
1212
1L1299
AUXILIARY INFORMATION CODE
TRANSIENT FLAG DECODING UNIT
TRANSIENT POSITION DECODING UNIT
TRANSIENT POSITION INFORMATION
TRANSIENT POSITION INFORMATION
QUANTIZED POSITION INFORMATION
1L999
TRANSIENT FLAG
TRANSIENT FLAG

Fig.49



AUDIO ENCODING UNIT
11
DECODED CORE SUBBAND POWER
12
AUXILIARY INFORMATION ENCODING UNIT
1210A
DIFFERENCE QUANTIZATION UNIT
1L107
DIFFERENTIAL SUBBAND POWER CODE
127
PARAMETER ENCODING UNIT
1L799
AUXILIARY INFORMATION CODE
128B
SUBBAND POWER CALCULATION UNIT
SUBFRAME POWER SPECTRUM
1L810
1L104
AUDIO SIGNAL IN TIME-FREQUENCY DOMAIN
TRANSIENT FLAG
1L48
TRANSIENT POSITION INFORMATION
124A
TRANSIENT DETECTION UNIT
1L004
AUDIO SIGNAL
TRANSIENT POSITION INFORMATION
1L45
125
TRANSIENT POSITION QUANTIZATION UNIT
QUANTIZED POSITION INFORMATION
1L57
TRANSIENT FLAG
1L47

Fig.50


AUDIO DECODING UNIT
42
DECODED CORE SUBBAND POWER
45
AUXILIARY INFORMATION DECODING UNIT
DIFFERENCE DECODING UNIT
1215
TRANSIENT POSITION DECODING UNIT
1212
TRANSIENT FLAG DECODING UNIT
129
DIFFERENTIAL SUBBAND POWER CODE
1L915
TRANSIENT POSITION INFORMATION
1L912
QUANTIZED POSITION INFORMATION
1L1599
1L1299
1L999
TRANSIENT FLAG
TRANSIENT POWER SPECTRUM
TRANSIENT POSITION INFORMATION
TRANSIENT FLAG
AUXILIARY INFORMATION CODE
4L001

REFERENCES CITED IN THE DESCRIPTION

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

Patent documents cited in the description

- WO 2007000988 A **[0010]**
- JP 2003316670 A **[0010]**
- JP 2008111991 A **[0010]**

Non-patent literature cited in the description

- Joint pre-echo control and frame erasure concealment of VOIP audio codecs. **GEISER BERNDT et al.** 17th EUROPENA SIGNAL PROCESSING CONFERENCE. IEEE, 24 August 2009, 1259-1263 **[0011]**
- Candidate proposal for IUT-T super-wideband speech and audio coding. **BERND GEISER et al.** ACOUTICS, SPEECH AND SIGNAL PROCESSING. ICASSP, 19 April 2009, 4121-4124 **[0011]**
- Enhanced aacPlus general audio codec General description. *3GPP TS26.401* **[0034]**