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

(11) EP 1 779 385 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
22.09.2010 Bulletin 2010/38

(21) Application number: **05774399.9**

(22) Date of filing: **08.07.2005**

(51) Int Cl.:
G11B 20/10 *(2006.01)* ***G10L 19/00*** *(2006.01)*
H04S 3/00 *(2006.01)*

(86) International application number:
PCT/KR2005/002213

(87) International publication number:
WO 2006/006809 (19.01.2006 Gazette 2006/03)

(54) **METHOD AND APPARATUS FOR ENCODING AND DECODING MULTI-CHANNEL AUDIO SIGNAL USING VIRTUAL SOURCE LOCATION INFORMATION**

VERFAHREN UND VORRICHTUNG ZUM CODIEREN UND DECODIEREN EINES MEHRKANALIGEN AUDIOSIGNALS UNTER VERWENDUNG VON VIRTUELLE-QUELLE-ORTSINFORMATIONEN

PROCEDE ET DISPOSITIF DESTINES A CODER ET DECODER UN SIGNAL AUDIO MULTICANAL AU MOYEN D'INFORMATIONS D'EMPLACEMENT DE SOURCE VIRTUELLE

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

(30) Priority: **09.07.2004 KR 2004053665**
12.10.2004 KR 2004081303
07.07.2005 KR 2005061425

(43) Date of publication of application:
02.05.2007 Bulletin 2007/18

(73) Proprietors:
- **Electronics and Telecommunications Research Institute**
  Daejeon 305-350 (KR)
- **Seoul National University Industry Foundation**
  Seoul 151-818 (KR)

(72) Inventors:
- **Seo, Jeong Il**
  Daejeon 305-728 (KR)
- **Moon, Han Gil**
  Yangcheon-gu,
  Seoul 158-073 (KR)
- **Beack, Seung Kwon,**
  Information and Communication
  Yuseong-gu,
  Daejeon 305-732 (KR)
- **Kang, Kyeong Ok**
  Yuseong-gu,
  Daejeon 305-727 (KR)
- **Jang, In Seon**
  Chungcheongbuk-do 360-815 (KR)
- **Sung, Koeng Mo**
  Gwanak-gu,
  Seoul 151-801 (KR)
- **Hahn, Min Soo**
  Yuseong-gu,
  Daejeon 305-761 (KR)
- **Hong, Jin Woo**
  Yuseong-gu,
  Daejeon 305-755 (KR)

(74) Representative: **Betten & Resch**
Patentanwälte
Theatinerstrasse 8
80333 München (DE)

(56) References cited:
WO-A-99/52326
WO-A-03/090208
WO-A-2005/101905
US-A- 5 946 352
US-A- 6 128 597
US-A1- 2003 014 243
US-A1- 2003 236 583

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

EP 1 779 385 B1

Description

Background Art

1. Field of the Invention

[0001] The present invention relates to a method and apparatus for encoding/decoding a multi-channel audio signal, and more particularly, to a method and apparatus for effectively encoding/decoding a multi-channel audio signal using Virtual Sound Location Information (VLSI).

2. Description of Related Art

[0002] Throughout the later half of the 1990s, Moving Picture Experts Group (MPEG) has performed research on compressing a multi-channel audio signal. Owing to the remarkable increase in multi-channel contents, increased demand for multi-channel contents, and increased need for a multi-channel audio services in a broadcasting communications environment, research on the multi-channel audio compression technology has been stepped up.

[0003] As a result, multi-channel audio compression technology such as MPEG-2 Backward Compatibility (BC), MPEG-2 Advanced Audio Coding (AAC), and MPEG-4 AAC, has been standardized in the MPEG. Also, multi-channel audio compression technology, such as AC-3 and Digital Theater System (DTS), has been commercialized.

[0004] In recent years, innovative multi-channel audio signal compression method such as typical Binaural Cue Coding (BCC) has been actively researched (C. Faller, 2002 & 2003; F. Baumgarte, 2001 & 2002). The goal of such research is the transfer of more realistic audio data.

[0005] BCC is technology for effectively compressing a multi-channel audio signal that has been developed on a basis of the fact that people can acoustically perceive space due to a binaural effect. BCC is based on the fact that a pair of ears perceives a location of a specific sound source using interaural level differences and/or interaural time differences.

[0006] Accordingly, in BCC, a multi-channel audio signal is downmixed to a monophonic or stereophonic signal and channel information is represented by binaural cue parameters such as Inter-channel Level Difference (ICLD) and Inter-channel Time Difference (ICTD).

[0007] However, there is a drawback in that a large number of bits are required to quantize the channel information such as ICLD and ICTD, and consequently, a wide bandwidth is required in transmitting the channel information.

[0008] Document WO 99/52326 discloses a spatial audio coding system, including an encoder and a decoder, operates at very low bit-rates and is useful for audio via the Internet. The listener or listeners preferably are located within a predictable listening area, for example, users of a personal computer or television viewers. An encoder produces a composite audio-information signal representing the soundfield to be reproduced and a directional vector or "steering control signal". The composite audio-information signal has its frequency spectrum broken into a number of subbands, preferably commensurate with the critical bands of the human ear. The steering control signal has a component relating to the dominant direction of the soundfield in each of the subbands. Because the system is based on the premise that only sound from a single diection is heard at any instant, decoder need not apply a signal to more than two sound transducers at any instant.

[0009] Document WO 03/090208 discloses a psycho-acoustically motivated, parametric representation of the spatial attributes of multichannel audio signals. The parametric representation allows strong bitrate reductions in audio coders, since only one monaural signal has to be transmitted, combined with (quantized) parameters which describe the spatial properties of the signal. The decoder can form the original amount of audio channels by applying the spatial parameters.

SUMMARY OF THE INVENTION

[0010] The present invention as defined by the independent claims 1, 10, 13, 22, 26 and 27, is directed to reproduction of a realistic audio signal by encoding/decoding a multi-channel audio signal using only a downmixed audio signal and a small amount of additional information.

[0011] The present invention is also directed to maximizing transmission efficiency by analyzing a per-channel sound source of a multi-channel audio signal, extracting a small amount of virtual source location information, and transmitting the extracted virtual source location information together with a downmixed audio signal.

Brief Description of the Drawings

[0012] The above and other features and advantages of the present invention will become more apparent to those of ordinary skill in the art by describing in detail exemplary embodiments of the invention with reference to the attached drawings in which:

FIG. 1 is a block diagram of an apparatus for encoding a multi-channel audio signal according to an exemplary embodiment of the present invention;

FIG. 2 is a conceptual diagram of a time-to-frequency lattice using an Equivalent Rectangular Bandwidth (ERB) filter bank;

FIG. 3 is a conceptual diagram of source location vectors estimated according to the preset invention, in the case where a downmixed multi-channel audio signal is monophonic;

FIG. 4 is a conceptual diagram of source location vectors estimated according to the preset invention, in the case where a downmixed multi-channel audio signal is stereophonic;

FIG. 5 is a conceptual diagram illustrating a process of estimating virtual source location information according to an exemplary embodiment of the present invention;

FIG. 6 shows an example of per-channel energy vectors when 5.1 channel speakers are used;

FIG. 7 is a conceptual diagram illustrating a process of estimating a Left Half-plane Vector (LHV) and a Right Half-plane Vector (RHV) according to the present invention;

FIG. 8 is a conceptual diagram illustrating a process of estimating a Left Subsequent Vector (LSV) and a Right Subsequent Vector (RSV) according to the present invention;

FIG. 9 is a conceptual diagram illustrating a process of estimating a Global Vector (GV) according to the present invention;

FIG. 10 illustrates azimuth angles, each of which represents the corresponding virtual source location information according to the present invention;

FIG. 11 is a block diagram of an apparatus for decoding an encoded multi-channel audio signal according to an exemplary embodiment of the present invention; and

FIG. 12 is a block diagram illustrating a process of calculating per-channel gains of a downmixed audio signal using Virtual Source Location Information (VSLI) according to an exemplary embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0013] The present invention will now be described more fully with reference to the accompanying drawings, in which exemplary embodiments of the invention are shown. This invention may, however, be embodied in different forms and should not be construed as limited to the exemplary embodiments set forth herein. Rather, these exemplary embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art.

[0014] FIG. 1 is a block diagram of an apparatus for encoding a multi-channel audio signal according to an exemplary embodiment of the present invention. As shown in FIG. 1, the multi-channel audio signal encoding apparatus includes a frame converter 100, a downmixer 110, an Advanced Audio Coding (AAC) encoder 120, a multiplexer 130, a quantizer 140, and a Virtual Source Location Information (VSLI) analyzer 150.

[0015] The frame converter 100 frames the multi-channel audio signal, using a window function such as a sine window, to process the multi-channel audio signal in each block. The downmixer 110 receives the framed multi-channel audio signal from the frame converter 100 and downmixes it into a monophonic signal or a stereophonic signal. The AAC encoder 120 compresses the downmixed audio signal received from the downmixer 110, to generate an AAC encoded signal. It then transmits the AAC encoded signal to the multiplexer 130.

[0016] The VSLI analyzer 150 extracts Virtual Source Location Information (VSLI) from the framed audio signal. Specifically, the VSLI analyzer 150 may include a time-to-frequency converter 151, an Equivalent Rectangular Bandwidth (ERB) filter bank 152, an energy vector detector 153, and a location estimator 154.

[0017] The time-to-frequency converter 151 performs a plurality of Fast Fourier Transforms (FFTs) to convert the framed audio signal into a frequency domain signal. The ERB filter bank 152 divides the converted frequency domain signal (spectrum) into per-band spectrums (for example, 20 bands). FIG. 2 is a conceptual diagram of a time-to-frequency lattice using the ERB filter bank 152.

[0018] The energy vector extractor 153 estimates per-channel energy vectors from the corresponding per-band spectrum.

[0019] The location estimator 154 estimates virtual source location information (VSLI) using the per-channel energy vectors estimated by the energy vector extractor 153. In one exemplary embodiment, the VSLI may be represented using azimuth angles between the source location vectors and a center channel. As described later, the VSLI estimated by the location estimator 154 can vary depending on whether the downmixed audio signal is monophonic or stereophonic.

[0020] FIG. 3 is a conceptual diagram illustrating the source location vectors estimated according to the present invention, in the case where the downmixed audio signal is monophonic. As shown in FIG. 3, the source location vectors estimated from the downmixed monophonic signal include a Left Half-plane Vector (LHV), a Right Half-plane Vector (RHV), a Left Subsequent Vector (LSV), a Right Subsequent Vector (RSV), and a Global Vector (GV). In the case where the downmixed multi-channel audio signal is monophonic, since it is not known whether channel gain is higher on the left or on the right, the GV is required.

