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

(11) EP 2 022 263 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
01.08.2012 Bulletin 2012/31

(21) Application number: **07746543.3**

(22) Date of filing: **16.05.2007**

(51) Int Cl.:
H04N 7/00 *(2011.01)* ***H04S 7/00*** *(2006.01)*
H04H 20/89 *(2008.01)*

(86) International application number:
PCT/KR2007/002395

(87) International publication number:
WO 2007/136187 (29.11.2007 Gazette 2007/48)

(54) **OBJECT-BASED 3-DIMENSIONAL AUDIO SERVICE SYSTEM USING PRESET AUDIO SCENES**

AUF OBJEKTEN BASIERENDES DREIDIMENSIONALES AUDIODIENSTSYSTEM MIT VOREINGESTELLTEN AUDIOSZENEN

SYSTEME DE SERVICE AUDIO TRIDIMENSIONNEL FONDE SUR L'OBJET UTILISANT DES SCENES AUDIO FIXEES PREALABLEMENT

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

(30) Priority: **19.05.2006 KR 20060045184**

(43) Date of publication of application:
11.02.2009 Bulletin 2009/07

(60) Divisional application:
11159156.6 / 2 369 836
12171933.0

(73) Proprietor: **Electronics and Telecommunications Research Institute**
Yuseong-gu
Daejon 305-350 (KR)

(72) Inventors:
- **LEE, Yong-Ju**
  Daejon 305-345 (KR)
- **LEE, Tae-Jin**
  Daejon 305-752 (KR)
- **YOO, Jae-Hyoun**
  Daejon 305-350 (KR)
- **KANG, Kyeong-Ok**
  Daejon 305-727 (KR)
- **HONG, Jin-Woo**
  Daejon 305-333 (KR)
- **JANG, In-Seon**
  Daejon 305-345 (KR)
- **SEO, Jeong-Il**
  Daejon 305-728 (KR)
- **JANG, Dae-Young**
  Daejon 305-768 (KR)

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

(56) References cited:
WO-A2-2004/051624 **JP-A- 2000 358 294**
US-A- 5 590 207 **US-A1- 2004 111 171**

- **TAEJIN LEE, GI YOON PARK, INSEON JANG, KYEONGOK KANG: "An Object-based 3D Audio Broadcasting System for Interactive Service." CONVENTION PAPER 6384 118TH CONVENTION AUDIO ENGINEERING SOCIETY 31 May 2005 (2005-05-31), XP002577516 Retrieved from the Internet: URL:http://www.aes.org/tmpFiles/elib/20100 413/13100.pdf> [retrieved on 2010-04-12]**
- **KALVA H ET AL: "DELIVERING OBJECT-BASED AUDIO-VISUAL SERVICES" IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, IEEE SERVICE CENTER, NEW YORK, NY, US LNKD- DOI:10.1109/30.809189, vol. 45, no. 4, 1 November 1999 (1999-11-01), pages 1108-1111, XP000928095 ISSN: 0098-3063**

Printed by Jouve, 75001 PARIS (FR)

(Cont. next page)

- "Information technology â Coding of audio-visual objects â Part 20 : LASeR (Lightweight Applications Scene Representation)" 3GPP DRAFT; 29N66983, 3RD GENERATION PARTNERSHIP PROJECT (3GPP), MOBILE COMPETENCE CENTRE ; 650, ROUTE DES LUCIOLES ; F-06921 SOPHIA-ANTIPOLIS CEDEX ; FRANCE, vol. SA WG4, no. San Diego, USA; 20050429, 29 April 2005 (2005-04-29), XP050287655 [retrieved on 2005-04-29]
- AVARO O ET AL: "The MPEG-4 systems and description languages: A way ahead in audio visual information representation" SIGNAL PROCESSING. IMAGE COMMUNICATION, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL LNKD- DOI: 10.1016/S0923-5965(97)00027-1, vol. 9, no. 4, 1 May 1997 (1997-05-01), pages 385-431, XP004075337 ISSN: 0923-5965

Description

TECHNICAL FIELD

[0001] The present invention relates to an object-based three dimensional (3-D) audio service system using preset audio scenes and a method thereof; and, more particularly, to an object-based 3-D audio service system using preset audio scenes and a method thereof for providing an interactive service that enables a user or a viewer to directly form an audio scene using a 3-D audio related technology for providing realistic broadcasting to a user or a viewer.

BACKGROUND ART

[0002] Fig. 1 is a diagram illustrating a conventional audio service system.

[0003] As shown in Fig. 1, the conventional audio service system includes an audio service providing apparatus 10 and an audio service reproducing apparatus 20. The audio service providing apparatus 10 includes an audio-capture unit 11 for capturing an audio signal such as sound, an editing/mixing unit 12 for editing and mixing the captured audio signal to transmit the audio signal to an audio service reproducing apparatus 20, and a storing/transmitting unit 13 for storing the mixed audio signal and transmitting the mixed audio signal to the audio service reproducing apparatus 20.

[0004] The audio service reproducing apparatus 20 includes a receiver 21 for receiving an audio signal transmitted from the audio service providing apparatus 10, a controller 22 for controlling the received audio signal, and a reproducer 23 for reproducing an audio signal.

[0005] An audio signal, which is provided through broadcasting services such as TV broadcasting, radio broadcasting, and Digital Multimedia Broadcasting (DMB) based on the conventional audio service system, is generally created by mixing a plurality of audio signals captured from various sound sources. For example, an audio signal provided through a soccer game broadcasting is created by mixing noises in a soccer stadium, yelling of a crowd, and a voice of an announcer.

[0006] Although a user or a viewer can control the volume of the overall audio signal, it is impossible to control the volume of each object such as the voice of an announcer, the yelling of a crowd, and the noises of the soccer stadium. It is because the audio signal is transmitted after a plurality of object audio signals are mixed into one audio signal in a general broadcasting service.

[0007] However, if a transmitter such as the audio service providing apparatus 10 independently transmits object audio signals of the sound sources without the object audio signals of the sound sources mixed to one audio signal, a receiver such as the audio service reproducing apparatus 20 can independently control the volumes of the object audio signals of the sound sources. An object-based audio service denotes such an audio service that allows a user or a viewer to control each of the object audio signals at a receiver by independently transmitting the object audio signals of the sound sources through a transmitter.

[0008] For example, if an audio signal of a soccer game broadcasting is provided based on an object-based 3-D audio service, a user or a viewer can control each of objects, such as the noises in the soccer stadium, the yelling of the crowd, and the voices of an announcer to obtain a desired audio setting. That is, a user or a viewer can control the noise of the soccer stadium loud, the yelling of the crowd soft, and the voice of the announcer loud. Or, a viewer can control the audio signal to reproduce only the noises of the soccer stadium and the voice of an announcer without the yelling of the crowd reproduced.

