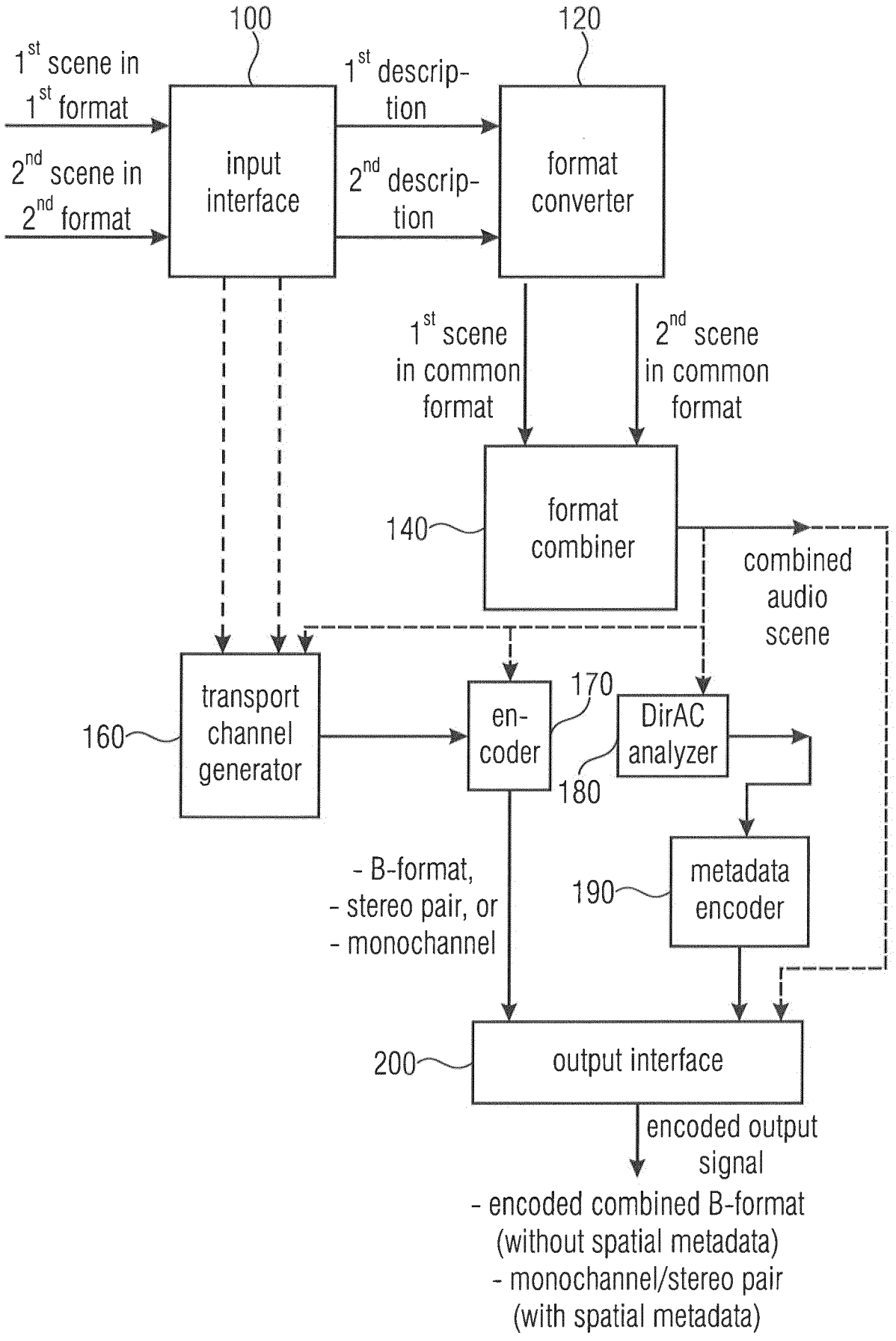
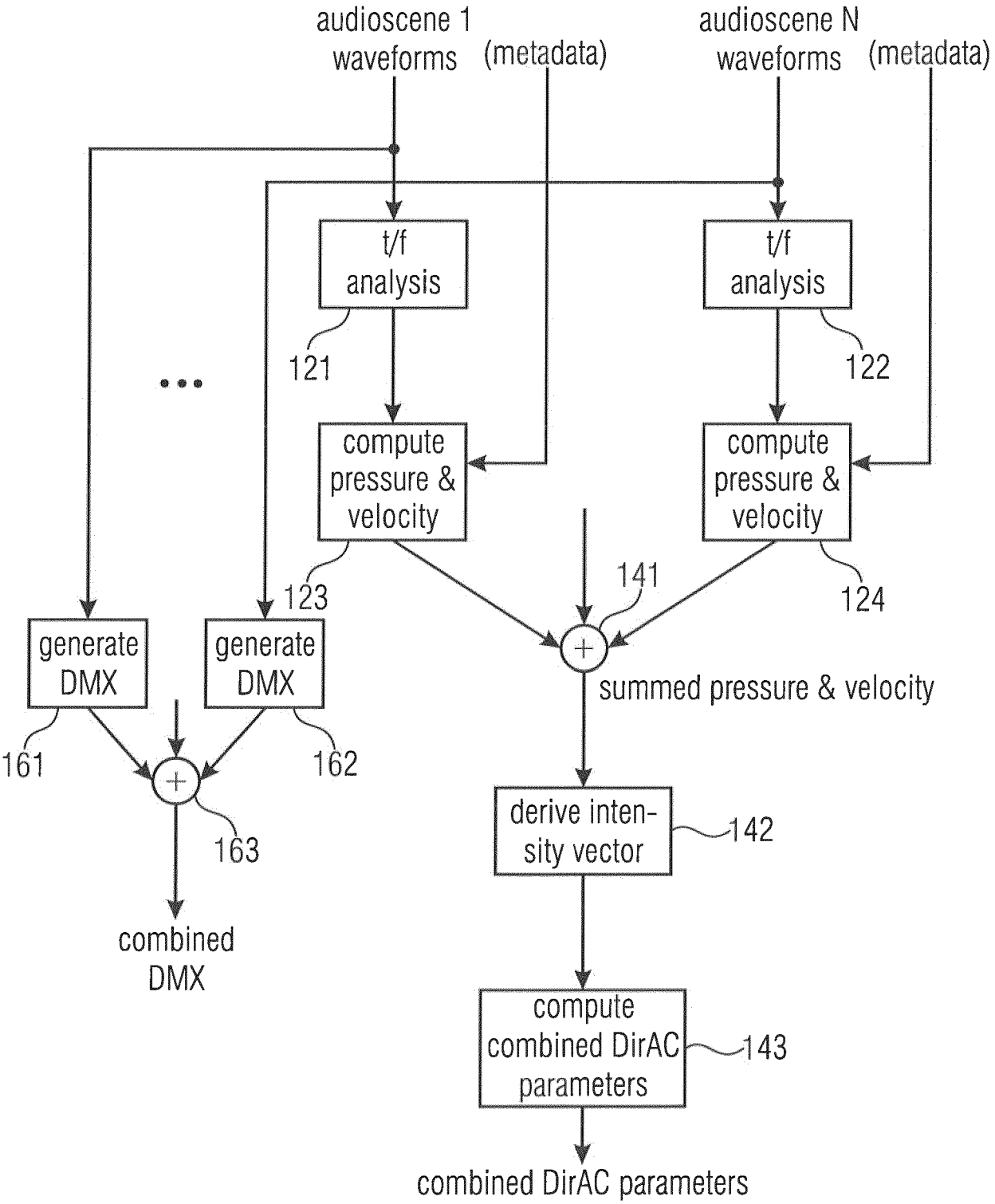
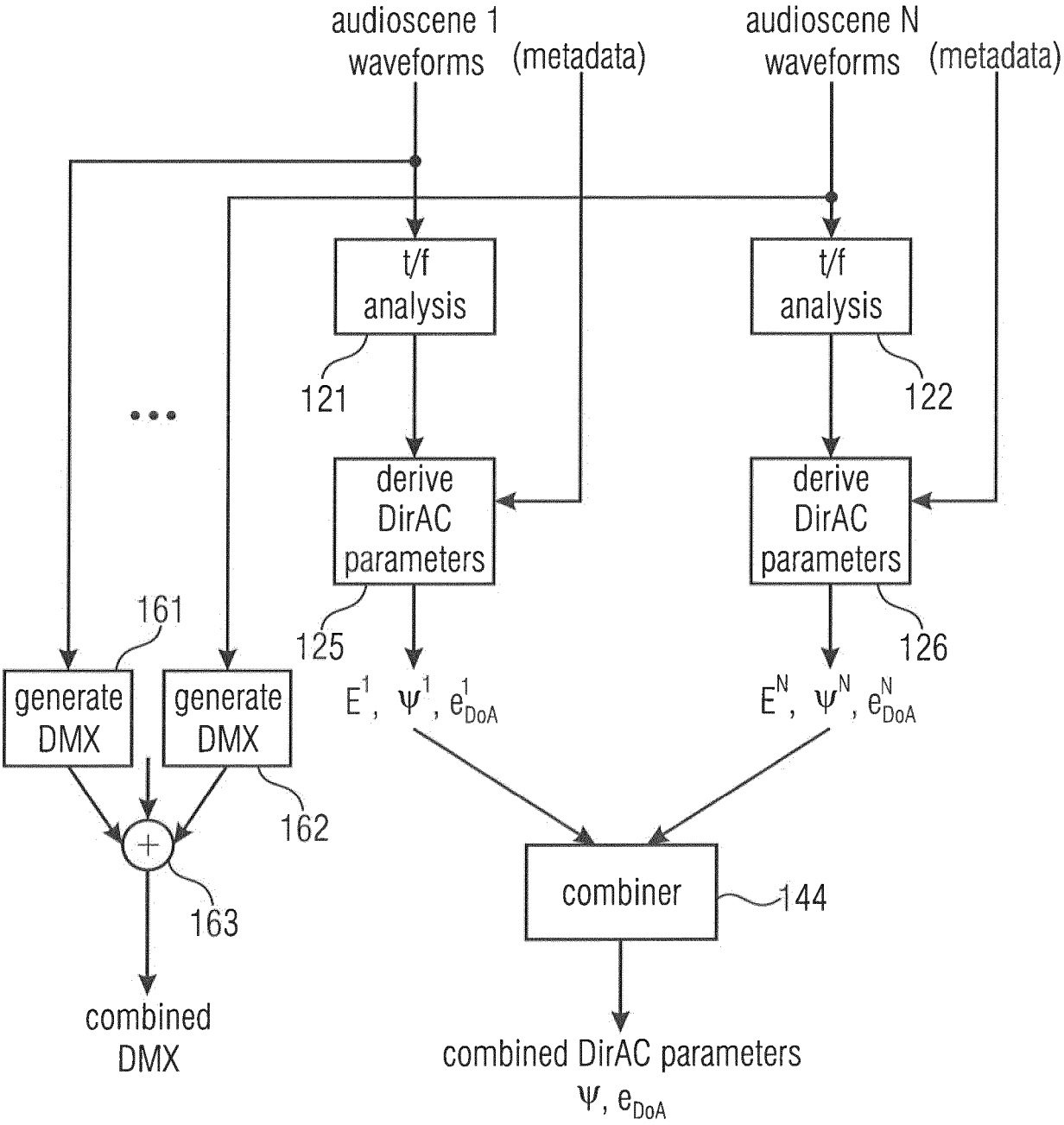