[0021] FIG. 4 is a conceptual diagram illustrating the source location vectors estimated according to the present invention, in the case where the downmixed multi-channel audio signal is stereophonic. As shown in FIG. 4, the source location vectors estimated from the downmixed monophonic signal include the LHV, the RHV, the LSV, and the RSV, but not the GV.

[0022] Referring again to FIG. 1, the quantizer 140 quantizes the VSLI (azimuth angles) received from the VSLI analyzer 150 and transmits the quantized VSLI signal to the multiplexer 130. The multiplexer 130 receives the AAC encoded signal from the AAC encoder 120 and the quantized VSLI signal from the quantizer 140 and multiplexes them to generate an encoded multi-channel audio signal (i.e., the AAC encoded signal +the VSLI signal).

[0023] FIG. 5 is a conceptual diagram illustrating a process of estimating the VSLI according to an exemplary embodiment of the present invention. As shown in FIG. 5, in the case where the input multi-channel audio signal is comprised of five channels including center (C), front left (L), front right (R), left subsequent (LS), and right subsequent (RS), the input signal is converted into the frequency axis signal through the plurality of FFTs and divided into N number of frequency bands (BAND 1, BAND 2,..., and BAND N) in the ERB filter bank 152.

[0024] Next, the per-channel energy vectors may be detected from the power of each of the five channels for each band (for example, C1 PWR, L1 PWR, R1 PWR, LS1 PWR, and RS1 PWR). Using Constant Power Panning (CPP) in which the magnitudes of signals of neighboring channels are adjusted for sound localization, the source location vectors may be estimated from the detected per-channel energy vectors and the azimuth angles between the source location vectors and the center channel, which represent VSLI, may be estimated.

[0025] FIG. 6 to 9 illustrate detailed processes of estimating the VSLI according to the present invention. In detail, as shown in FIG. 6, it is assumed that the per-channel energy vectors estimated using the energy vector estimator are a center channel energy vector (C), a front left channel energy vector (L), a left subsequent channel energy vector (LS), a front right channel energy vector (R), and a right subsequent channel energy vector (RS). The LHV is estimated using the front left channel energy vector (L) and the left subsequent channel energy vector (LS), and the RHV is estimated using the front right channel energy vector (R) and the right subsequent channel energy vector (RS) (Refer to FIG. 7).

[0026] The LSV and RSV may be estimated using the LHV, the RHV, and the center channel energy vector (C) (Refer to FIG. 8).

[0027] In the case where the downmixed audio signal is stereophonic, the gain of each channel can be calculated using only the LHV, RHV, LSV, and RSV. However, in the case where the downmixed audio signal is the monophonic signal, it is not known whether the channel gain is higher on the left or on the right, and therefore the GV is required. The GV can be calculated using the LSV and RSV (Refer to FIG. 9). The magnitude of the GV is set to the magnitude of the downmixed audio signal.

[0028] The source location vectors extracted using the above method may be expressed using the azimuth angles between themselves and the center channel. FIG. 10 illustrates the azimuth angles of the source location vectors extracted by the processes shown in FIGS. 6 to 9. As shown, the VSLI may be expressed using five azimuth angles, which include a Left Half-plane vector angle (LHa), a Right Half-plane vector angle (RHa), a Left Subsequent vector angle (LSa), and a Right Subsequent vector angle (RSa), and further include a Global vector angle (Ga) in the case where the downmixed audio signal is monophonic. Since each value has a limited dynamic range, quantization can be performed using fewer bits than Inter-Channel Level Difference (ICLD).

[0029] To quantize the VSLI information, a linear quantization method in which quantization is performed in uniform intervals or a nonlinear quantization method in which quantization is performed in non-uniform intervals may be used.

[0030] In one exemplary embodiment, the linear quantization method is based on Equation 1 below:

[Equation 1]

$$I_{i,b}=\left[\frac{\Delta\theta_{i,b}(Q-1)}{2\Delta\theta_{i,\max}}+\frac{1}{2}\right]+\frac{Q-1}{2},\ i=1,K,5$$

, wherein "θ"represents the magnitude of an angle to be quantized and the corresponding quantization index can be obtained from quantization level Q. "i" represents angle index (Ga:i=1, RHa:i=2, LHa:i=3, LSa:i=4, RSa:i=5) and "b" represents sub-band index. "$\Delta\theta_{i,max}$ represents the maximal variance level of each angle. For example, $\Delta\theta_{1,max}$ equals 180° $\Delta\theta_{2,max}$ and $\Delta\theta_{3,max}$ equal 15° and $\Delta\theta_{4,max}$ and $\Delta\theta_{5,\ max}$ equal 55°. As mentioned above, a maximal variance interval of each angle magnitude is limited, and therefore more effective and higher resolution quantization can be provided.

[0031] In general, statistical information on generation frequency with respect to the RHa, LHa, LSa, and RSa is inconclusive. However, the Ga has a generation frequency with a roughly symmetrical distribution centered on a center speaker. In other words, since the Ga varies evenly about the center speaker, it can be assumed that the generation distribution has an average expectation value of 0°. Accordingly, for the Ga, a more effective quantization level can be obtained when quantization is performed using the nonlinear quantization method.

[0032] Typically, the nonlinear quantization method is performed in a general m-law scheme, and m value can be determined depending on a resolution of the quantization level. For example, when the resolution is low, a relatively large m value may be used ($15< \mu\leq 255$), and when the resolution is high, a smaller m value ($0\leq\mu\leq5$) may be used to perform the nonlinear quantization.

[0033] FIG. 11 is a block diagram illustrating an apparatus for decoding an encoded multi-channel audio signal according to an exemplary embodiment of the present invention. As shown, the multi-channel audio signal decoding apparatus includes a signal distributor 1110, an AAC decoder 1120, a time-to-frequency converter 1130, an inverse quantizer 1140, a per-band channel gain distributor 1150, a multi-channel spectrum synthesizer 1160, and a frequency-to-time converter 1170.

[0034] The signal distributor 1110 separates the encoded multi-channel audio signal back into the AAC encoded signal and the VLSI encoded signal, respectively. The AAC decoder 1120 converts the AAC encoded signal back into the downmixed audio signal (monophonic or stereophonic signal). The converted downmixed audio signal can be used to produce monophonic or stereophonic sound. The time-to-frequency converter 1130 converts the downmixed audio signal into a frequency axis signal and transmits it to the multi-channel spectrum synthesizer 1160.

[0035] The inverse quantizer 1140 receives the separated VSLI encoded signal from the signal distributor 1110 and produces per-band source location vector information from the received VSLI encoded signal. In the encoding process, as described above, the VSLI includes azimuth angle information (for example, LHa, RHa, LSa, RSa, and Ga in the case where the downmixed audio signal is monophonic), each of which represents the corresponding per-band source location vector. The source location vector is produced from the VSLI.

[0036] The per-band channel gain distributor 1150 calculates the gain per channel using the per-band VSLI signal converted by the inverse quantizer 1140, and transmits the calculated gain to the multi-channel spectrum synthesizer 1160.

[0037] The multi-channel spectrum synthesizer 1160 receives a spectrum of the downmixed audio signal from the time-to-frequency converter 1130, separates the received spectrum into per-band spectrums using the ERB filter bank, and restores the spectrum of the multi-channel signal using per-band channel gains output from the per-band channel gain distributor 1150. The frequency-to-time converter 1170 (for example, IFFF) converts the spectrum of the restored multi-channel signal into a time axis signal to generate the multi-channel audio signal.

[0038] FIG. 12 is a block diagram illustrating a process of calculating the per-channel gain of the downmixed audio signal using the VSLI according to an exemplary embodiment of the present invention. Here, the case in which the downmixed audio signal is monophonic is illustrated. In the case where the downmixed audio signal is stereophonic, block 1210 is omitted.

[0039] In block 1210, magnitudes of the LSV and the RSV are calculated using the magnitude of the downmixed monophonic signal, which is the magnitude of the GV, and the angle (Ga) of the GV. Next, magnitudes of the LHV and the first gain of the center channel (C) are calculated using the magnitude and angle (LSa) of the LSV (Block 1220). The gain of the center channel (C) is obtained by summing the first gain and the second gain calculated in the above process (block 1240).

[0040] Last, gains of the front left channel (L) and the left subsequent channel (LS) are calculated using the magnitude of the LHV and the corresponding angle (LHa) (block 1250), and gains of the front right channel (R) and the right subsequent channel (RS) are calculated using the magnitude of the RHV and the corresponding angle (RHa) (block 1260). According to the above processes, the gains of all channels can be calculated.

[0041] According to the present invention, a multi-channel audio signal can be more effectively encoded/decoded using virtual source location information, and more realistic audio signal reproduction in a multi-channel environment can be realized.

[0042] While the invention has been shown and described with reference to certain exemplary embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention as defined by the appended claims and their equivalents.

Claims

1. An apparatus for encoding a multi-channel audio signal, the apparatus comprising:

   a frame converter (100) for converting the multi-channel audio signal into a framed audio signal;
   means for downmixing (110) the framed audio signal;
   means for encoding (120) the downmixed audio signal;

a source location information estimator (150) for estimating virtual source location information (VSLI) from the framed audio signal;
means for quantizing (140) the estimated VSLI; and
means for multiplexing (130) the encoded audio signal and the quantized VSLI to generate an encoded multi-channel audio signal;

wherein:

the VSLI includes a Left Half-plane vector angle (LHa), a Right Half-phane vector angle (RHa), a Left Subsequent vector angle (LSa), and a Right Subsequent vector angle (RSa).

2. The apparatus according to claim 1, wherein said downmixing means (110) is adapted to downmixes the framed audio signal as either one of monophonic or stereophonic signal.

3. The apparatus according to claim 2, wherein when the downmixed audio signal is the monophonic signal, the source location information estimator (150) estimates a Left Half-plane Vector (LHV), a Right Half-plane Vector (RHV), a Left Subsequent Vector (LSV), a Right Subsequent Vector (RSV), and a Global Vector (GV).

4. The apparatus according to claim 2, wherein when the downmixed audio signal is the stereophonic signal, the source location information estimator (150) estimates a Left Half-plane Vector (LHV), a Right Half-plane Vector (RHV), a Left Subsequent Vector (LSV), and a Right Subsequent Vector (RSV).

5. The apparatus according to claim 1, wherein said source location information estimator (150) comprises:

a time-to-frequency converter (151) for converting the framed audio signal into a spectrum;
a separator (152) for separating per-band spectrums;
an energy vector detector (153) for detecting per-channel energy vectors from the corresponding per-band spectrum; and
a VSLI estimator (154) for estimating the VSLI using the detected per-channel energy vector detected by the energy vector detector (153).