[0009] Therefore, there is a great demand for developing a method for providing an object-based 3-D audio service that enables a user to control each of object audio signals of sound sources, which can be applied to all broadcasting services and multimedia services providing audio such as digital broadcasting, radio broadcasting, Digital Multimedia Broadcasting, Internet broadcasting, digital movie, DVD, moving picture contents.

[0010] Although a conventional object-based 3-D audio system and a control method thereof was introduced in Korean Patent Publication No. 10-2004-0037437, published on May 7th 2004, the conventional object-based 3-D audio system requires a user to control each of object audio signals of sound sources to set the audio signals according to user's preference. Therefore, it is very annoying to a user or a viewer.

[0011] Taejin Lee et al., "An Object-based 3D Audio Broadcasting System for Interactive Services", Audio Engineering Society, Convention paper 6384, 2005, May 28-31, Barcelona, Spain, disclose an object-based 3D audio broadcasting system. This system consists of an authoring tool, a streaming server, and a client. The authoring tool generates an MPEG-4 file, made of multiple audio objects, after adapting several kinds of acoustical effects to the audio objects. The streaming server generates transmission packets and sends them to the client through the Internet. The client reconstructs the bitstream and plays the 3D audio with a user interaction. The authors present the design and implementation method of an object-based 3D audio system and describe the simulation result and applications thereof.

DISCLOSURE

TECHNICAL PROBLEM

[0012] An embodiment of the present invention is directed to providing an object-based three dimensional (3-D) audio service system and a method thereof for enabling a user to easily and conveniently watch and listen an object-based 3-D audio service by eliminating inconvenience that requires a user to control each of object

audio signals of sound sources.

[0013] Other objects and advantages of the present invention can be understood by the following description, and become apparent with reference to the embodiments of the present invention. Also, it is obvious to those skilled in the art of the present invention that the objects and advantages of the present invention can be realized by the means as claimed and combinations thereof.

TECHNICAL SOLUTION

[0014] The present invention is defined in the independent claims. The dependent claims define embodiments thereof.

[0015] In accordance with an embodiment , there is provided an object-based three dimensional (3-D) audio service providing apparatus using preset audio scenes, including: audio input means for inputting an audio signal; preset audio scene generating means for extracting object audio signals from the audio signal inputted through the audio input means and generating more than one of 3-D audio scene information by arranging the extracted object audio signals in a 3-D space and editing features of each object; and encoding means for encoding and multiplexing the audio signal and the 3-D audio scene information for each object audio signal.

[0016] In accordance with another embodiment , there is provided an object-based 3-D audio service reproducing apparatus using preset audio scenes including: decoding means for de-multiplexing and decoding object-based 3-D audio contents; audio scene forming means for forming 3-D audio scene information according to one selected from a plurality of 3-D audio scene information in the de-multiplexed and decoded object-based 3-D audio contents by a user including a viewer; audio signal mixing means for controlling features of objects in an audio signal of the de-multiplexed and decoded object-based 3-D audio contents according to the formed 3-D audio scene information; and reproducing means for reproducing the audio signal with one of the features controlled.

[0017] In accordance with another embodiment , there is provided a method for providing an object-based 3-D audio service using preset audio scenes, including the steps of: inputting an audio signal; extracting object audio signals from the inputted audio signal and generating more than one of 3-D audio scene information by arranging the extracted object audio signals in a 3-D space and editing features of each object; and encoding and multiplexing the audio signal and the 3-D audio scene information for each object audio signal.

[0018] In accordance with another embodiment , there is provided a method for reproducing object-based 3-D audio service using preset audio scenes including the steps of: de-multiplexing and decoding object-based 3-D audio contents; forming 3-D audio scene information according to one selected from a plurality of 3-D audio scene information in the de-multiplexed and decoded object-based 3-D audio contents by a user including a viewer; controlling features of objects in an audio signal of the de-multiplexed and decoded object-based 3-D audio contents according to the formed 3-D audio scene information; and reproducing the audio signal with one of the features controlled.

ADVANTAGEOUS EFFECTS

[0019] An object-based three dimensional (3-D) audio service system and a method thereof according to the present invention provides previously generated preset audio scenes to a user or a viewer with an object-based 3-D audio service applied to all broadcasting services and multimedia services providing audio, such as digital broadcasting, radio broadcasting, Digital Multimedia Broadcasting (DMB), Internet broadcasting, digital movies, Digital Video Disk (DVD), and moving picture contents. Therefore, the object-based 3-D audio service system and a method thereof according to the present invention eliminates the inconvenience of a user to control each of object audio signals of sound sources and enables the user to easily and conveniently watch and listen the object-based 3-D audio service.

[0020] The present invention can be applied to broadcasting services and multimedia services providing audio, such as digital broadcasting, radio broadcasting, DMB, Internet broadcasting, digital movies, DVD, and moving picture contents, and the present invention is not limited to the types of mediums for transmitting and storing object-based audio contents for broadcasting and multimedia services providing audio.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021]

Fig. 1 is a diagram illustrating a conventional audio service system.

Fig. 2 is a block diagram illustrating an object-based three-dimensional (3-D) audio service system using preset audio scenes in accordance with an embodiment of the present invention.

Fig. 3 is a flowchart illustrating a method for providing an object-based 3-D audio service using preset audio scenes in accordance with an embodiment of the present invention.

Fig. 4 is a flowchart illustrating a method for reproducing an object-based 3-D audio service using preset audio scenes in accordance with the embodiment of the present invention.

BEST MODE FOR THE INVENTION

[0022] The advantages, features and aspects of the invention will become apparent from the following description of the embodiments with reference to the accompanying drawings, which is set forth hereinafter.

[0023] Fig. 2 is a block diagram illustrating an object-based three-dimensional (3-D) audio service system using preset audio scenes in accordance with an embodiment of the present invention.

[0024] As shown in Fig. 2, the object-based 3-D audio service system includes an object-based 3-D audio service providing apparatus 30, a transmitting medium 50, and an object-based 3-D audio service reproducing apparatus 40. The 3-D service providing apparatus 30 receives an audio signal through various input devices, creates more than one of object-based 3-D audio scene information which can be selected by a user or a viewer, and transmits the created object-based 3-D audio scene information to the object-based 3-D audio service reproducing apparatus 40. The transmitting medium 50 is a medium such as a digital broadcasting network or an Internet network for connecting the object-based 3-D audio service providing apparatus 30 and the object-based 3-D audio service reproducing apparatus 40 through a network. The object-based 3-D audio service reproducing apparatus 40 generates more than one of 3-D audio scenes based on the object-based 3-D audio scene information transmitted from the object-based 3-D audio service providing apparatus 30.