100
120
1st scene in 1st format
2nd scene in 2nd format
input interface
1st description
2nd description
format converter
1st scene in common format
2nd scene in common format
140
format combiner
combined audio scene
160
transport channel generator
en-coder
170
DirAC analyzer
180
- B-format,
- stereo pair, or
- monochannel
190
metadata encoder
200
output interface
encoded output signal
- encoded combined B-format (without spatial metadata)
- monochannel/stereo pair (with spatial metadata)

Fig. 1a



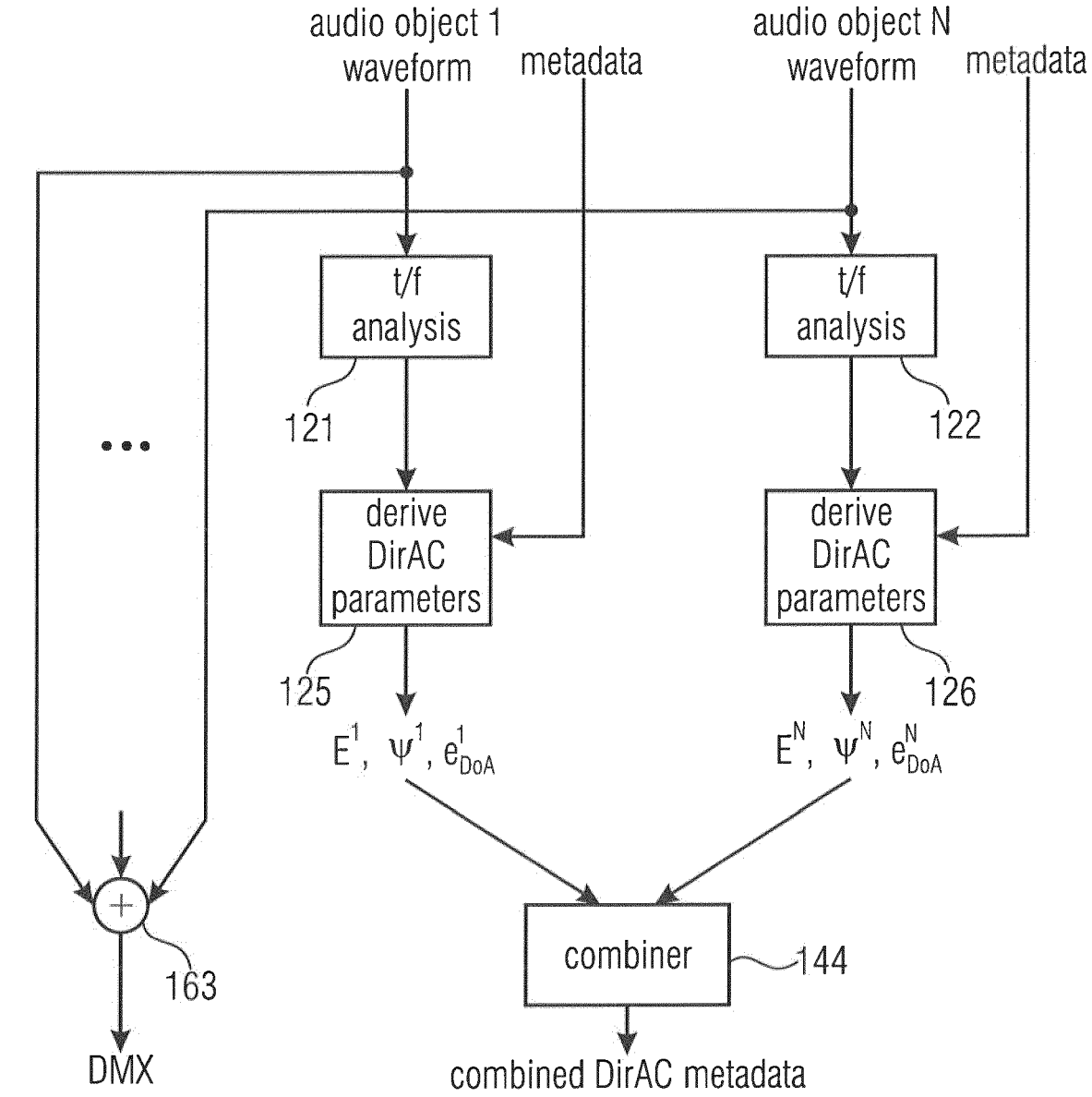
- block diagram for combining different audio scenes in different formats in the pressure/velocity domain

Fig. 1b



- block diagram for combining
≠ audio scenes in different formats in the
DirAC parameter domain

Fig. 1c



combiner could be:

alt. ①

$$\psi = \frac{1}{\sum E^i} \sum_{i=1}^{N} E^i \psi^i$$

$$e_{DoA} = \frac{1}{\sum (1-\psi^i E^i)} \sum_{i=1}^{N} (1-\psi^i E^i) E^i e^i_{DoA}$$

alt. ②

$$\psi = 0 \quad \text{(since audio objects have no diffuseness usually)}$$

$$e_{DoA} = e_{DoA}^{\operatorname{argmax}(E^i)}$$

Fig. 1d

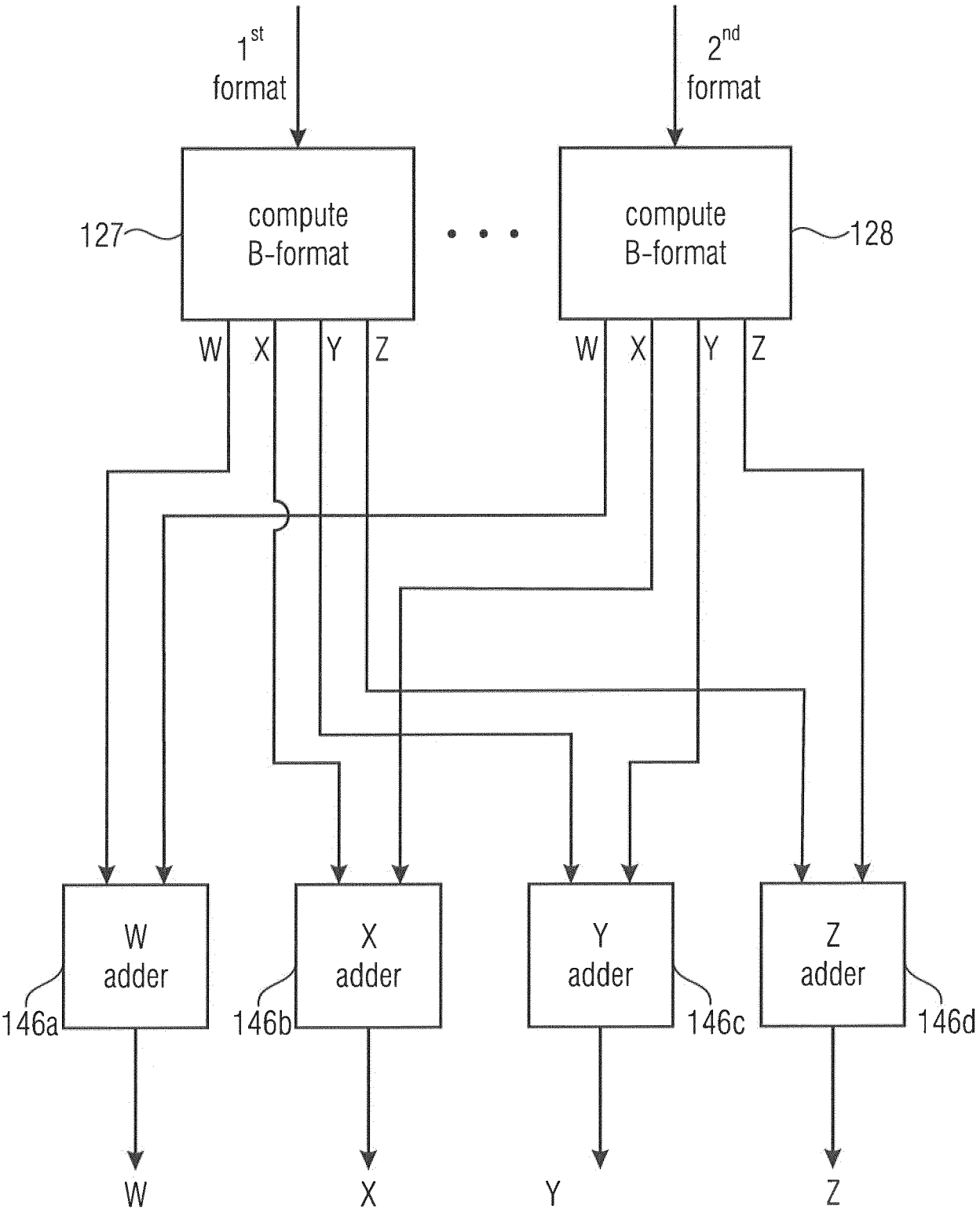
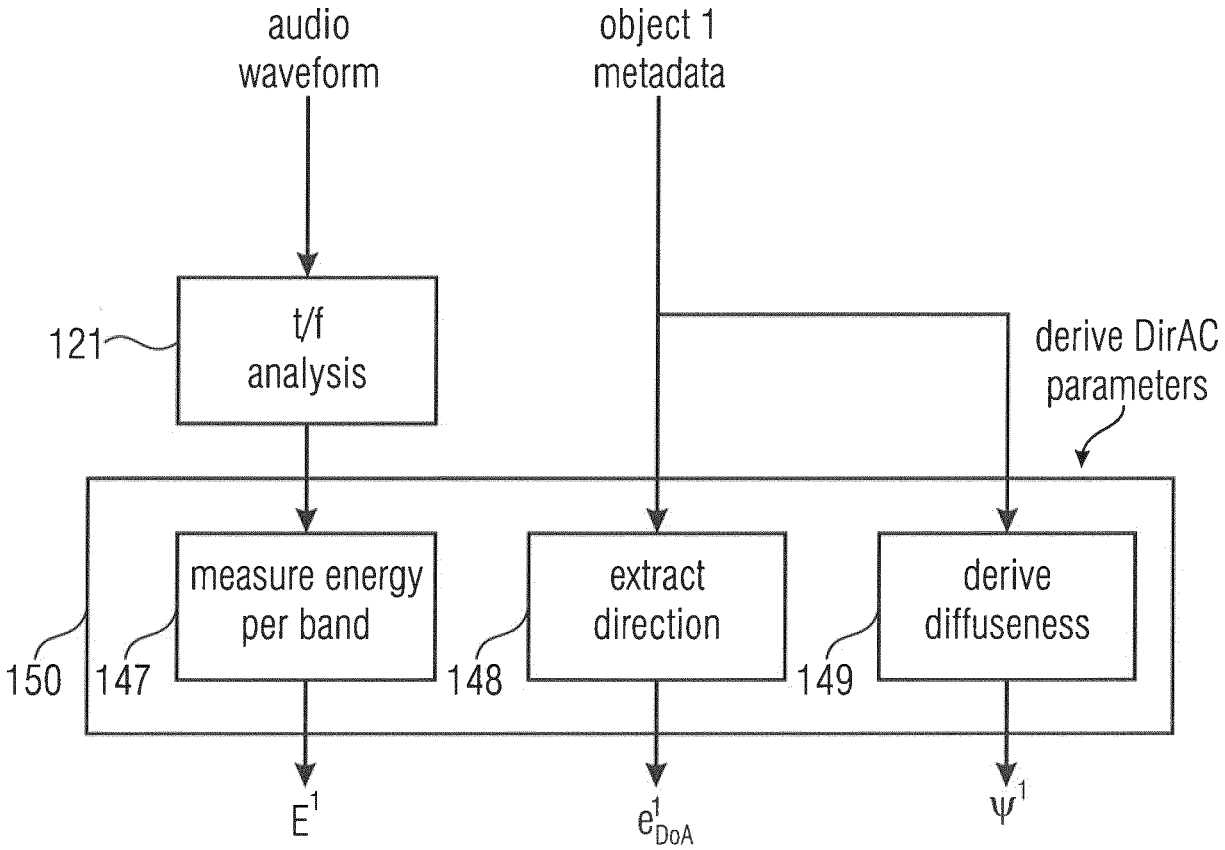