6. The apparatus according to claim 5, wherein said time-to-frequency converter (151) is adapted to converts the framed audio signal into the spectrum using a plurality of Fast Fourier Transforms (FFTs).

7. The apparatus according to claim 5, wherein the separator (152) is adapted to separates the spectrum using an Equivalent Rectangular Bandwidth (ERB) filter bank.

8. The apparatus according to claim 5, wherein the detected per-channel energy vector includes a center channel energy vector (C), a front left channel energy vector (L), a left subsequent channel energy vector (LS), a front right channel energy vector (R), and a right subsequent channel energy vector (RS).

9. The apparatus according to claim 9, wherein when the downmixed audio signal is the monophonic signal, the VSLI further includes a Global vector angle (Ga).

10. An apparatus for decoding a multi-channel audio signal, the apparatus comprising:

means for receiving the multi-channel audio signal;
a signal distributor (1110) for separating the received multi-channel audio signal into an encoded downmixed audio signal and a quantized virtual source location information (VSLI) signal;
means for decoding (1120) the encoded downmixed audio signal,
means for converting (1130) the decoded downmixed audio signal into a frequency axis signal;
a VSLI extractor (1140) for extracting per-band VSLI from the quantized VSLI signal;
a channel gain calculator (1150) for calculating per-band channel gains using the extracted per-band VSLI;
means for synthesizing (1160) a multi-channel audio signal spectrum using the converted frequency axis signal and the calculated per-band channel gains; and
means for generating (1170) a multi-channel audio signal from the synthesized multi-channel spectrum wherein the per-band VSLI includes a Left Half-plane vector angle (LHa), a Right Half-plane vector angle (RHa), a Left Subsequent vector angle (LSa), and a Right Subsequent vector angle (RSa), for each band.

11. The apparatus according to claim 10, wherein the VSLI extractor (1140) is adapted to produce a Left Half-plane Vector (LHV), a Right Half-plane Vector (RHV), a Left Subsequent Vector (LSV), and a Right Subsequent Vector (RSV).

12. The apparatus according to claim 10, wherein when the encoded downmixed audio signal is monophonic, and the VSLI further includes a Global vector angle (Ga), and a Global Vector (GV) is produced from the Ga.

13. A method of encoding a multi-channel audio signal, comprising the steps of:

converting the multi-channel audio signal into a framed audio signal;
downmixing the framed audio signal;
encoding the downmixed audio signal;
estimating VSLI from the framed audio signal;
quantizing the estimated VSLI; and
multiplexing the encoded downmixed audio signal and the quantized VSLI to generate an encoded multi-channel audio signal;
the VSLI includes a Left Half-plane vector angle (LHa), a Right Half-plane vector angle (RHa), a Left Subsequent vector angle (LSa), and a Right Subsequent vector angle (RSa).

14. The method according to claim 13, wherein the framed audio signal is downmixed into either one of a monophonic signal and a stereophonic signal.

15. The method according to claim 14, wherein when the downmixed audio signal is the monophonic signal, the VSLI further includes Ga and a Left Half-plane Vector (LHV), a Right Half-plane Vector (RHV), a Left Subsequent Vector (LSV), a Right Subsequent Vector (RSV), and a Global Vector (GV) are produced from the VSLI.

16. The method according to claim 14, wherein when the downmixed audio signal is the stereophonic signal, a Left Half-plane Vector (LHV), a Right Half-plane Vector (RHV), a Left Subsequent Vector (LSV), and a Right Subsequent Vector (RSV), are produced from the VSLI.

17. The method according to claim 13, wherein the step of estimating the VSLI comprises the steps of:

converting the framed audio signal into a spectrum:
separating the spectrum into per-band spectrums;
detecting per-channel energy vectors from the per-band spectrums; and
estimating the VSLI using the detected per-channel energy vectors.

18. The method according to claim 17, wherein the detected per-channel energy vectors include a center channel energy vector (C), a front left channel energy vector (L), a left subsequent channel energy vector (LS), a front right channel energy vector (R), and a right subsequent channel energy vector (RS).

19. The method according to claim 17, wherein the step of estimating the VSLI comprises the steps of:

estimating an LHV using the front left channel energy vector (L) and the left subsequent channel energy vector (LS);
estimating an RHV using the front right channel energy vector (R) and the right subsequent channel energy vector (RS);
estimating an LSV using the estimated LHV and the center channel energy vector (C); and
estimating an RSV using the estimated RHV and the center channel energy vector (C).

20. The method according to claim 19, wherein when the downmixed audio signal is the monophonic signal, the estimated VLSI further includes a GV, and the estimating of the VSLI further comprises the step of estimating the GV using the estimated LSV and RSV.

21. The method according to claim 17, wherein when the downmixed audio signal is the monophonic signal, the azimuth angle information further includes.

22. A method of decoding a multi-channel audio signal, comprising the steps of:

receiving the multi-channel audio signal;
separating the received multi-channel audio signal into an encoded downmixed audio signal and a quantized VSLI signal;
decoding the encoded downmixed audio signal;
converting the decoded downmixed audio signal into a frequency axis signal;
analyzing the quantized VSLI signal and extracting per-band VSLI therefrom;
calculating per-band channel gains from the extracted per-band VSLI;
synthesizing a multi-channel audio signal spectrum using the converted frequency axis signal and the calculated per-band channel gains; and
producing a multi-channel audio signal from the synthesized multi-channel spectrum;
wherein the per-band VSLI includes a Left Half-plane vector angle (LHa), a Right Half-plane vector angle (RHa), a Left Subsequent vector angle (LSa), and a Right Subsequent vector angle (RSa) for each band.

23. The method according to claim 22, wherein a Left Half-plane Vector (LHV), a Right half-plane Vector (RHV), a Left Subsequent Vector (LSV), and a Right Subsequent Vector (RSV) are produced from the VSLI.

24. The method according to claim 22, wherein when the encoded downmixed audio signal is monophonic, the VSLI further includes a Global vector angle (Ga), and a Global Vector (GV) is produced from the Ga.

25. The method according to claim 23, wherein said step of calculating the channel gain comprises, for each band, the steps of:

calculating magnitudes of the LSV and the RSV using a magnitude of the downmixed audio signal;
calculating a first gain of a center channel (C) and a magnitude of the LHV using the magnitude of the LSV and the LSa;
calculating a second gain of a center channel (C) and a magnitude of the RHV using the magnitude of the RSV and the RSa;
summing the first and second gains of the center channel (C) to produce a gain of the center channel (C);
calculating gains of a front left channel (L) and a left subsequent channel (LS) using the magnitude of the LHV and the LHa; and
calculating gains of a front right channel (R) and a right subsequent channel (RS) using the magnitude of the RHV and the RHa.

26. A computer-readable recording medium storing a computer program for performing the method for encoding a multi-channel audio signal according to any one of claims 13 to 21.

27. A computer-readable recording medium storing a computer program for performing the method for decoding a multi-channel audio signal according to any one of claims 22 to 25.

Patentansprüche

1. Vorrichtung zum Codieren eines Mehrkanal-Audiosignals, wobei die Vorrichtung umfasst:

einen Rahmenumsetzer (100), um das Mehrkanal-Audiosignal in ein Rahmen-Audiosignal umzusetzen;
Mittel (110) zum Heruntermischen des Rahmen-Audiosignals;
Mittel (120) zum Codieren des heruntergemischten Audiosignals;
eine Quellenortinformations-Schätzeinrichtung (150) zum Schätzen von Informationen (VSLI) bezüglich eines virtuellen Quellenorts aus dem Rahmen-Audiosignal;
Mittel (140) zum Quantisieren der geschätzten VSLI; und
Mittel (130) zum Multiplexieren des codierten Audiosignals und der quantisierten VSLI, um ein codiertes Mehrkanal-Audiosignal zu erzeugen;

wobei:

die VSLI einen Vektorwinkel der linken Halbebene (LHa), einen Vektorwinkel der rechten Halbebene (RHa), einen nachfolgenden linken Vektorwinkel (LSa), und einen nachfolgenden rechten Vektorwinkel (RSa) enthalten.

2. Vorrichtung nach Anspruch 1, wobei die Herabmischungsmittel (110) dazu ausgelegt sind, das Rahmen-Audiosignal entweder als Mono- oder als Stereosignal herunterzumischen.

3. Vorrichtung nach Anspruch 2, wobei dann, wenn das heruntergemischte Audiosignal das Monosignal ist, die Quellenortinformations-Schätzeinrichtung (150) einen Vektor der linken Halbebene (LHV), einen Vektor der rechten Halbebene (RHV), einen nachfolgenden linken Vektor (LSV), einen nachfolgenden rechten Vektor (RSV) und einen globalen Vektor (GV) schätzt.

4. Vorrichtung nach Anspruch 2, wobei dann, wenn das heruntergemischte Audiosignal das Stereosignal ist, die Quellenortinformations-Schätzeinrichtung (150) einen Vektor der linken Halbebene (LHV), einen Vektor der rechten Halbebene (RHV), einen nachfolgenden linken Vektor (LSV) und einen nachfolgenden rechten Vektor (RSV) schätzt.

5. Vorrichtung nach Anspruch 1, wobei die Quellenortinformations-Schätzeinrichtung (150) umfasst:

einen Zeit/Frequenz-Umsetzer (151), um das Rahmen-Audiosignal in ein Spektrum umzusetzen;
einen Separator (152), um Spektren pro Band zu separieren;
einen Energievektor-Detektor (153), um Energievektoren pro Kanal aus dem entsprechenden Spektrum pro Band zu detektieren; und
eine VSLI-Schätzeinrichtung (154), um die VSLI unter Verwendung des detektierten Energievektors pro Kanal, der durch den Energievektor-Detektor (153) detektiert wird, zu schätzen.

6. Vorrichtung nach Anspruch 5, wobei der Zeit/Frequenz-Umsetzer (151) dazu ausgelegt ist, das Rahmen-Audiosignal in das Spektrum unter Verwendung mehrerer schneller Fourier-Transformationen (FFTs) umzusetzen.

7. Vorrichtung nach Anspruch 5, wobei der Separator (152) dazu ausgelegt ist, das Spektrum unter Verwendung einer Filterbank für äquivalente Rechteckbandbreite (ERB-Filterbank) zu separieren.

8. Vorrichtung nach Anspruch 5, wobei der detektierte Energievektor pro Kanal einen Energievektor (C) für den Mittelkanal, einen Energievektor (L) für den vorderen linken Kanal, einen Energievektor für den nachfolgenden linken Kanal (LS), einen Energievektor (R) für den vorderen rechten Kanal und einen Energievektor (RS) für den nachfolgenden rechten Kanal enthält.