[0025] Hereinafter, the constituent elements of the object-based 3-D audio service system using preset audio scenes according to the present embodiment will be described in detail.

[0026] The object-based 3-D audio service providing apparatus 30 includes an input unit 31, a preset audio scene generator 32, an encoder 33, and a transmitter 34. The input unit 31 receives audio signals through various input devices. The preset audio scene generator 32 extracts object-based audio signals (hereinafter, object audio signals) from the audio signal received through the input unit 31, arranges the extracted object audio signals in a three dimensional space, and creates more than one of 3-D audio scene information by editing features such as a location, a size, a direction, and a sound field environment of each object. The encoder 33 encodes and multiplexes the audio signal inputted through the input unit 31 and the object-based 3-D audio scene information created by the preset audio scene generating unit 32 for transmitting the input audio signal and the generated preset audio scene information to the object-based 3-D audio service reproducing apparatus 40. For example, the input audio signals and the generated preset audio scene information are multiplexed to a moving picture experts group 4 (MPEG-4) file format in a digital broadcasting network. The transmitter 34 transforms the multiplexed object-based audio contents including the input audio signal and the created object-based 3-D audio scene information from the encoding unit 33 to a transport format. For example, the transmitter 34 transforms the multiplexed object-based audio contents to a MPEG-2 transport stream (TS) for a digital broadcasting network.

[0027] The transformed object-based audio contents including the input audio signal and the generated object-based 3-D audio scene information may be transmitted to the object-based 3-D audio reproducing apparatus 40 and may be stored in a storing medium.

[0028] The transmitter 34 may transmit the object-based audio contents including the input audio signal and the object-based 3-D audio scene information to the object-based 3-D audio reproducing apparatus 40 through a digital broadcasting network such as a terrestrial DMB channel 50.

[0029] If the sound source of the audio signal inputted to the input unit 31 is a mixed sound source, the preset audio scene generator 32 uses a Convolutive Blind Source Separation technique to extract object audio signals. Especially, the preset audio scene generator 32 forms more than one of object-based 3-D audio scene information by controlling a ratio of each object-based the audio scene information of each object audio signal, which is set according to the control of a user such as an editor.

[0030] The object-based 3-D audio service reproducing apparatus 40 includes a decoder 42, an audio scene information forming unit 43, an audio signal mixer 44, and an audio signal reproducer 45. The decoder 42 de-multiplexes and decodes object-based audio contents including an audio signal and object-based 3-D audio scene information for reproducing. The audio scene information forming unit 43 provides the object-based 3-D audio scene information of the object-based 3-D audio contents, which is de-multiplexed and decoded by the decoder 42, to a user such as a viewer to select, and forms the object-based 3-D audio scene information according to the user selection. The audio signal mixer 44 mixes object audio signals of the audio signal of the de-multiplexed and decoded object-based 3-D audio contents from the decoder 42 by controlling features of each object, such as a location, a direction, a size, and a sound field of each object according to the object-based 3-D audio scene information formed by the audio scene information forming unit 43. The audio signal reproducer 45 reproduces the audio signal mixed to one object-based 3-D audio scene by the audio signal mixer 44.

[0031] The object-based audio contents including the audio signal and the object-based 3-D audio scene information may be provided through a broadcasting service or a multimedia service such as digital broadcasting, radio broadcasting, Digital Multimedia Broadcasting (DMB), Internet broadcasting, digital movies, Digital Video Disk (DVD), and moving picture contents. Although the object-based audio contents may be received through the receiver 41 in the present embodiment, the present invention is not limited thereto. That is, the object-based audio contents may be provided through a transmission medium or a storage medium that can provide a broadcasting service or a multimedia service that provides an audio.

[0032] The audio scene information forming unit 43 enables a user or a viewer to select features of objects such as a location, a direction, a volume, and a sound field

environment of each object and forms new object-based 3-D audio scene information according to the features including a location, a direction, a volume, and a sound field environment of each object set by the user.

[0033] A user or a viewer can control features of a 3-D audio space by changing a reverberation time of a 3-D space through controlling a volume and a delay time of an initial reflected sound through the audio scene information forming unit 43.

[0034] That is, the object-based 3-D audio service system using the preset audio scene according to the present embodiment previously generates object-based 3-D audio scenes that are expected to be frequently used and provides the generated object-based 3-D audio scenes as preset audio scenes to a user or a viewer. That is, the object-based 3-D audio service system according to the present embodiment enables a user or a viewer to select one of the preset audio scenes in order to make a user to conveniently watch and listen a broadcasting program with the desired audio preference.

[0035] For example, noises of a soccer stadium, yelling of a crowd, a voice of an announcer are defined as audio objects for a soccer game broadcasting, and the defined audio objects are transmitted independently. With the audio objects, a first audio scene having information about volume of the noises of soccer stadium, the yelling of a crowd, and the voice of an announcer set to 1:1:1, a second audio scene having information about volume of the noises of a soccer stadium, the yelling of a crowd, and the voice of an announcer set to 1:0.5:1, and an audio scene having information about volume of the noises of a soccer stadium, the yelling of a crowd, and the voice of an announcer set to 1:0:1 are transmitted as the preset audio scenes. Then, a user or a viewer selects one of the preset audio scenes to watch and listen the soccer game broadcasting with the desired audio preference.

[0036] A user may directly control each of the audio objects if the user cannot find a desired audio scene from the provided audio scenes. However, it is preferable to provide a large number of preset audio scenes to a user in order to enable the user to find a desired audio scene from the provided preset audio scenes.

[0037] Fig. 3 is a flowchart illustrating a method for providing an object-based 3-D audio service using preset audio scenes in accordance with an embodiment of the present invention.

[0038] Referring to Fig. 3, the input unit 31 of the object-based 3-D audio service providing apparatus 30 receives an object-based audio signal through various input devices at step S301.

[0039] The preset audio scene generator 32 extracts object-based audio signals, that is, object audio signals, from the audio signal inputted through the input unit 31 at step S302. Then, the preset audio scene generator 32 generates more than one of object-based 3-D audio scene information at step S304 by arranging the extracted object audio signals in a 3-D space and editing the features of each object audio signal such as a location, a direction, a volume, and a sound field environment of the audio object at step S303. The encoder 33 encodes and multiplexes the audio signal inputted through the input unit 31 and the object-based 3-D audio scene information generated by the preset audio scene generator 32 at step S305. For example, the encoder 33 encodes and multiplexes the audio signal and the object-based 3-D audio scene information into MPEG-4 file format for a digital broadcasting network.

[0040] Then, the transmitter 34 transforms the multiplexed object-based audio contents including the audio signal and the object-based 3-D audio scene information to be proper to a transport format and transmits the transformed object-based audio contents at step S306. For example, the multiplexed object-based audio contents are transformed to a MPEG-2 TS in a digital broadcasting network.

