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

(11) EP 2 369 836 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
23.04.2014 Bulletin 2014/17

(21) Application number: **11159156.6**

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

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

(54) **Object-based 3-dimensional audio service system using preset audio scenes**

Objektbasiertes dreidimensionales Audiodienstsystem mit im Voraus eingestellten Audioszenen

Système de service audio tridimensionnel à base d'objet utilisant des scènes audio prédéfinies

(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:
28.09.2011 Bulletin 2011/39

(60) Divisional application:
12171933.0 / 2 501 128

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
07746543.3 / 2 022 263

(73) Proprietor: **Electronics and Telecommunications Research Institute**
Daejeon 305-350 (KR)

(72) Inventors:
- **Lee, Yong-Ju**
  305-345, Daejon (KR)
- **Lee, Tae-Jin**
  305-752, Daejon (KR)
- **Yoo, Jae-Hyoun**
  305-350, Daejon (KR)
- **Kang, Kyeong-Ok**
  305-727, Daejon (KR)
- **Hong, Jin-Woo**
  305-333, Daejon (KR)
- **Jang, In-Seon**
  305-345, Daejon (KR)
- **Seo, Jeong-Il**
  305-728, Daejon (KR)
- **Jang, Dae-Young**
  305-768, Daejon (KR)

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

(56) References cited:
US-B1- 6 665 318

- **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**
- **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**
- **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]**

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

EP 2 369 836 B1

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 in, 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 two user's preference. Therefore, it is very annoying to a user or a viewer.

[0011] Kalva H. et al., "DELIVERING OBJECT-BASED AUDIO-VISUAL SERVICES", IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, IEEE SERVICE CENTER, NEW YORK, US, vol. 45, no,4, 1 November 1999 (1999-11-01), pages 1108-1111, discloses a method for delivering audio-visual services based on MPEG-4. In particular, algorithms to schedule the delivery of object-based audio-visual presentations in general and MPEG-4 presentations are disclosed.

[0012] The document US 6,665, 318 B1 describes a video stream decoder which includes a demultiplexing unit for demultiplexing a video stream containing at least one or more object encoded visual or audio data and one or more scene descriptions which express scene contents by object encoded data; a decoder unit for decoding the object encoded visual data; a decoder unit for decoding the object encoded audio data; a visual synthesizing unit for synthesizing images corresponding to the object encoded visual data; an audio synthesizing unit for synthesizing sounds corresponding to the object encoded audio data; an analyzing unit for analyzing each scene description; and a selector for selecting one of at least two or more scene descriptions contained in the video stream.

DISCLOSURE

TECHNICAL PROBLEM

[0013] 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.

[0014] 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

[0015] The present invention is defined in the independent claims. The dependent claims define embodiments thereof.

[0016] In accordance with an aspect of the present invention, 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.

[0017] In accordance with another aspect of the present invention, 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.

[0018] In accordance with another aspect of the present invention, 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.

[0019] In accordance with another aspect of the present invention, 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

[0020] 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.

[0021] 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

[0022]

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

[0023] 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.

[0024] 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.

[0025] 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.

[0026] 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.

[0027] 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.

[0028] 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.

[0029] 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.

[0030] 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.

[0031] 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 scenes by the audio signal mixer 44.

[0032] 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 Vid-

eo 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.

[0033] 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.

[0034] 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.

[0035] 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.

[0036] 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 halving 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.

[0037] 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.

[0038] Fig. 3 is a flowchart illustrating a method for providing an object-based audio service using preset audio scenes in accordance with an embodiment of the present invention.

[0039] 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 carious input device at step S301.

[0040] The preset audio scene generator 32 extracts object-based audio signals, that is, object audio signals, front 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 multiplexers 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.

[0041] 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 MFEG-2 TS in a digital broadcasting network.

[0042] 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.

[0043] 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.

[0044] 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.

[0045] 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.

[0046] 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.

[0047] Then, the audio signal mixer 44 mixers 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 forming 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.

[0048] 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.