9. Vorrichtung nach Anspruch 1, wobei dann, wenn das heruntergemischte Audiosignal das Monosignal ist, die VSLI ferner einen globalen Vektorwinkel (Ga) enthalten.

10. Vorrichtung zum Decodieren eines Mehrkanal-Audiosignals, wobei die Vorrichtung umfasst:

Mittel, um das Mehrkanal-Audiosignal zu empfangen;
einen Signalverteiler (1110), um das empfangene Mehrkanal-Audiosignal in ein codiertes heruntergemischtes Audiosignal und in ein Signal quantisierter Informationen (VSLI) bezüglich eines virtuellen Quellenortes zu separieren;
Mittel (1120), um das codierte heruntergemischte Audiosignal zu decodieren;
Mittel (1130), um das decodierte heruntergemischte Audiosignal in ein Frequenzachsensignal umzusetzen;
eine VSLI-Extraktionseinrichtung (1140), um die VSLI pro Band aus dem quantisierten VSLI-Signal zu extrahieren;
einen Kanalverstärkungsrechner (1150), um Kanalverstärkungen pro Band unter Verwendung der extrahierten VSLI pro Band zu berechnen;
Mittel (1160), um ein Mehrkanal-Audiosignalspektrum unter Verwendung des umgesetzten Frequenzachsensignals und der berechneten Kanalverstärkungen pro Band zu synthetisieren; und
Mittel (1170), um ein Mehrkanal-Audiosignal aus dem synthetisierten Mehrkanalspektrum zu erzeugen,
wobei die VSLI pro Band einen Vektorwinkel der linken Halbebene (LHa), einen Vektorwinkel der rechten Halbebene (RHa), einen nachfolgenden linken Vektorwinkel (LSa) und einen nachfolgenden rechten Vektorwinkel (RSa) für jedes Band enthalten.

11. Vorrichtung nach Anspruch 10, wobei die VSLI-Extraktionseinrichtung (1140) dazu ausgelegt ist, einen Vektor der linken Halbebene (LHV), einen Vektor der rechten Halbebene (RHV), einen nachfolgenden linken Vektor (LSV) und einen nachfolgenden rechten Vektor (RSV) zu erzeugen.

12. Vorrichtung nach Anspruch 10, wobei dann, wenn das codierte heruntergemischte Signal das Monosignal ist und die VSLI ferner einen globalen Vektorwinkel (Ga) enthalten, ein globaler Vektor (GV) aus dem Ga erzeugt wird.

13. Verfahren zum Codieren eines Mehrkanal-Audiosignals, das die folgenden Schritte umfasst:

Umsetzen des Mehrkanal-Audiosignals in ein Rahmen-Audiosignal;
Heruntermischen des Rahmen-Audiosignals;
Codieren des heruntergemischten Audiosignals;
Schätzen von VSLI aus dem Rahmen-Audiosignal;
Quantisieren der geschätzten VSLI; und
Multiplexieren des codierten heruntergemischten Audiosignals und der quantisierten VSLI, um ein codiertes Mehrkanal-Audiosignal zu erzeugen;
wobei die VSLI einen Vektorwinkel der linken Halbebene (LHa), einen Vektorwinkel der rechten Halbebene (RHa), einen nachfolgenden linken Vektorwinkel (LSa) und einen nachfolgenden rechten Vektorwinkel (RSa) enthalten.

14. Verfahren nach Anspruch 13, wobei das Rahmen-Audiosignal entweder in ein Monosignal oder in ein Stereosignal heruntergemischt wird.

15. Verfahren nach Anspruch 14, wobei dann, wenn das heruntergemischte Audiosignal das Monosignal ist, die VSLI ferner einen Ga enthalten und ein Vektor der linken Halbebene (LHV), ein Vektor der rechten Halbebene (RHV), ein nachfolgender linker Vektor (LSV), ein nachfolgender rechter Vektor (RSV) und ein globaler Vektor (GV) aus den VSLI erzeugt werden.

16. Verfahren nach Anspruch 14, wobei dann, wenn das heruntergemischte Audiosignal das Stereosignal ist, ein Vektor der linken Halbebene (LHV), ein Vektor der rechten Halbebene (RHV), ein nachfolgender linker Vektor (LSV) und ein nachfolgender rechter Vektor (RSV) aus den VSLI erzeugt werden.

17. Verfahren nach Anspruch 13, wobei der Schritt des Schätzens der VSLI die folgenden Schritte umfasst:

Umsetzen des Rahmen-Audiosignals in ein Spektrum;
Separieren des Spektrums in Spektren pro Band;
Detektieren von Energievektoren pro Kanal aus den Spektren pro Band; und
Schätzen der VSLI unter Verwendung der detektierten Energievektoren pro Kanal.

18. Verfahren nach Anspruch 17, wobei die detektierten Energievektoren pro Kanal einen Energievektor (C) des Mittelkanals, einen Energievektor (L) des vorderen linken Kanals, einen Energievektor (LS) des nachfolgenden linken Kanals, einen Energievektor (R) des vorderen rechten Kanals und einen Energievektor (RS) des nachfolgenden rechten Kanals enthalten.

19. Verfahren nach Anspruch 17, wobei der Schritt des Schätzens der VSLI die folgenden Schritte umfasst:

Schätzen eines LHV unter Verwendung des Energievektors (L) des vorderen linken Kanals und des Energievektors (LS) des nachfolgenden linken Kanals;
Schätzen eines RHV unter Verwendung des Energievektors (R) des vorderen rechten Kanals und des Energievektors (RS) des nachfolgenden rechten Kanals;
Schätzen eines LSV unter Verwendung des geschätzten LHV und des Energievektors (C) des Mittelkanals; und
Schätzen eines RSV unter Verwendung des geschätzten RHV und des Energievektors (C) des Mittelkanals.

20. Verfahren nach Anspruch 19, wobei dann, wenn das heruntergemischte Audiosignal das Monosignal ist, die geschätzten VSLI ferner einen GV enthalten und das Schätzen der VSLI ferner den Schritt des Schätzens des GV unter Verwendung des geschätzten LSV und des geschätzten RSV umfasst.

21. Verfahren nach Anspruch 17, wobei dann, wenn das heruntergemischte Audiosignal das Monosignal ist, die Azimutwinkelinformationen ferner einen Ga enthalten.

22. Verfahren zum Decodieren eines Mehrkanal-Audiosignals, das die folgenden Schritte umfasst:

Empfangen des Mehrkanal-Audiosignals;
Separieren des empfangenen Mehrkanal-Audiosignals in ein codiertes heruntergemischtes Audiosignal und ein Signal quantisierter VSLI;
Decodieren des codierten heruntergemischten Audiosignals;
Umsetzen des decodierten heruntergemischten Audiosignals in ein Frequenzachsensignal;
Analysieren des Signals quantisierter VSLI und daraus Extrahieren von VSLI pro Band;
Berechnen von Kanalverstärkungen pro Band aus den extrahierten VSLI pro Band;
Synthetisieren eines Mehrkanal-Audiosignalspektrums unter Verwendung des umgesetzten Frequenzachsensignals und der berechneten Kanalverstärkungen pro Band; und
Erzeugen eines Mehrkanal-Audiosignals aus dem synthetisierten Mehrkanalspektrum;
wobei die VSLI pro Band einen Vektorwinkel der linken Halbebene (LHa), einen Vektorwinkel der rechten Halbebene (RHa), einen nachfolgenden linken Vektorwinkel (LSa) und einen nachfolgenden rechten Vektorwinkel (RSa) für jedes Band enthalten.

23. Verfahren nach Anspruch 22, wobei aus den VSLI ein Vektor der linken Halbebene (LHV), ein Vektor der rechten Halbebene (RHV), ein nachfolgender linker Vektorwinkel (LSV) und ein nachfolgender rechter Vektor (RSV) erzeugt werden.

24. Verfahren nach Anspruch 22, wobei dann, wenn das codierte heruntergemischte Audiosignal das Monosignal ist, die VSLI ferner einen globalen Vektorwinkel (Ga) enthalten und aus dem Ga ein globaler Vektor (GV) erzeugt wird.

25. Verfahren nach Anspruch 23, wobei der Schritt des Berechnens der Kanalverstärkung für jedes Band die folgenden Schritte umfasst:

Berechnen von Größen des LSV und des RSV unter Verwendung einer Größe des heruntergemischten Audiosignals;
Berechnen einer ersten Verstärkung eines Mittelkanals (C) und einer Größe des LHV unter Verwendung der Größe des LSV und des LSa;
Berechnen einer zweiten Verstärkung eines Mittelkanals (C) und einer Größe des RHV unter Verwendung der Größe des RSV und des RSa;
Summieren der ersten und der zweiten Verstärkung des Mittelkanals (C), um eine Verstärkung des Mittelkanals (C) zu erzeugen;
Berechnen von Verstärkungen eines vorderen linken Kanals (L) und eines nachfolgenden linken Kanals (LS) unter Verwendung der Größe des LHV und des LHa; und
Berechnen von Verstärkungen eines vorderen rechten Kanals (R) und eines nachfolgenden rechten Kanals (RS) unter Verwendung der Größe des RHV und des RHa.

26. Computerlesbares Aufzeichnungsmedium zum Speichern eines Computerprogramms, um das Verfahren zum Codieren eines Mehrkanal-Audiosignals nach einem der Ansprüche 13 bis 21 auszuführen.

27. Computerlesbares Aufzeichnungsmedium zum Speichern eines Computerprogramms, um das Verfahren zum Decodieren eines Mehrkanal-Audiosignals nach einem der Ansprüche 22 bis 25 auszuführen.

Revendications

1. Dispositif d'encodage d'un signal audio multicanal, le dispositif comprenant :

un convertisseur de trame (100) pour convertir le signal audio multicanal en un signal audio sous forme de trames ;
des moyens pour mélanger-abaisser (110) le signal audio sous forme de trames ;
des moyens pour encoder (120) le signal audio mélangé-abaissé;
un estimateur d'informations de position de source (150) pour estimer des informations de position de source virtuelle (VSLI) à partir du signal audio sous forme de trames ;
des moyens pour quantifier (140) les informations VSLI estimées ; et
des moyens pour multiplexer (130) le signal audio encodé et les informations VSLI quantifiées pour générer un signal audio multicanal encodé ;

dans lequel :

les informations VSLI comprennent un angle de vecteur de demi-plan gauche (LHa), un angle de vecteur de demi-plan droit (RHa), un angle de vecteur gauche suivant (LSa) et un angle de vecteur droit suivant (RSa).

2. Dispositif selon la revendication 1, dans lequel lesdits moyens de mélange-abaissement (110) sont adaptés pour mélanger-abaisser le signal audio sous forme de trames en tant que l'un quelconque d'un signal monophonique ou d'un signal stéréophonique.