[0041] For example, the transmitter 34 transmits the transformed object-based audio contents including the audio signal and the object-based 3-D audio scene information to the object-based 3-D audio reproducing apparatus 40 through a digital broadcasting network such as a terrestrial DMB channel. The transformed object-based audio contents including the audio signal and the object-based 3-D audio scene information may be stored in a storing medium.

[0042] Fig. 4 is a flowchart illustrating a method for reproducing an object-based 3-D audio service using preset audio scenes in accordance with an embodiment of the present invention.

[0043] Referring to Fig. 4, the receiver 41 of the object-based 3-D audio service reproducing apparatus 40 receives the object-based audio contents including an audio signal and object-based 3-D audio information through, for example, a digital broadcasting network such as a terrestrial DMB channel 50 or the Internet network at step S401.

[0044] The receiver 41 may receive the object-based audio contents through a transmission medium that can provide a broadcasting service or a multimedia service that provides an audio. Or, the object-based audio contents may be inputted through the storing medium.

[0045] The decoder 42 de-multiplexes and decodes the received or inputted object-based audio contents including the audio signal and the object-based 3-D audio scene information at step S402. The audio scene information forming unit 43 provides the object-based 3-D audio scene information of the de-multiplexed and decoded object-based 3-D audio contents to a user or a viewer to select, and forms object-based 3-D audio scene information according to the user selection at step S403.

[0046] Then, the audio signal mixer 44 mixes object audio signals by controlling features of objects in the audio signal of the de-multiplexed and decoded object-based 3-D audio contents, such as a location, a direction, a volume, and a sound field environment of each audio object, according to the object-based 3-D audio scene information formed by the audio scene information form-

ing unit 43 at step S404. Finally, the audio signal reproducer 45 reproduces the audio signal mixed based on one of the object-based 3-D audio scenes by the audio signal mixer 44 at step S405.

[0047] The above described method according to the present invention can be embodied as a program and stored on a computer readable recording medium. The computer readable recording medium is any data storage device that can store data which can be thereafter read by the computer system. The computer readable recording medium includes a read-only memory (ROM), a random-access memory (RAM), a CD-ROM, a floppy disk, a hard disk and an optical magnetic disk.

[0048] While the present invention has been described with respect to certain preferred embodiments, it will be apparent to those skilled in the art that various changes and modifications may be made without departing from the scope of the invention as defined in the following claims.

Claims

1. An object-based three dimensional (3-D) audio service providing apparatus (30) using preset audio scenes, comprising:

   audio input means (31) for inputting an audio signal;
   preset audio scene generating means (32) for extracting object audio signals from the audio signal inputted through the audio input means and generating more than one of 3-D audio scene information by arranging the extracted object audio signals in a 3-D space and editing features of each object based on the preset audio scenes; and
   encoding means (33) for encoding and multiplexing the audio signal and the 3-D audio scene information for each object audio signal,
   wherein the preset audio scene generating means (32) is adapted to generate more than one of the 3-D audio scene information by setting the ratios between the volumes of the object audio signals, based on audio scene information of each object audio signal, which ratios are set according to control of a user including a viewer.

2. The object-based 3-D audio service providing apparatus (30) of claim 1, further comprising processing means for processing the encoded and multiplexed object-based 3-D audio contents.

3. The object-based 3-D audio service providing apparatus (30) of claim 2, wherein the processing means is adapted to transmit the encoded and multiplexed object-based 3-D audio contents to an audio reproducing terminal through a digital broadcasting network.

4. The object-based 3-D audio service providing apparatus (30) of claim 1, wherein the audio scene generating means (32) is adapted to extract object audio signals using a Convolutive Blind Source Separation technique when a sound source of the input audio signal is a mixed sound source.

5. An object-based three dimensional (3-D) audio service reproducing apparatus (40) for reproducing object-based 3-D audio service provided by the object-based 3-D audio service providing apparatus (30) according to any one of claims 1 to 4, using preset audio scenes comprising:

   decoding means (42) for de-multiplexing and decoding object-based 3-D audio contents;
   audio scene selecting means (43) for selecting one 3-D audio scene information from a plurality of 3-D audio scene information based on the preset audio scenes in the de-multiplexed and decoded object-based 3-D audio contents by a user including a viewer; audio scene forming means for forming 3-D audio scene information according to the selected 3-D audio scene information;
   audio signal mixing means (44) for controlling features of objects in an audio signal of the de-multiplexed and decoded object-based 3-D audio contents according to the formed 3-D audio scene information; and
   reproducing means (45) for reproducing the audio signal with one of the features controlled.

6. A method for providing an object-based three dimensional (3-D) audio service using preset audio scenes, comprising the steps of:

   inputting an audio signal;
   generating (S304) more than one of 3-D audio scene information by extracting object audio signals (S302) from the inputted audio signal, arranging the extracted object audio signals in a 3-D space and editing (S303) features of each object based on the preset audio scenes; and
   encoding and multiplexing (S305) the audio signal and the 3-D audio scene information for each object audio signal,
   wherein in the step of generating (S304) 3-D audio scene information, more than one of the 3-D audio scene information is generated by setting the ratios between the volumes of the object audio signals, based on audio scene information of each object audio signal, which ratios are set according to control of a user including a viewer.

7. The method of claim 6, further comprising the step of:

processing the encoded and multiplexed object-based 3-D audio contents.

8. The method of claim 7, wherein in the step of processing the object-based 3-D audio contents, the encoded and multiplexed object-based 3-D audio contents are transmitted through a digital broadcasting network,

9. A method for reproducing object-based three dimensional (3-D) audio service provided by the object-based 3-D audio service providing method according to any one of claims 6 to 8, using preset audio scenes comprising the steps of :

de-multiplexing and decoding (S402) object-based 3-D audio contents;
selecting (S403) one 3-D audio scene information from a plurality of 3-D audio scene information based on the preset audio scenes in the de-multiplexed and decoded object-based 3-D audio contents by a user including a viewer; forming (S403) 3-D audio scene information according to the selected 3-D audio scene information;
controlling (S404) features of objects in an audio signal of the de-multiplexed and decoded object-based 3-D audio contents according to the formed 3-D audio scene information; and
reproducing (S405) the audio signal with one of the features controlled.

10. The method of claim 9 or apparatus (40) of claim 5, wherein in the step of forming (S403) 3-D audio scene information, 3-D audio scene information is formed according to features of each object, which are set by a user including a viewer.

11. The apparatus (30; 40) of claim 1 or 5, or the method of claim 6 or 9, wherein the features include at least one among a location, a volume, a direction, and a sound field environment of each object.