1st format
2nd format
compute B-format
127
compute B-format
128
W
X
Y
Z
W
X
Y
Z
W adder
X adder
Y adder
Z adder
146a
146b
146c
146d
W
X
Y
Z

Fig. 1e



- block diagram for deriving DirAC parameters from an audio object

Fig. 1f

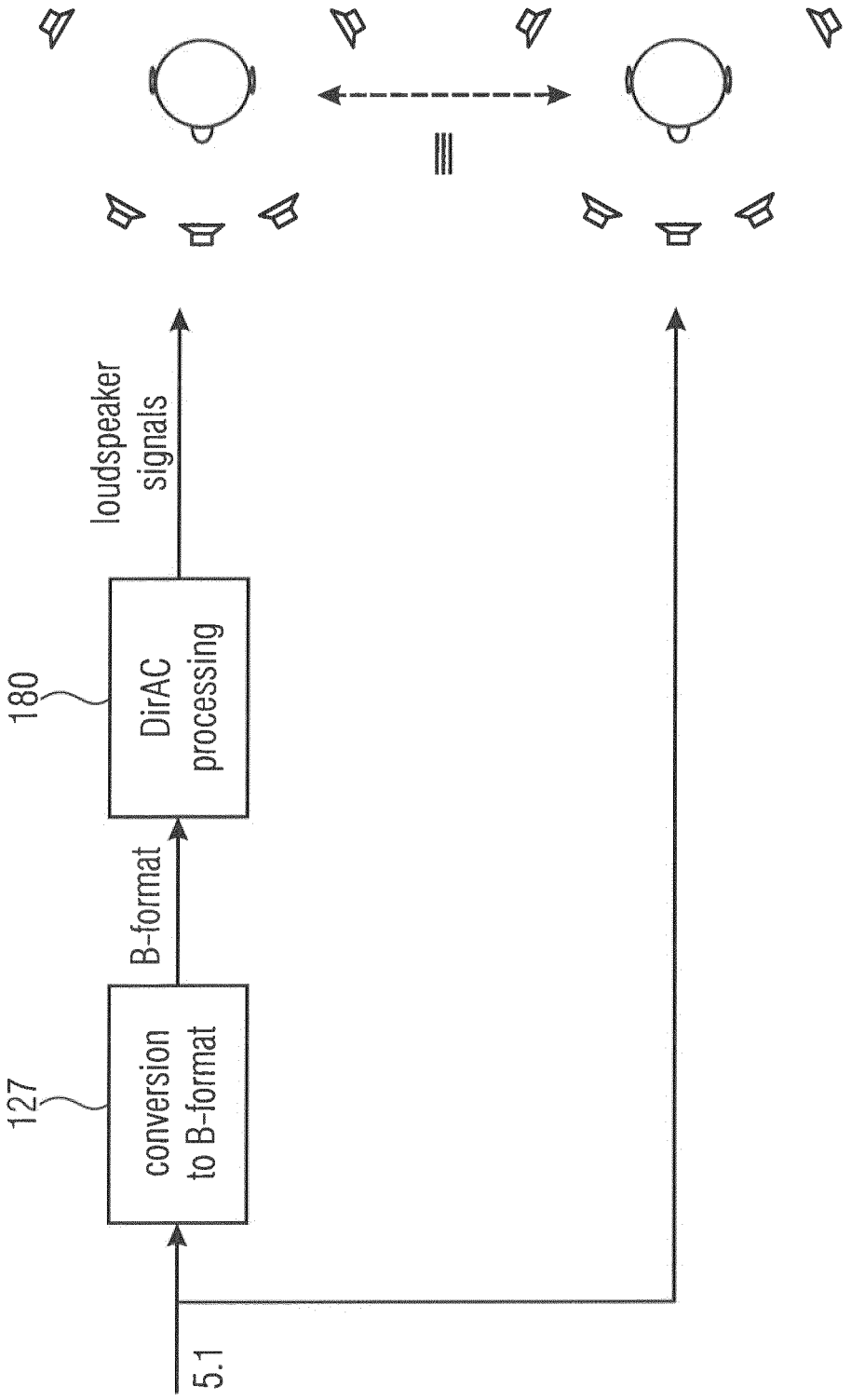

5.1
127
conversion to B-format
B-format
180
DirAC processing
loudspeaker signals

Fig. 1g

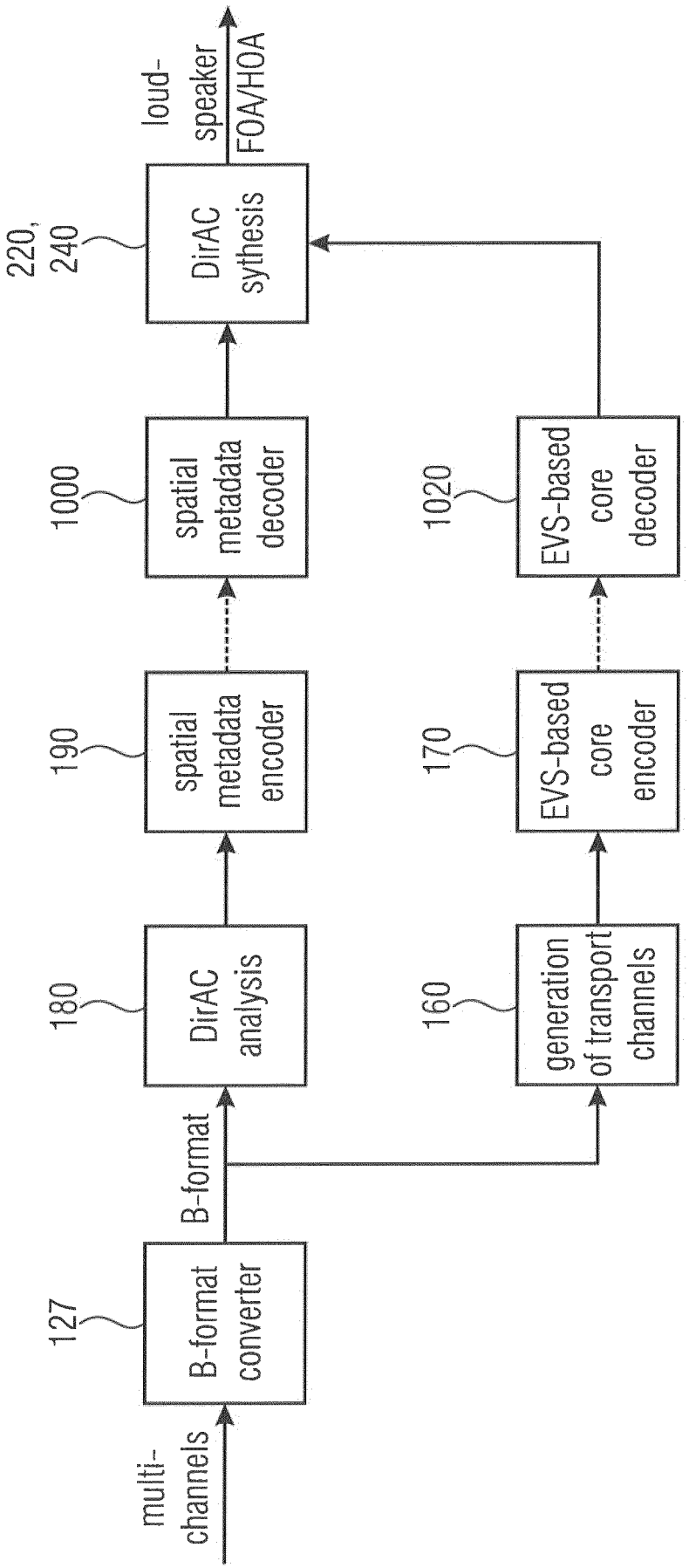

multi-channels
127
B-format converter
B-format
180
DirAC analysis
190
spatial metadata encoder
1000
spatial metadata decoder
220, 240
DirAC sythesis
loud-speaker FOA/HOA
160
generation of transport channels
170
EVS-based core encoder
1020
EVS-based core decoder