3. Dispositif selon la revendication 2, dans lequel, lorsque le signal audio mélangé-abaissé est le signal monophonique, l'estimateur d'informations de position de source (150) estime un vecteur de demi-plan gauche (LHV), un vecteur de demi-plan droit (RHV), un vecteur gauche suivant (LSV), un vecteur droit suivant (RSV) et un vecteur global (GV).

4. Dispositif selon la revendication 2, dans lequel, lorsque le signal audio mélangé-abaissé est le signal stéréophonique, l'estimateur d'informations de position de source (150) estime un vecteur de demi-plan gauche (LHV), un vecteur de demi-plan droit (RHV), un vecteur gauche suivant (LSV) et un vecteur droit suivant (RSV).

5. Dispositif selon la revendication 1, dans lequel ledit estimateur d'informations de position de source (150) comprend :

un convertisseur temps-fréquence (151) pour convertir le signal audio sous forme de trames en un spectre ;
un séparateur (152) pour séparer en spectres par bande ;
un détecteur de vecteur d'énergie (153) pour détecter les vecteurs d'énergie par canal dans le spectre par bande correspondant ; et
un estimateur d'informations VSLI (154) pour estimer les informations VSLI en utilisant le vecteur d'énergie par canal détecté qui est détecté par le détecteur de vecteur d'énergie (153).

6. Dispositif selon la revendication 5, dans lequel ledit convertisseur temps-fréquence (151) est adapté pour convertir le signal audio sous forme de trames en le spectre en utilisant une pluralité de transformations de Fourier rapide (FFT).

7. Dispositif selon la revendication 5, dans lequel le séparateur (152) est adapté pour séparer le spectre en utilisant une batterie de filtres à largeur de bande rectangulaire équivalente (ERB).

8. Dispositif selon la revendication 5, dans lequel le vecteur d'énergie par canal détecté comprend un vecteur d'énergie de canal central (C), un vecteur d'énergie de canal gauche avant (L), un vecteur d'énergie de canal gauche suivant (LS), un vecteur d'énergie de canal droit avant (R) et un vecteur d'énergie de canal droit suivant (RS).

9. Dispositif selon la revendication 1, dans lequel, lorsque le signal audio mélangé-abaissé est le signal monophonique, les informations VSLI comprennent en outre un angle de vecteur global (Ga).

10. Dispositif de décodage d'un signal audio multicanal, le dispositif comprenant :

des moyens pour recevoir le signal audio multicanal ;
un distributeur de signal (1110) pour séparer le signal audio multicanal reçu en un signal audio mélangé-abaissé encodé et un signal d'informations de position de source virtuelle (VSLI) quantifiées ;
des moyens pour décoder (1120) le signal audio mélangé-abaissé encodé ;
des moyens pour convertir (1130) le signal audio mélangé-abaissé décodé en un signal d'axe de fréquence ;
un extracteur d'informations VSLI (1140) pour extraire les informations VSLI par bande du signal d'informations VSLI quantifiées ;
un calculateur de gain de canal (1150) pour calculer des gains de canal par bande en utilisant les informations VSLI par bande extraites ;
des moyens pour synthétiser (1160) un spectre de signal audio multicanal en utilisant le signal d'axe de fréquence converti et les gains de canal par bande calculés ; et
des moyens pour générer (1170) un signal audio multicanal à partir du spectre multicanal synthétisé,
dans lequel les informations VSLI par bande comprennent un angle de vecteur de demi-plan gauche (LHa), un angle de vecteur de demi-plan droit (RHa), un angle de vecteur gauche suivant (LSa) et un angle de vecteur droit suivant (RSa) pour chaque bande.

11. Dispositif selon la revendication 10, dans lequel l'extracteur d'informations VSLI (1140) est adapté pour produire un vecteur de demi-plan gauche (LHV), un vecteur de demi-plan droit (RHV), un vecteur gauche suivant (LSV) et un vecteur droit suivant (RSV).

12. Dispositif selon la revendication 10, dans lequel, lorsque le signal audio mélangé-abaissé encodé est monophonique, les informations VSLI comprennent en outre un angle de vecteur global (Ga), et un vecteur global (GV) est produit à partir de Ga.

13. Procédé d'encodage d'un signal audio multicanal, comprenant les étapes consistant à :

convertir le signal audio multicanal en un signal audio sous forme de trames ;
mélanger-abaisser le signal audio sous forme de trames ;
encoder le signal audio mélangé-abaissé;
estimer des informations VSLI à partir du signal audio sous forme de trames ;
quantifier les informations VSLI estimées ; et
multiplexer le signal audio mélangé-abaissé encodé et les informations VSLI quantifiées pour générer un signal audio multicanal encodé ;
les informations VSLI comprenant un angle de vecteur de demi-plan gauche (LHa), un angle de vecteur de demi-plan droit (RHa), un angle de vecteur gauche suivant (LSa) et un angle de vecteur droit suivant (RSa).

14. Procédé selon la revendication 13, dans lequel le signal audio sous forme de trames est mélangé-abaissé en l'un quelconque d'un signal monophonique et d'un signal stéréophonique.

15. Procédé selon la revendication 14, dans lequel, lorsque le signal audio mélangé-abaissé est le signal monophonique, les informations VSLI comprennent en outre un angle de vecteur global (Ga), et un vecteur de demi-plan gauche (LHV), un vecteur de demi-plan droit (RHV), un vecteur gauche suivant (LSV), un vecteur droit suivant (RSV) et un vecteur global (GV) sont produits à partir des informations VSLI.

16. Procédé selon la revendication 14, dans lequel, lorsque le signal audio mélangé-abaissé est le signal stéréophonique, un vecteur de demi-plan gauche (LHV), un vecteur de demi-plan droit (RHV), un vecteur gauche suivant (LSV) et un vecteur droit suivant (RSV) sont produits à partir des informations VSLI.

17. Procédé selon la revendication 13, dans lequel l'étape d'estimation des informations VSLI comprend les étapes consistant à :

convertir le signal audio sous forme de trames en un spectre ;
séparer le spectre en spectres par bande ;
détecter des vecteurs d'énergie par canal dans les spectres par bande ; et
estimer les informations VSLI en utilisant les vecteurs d'énergie par canal détectés.

18. Procédé selon la revendication 17, dans lequel les vecteurs d'énergie par canal détectés comprennent un vecteur d'énergie de canal central (C), un vecteur d'énergie de canal gauche avant (L), un vecteur d'énergie de canal gauche suivant (LS), un vecteur d'énergie de canal droit avant (R) et un vecteur d'énergie de canal droit suivant (RS).

19. Procédé selon la revendication 17, dans lequel l'étape d'estimation des informations VSLI comprend les étapes consistant à :

estimer un vecteur LHV en utilisant le vecteur d'énergie de canal gauche avant (L) et le vecteur d'énergie de canal gauche suivant (LS) ;
estimer un vecteur RHV en utilisant le vecteur d'énergie de canal droit avant (R) et le vecteur d'énergie de canal droit suivant (RS) ;
estimer un vecteur LSV en utilisant le vecteur LHV estimé et le vecteur d'énergie de canal central (C) ; et
estimer un vecteur RSV en utilisant le vecteur RHV estimé et le vecteur d'énergie de canal central (C).

20. Procédé selon la revendication 19, dans lequel, lorsque le signal audio mélangé-abaissé est le signal monophonique, les informations VLSI estimées comprennent en outre un vecteur GV, et l'estimation des informations VSLI comprend en outre l'étape d'estimation du vecteur GV en utilisant le vecteur LSV et le vecteur RSV estimés.

21. Procédé selon la revendication 17, dans lequel, lorsque le signal audio mélangé-abaissé est le signal monophonique, les informations d'angle azimutal comprennent en outre un Ga.

22. Procédé de décodage d'un signal audio multicanal comprenant les étapes consistant à :

recevoir le signal audio multicanal ;
séparer le signal audio multicanal reçu en un signal audio mélangé-abaissé encodé et un signal d'informations VSLI quantifiées ;
décoder le signal audio mélangé-abaissé encodé ;
convertir le signal audio mélangé-abaissé décodé en un signal d'axe de fréquence ;
analyser le signal d'informations VSLI quantifiées et extraire les informations VSLI par bande de celui-ci ;
calculer des gains de canal par bande à partir des informations VSLI par bande extraites ;
synthétiser un spectre de signal audio multicanal en utilisant le signal d'axe de fréquence converti et les gains de canal par bande calculés ; et
produire un signal audio multicanal à partir du spectre multicanal synthétisé,
dans lequel les informations VSLI par bande comprennent un angle de vecteur de demi-plan gauche (LHa), un angle de vecteur de demi-plan droit (RHa), un angle de vecteur gauche suivant (LSa) et un angle de vecteur droit suivant (RSa) pour chaque bande.

23. Procédé selon la revendication 22, dans lequel un vecteur de demi-plan gauche (LHV), un vecteur de demi-plan droit (RHV), un vecteur gauche suivant (LSV) et un vecteur droit suivant (RSV) sont produits à partir des informations VSLI.

24. Procédé selon la revendication 22, dans lequel, lorsque le signal audio mélangé-abaissé encodé est monophonique, les informations VSLI comprennent en outre un angle de vecteur global (Ga), et un vecteur global (GV) est produit à partir de Ga.