12. The apparatus (40) of claim 5 or the method of claim 9, wherein a feature of a 3-D audio space is controlled by changing a reverberation time of a 3-D audio space through controlling a volume and a delay time of an initial reflected sound.

Patentansprüche

1. Eine objekt-basierte dreidimensionale (3-D) Audio-Service-Lieferungsvorrichtung (30), die voreingestellte Audioszenen verwendet, aufweisend:

Audio-Eingabemittel (31) zum Eingeben eines Audiosignals;
Erzeugungsmittel (32) zur Erzeugung voreingestellter Audioszenen zum Extrahieren von Objektaudiosignalen aus dem Audiosignal, das durch die Audioeingabemittel eingegeben wurde und zum Erzeugen von mehr als einer 3-D-Audioszeneninformation durch Anordnen der extrahierten Objektaudiosignale in einem 3-D-Raum und Editieren von Merkmalen von jedem Objekt basierend auf den voreingestellten Audio-Szenen; und
Codiermittel (33) zum Codieren und Multiplexen des Audiosignals und der 3-D-Audioszeneninformation für jedes Objektaudiosignal,
wobei die Erzeugungseinrichtung (32) für voreingestellte Audioszenen angepasst ist, mehr als eine der 3-D-Audioszeneninformation zu erzeugen durch Setzen des Verhältnisses zwischen den Lautstärken der Objektaudiosignale basierend auf Audioszeneninformation von jedem Objektaudiosignal,
wobei die Verhältnisse gesetzt werden in Übereinstimmung mit der Steuerung eines Benutzers einschließend einen Betrachter.

2. Die objekt-basierte 3-D-Audio-Service-Lieferungsvorrichtung (30) von Anspruch 1, ferner aufweisend: Verarbeitungsmittel zum Verarbeiten der codierten und gemultiplexten objektbasierten 3-D-Audioinhalte.

3. Die objekt-basierte 3-D-Audio-Service-Lieferungsvorrichtung (30) nach Anspruch 2, wobei die Verarbeitungsmittel angepasst sind, um die codierten und gemultiplexten objektbasierten 3-D-Audioinhalte an ein Audio-Wiedergabeterminal über ein digitales Broadcasting-Netzwerk zu übertragen.

4. Die objekt-basierte 3-D-Audio-Service-Lieferungsvorrichtung (30) nach Anspruch 1, wobei die Audioszenen-Erzeugungsmittel (32) angepasst sind, Objektaudiosignale unter Verwendung einer konvolutiven Blindquellenseparationstechnik zu extrahieren, wenn eine Schallquelle des Audiosignals eine gemischte Schallquelle ist.

5. Eine objekt-basierte dreidimensionale (3-D)-Audio-Service-Wiedergabe-vorrichtung (40) zum Wiedergeben von objektbasiertem 3-D Audioservice, der geliefert wird durch die objektbasierte 3-D-Audio-Service-Lieferungsvorrichtung (30) gemäß einem der Ansprüche 1 - 4, welche voreingestellte Audioszenen verwendet, aufweisend:

Decodiermittel (42) zum Demultiplexen und Decodieren objektbasierter 3-D-Audioinhalte;
Audioszenen-Auswahlmittel (43) zum Auswählen einer 3-D-Audioszeneninformation aus einer Mehrzahl von 3-D-Audioszeneninformationen basierend auf den voreingestellten Audioszе-

nen in den gedemultiplexten und decodierten objektbasierten 3-D-Audioinhalten durch einen Benutzer einschließend einen Betrachter;
Audioszenen-Formungsmittel zum Formen von 3-D-Audioszeneninformation gemäß der ausgewählten 3-D-Audioszeneninformation;
Audiosignal-Mischmittel (44) zum Steuern von Merkmalen von Objekten in einem Audiosignal der gedemultiplexten und decodierten objektbasierten 3-D-Audioinhalte gemäß den geformten 3-D-Audioszeneninformationen; und
Wiedergabemittel (45) zum Wiedergeben des Audiosignals mit einem der gesteuerten Merkmale.

6. Ein Verfahren zum Liefern eines objektbasierten dreidimensionalen (3-D) Audioservice unter Verwendung voreingestellter Audioszenen, aufweisend die Schritte:

Eingeben eines Audiosignals;
Erzeugen (S304) von mehr als einer 3-D-Audioszeneninformation durch Extrahieren von Objektaudiosignalen (S302) von dem eingegebenen Audiosignal, Anordnen der extrahierten Objektaudiosignale in einem 3-D-Raum und Editieren (S303) von Merkmalen von jedem Objekt basierend auf den voreingestellten Audioszenen; und
Codieren und Multiplexen (S305) des Audiosignals und der 3-D-Audioszeneninformation für jedes Objektaudiosignal,
wobei in dem Schritt des Erzeugens (S304) von 3-D-Audioszeneninformationen mehr als eine der 3-D-Audioszeneninformation erzeugt wird durch Setzen des Verhältnisses zwischen den Lautstärken der Objektaudiosignale basierend auf Audioszeneninformation von jedem Objektaudiosignal, wobei die Verhältnisse gesetzt werden in Übereinstimmung mit der Steuerung eines Benutzers einschließend einen Betrachter.

7. Das Verfahren nach Anspruch 6, ferner aufweisend den Schritt:

Verarbeiten der codierten und gemultiplexten objektbasierten 3-D-Audioinhalte.

8. Das Verfahren nach Anspruch 7, wobei in dem Schritt des Verarbeitens der objektbasierten 3-D-Audioinhalte die codierten und gemultiplexten objektbasierten 3-D-Audioinhalte über ein digitales Broadcasting-Netzwerk übertragen werden.

9. Ein Verfahren zum Wiedergeben von objektbasiertem dreidimensionalem (3-D) Audioservice, der geliefert wird durch das objektbasierte 3-D-Audioservice-Lieferungsverfahren gemäß einem der Ansprüche 6 bis 8, unter Verwendung voreingestellter Audioszenen, aufweisend die Schritte:

Demultiplexen und Decodieren (S402) von objektbasierten 3-D-Audioinhalten;
Auswählen (S403) von einer 3-D-Audioszeneninformation aus einer Mehrzahl von 3-D-Audioszeneninformationen basierend auf den voreingestellten Audioszenen in dem gedemultiplexten und decodierten objektbasierten 3-D-Audioinhalten durch einen Benutzer einschließend eine Betrachter;
Bilden (S403) von 3-D-Audioszeneninformationen gemäß der ausgewählten 3-D-Audioszeneninformation;
Steuren (S404) von Merkmalen von Objekten in einem Audiosignal der gedemultiplexten und decodierten objektbasierten 3-D-Audioinhalte in Übereinstimmung mit der gebildeten 3-D-Audioszeneninformation; und
Wiedergeben (S405) des Audiosignals mit einem der gesteuerten Merkmale.