[0049] 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 audio service providing apparatus (30), comprising:

   preset audio scene generating means (32) for generating a plurality of preset audio scenes, wherein each preset audio scene is based on 3D audio scene information which indicates features of a plurality of audio objects; and
   encoding means (33) for encoding and multiplexing the formed audio objects and the plurality of preset audio scenes,
   wherein the 3D audio scene information includes at least one of a location, a size, a volume, a direction, and a sound field environment of each of the plurality of audio objects in 3D space,

2. The object-based audio service providing apparatus (30) of claim 1, further comprising processing means (34) for processing the encoded and multiplexed audio objects.

3. The object-based audio service providing apparatus of claim 1, wherein the processing means (34) transmits the encoded and multiplexed audio objects to an audio reproducing terminal (40) through a digital broadcasting network.

4. An object-based audio service reproducing apparatus (40), comprising:

   decoding means (42) for de-multiplexing and decoding object-based audio contents including a plurality of audio objects and a plurality of present audio scenes;
   audio scene forming means (43) for forming one of said plurality of preset audio scenes , wherein each preset audio scene is based on 3D audio scene information which indicates features of said plurality of audio objects based on a selection of a user;
   audio signal mixing means (44) for mixing the plurality of audio objects according to the formed preset audio scene ; and
   reproducing means (45) for reproducing the mixed plurality of audio objects according to the formed preset audio scene,
   wherein the 3D audio scene information includes at least one of a location, a size, a volume, a direction, and a sound field environment of each of the plurality of audio objects in 3D space.

5. An object-based audio service providing method, comprising:

   generating (S304) a plurality of preset audio scenes, wherein each preset audio scene is based on 3D audio scene information which indicate features of a plurality of audio objects; and
   encoding and multiplexing (S305) the plurality of audio objects and the plurality of preset audio scenes,
   wherein the 3D audio scene information includes at least one of a location, a size, a volume, a direction, and a sound field environment of each of the plurality of audio objects in 3D space,

6. The method of claim 5, further comprising the step of:

   processing the encoded and multiplexed audio objects.

7. The method of claim 6, wherein in the step of processing the audio objects, the encoded and multiplexed audio objects are transmitted through a digital broadcasting network.

8. An object-based audio service reproducing method, comprising:

de-multiplexing and decoding (S402) object-based audio contents including a plurality of audio objects and a plurality of present audio scenes;
forming (S403) one of said plurality of preset audio scenes, wherein each preset audio scene is based on 3D audio scene information which indicates features of said plurality of audio objects, based on a selection of a user;
mixing (S404) the plurality of audio objects according to the formed preset audio scene; and
reproducing (S405) the mixed plurality of audio objects according to the formed preset audio scene,
wherein the 3D audio scene information includes at least one of a location, a size, a volume, a direction, and a sound field environment of each of the plurality of audio objects in the 3D space.

Patentansprüche

1. Eine Vorrichtung (30) zum Liefern eines objektbasierten Audioservice, aufweisend:

Erzeugungsmiftel (32) für voreingestellte Audioszenen zum Erzeugen einer Mehrzahl von voreingestellten Audioszenen, wobei jede voreingestellte Audioszene basiert auf 3D-Audioszeneninformation, die Merkmale einer Mehrzahl von Audioobjekten angeben; und
Codiermittel (33) zum Codieren und Multiplexen der Mehrzahl von Audioobjekten und der Mehrzahl von voreingestellten Audioszenen;
wobei die 3D-Audioszeneninformation zumindest eines einschließt aus: Einer Lokation, einer Größe, einer Lautstärke, einer Richtung und einer Schallfeldumgebung von jedem der Mehrzahl von Audioobjekten im 3D-Raum.

2. Die Vorrichtung (30) zum Liefern von objektbasierten Audioservices nach Anspruch 1, ferner aufweisend: Verarbeitungsmittel (34) zum Verarbeiten der codierten und gemultiplexten Audioobjekte.