Fig. 1h

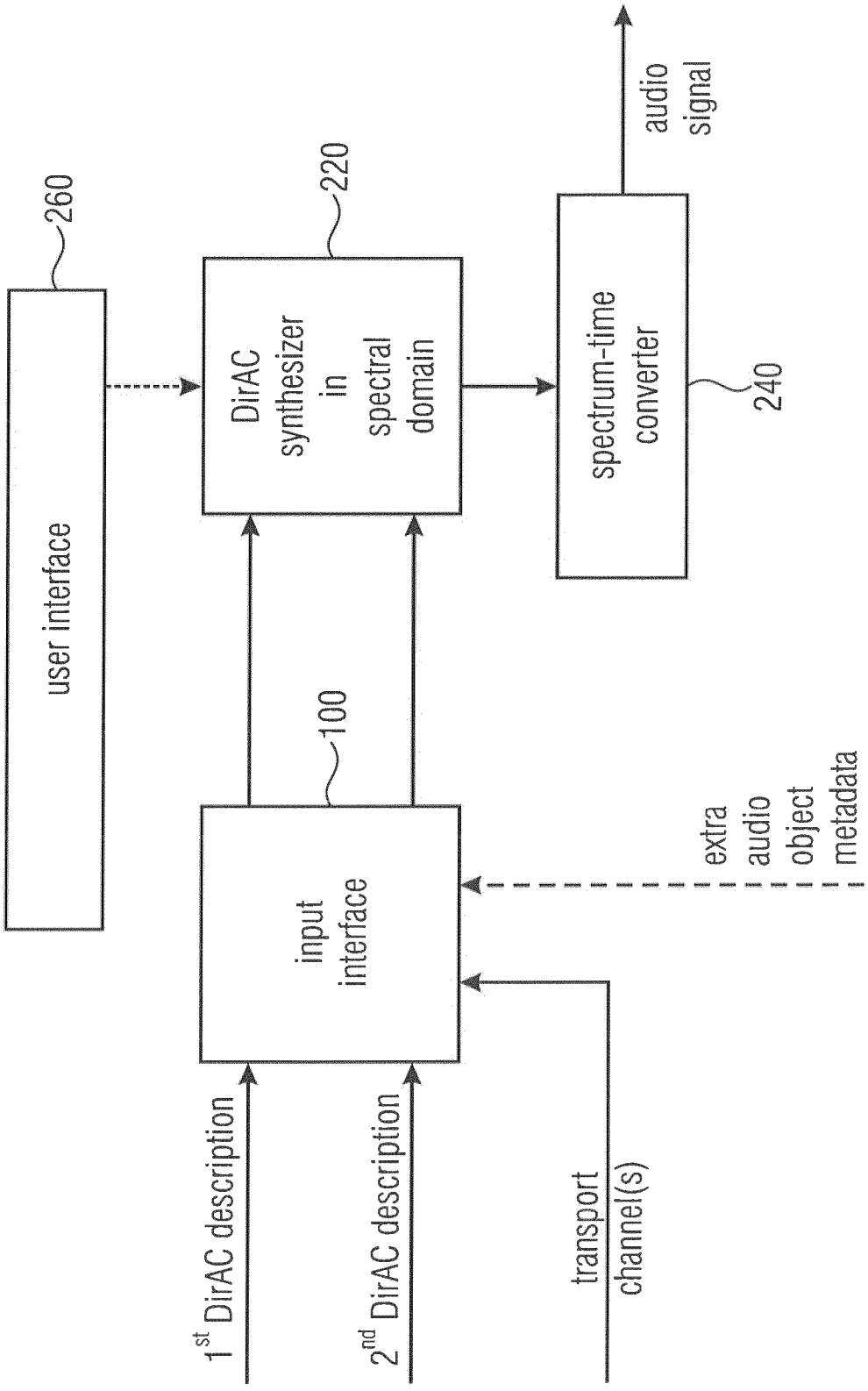
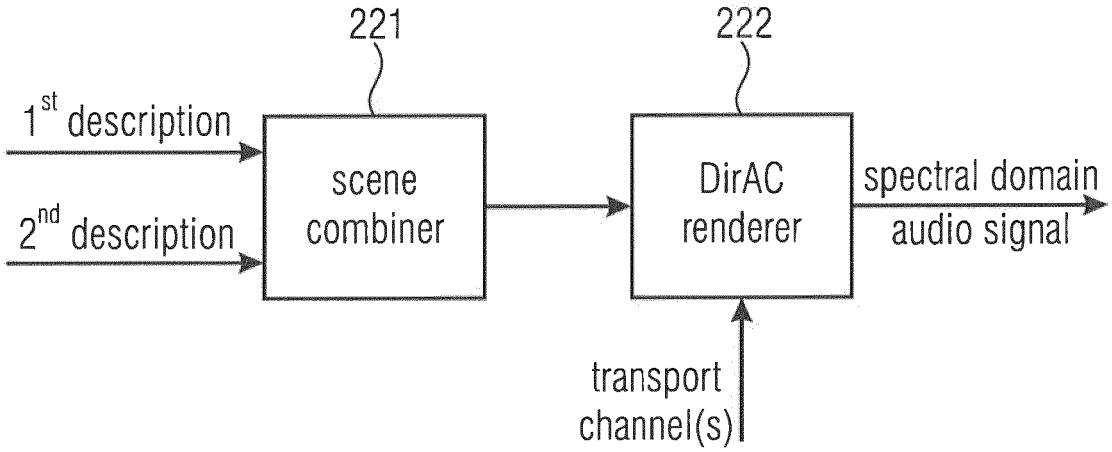
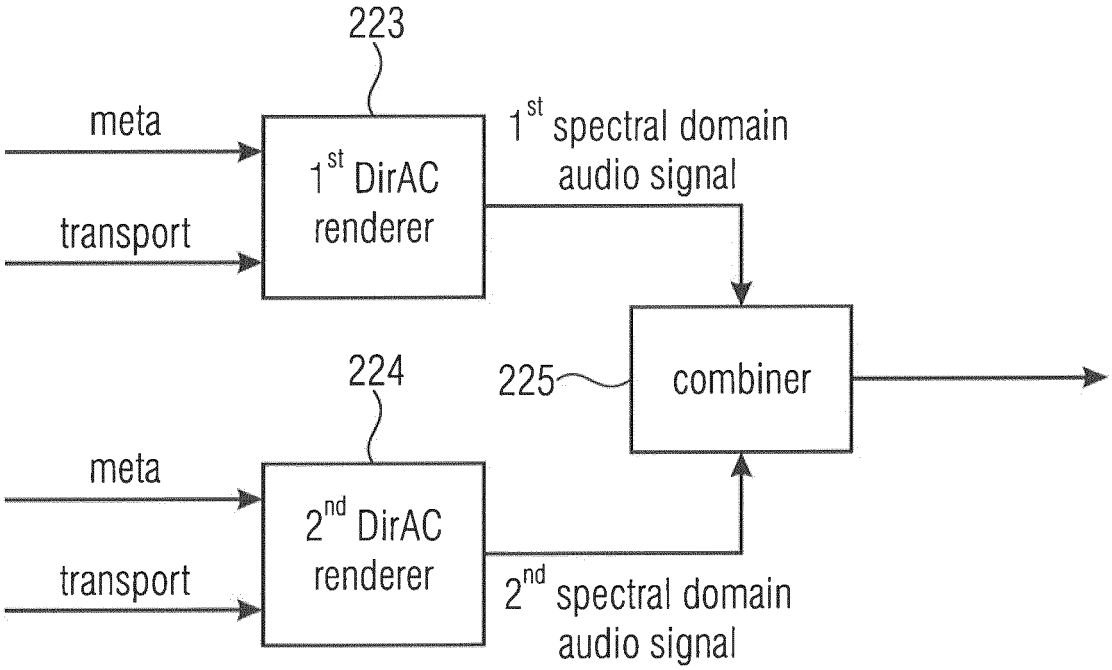

260
user interface
220
DirAC synthesizer in spectral domain
spectrum-time converter
240
audio signal
100
input interface
extra audio object metadata
1st DirAC description
2nd DirAC description
transport channel(s)

Fig. 2a



DirAC synthesizer

Fig. 2b



DirAC synthesizer

Fig. 2c

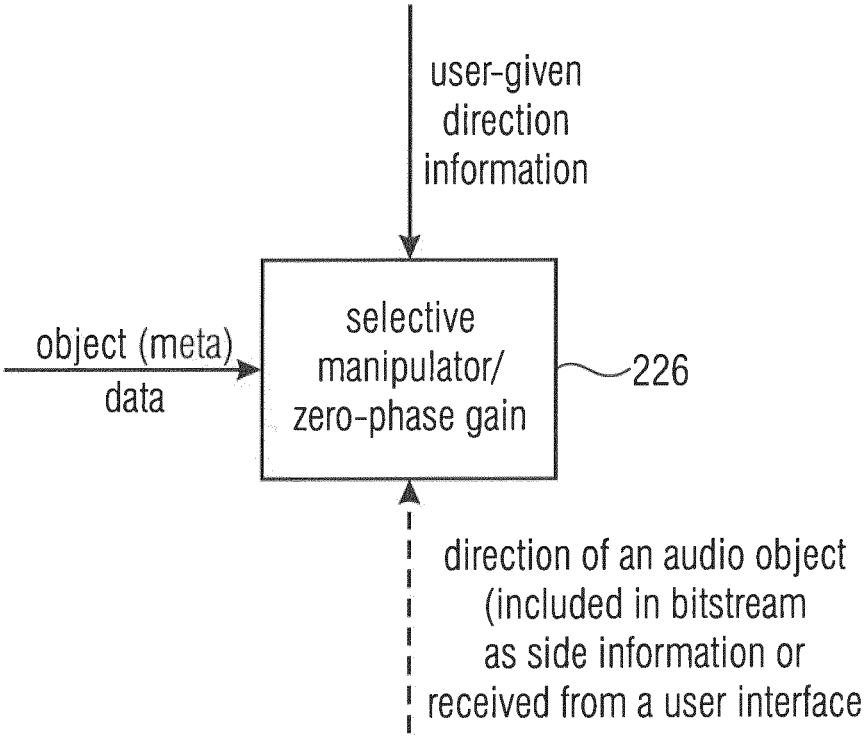

user-given
direction
information
object (meta)
data
selective
manipulator/
zero-phase gain
226
direction of an audio object
(included in bitstream
as side information or
received from a user interface

Fig. 2d

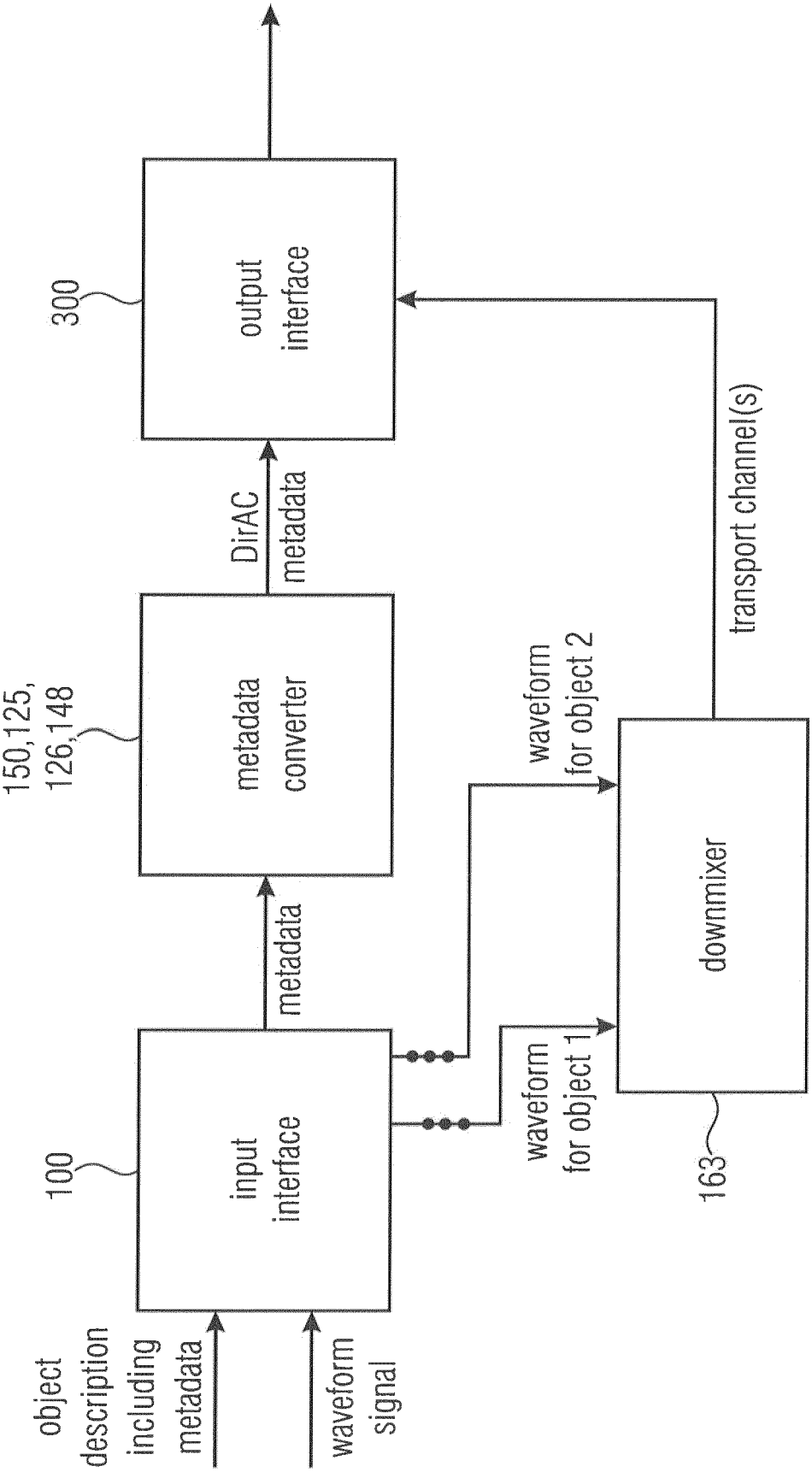

object description including metadata
waveform signal
100
input interface
metadata
150,125, 126,148
metadata converter
DirAC metadata
300
output interface
waveform for object 1
waveform for object 2
downmixer
163
transport channel(s)

Fig. 3a

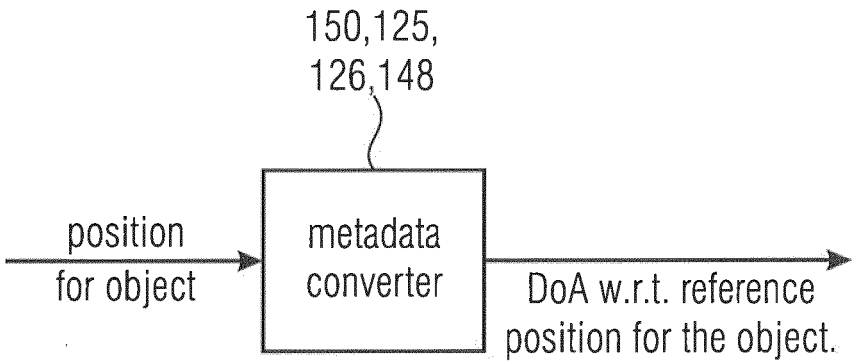

150,125,
126,148
position
for object
metadata
converter
DoA w.r.t. reference
position for the object.