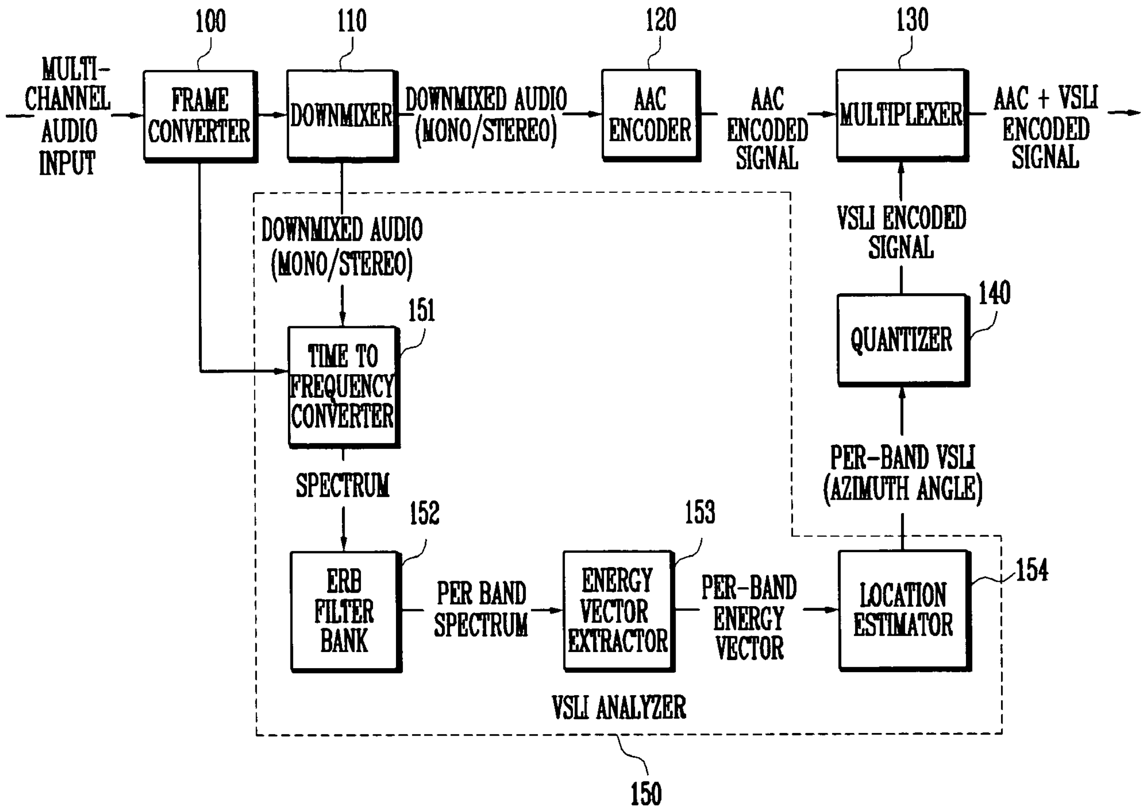
25. Procédé selon la revendication 23, dans lequel ladite étape de calcul du gain de canal comprend, pour chaque bande, les étapes consistant à :

calculer les amplitudes du vecteur LSV et du vecteur RSV en utilisant une amplitude du signal audio mélangé-abaissé;
calculer un premier gain d'un canal central (C) et une amplitude du vecteur LHV en utilisant l'amplitude du vecteur LSV et l'angle LSa ;
calculer un deuxième gain d'un canal central (C) et une amplitude du vecteur RHV en utilisant l'amplitude du vecteur RSV et l'angle RSa ;
sommer les premier et deuxième gains du canal central (C) pour produire un gain du canal central (C) ;
calculer les gains d'un canal gauche avant (L) et d'un canal gauche suivant (LS) en utilisant l'amplitude du vecteur LHV et l'angle LHa ; et
calculer les gains d'un canal droit avant (R) et d'un canal droit suivant (RS) en utilisant l'amplitude du vecteur RHV et l'angle RHa.

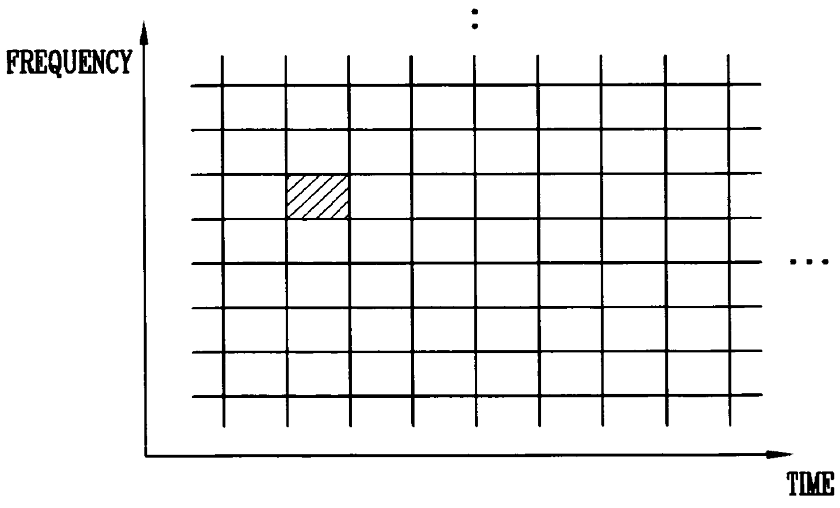
26. Support d'enregistrement pouvant être lu par un ordinateur mémorisant un programme d'ordinateur pour effectuer le procédé d'encodage d'un signal audio multicanal selon l'une quelconque des revendications 13 à 21.

27. Support d'enregistrement pouvant être lu par un ordinateur mémorisant un programme d'ordinateur pour effectuer le procédé de décodage d'un signal audio multicanal selon l'une quelconque des revendications 22 à 25.

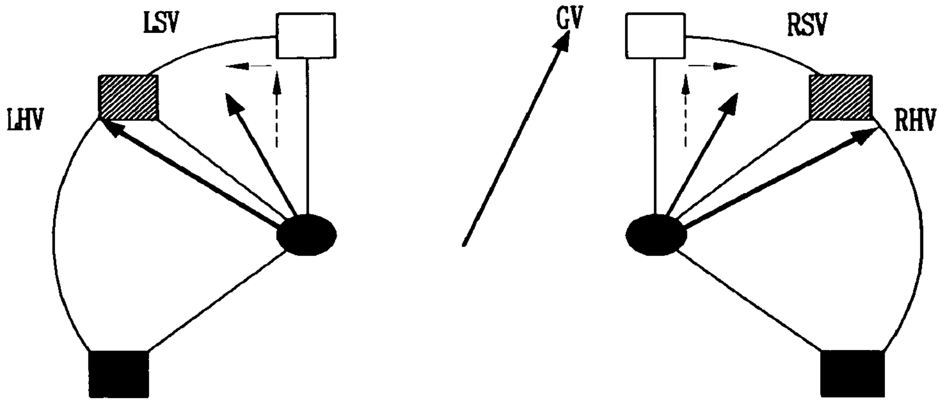
[Fig. 1]


100
110
120
130
MULTI-CHANNEL AUDIO INPUT
FRAME CONVERTER
DOWNMIXER
DOWNMIXED AUDIO (MONO/STEREO)
AAC ENCODER
AAC ENCODED SIGNAL
MULTIPLEXER
AAC + VSLI ENCODED SIGNAL
DOWNMIXED AUDIO (MONO/STEREO)
VSLI ENCODED SIGNAL
151
TIME TO FREQUENCY CONVERTER
QUANTIZER
140
SPECTRUM
PER-BAND VSLI (AZIMUTH ANGLE)
152
153
ERB FILTER BANK
PER BAND SPECTRUM
ENERGY VECTOR EXTRACTOR
PER-BAND ENERGY VECTOR
LOCATION ESTIMATOR
154
VSLI ANALYZER
150

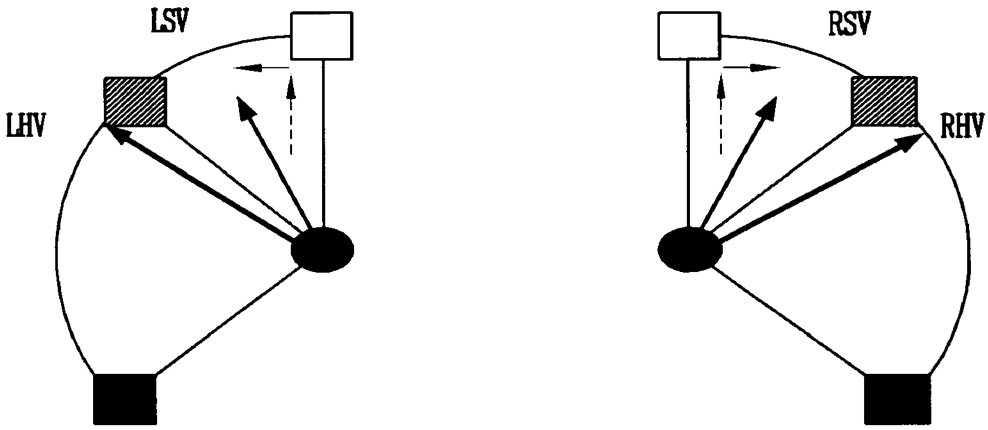
[Fig. 2]


FREQUENCY
TIME

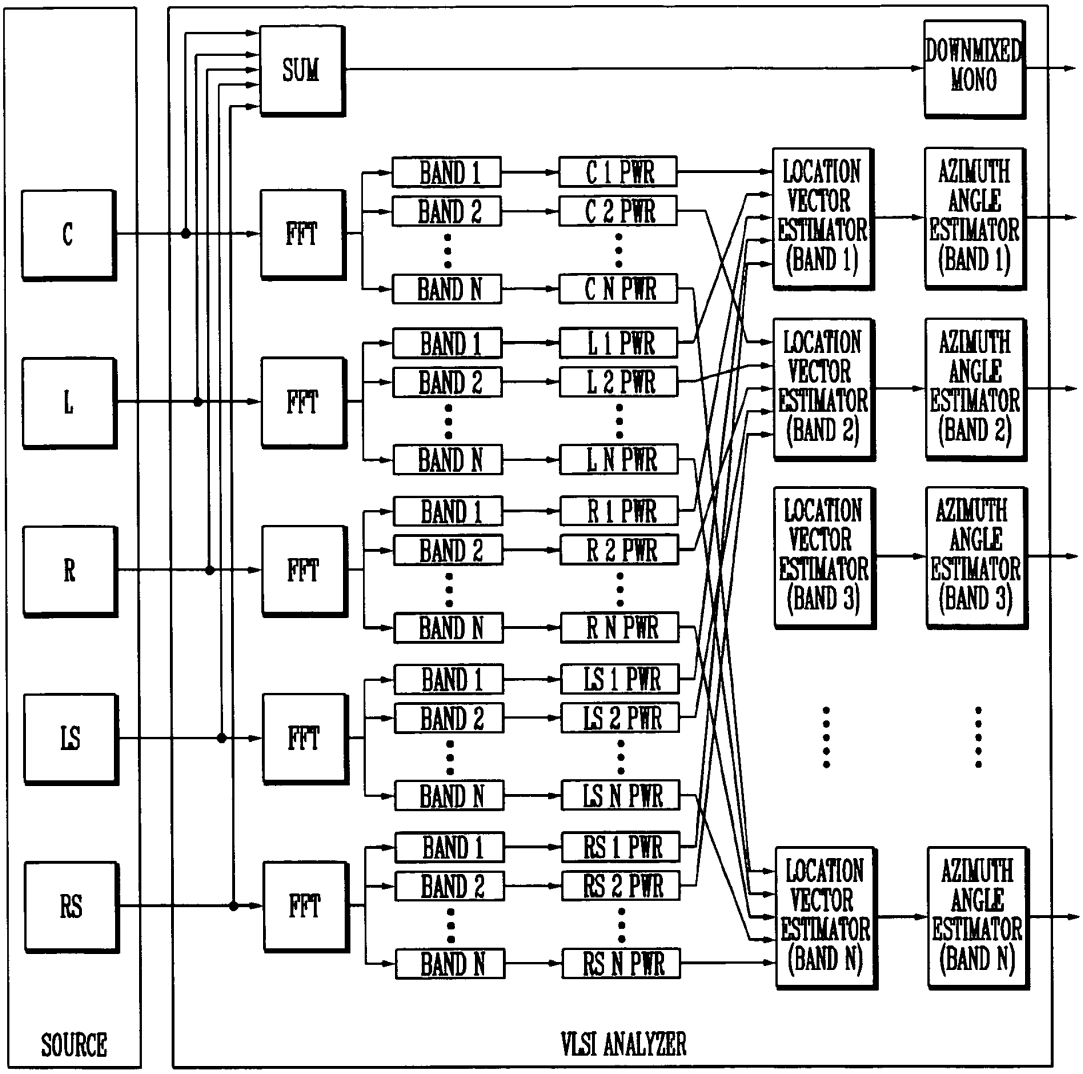
[Fig. 3]