3. Die Vorrichtung zum Liefern von objektbasierten Audioservices nach Anspruch 1, wobei die Verarbeitungsmittel (34) die codierten und gemultiplexten Audioobjekte an ein Audiowiedergabeterminal (40) über ein digitales Übertragungsnetzwerk übertragen.

4. Eine Vorrichtung (40) zum Liefern objektbasierter Audioservices, aufweisend:

Decodiermittel (42) zum Demultiplexen und Decodieren objektbasierter Audioinhalte einschließend eine Mehrzahl von Audioobjekten und eine Mehrzahl von voreingestellten Audioszenen;
Audioszenen-Bildungsmittel (43) zum Bilden einer der mehrzahl von voreingestellten Audioszenen, wobei jede voreingestellte Audioszene basiert auf 3D-Audioszeneninformationen, die Merkmale der Mehrzahl von Audioobjekten angeben und auf einer Auswahl eines Benutzers basieren;
Audiosignal-Mischmittel (44) zum Mischen der Mehrzahl von Audioobjekten in Übereinstimmung mit der gebildeten voreingestellten Audioszene; und
Wiedergabemittel (45) zum Wiedergeben der gemischten Mehrzahl von Audioobjekten in Übereinstimmung mit der gebildeten voreingestellten Audioszene,
wobei die 3D-Audioszeneninformationen zumindest eines enschließen aus: Einer Lokation, einer Größe, einer Lautstärke, einer Richtung und einer Schallfeldumgebung von jedem der Mehrzahl von Audioobjekten im 3D-Raum.

5. Ein Verfahren zum Liefern eines objektbasierten Audioservices, aufweisend:

Erzeugen (S304) einer Mehrzahl von voreingestellten Audioszenen, wobei jede voreingestellte Audioszene basiert auf 3D-Audioszeneninformationen, die Merkmale einer Mehrzahl von Audioobjekten angeben; und
Codieren und Multiplexen (S305) der Mehrzahl von Audioobjekten und der Mehrzahl von voreingestellten Audioszenen,
wobei die 3D-Audioszeneninformationen zumindest eines enschließen aus: Einer Lokation, einer Größe, einer Lautstärke, einer Richtung und einer Schallfeldumgebung von jeder der Mehrzahl von Audioobjekten im 3D-Raum.

6. Das Verfahren nach Anspruch 5, ferner aufweisend den Schritt:

Verarbeiten der codierten und gemultiplexten Audioobjekte. Verarbeiten der codierten und gemultiplexten Audioobjekte

7. Das Verfahren nach Anspruch 6, wobei beim Schritt des Verarbeitens der Audioobjekte die codierten und gemultiplexten Audioobjekte übertragen werden durch ein digitales Übertragungsnetzwerk.

8. Ein objektbasiertes Audioservice-Wiedergabeverfahren, aufweisend:

Demultiplexen und Decodieren (5402) objektbasierter Audioinhalte einschließend eine Mehrzahl von Audioobjekten und eine Mehrzahl von

voreingestellten Audioszenen;
Bilden (S403) einer der Mehrzahl von Voreingestellten Audioszenen, wobei jede voreingestellte Audioszene basiert auf 3D-Audioszeneninformationen, die Merkmale der Mehrzahl von Audioobjekten basierend auf einer Auswahl eines Benutzers angeben;
Mischen (S404) der Mehrzahl von Audioobjekten in Übereinstimmung mit der gebildeten voreingestellten Audioszene; und
Wiedergeben (5405) der gemischten Mehrzahl von Audioobjekten in Übereinstimmung mit der gebildeten voreingestellten Audioszene;
wobei die 3D-Audioszeneninformationen zumindest eines emschließen von: Einer Lokation, einer Größe, einer Lautstärke, einer Richtung und einer Schallfeldumgebung jeder der Mehrzahl von Audioobjekten im 3D-Raum.