10. Das Verfahren nach Anspruch 9 oder die Vorrichtung (40) nach Anspruch 5, wobei in dem Schritt des Bildens (S403) von 3-D-Audioszeneninformation 3-D-Audioszenen-Information gebildet wird in Übereinstimmung mit Merkmalen von jedem Objekt, die durch einen Benutzer einschließend einen Betrachter gesetzt werden.

11. Die Vorrichtung (30; 40) nach Anspruch 1 oder 5, oder das Verfahren nach Anspruch 6 oder 9, wobei die Merkmale einschließen: zumindest eines aus einem Ort, einer Lautstärke, einer Richtung und einer Schallfeldumgebung jedes Objekts.

12. Die Vorrichtung nach Anspruch 5 oder das Verfahren nach Anspruch 9, wobei ein Merkmal eines 3-D-Audioraums gesteuert wird durch Verändern einer Hallzeit eines 3-D-Audioraums durch Steuern einer Lautstärke und einer Verzögerungszeit eines initialreflektierten Schalls.

Revendications

1. Appareil de fourniture de service audio tridimensionnel (3D) à base d'objet utilisant des scènes audio préétablies, comprenant :

des moyens d'entrée audio (31) pour entrer un signal audio ;
des moyens de génération de scène audio préétablie (32) pour extraire des signaux audio d'objet du signal audio entré via les moyens d'entrée audio et pour générer plus d'une information de

scène audio 3-D en agençant les signaux audio d'objet extraits dans un espace 3-D et en éditant des caractéristiques de chaque objet d'après les scènes audio préétablies ; et
des moyens de codage (33) pour coder et multiplexer le signal audio et les informations de scène audio 3-D pour chaque signal audio d'objet,
dans lequel les moyens de génération de scène audio préétablie (32) sont adaptés pour générer plus d'une des informations de scène audio 3-D en établissant les rapports entre les volumes des signaux audio d'objet, d'après des informations de scène audio de chaque signal audio d'objet, lesquels rapports sont établis selon la commande d'un utilisateur incluant un spectateur.

2. Appareil de fourniture de service audio 3-D à base d'objet (30) selon la revendication 1, comprenant en outre des moyens de traitement pour traiter les contenus audio 3-D à base d'objet codés et multiplexés.

3. Appareil de fourniture de service audio 3-D à base d'objet (30) selon la revendication 2, dans lequel les moyens de traitement sont adaptés pour transmettre les contenus audio 3-D à base d'objet codés et multiplexés à un terminal de reproduction audio via un réseau de diffusion numérique.

4. Appareil de fourniture de service audio 3-D à base d'objet (30) selon la revendication 1, dans lequel les moyens de génération de scène audio (32) sont adaptés pour extraire des signaux audio d'objet à l'aide d'une technique de séparation de source aveugle convolutive lorsqu'une source sonore du signal audio entré est une source sonore mélangée.

5. Appareil de reproduction de service audio tridimensionnel (3-D) à base d'objet (40) pour reproduire un service audio 3-D à base d'objet fourni par l'appareil de fourniture de service audio 3-D à base d'objet (30) selon l'une quelconque des revendications 1 à 4, utilisant des scènes audio préétablies comprenant :

des moyens de décodage (42) pour démultiplexer et décoder des contenus audio 3-D à base d'objet ;
des moyens de sélection de scène audio (43) pour sélectionner une information de scène audio 3-D parmi une pluralité d'informations de scène audio 3-D d'après les scènes audio préétablies dans les contenus audio 3-D à base d'objet démultiplexés et décodés par un utilisateur incluant un spectateur ;
des moyens de formation de scène audio pour former les informations de scène audio 3-D selon les informations de scène audio 3-D sélectionnées ;
des moyens de mélange de signal audio (44) pour contrôler les caractéristiques d'objets dans un signal audio des contenus audio 3-D à base d'objet démultiplexés et décodés selon les informations de scène audio 3-D formées ; et
des moyens de reproduction (45) pour reproduire le signal audio avec une des caractéristiques commandées.

6. Procédé de fourniture d'un service audio tridimensionnel (3-D) à base d'objet utilisant des scènes audio préétablies, comprenant les étapes de :

fourniture en entrée d'un signal audio ;
génération (S304) de plus d'une information de scène audio 3-D en extrayant des signaux audio d'objet (S302) du signal audio fourni en entrée,
agencement des signaux audio d'objet extraits dans un espace 3-D et édition (S303) de caractéristiques de chaque objet d'après les scènes audio préétablies ; et
codage et multiplexage (S305) du signal audio et des informations de scène audio 3-D pour chaque signal audio d'objet,
dans lequel dans l'étape de génération (S304) d'information de scène audio 3-D, plus d'une des informations de scène audio 3-D est générée en établissant les rapports entre les volumes des signaux audio d'objet, basés sur les informations de scène audio de chaque signal audio d'objet, lesquels rapports sont établis selon une commande d'un utilisateur incluant un spectateur.

7. Procédé selon la revendication 6, comprenant en outre l'étape de :

traitement des contenus audio 3-D à base d'objet codés et multiplexés.

8. Procédé selon la revendication 7, dans lequel dans l'étape de traitement des contenus audio 3-D à à base d'objet, les contenus audio 3-D à base d'objet codés et multiplexés sont transmis via un réseau de diffusion numérique.

9. Procédé de reproduction d'un service audio tridimensionnel (3-D) à base d'objet fourni par le procédé de fourniture de service audio 3-D à base d'objet selon l'une quelconque des revendications 6 à 8, utilisant des scènes audio préétablies comprenant les étapes de :

démultiplexage et décodage (S402) de contenus audio 3-D à base d'objet ;
sélection (S403) d'une information de scène

audio 3-D parmi une pluralité d'informations de scène audio 3-D d'après les scènes audio préétablies dans les contenus audio 3-D à base d'objet démultiplexés et décodés par un utilisateur incluant un spectateur ;
formation (S403) d'information de scène audio 3-D selon les informations de scène audio 3-D sélectionnées ;
commande (S404) des caractéristiques d'objet dans un signal audio des contenus audio 3-D à base d'objet démultiplexés et décodés selon les informations de scène audio 3-D formées ; et
reproduction (S405) du signal audio avec l'une des caractéristiques commandées.

10. Procédé selon la revendication 9 ou appareil (40) selon la revendication 5, où dans l'étape de formation (S403) d'information de scène audio 3-D, l'information de scène audio 3-D est formée selon les caractéristiques de chaque objet, qui sont établies par un utilisateur incluant un spectateur.