Fig. 3b

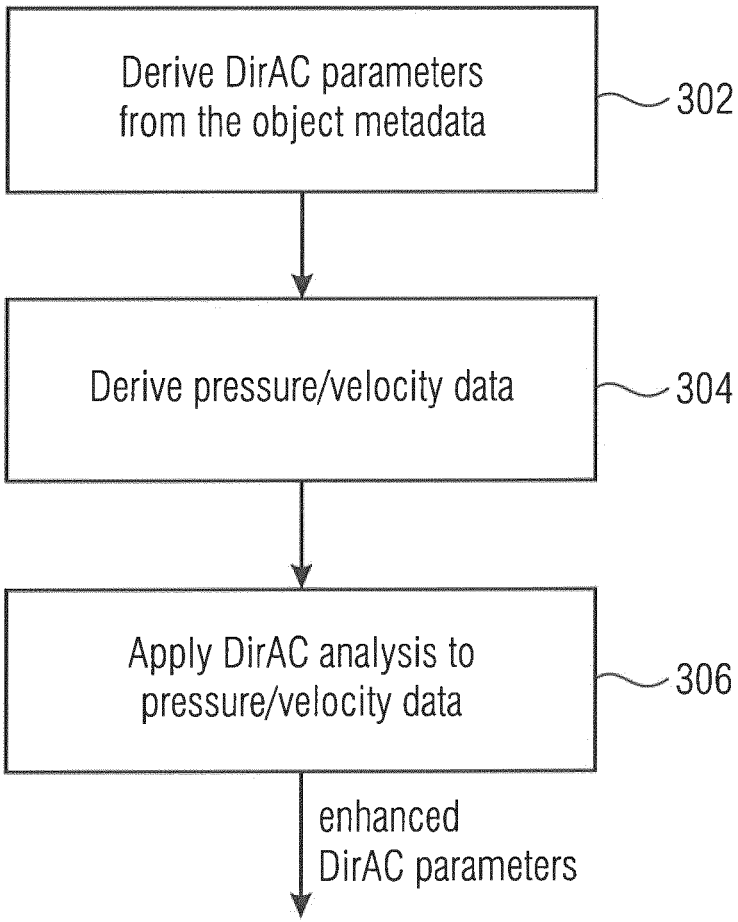

Derive DirAC parameters
from the object metadata
302
Derive pressure/velocity data
304
Apply DirAC analysis to
pressure/velocity data
306
enhanced
DirAC parameters

Fig. 3c

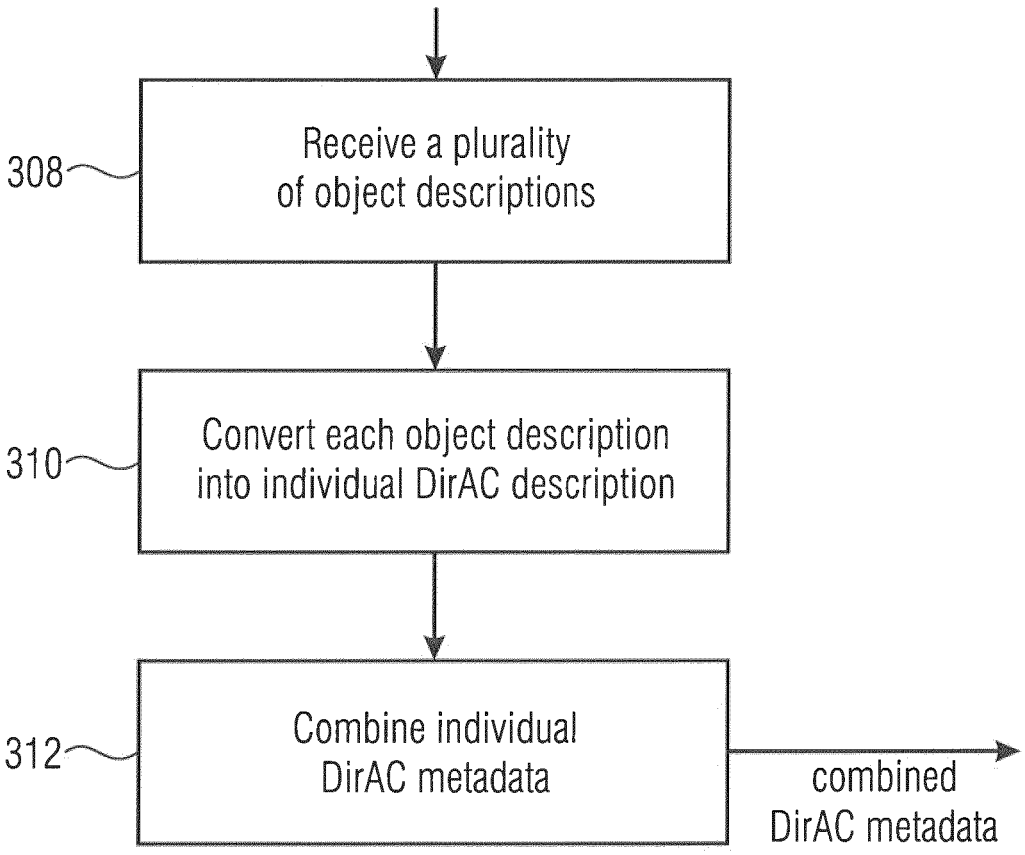

308
Receive a plurality of object descriptions
310
Convert each object description into individual DirAC description
312
Combine individual DirAC metadata
combined DirAC metadata

Fig. 3d

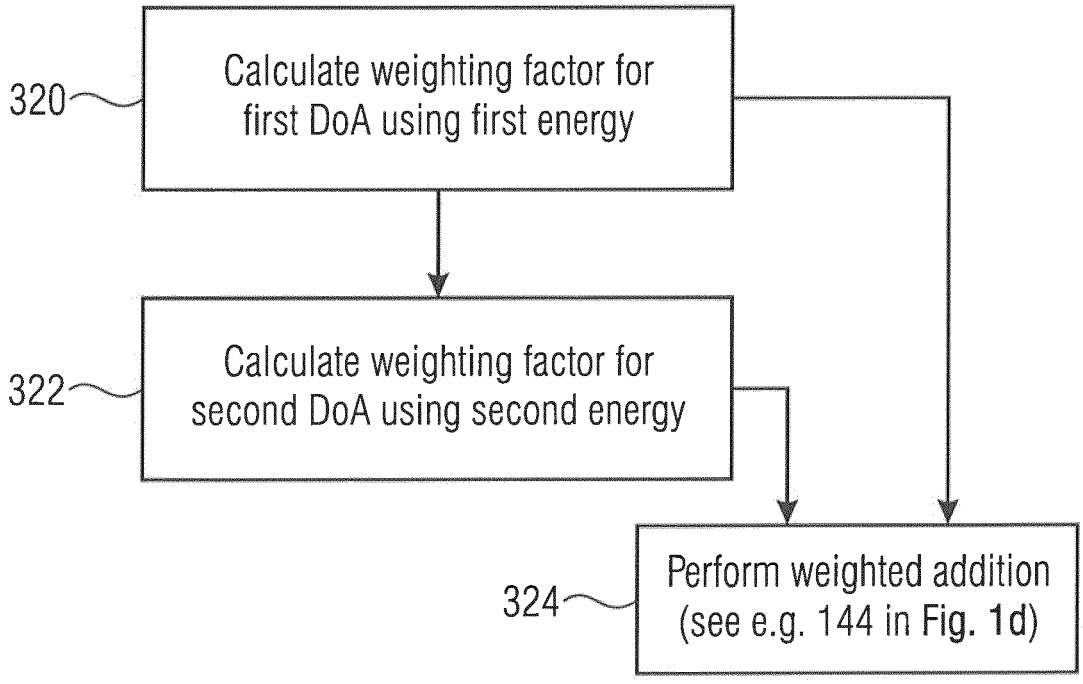
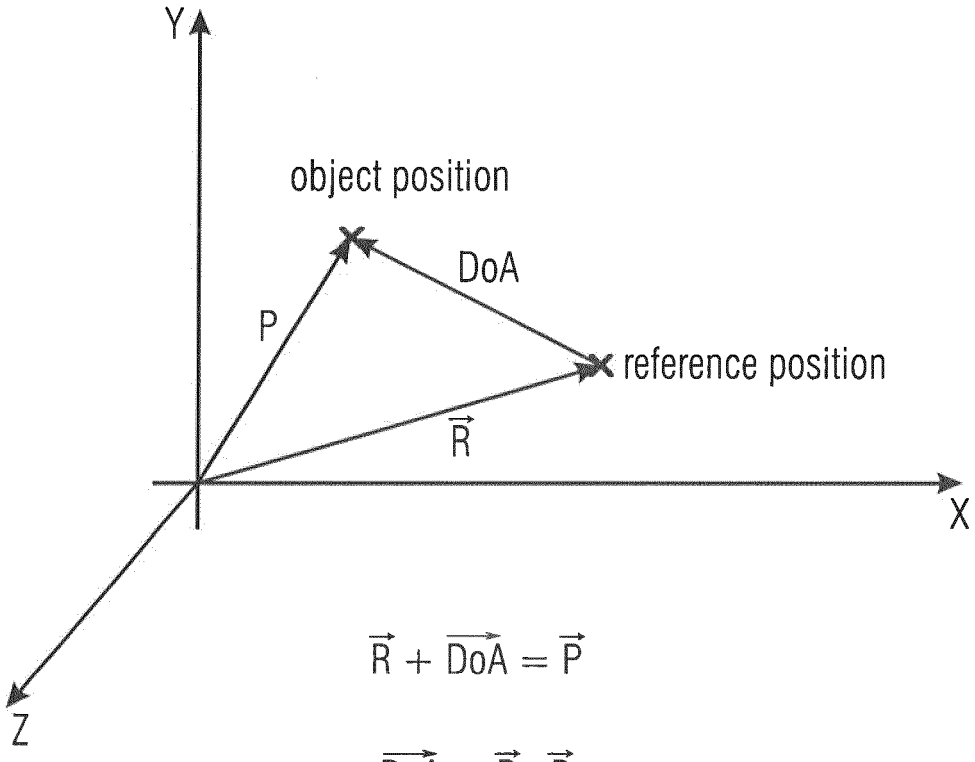