Revendications

1. Appareil fournissant un service audio basé sur des objets (30), comprenant :

des moyens de génération de scène audio préétablie (32) destinés à générer une pluralité de scènes audio préétablies, dans lequel chaque scène audio préétablie est basée sur des informations de scène audio 3D qui indique des particularités d'une pluralité d'objets audio ; et
des moyens de codage (33) destiné à coder et multiplexer la pluralité d'objets audio et la pluralité de scènes audio préétablies,
dans lequel les informations de scène audio 3D comprennent au moins l'un d'un emplacement, d'une taille, d'un volume, d'une direction et d'un environnement de champ sonore de chacun de la pluralité d'objets audio dans l'espace 3D.

2. Appareil fournissant un service audio basé sur des objets (30) selon la revendication 1, comprenant en outre des moyens de traitement (34) destinés à traiter les objets audio codés et multiplexés.

3. Appareil fournissant un service audio basé sur des objets selon la revendication 1, dans lequel les moyens de traitement (34) transmettent les objets audio codés et multiplexés à un terminal de reproduction audio (40) par l'intermédiaire d'un réseau de radiodiffusion numérique.

4. Appareil de reproduction de service audio basé sur des objets (40) comprenant
des moyens de décodage (42) destinés à démultiplexer et décoder des contenus audio basés sur les objets comprenant une pluralité d'objets audio et une pluralité de scènes audio préétablies ;
des moyens de formation de scène audio (43) destinés à former l'une de ladite pluralité de scènes audio préétablies, dans lequel chaque scène audio préétablie est basée sur des informations de scène audio 3D qui indiquent des particularités de ladite pluralité d'objets audio basée sur une sélection d'un utilisateur ;
des moyens de mixage de signal audio (44) destiné à mixer la pluralité d'objets audio selon la scène audio préétablie formée ; et
des moyens de reproduction (45) destinés à reproduire la pluralité mixée d'objets audio selon la scène audio préétablie formée,
dans lequel les informations de scène audio 3D comprennent au moins l'un d'un emplacement, d'une taille, d'un volume, d'une direction et d'un environnement de champ sonore de chacun de la pluralité d'objets audio dans l'espace 3D.

5. Procédé de fourniture de service audio basé sur des objets, comprenant :

la génération (S304) d'une pluralité de scènes audio préétablies, où chaque scène audio préétablie est basée sur des informations de scène audio 3D qui indiquent des particularités d'une pluralité d'objets audio ; et
le codage et le multiplexage (S305) de la pluralité d'objets audio et de la pluralité de scènes audio préétablies,
dans lequel les informations de scène audio 3D comprennent au moins l'un d'un emplacement, d'une taille, d'un volume, d'une direction et d'un environnement de champ sonore de chacun de la pluralité d'objets audio dans l'espace 3D.

6. Procédé selon la revendication 5, comprenant en outre l'étape de :

traitement des objets audio codés et multiplexés.

7. Procédé selon la revendication 6, dans lequel dans l'étape de traitement des objets audio, les objets audio codés et multiplexés sont transmis par l'intermédiaire d'un réseau de radiodiffusion numérique.

8. Procédé de reproduction de service audio basé sur des objets, comprenant :

le démultiplexage et le décodage (S402) de contenus audio basés sur les objets comprenant une pluralité d'objets audio et une pluralité de scènes audio préétablies ;
la formation (S403) de l'une de ladite pluralité de scènes audio préétablies, où chaque scène audio préétablie est basée sur des informations de scène audio 3D qui indiquent des particula-

rités de ladite pluralité d'objets audio basée sur une sélection d'un utilisateur ;
le mixage (S404) de la pluralité d'objets audio selon la scène audio préétablie formée ; et
la reproduction (S405) de la pluralité mixée d'objets audio selon la scène audio préétablie formée,
dans lequel les informations de scène audio 3D comprennent au moins l'un d'un emplacement, d'une taille, d'un volume, d'une direction et d'un environnement de champ sonore de chacun de la pluralité d'objets audio dans l'espace 3D.