11. Appareil (30 ; 40) selon la revendication 1 ou 5, ou procédé selon la revendication 6 ou 9, où les caractéristiques comprennent au moins un élément parmi un emplacement, un volume, une direction, et un environnement de champ acoustique de chaque objet.

12. Appareil (40) selon la revendication 5 ou procédé selon la revendication 9, où une caractéristique d'un espace audio 3-D est commandée en changeant un temps de réverbération d'un espace audio 3-D par la commande d'un volume et d'un délai d'attente d'un son réfléchi initial.

FIG. 1

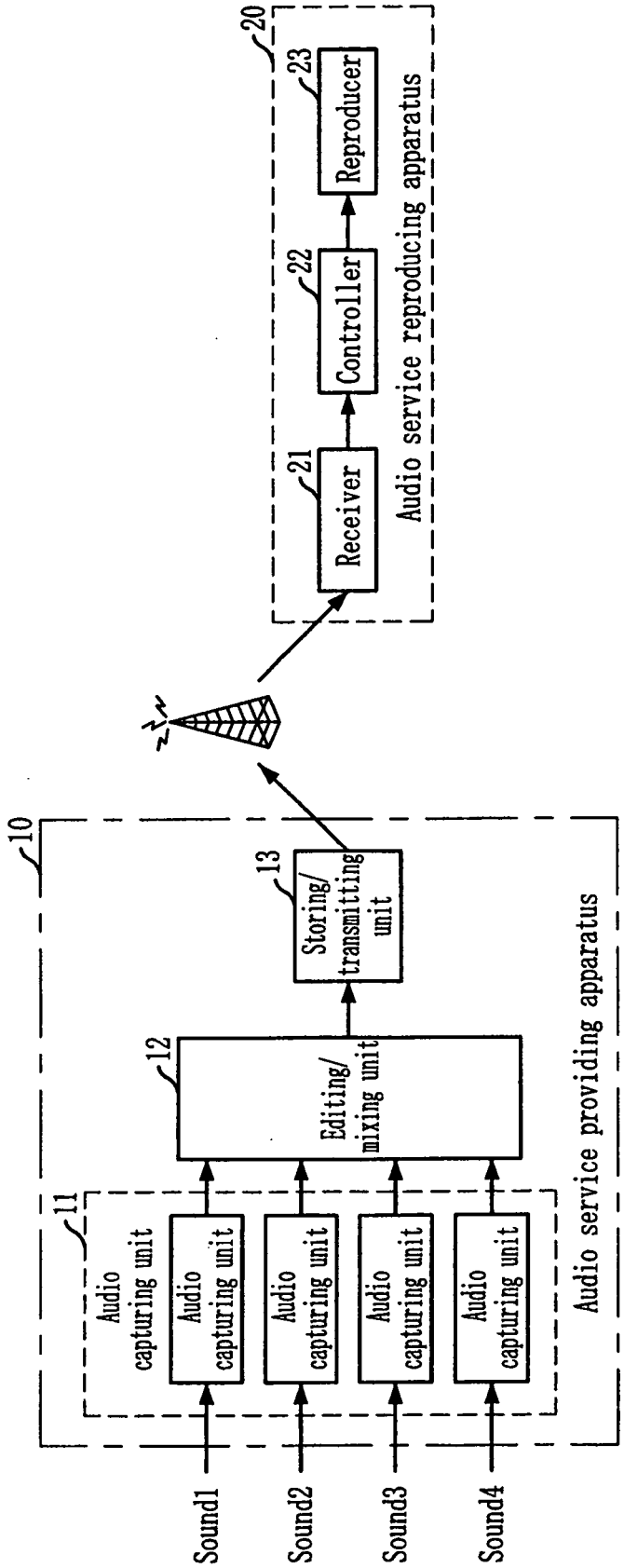

Sound1
Sound2
Sound3
Sound4
Audio service providing apparatus
10
11
Audio capturing unit
Audio capturing unit
Audio capturing unit
Audio capturing unit
Audio capturing unit
12
Editing/ mixing unit
13
Storing/ transmitting unit
Audio service reproducing apparatus
20
21
Receiver
22
Controller
23
Reproducer