320
Calculate weighting factor for first DoA using first energy
322
Calculate weighting factor for second DoA using second energy
324
Perform weighted addition (see e.g. 144 in Fig. 1d)

Fig. 3e



$$\vec{R} + \overrightarrow{DoA} = \vec{P}$$

$$\overrightarrow{DoA} = \vec{P} - \vec{R}$$

$$\overrightarrow{doa} = \frac{\vec{P} - \vec{R}}{|\overrightarrow{DoA}|}$$

Fig. 3f

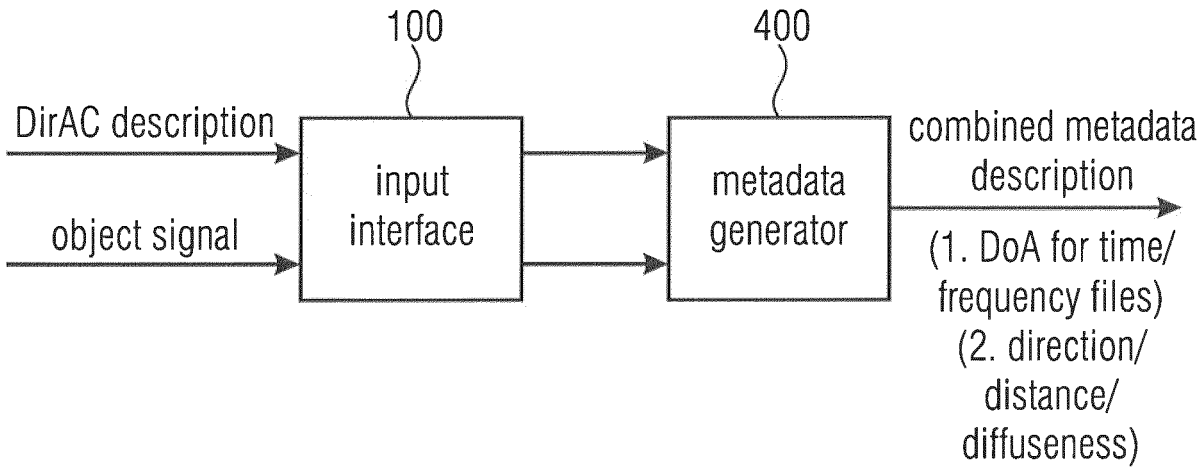

100
400
DirAC description
object signal
input interface
metadata generator
combined metadata description
(1. DoA for time/ frequency files)
(2. direction/ distance/ diffuseness)

Fig. 4a

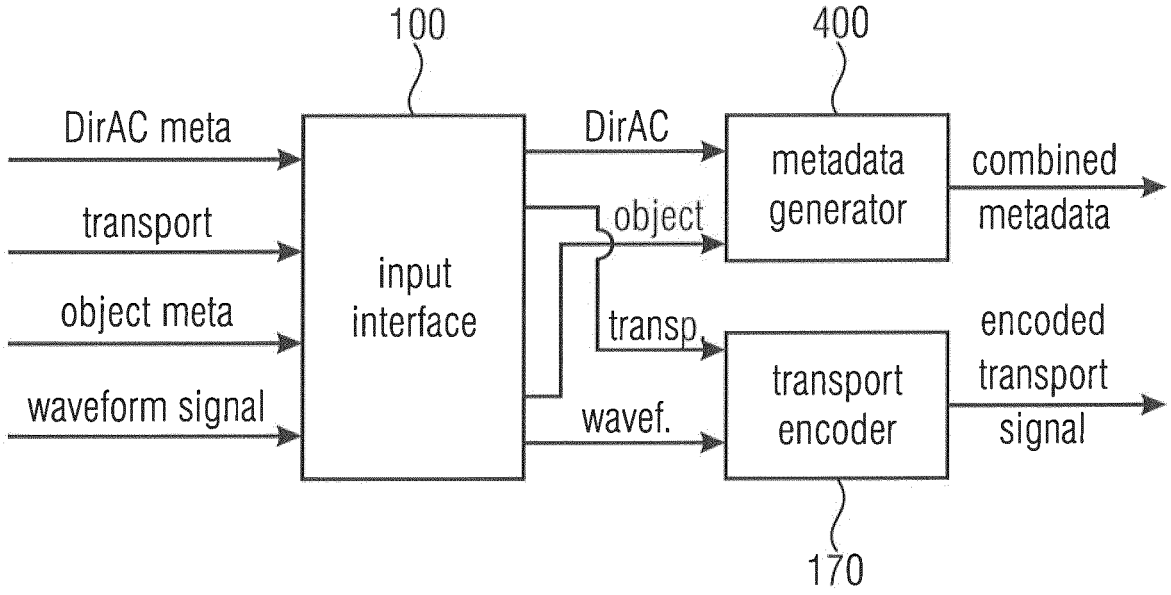

100
400
DirAC meta
transport
object meta
waveform signal
input interface
DirAC
object
transp.
wavef.
metadata generator
combined metadata
transport encoder
encoded transport signal
170

Fig. 4b

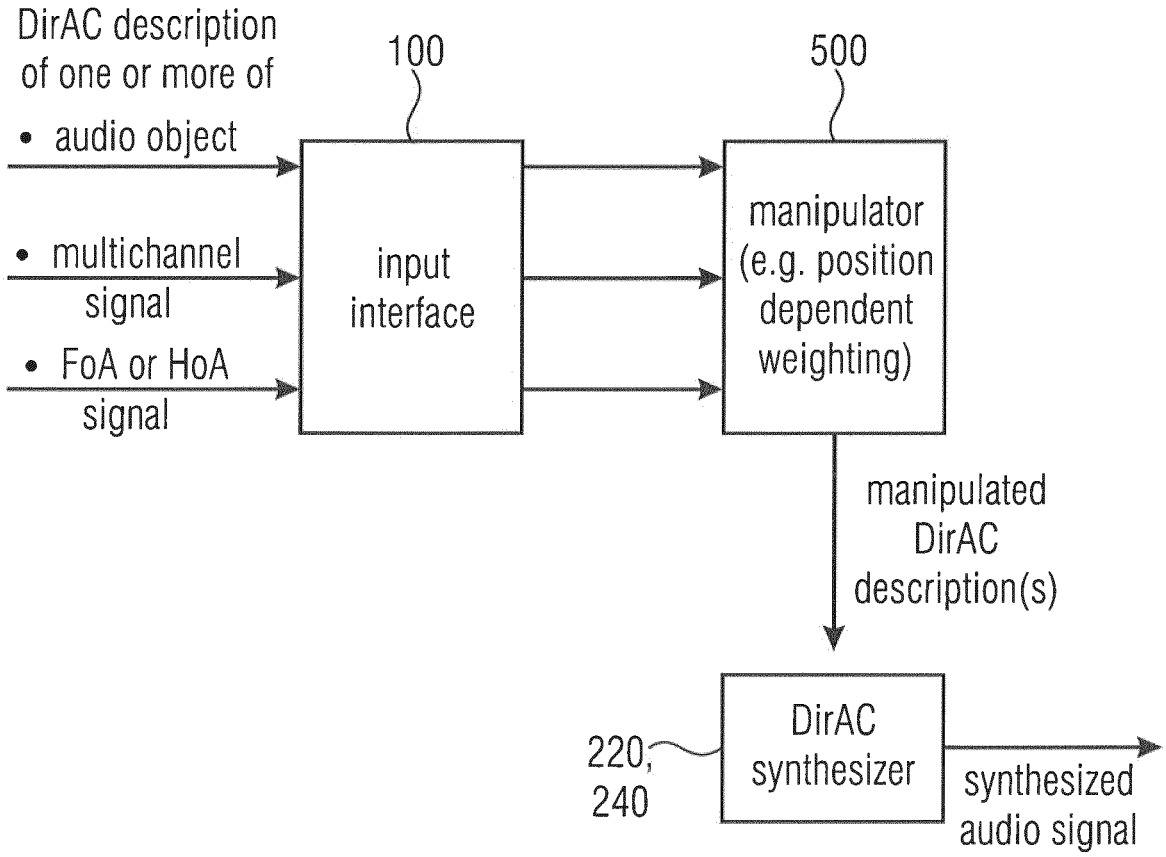

DirAC description
of one or more of
• audio object
• multichannel signal
• FoA or HoA signal
100
input interface
500
manipulator (e.g. position dependent weighting)
manipulated DirAC description(s)
220; 240
DirAC synthesizer
synthesized audio signal