FIG. 1

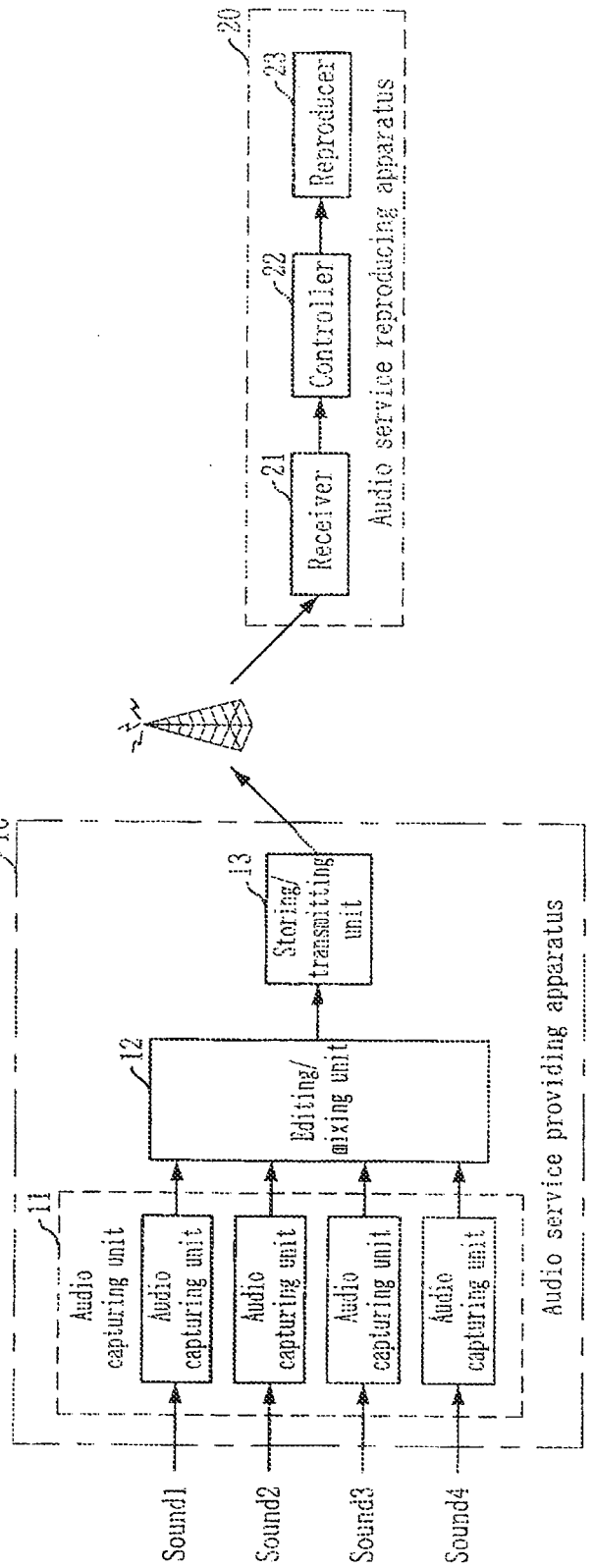

Sound1
Sound2
Sound3
Sound4
Audio capturing unit
Audio capturing unit
Audio capturing unit
Audio capturing unit
11
Editing/ mixing unit
12
Storing/ transmitting unit
13
10
Audio service providing apparatus
Receiver
21
Controller
22
Reproducer
23
20
Audio service reproducing apparatus