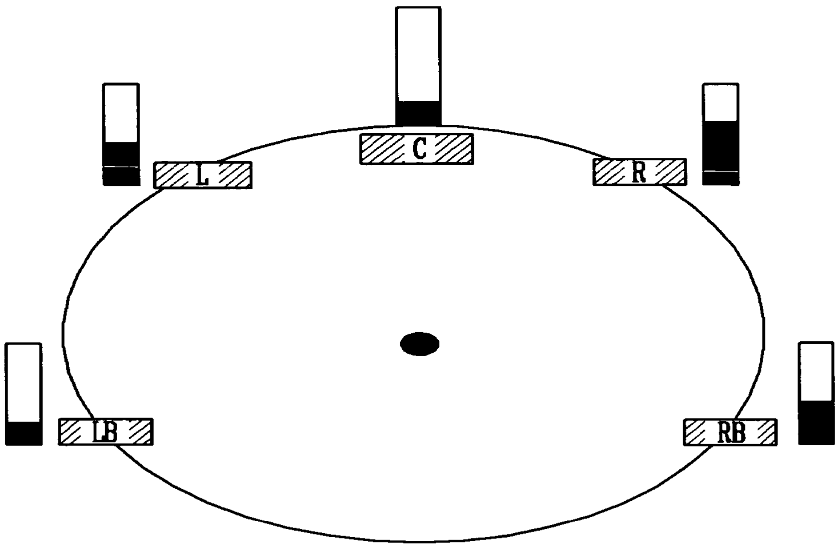
[Fig. 4]



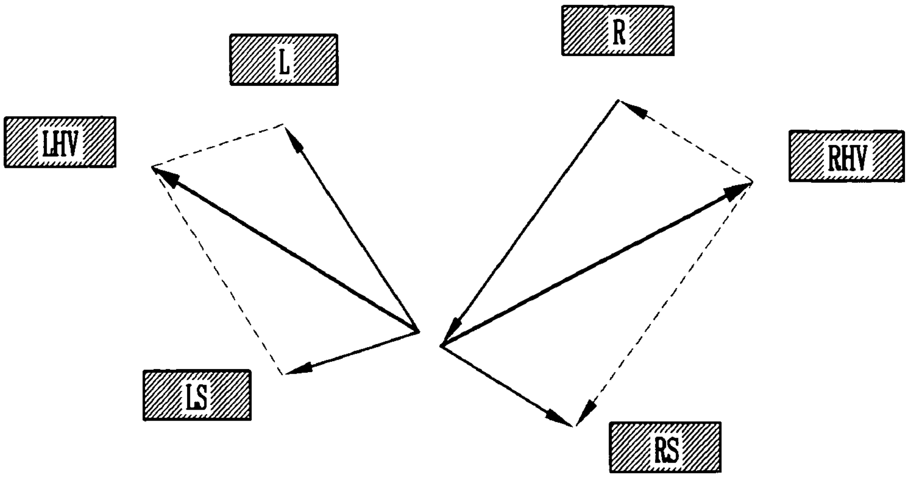
[Fig. 5]


SUM
DOWNMIXED MONO
C
L
R
LS
RS
FFT
BAND 1
BAND 2
BAND N
C 1 PWR
C 2 PWR
C N PWR
L 1 PWR
L 2 PWR
L N PWR
R 1 PWR
R 2 PWR
R N PWR
LS 1 PWR
LS 2 PWR
LS N PWR
RS 1 PWR
RS 2 PWR
RS N PWR
LOCATION VECTOR ESTIMATOR (BAND 1)
LOCATION VECTOR ESTIMATOR (BAND 2)
LOCATION VECTOR ESTIMATOR (BAND 3)
LOCATION VECTOR ESTIMATOR (BAND N)
AZIMUTH ANGLE ESTIMATOR (BAND 1)
AZIMUTH ANGLE ESTIMATOR (BAND 2)
AZIMUTH ANGLE ESTIMATOR (BAND 3)
AZIMUTH ANGLE ESTIMATOR (BAND N)
SOURCE
VLSI ANALYZER

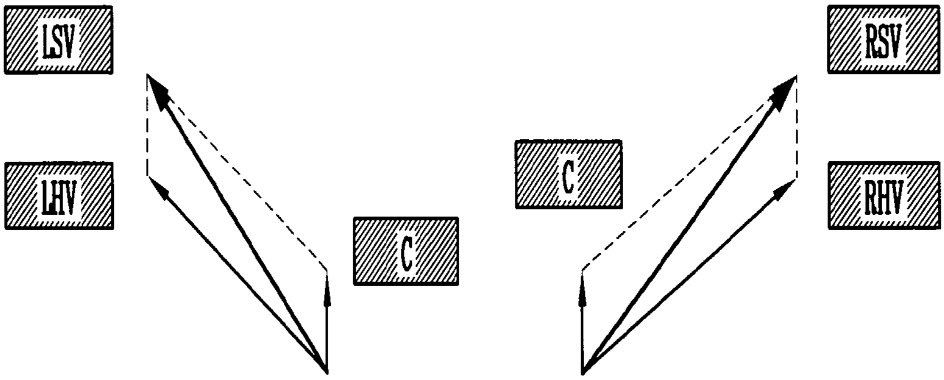
[Fig. 6]


C
L
R
LB
RB

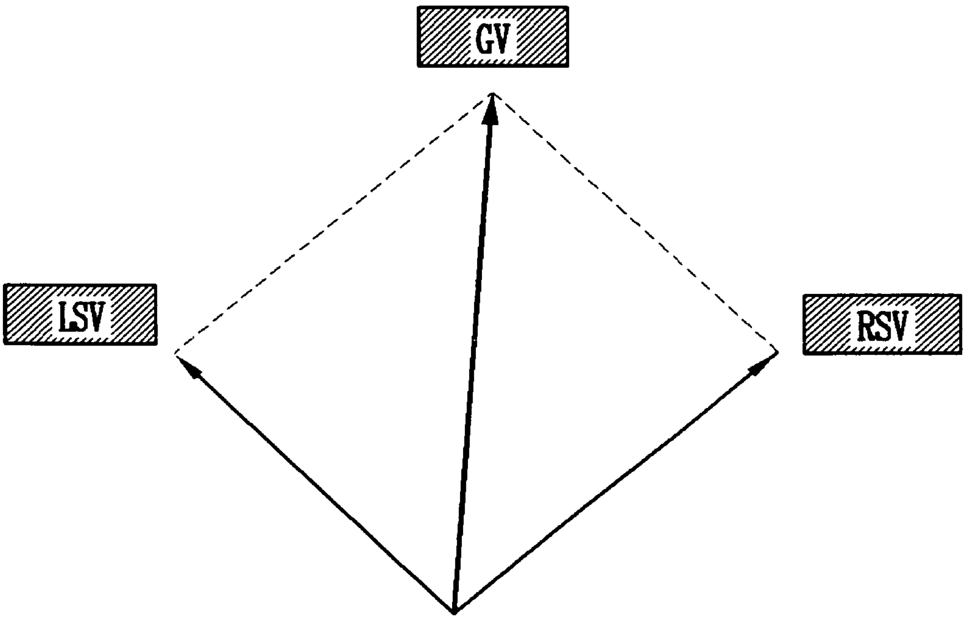
[Fig. 7]



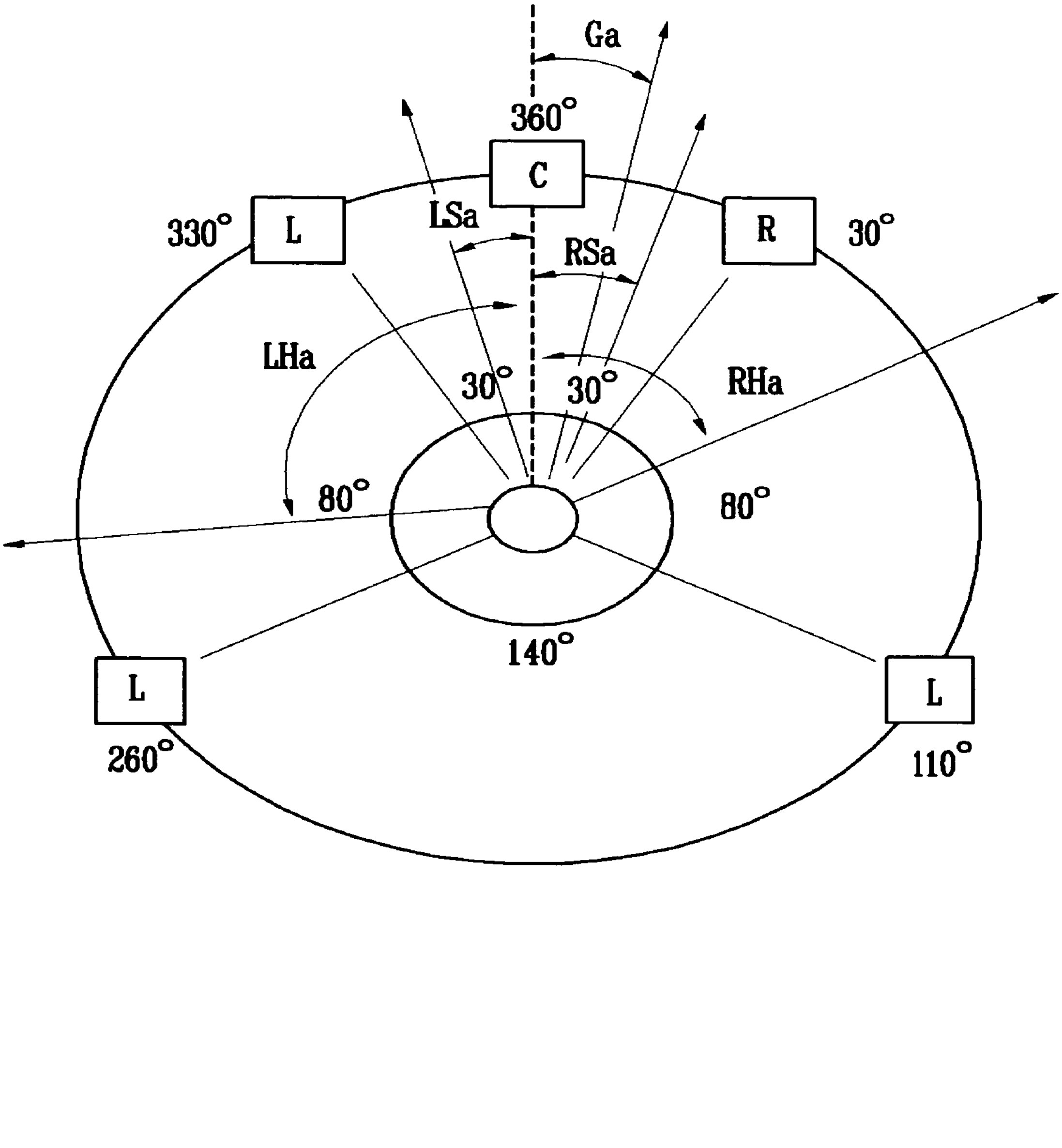
[Fig. 8]