Fig. 5a

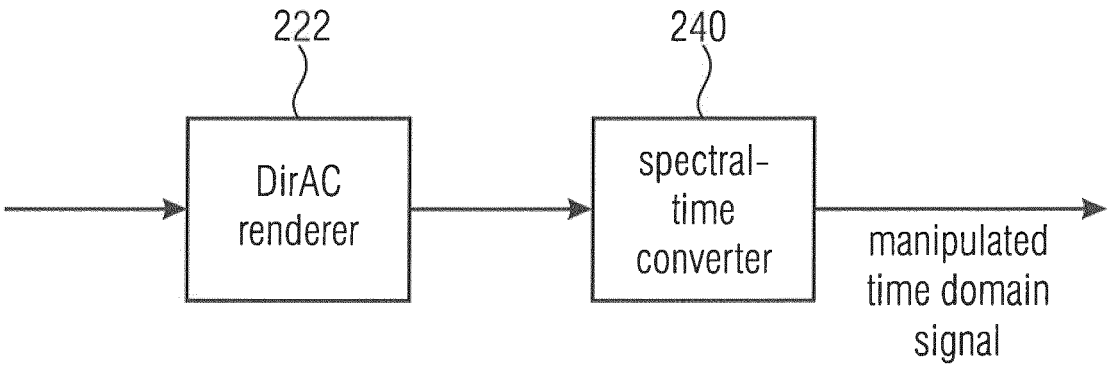

222
DirAC renderer
240
spectral-time converter
manipulated time domain signal

Fig. 5b

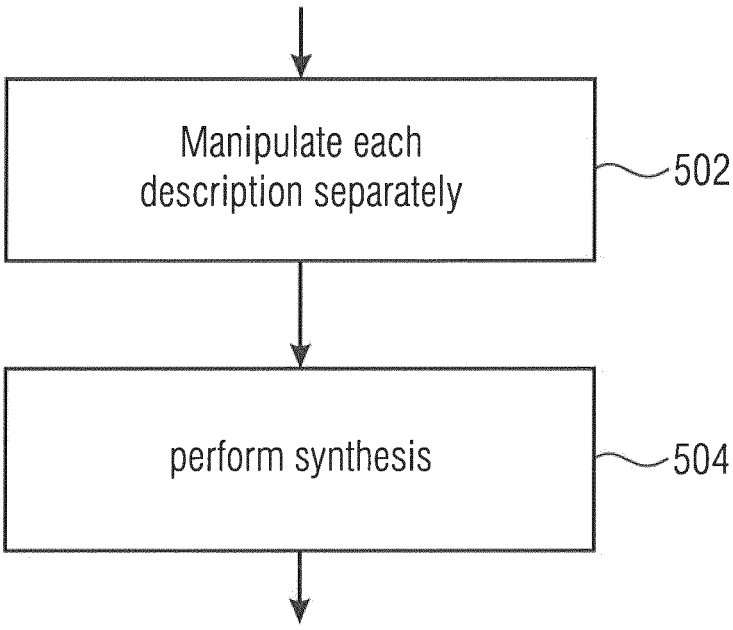

Manipulate each description separately
502
perform synthesis
504

Fig. 5c

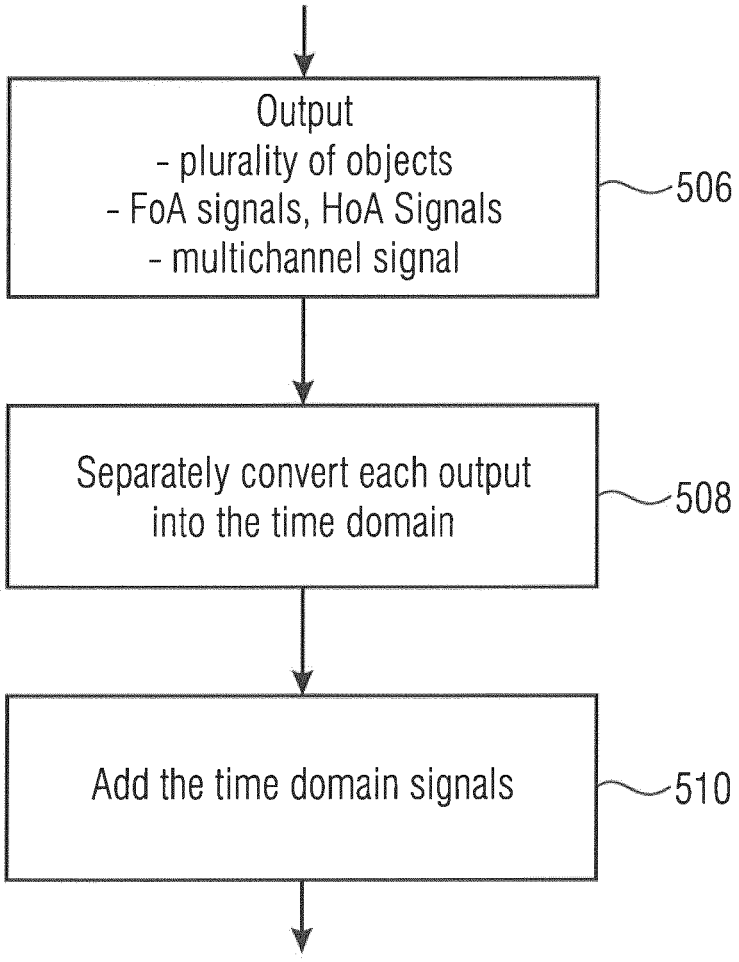
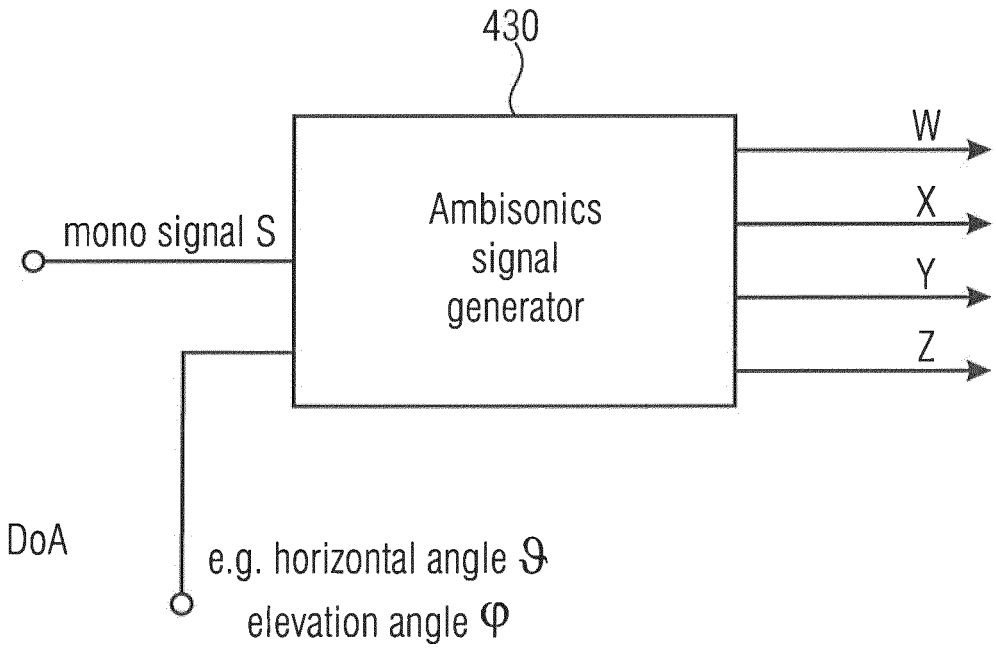

Output
- plurality of objects
- FoA signals, HoA Signals
- multichannel signal
506
Separately convert each output into the time domain
508
Add the time domain signals
510

Fig. 5d



$$\text{omnidir. comp } W = S \cdot \frac{1}{\sqrt{2}}$$

$$\text{directional components} \begin{cases} X = S \cdot \cos\vartheta \cos\varphi \\ Y = S \cdot \sin\vartheta \cos\varphi \\ Z = S \cdot \sin\varphi \end{cases}$$

Fig. 6

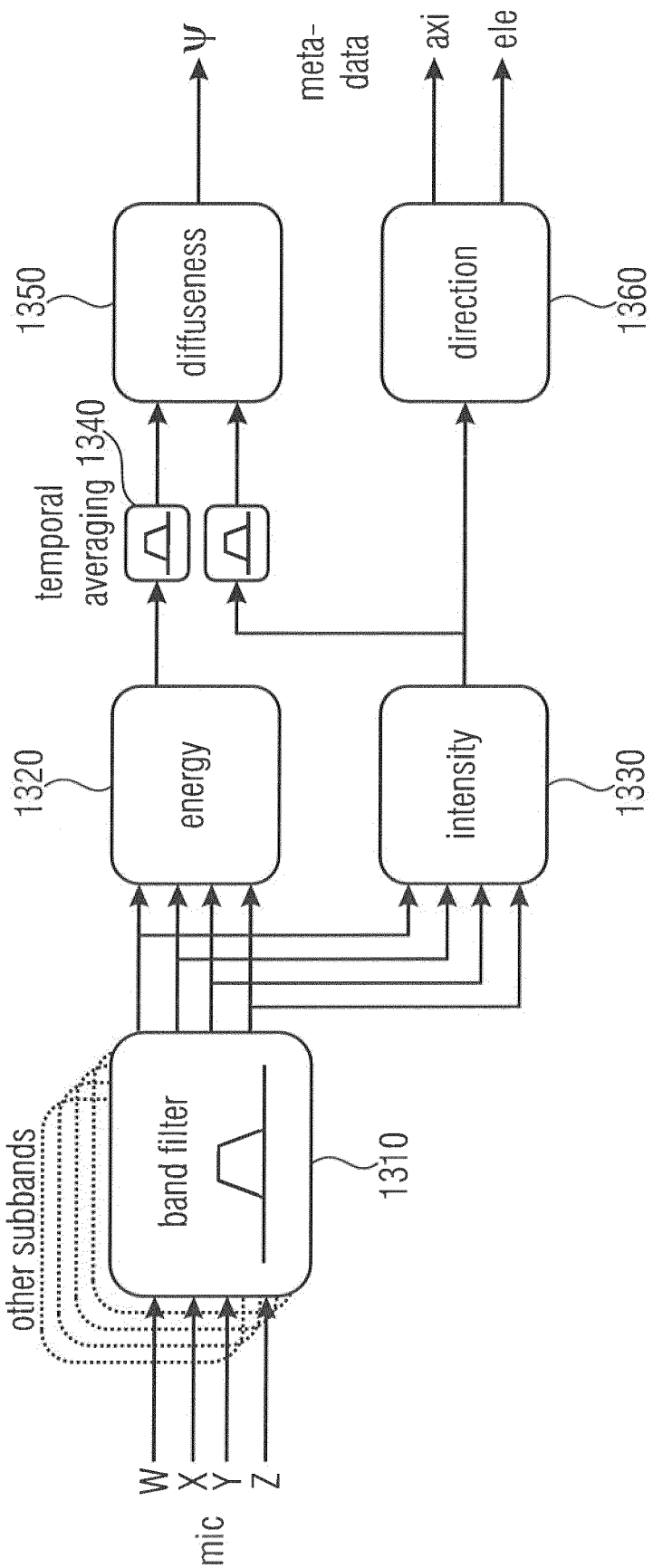
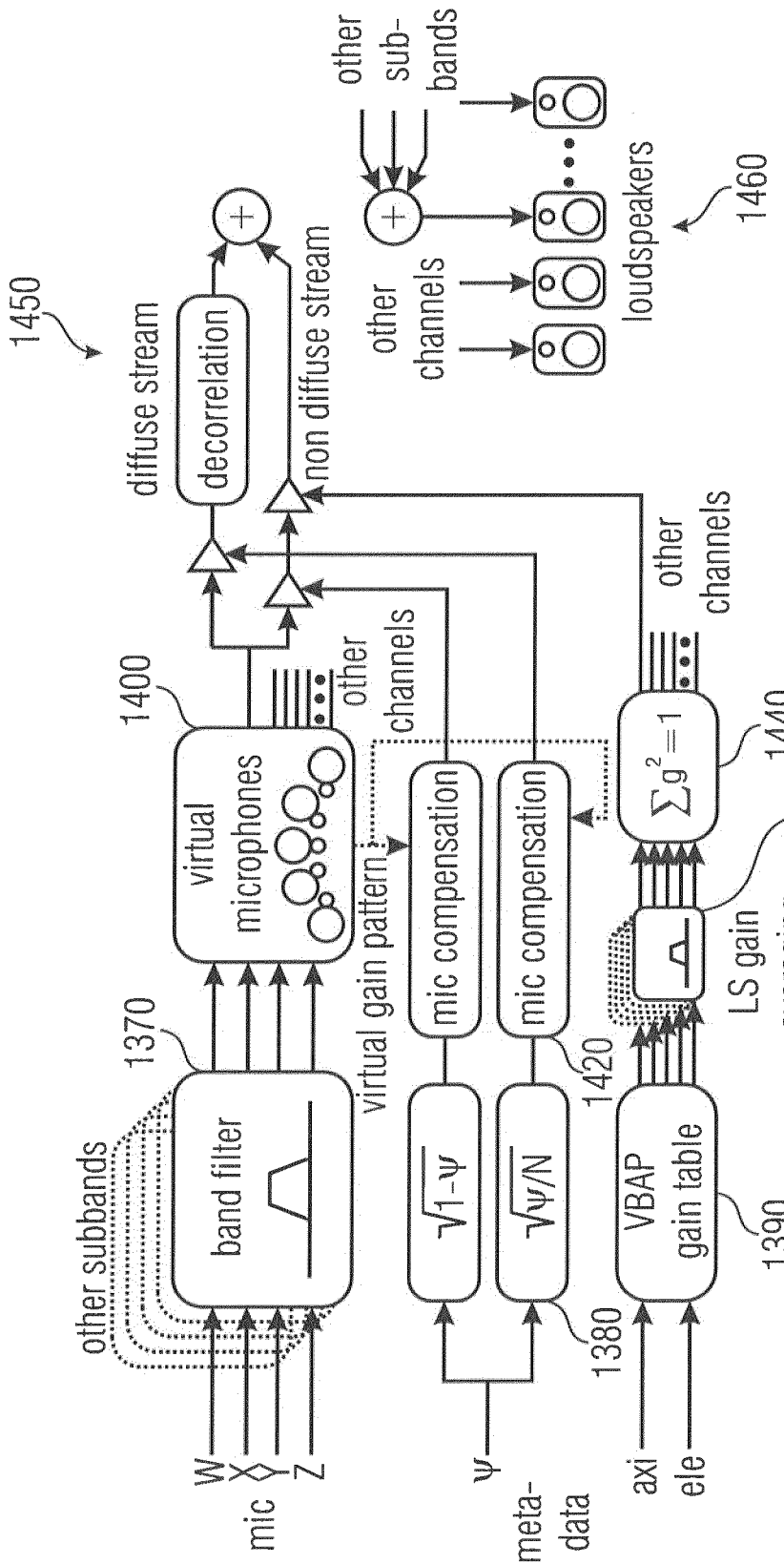