FIG. 2

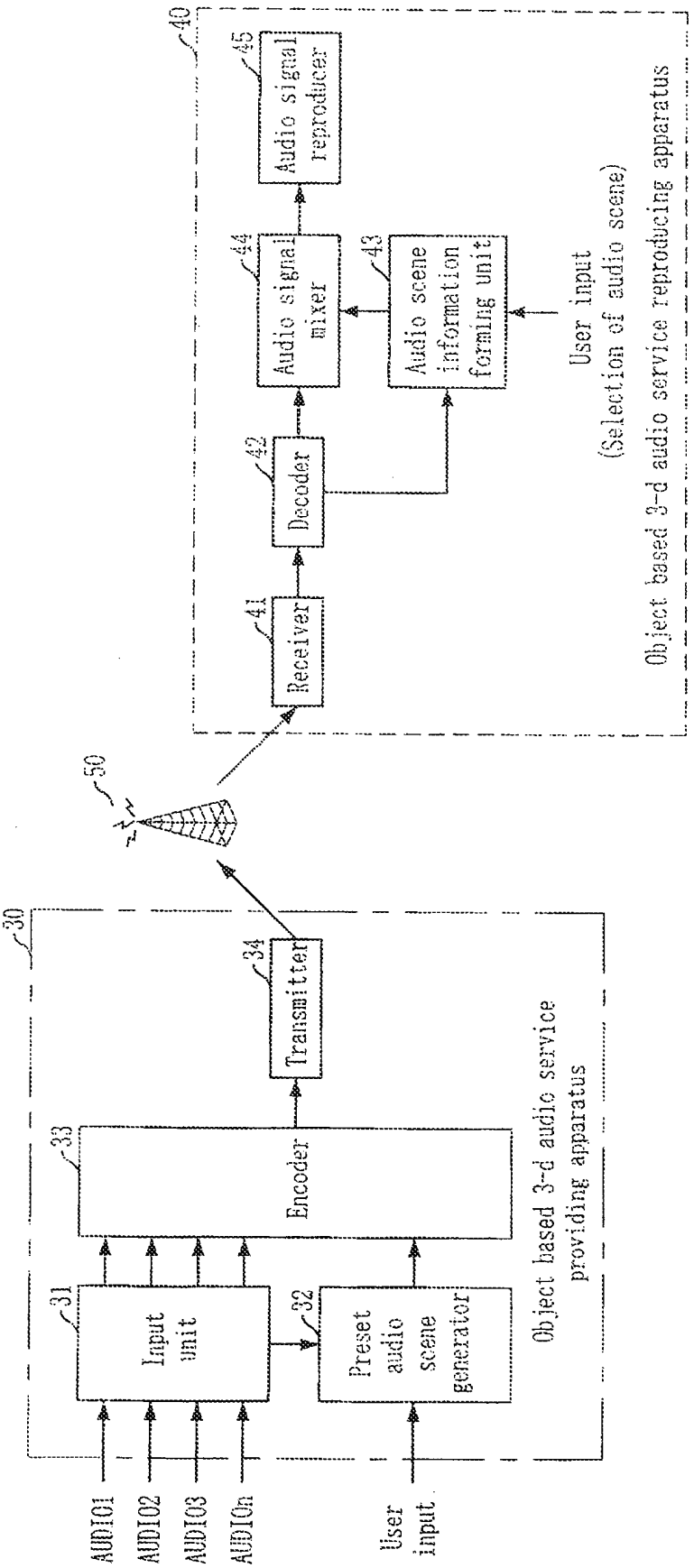

AUDIO1
AUDIO2
AUDIO3
AUDIOn
User input
Input unit
31
Preset audio scene generator
32
Encoder
33
Transmitter
34
30
Object based 3-d audio service providing apparatus
50
Receiver
41
Decoder
42
Audio scene information forming unit
43
Audio signal mixer
44
Audio signal reproducer
45
User input
(Selection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