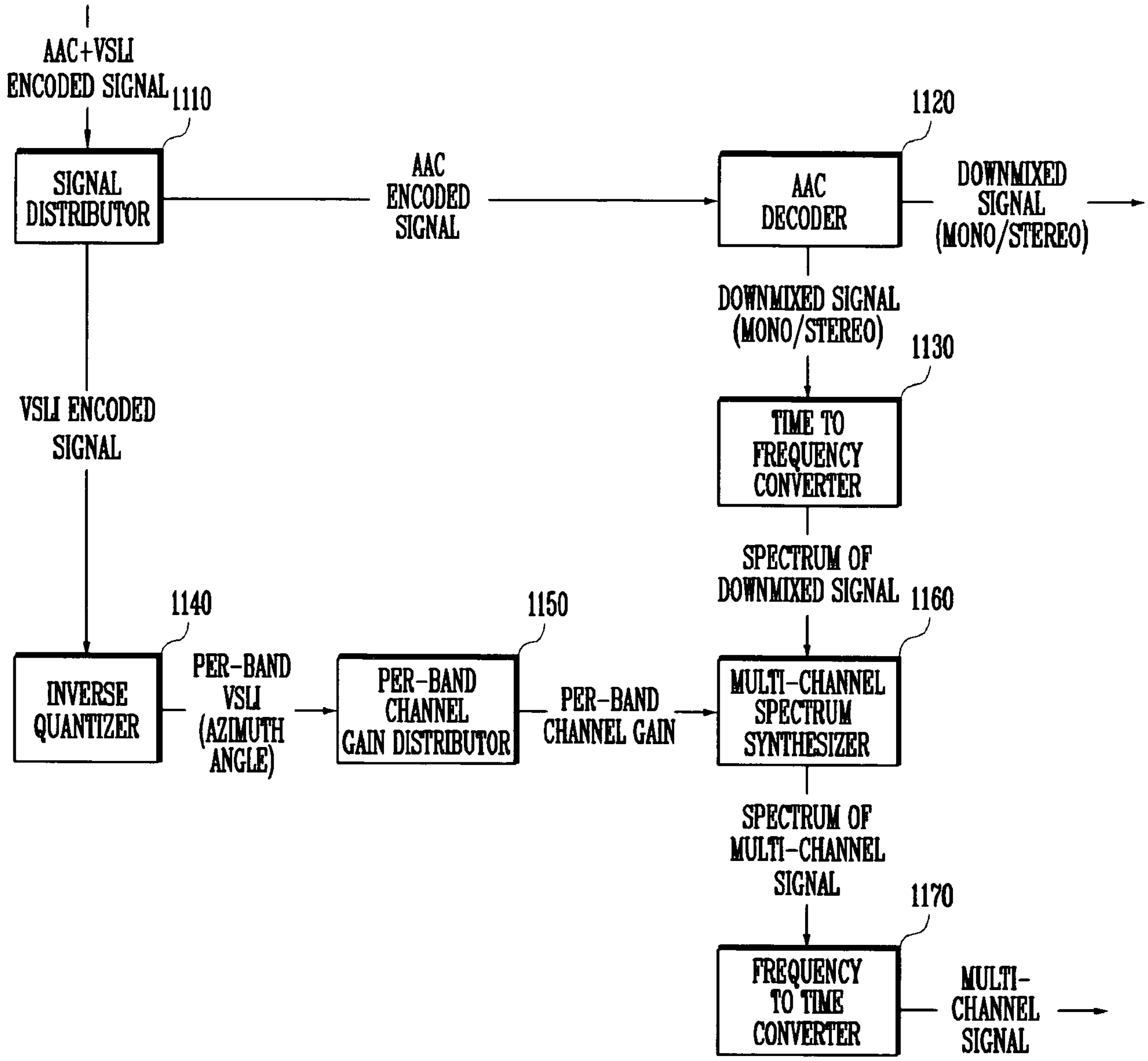
[Fig. 9]



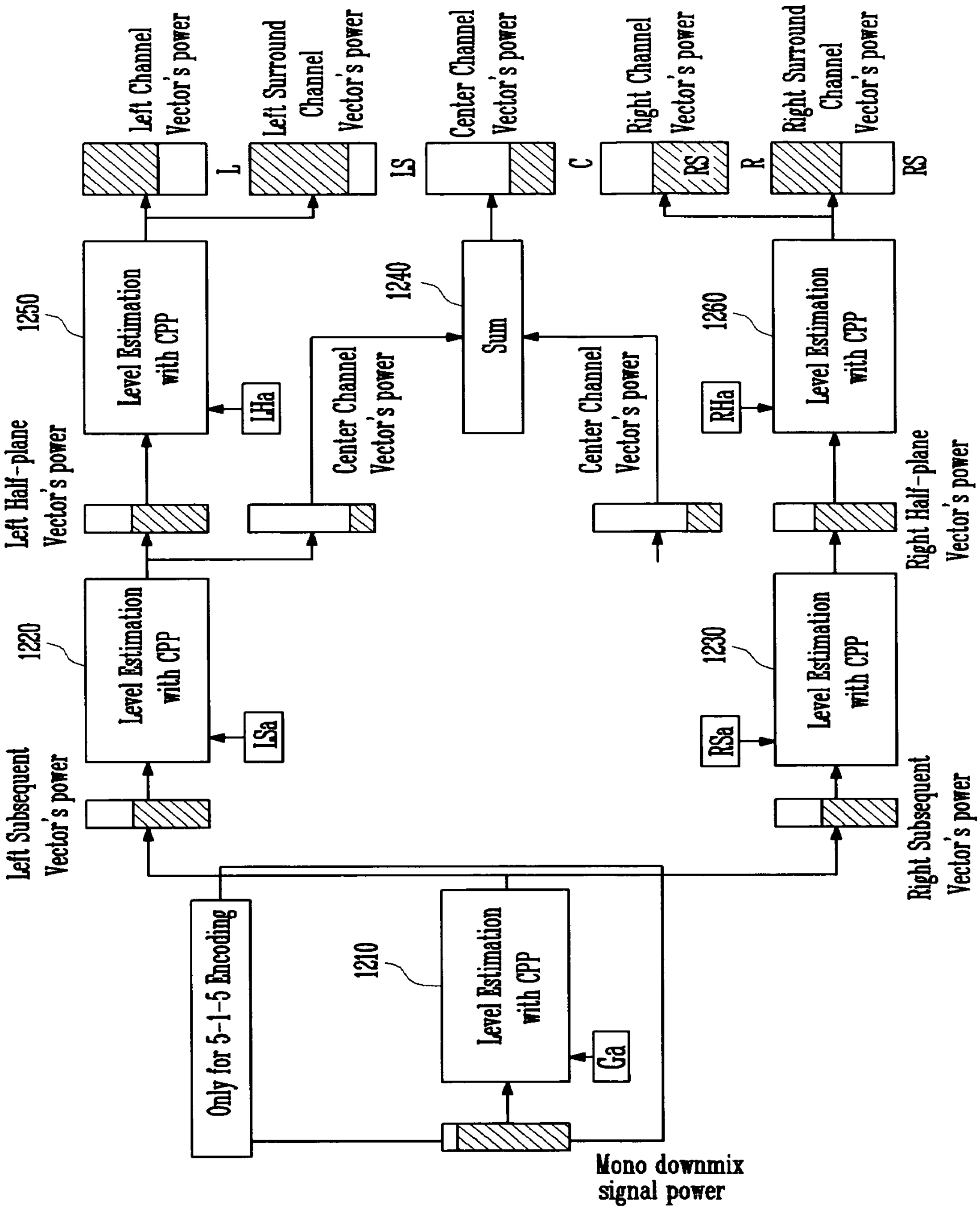
[Fig. 10]



[Fig. 11]


AAC+VSLI ENCODED SIGNAL
1110
SIGNAL DISTRIBUTOR
AAC ENCODED SIGNAL
1120
AAC DECODER
DOWNMIXED SIGNAL (MONO/STEREO)
DOWNMIXED SIGNAL (MONO/STEREO)
1130
TIME TO FREQUENCY CONVERTER
VSLI ENCODED SIGNAL
SPECTRUM OF DOWNMIXED SIGNAL
1140
INVERSE QUANTIZER
PER-BAND VSLI (AZIMUTH ANGLE)
1150
PER-BAND CHANNEL GAIN DISTRIBUTOR
PER-BAND CHANNEL GAIN
1160
MULTI-CHANNEL SPECTRUM SYNTHESIZER
SPECTRUM OF MULTI-CHANNEL SIGNAL
1170
FREQUENCY TO TIME CONVERTER
MULTI-CHANNEL SIGNAL

[Fig. 12]


Left Channel Vector's power
Left Surround Channel Vector's power
Center Channel Vector's power
Right Channel Vector's power
Right Surround Channel Vector's power
L
LS
C
RS
R
RS
1250
Level Estimation with CPP
LHa
1240
Sum
Center Channel Vector's power
Center Channel Vector's power
1260
Level Estimation with CPP
RHa
Left Half-plane Vector's power
Right Half-plane Vector's power
1220
Level Estimation with CPP
LSa
1230
Level Estimation with CPP
RSa
Left Subsequent Vector's power
Right Subsequent Vector's power
Only for 5-1-5 Encoding
1210
Level Estimation with CPP
Ga
Mono downmix signal power

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 9952326 A **[0008]**
- WO 03090208 A **[0009]**