FIG. 2

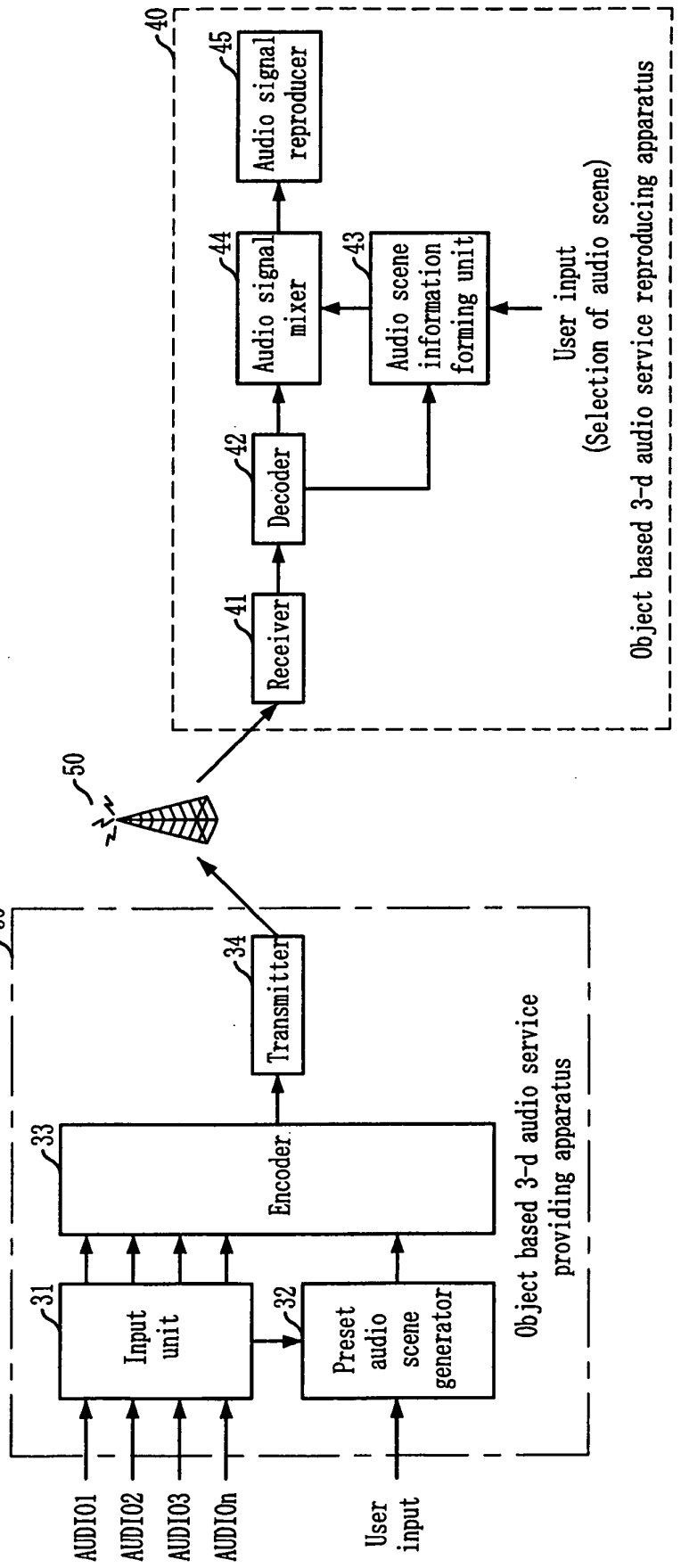

AUDIO1
AUDIO2
AUDIO3
AUDIOn
User input
30
Object based 3-d audio service providing apparatus
31
Input unit
32
Preset audio scene generator
33
Encoder
34
Transmitter
50
40
Object based 3-d audio service reproducing apparatus
41
Receiver
42
Decoder
43
Audio scene information forming unit
44
Audio signal mixer
45
Audio signal reproducer
User input
(Selection of audio scene)

FIG. 3

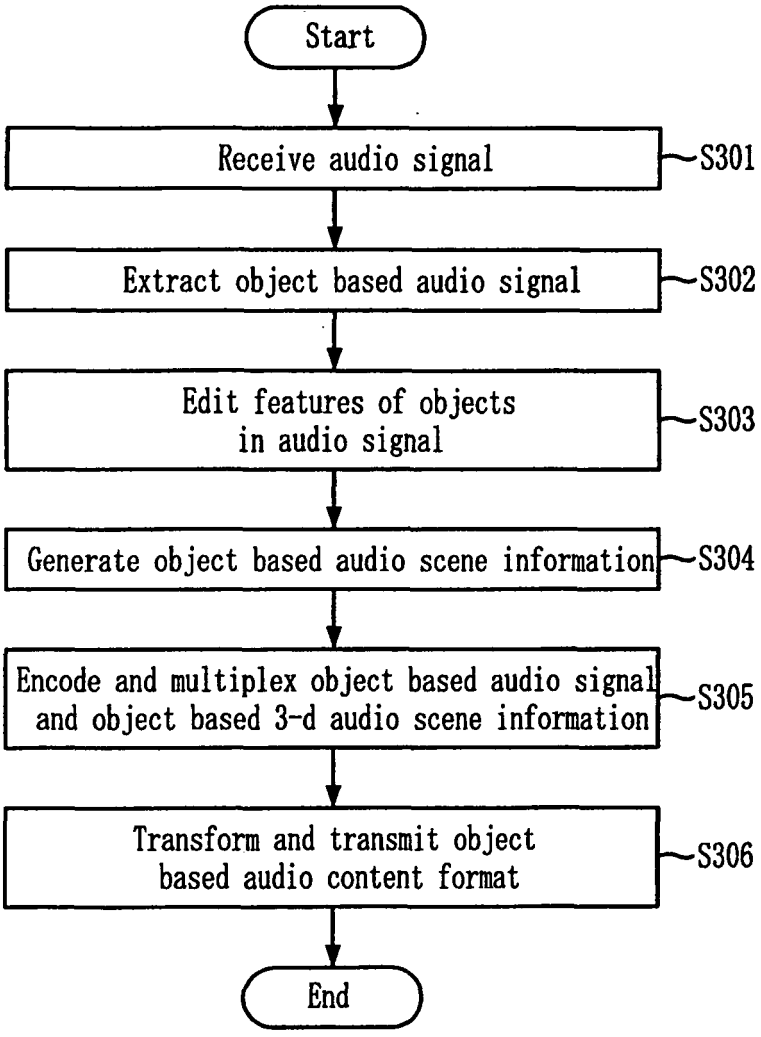
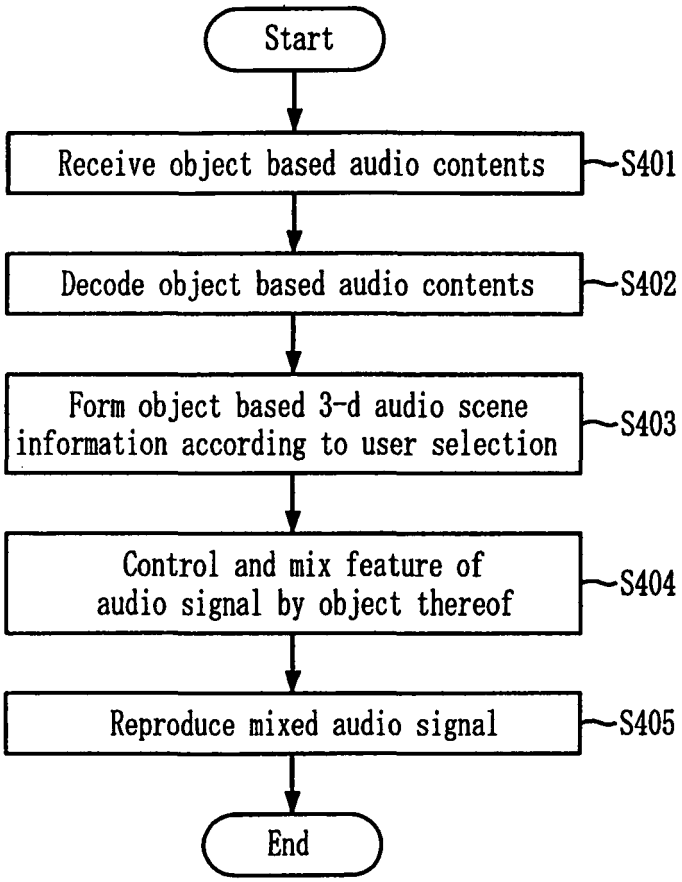

Start
Receive audio signal
S301
Extract object based audio signal
S302
Edit features of objects in audio signal
S303
Generate object based audio scene information
S304
Encode and multiplex object based audio signal and object based 3-d audio scene information
S305
Transform and transmit object based audio content format
S306
End

FIG. 4


Start
Receive object based audio contents
S401
Decode object based audio contents
S402
Form object based 3-d audio scene information according to user selection
S403
Control and mix feature of audio signal by object thereof
S404
Reproduce mixed audio signal
S405
End

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

- KR 1020040037437 **[0010]**

Non-patent literature cited in the description

- **TAEJIN LEE et al.** An Object-based 3D Audio Broadcasting System for Interactive Services. *Audio Engineering Society, Convention paper 6384,* 28 May 2005 **[0011]**