FIG. 3

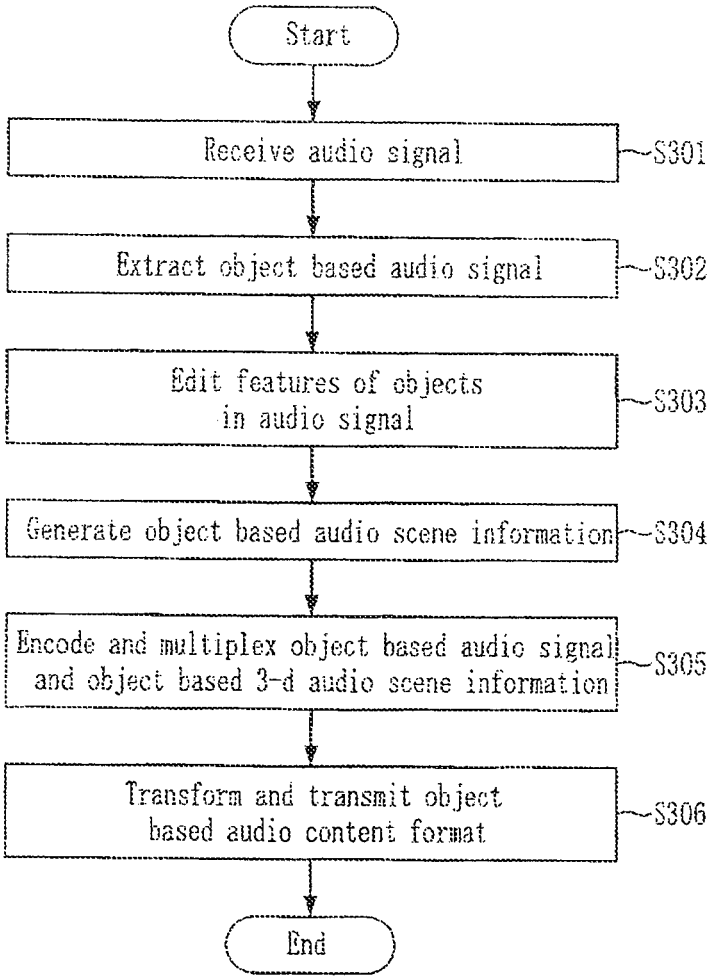
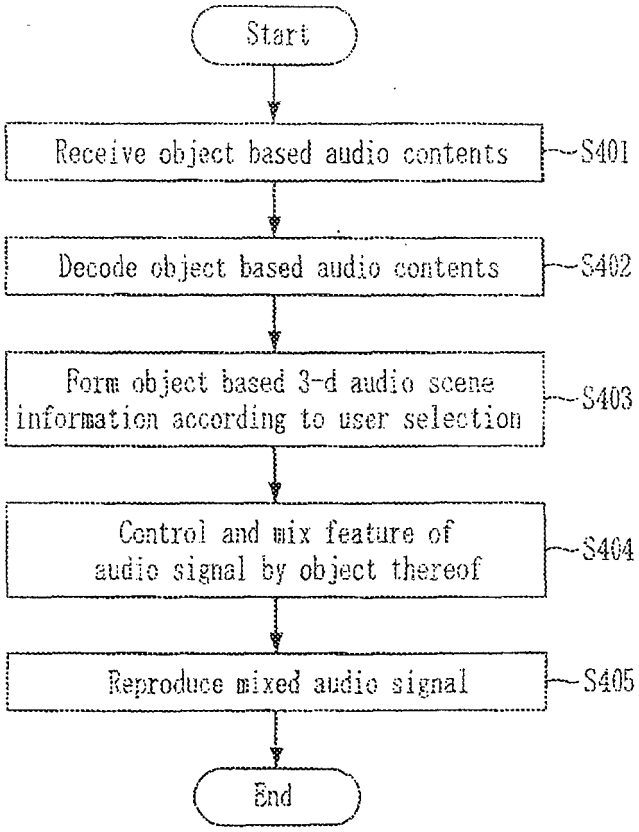

Start
Receive audio signal
S301
Extract object based audio signal
S302
Edit features of objects in audio signal
S303
Generate object based audio scene information
S304
Encode and multiplex object based audio signal and object based 3-d audio scene information
S305
Transform and transmit object based 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]**
- US 6665 A **[0012]**
- US 318 B1 **[0012]**

Non-patent literature cited in the description

- **KALVA H. et al.** DELIVERING OBJECT-BASED AUDIO-VISUAL SERVICES. *IEEE TRANSACTIONS ON CONSUMER ELECTRONICS,* 01 November 1999, vol. 45 (4), 1108-1111 **[0011]**