Fig. 7a
(PRIOR ART)



DirAC synthesis

Fig. 7b
(PRIOR ART)

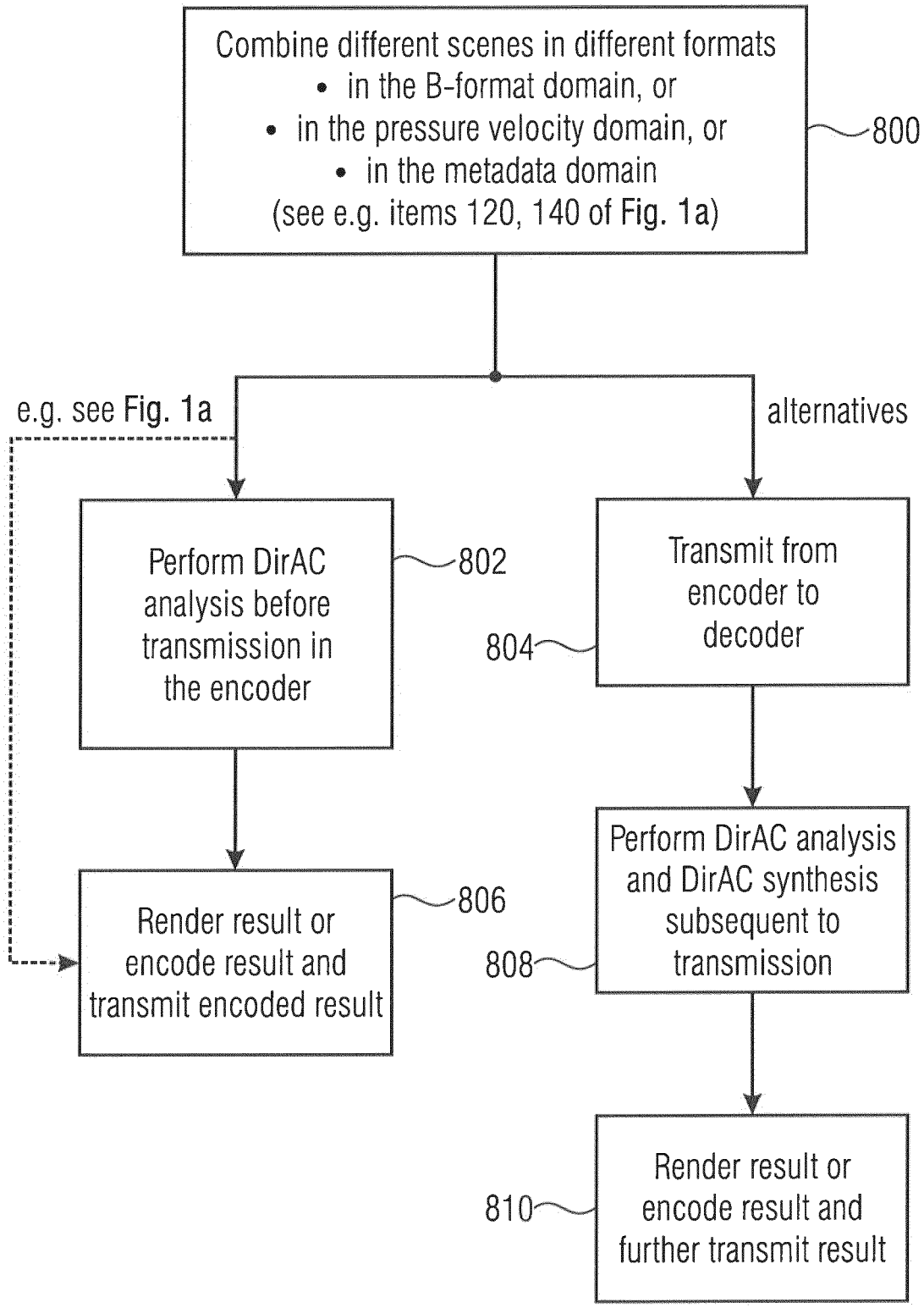

Combine different scenes in different formats
• in the B-format domain, or
• in the pressure velocity domain, or
• in the metadata domain
(see e.g. items 120, 140 of Fig. 1a)
800
e.g. see Fig. 1a
alternatives
Perform DirAC analysis before transmission in the encoder
802
Transmit from encoder to decoder
804
Render result or encode result and transmit encoded result
806
Perform DirAC analysis and DirAC synthesis subsequent to transmission
808
Render result or encode result and further transmit result
810

Fig. 8

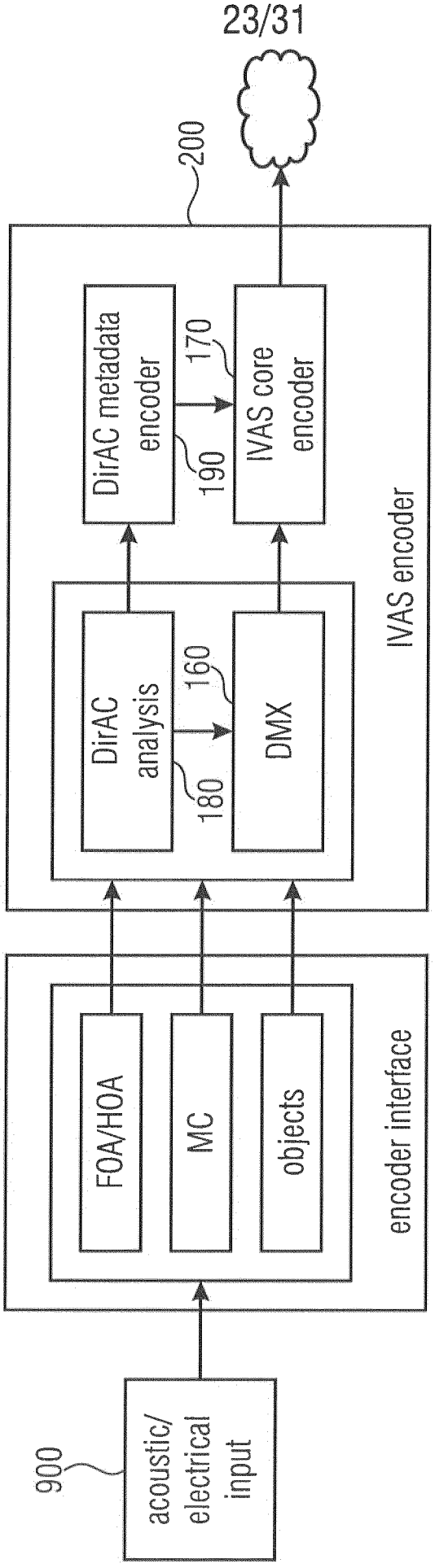

200
DirAC metadata encoder
170
IVAS core encoder
190
IVAS encoder
DirAC analysis
160
DMX
180
FOA/HOA
MC
objects
encoder interface
900
acoustic/ electrical input

Fig. 9

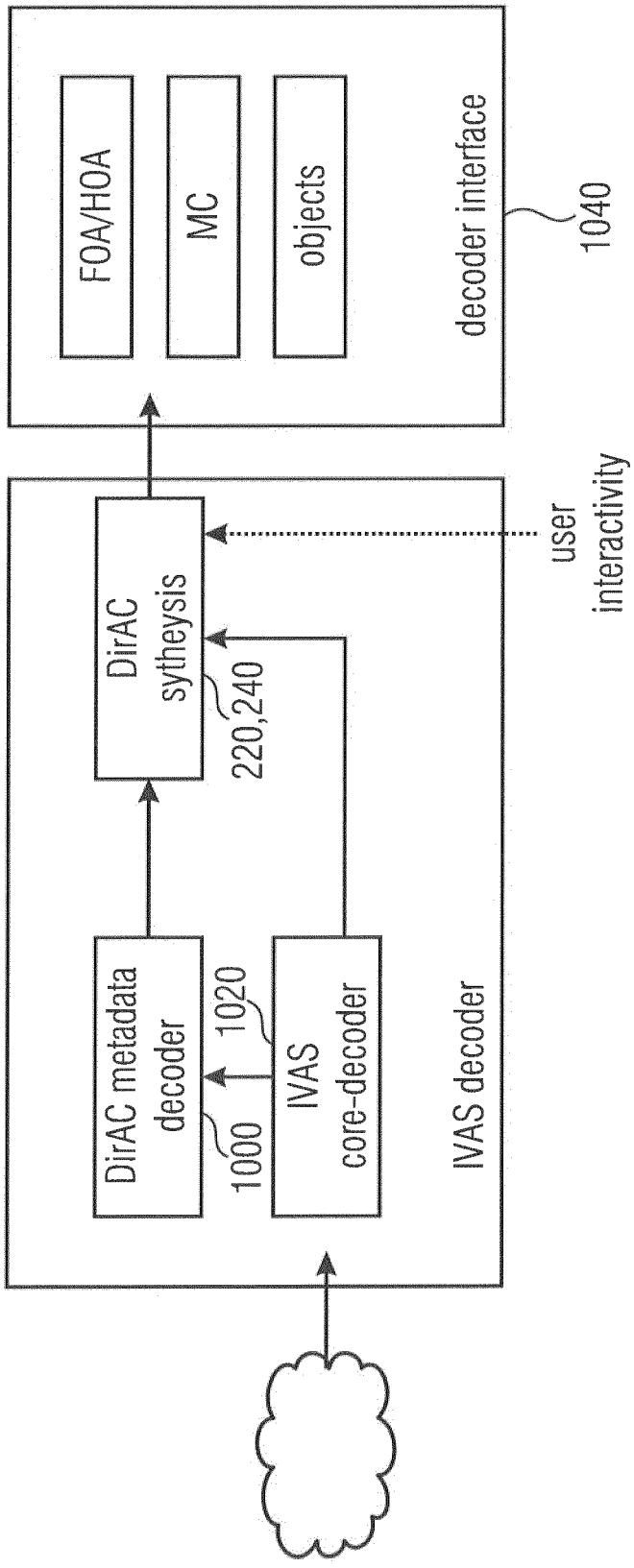

FOA/HOA
MC
objects
decoder interface
1040
user interactivity
DirAC sytheysis
220,240
DirAC metadata decoder
1000
1020
IVAS core-decoder
IVAS decoder

Fig. 10

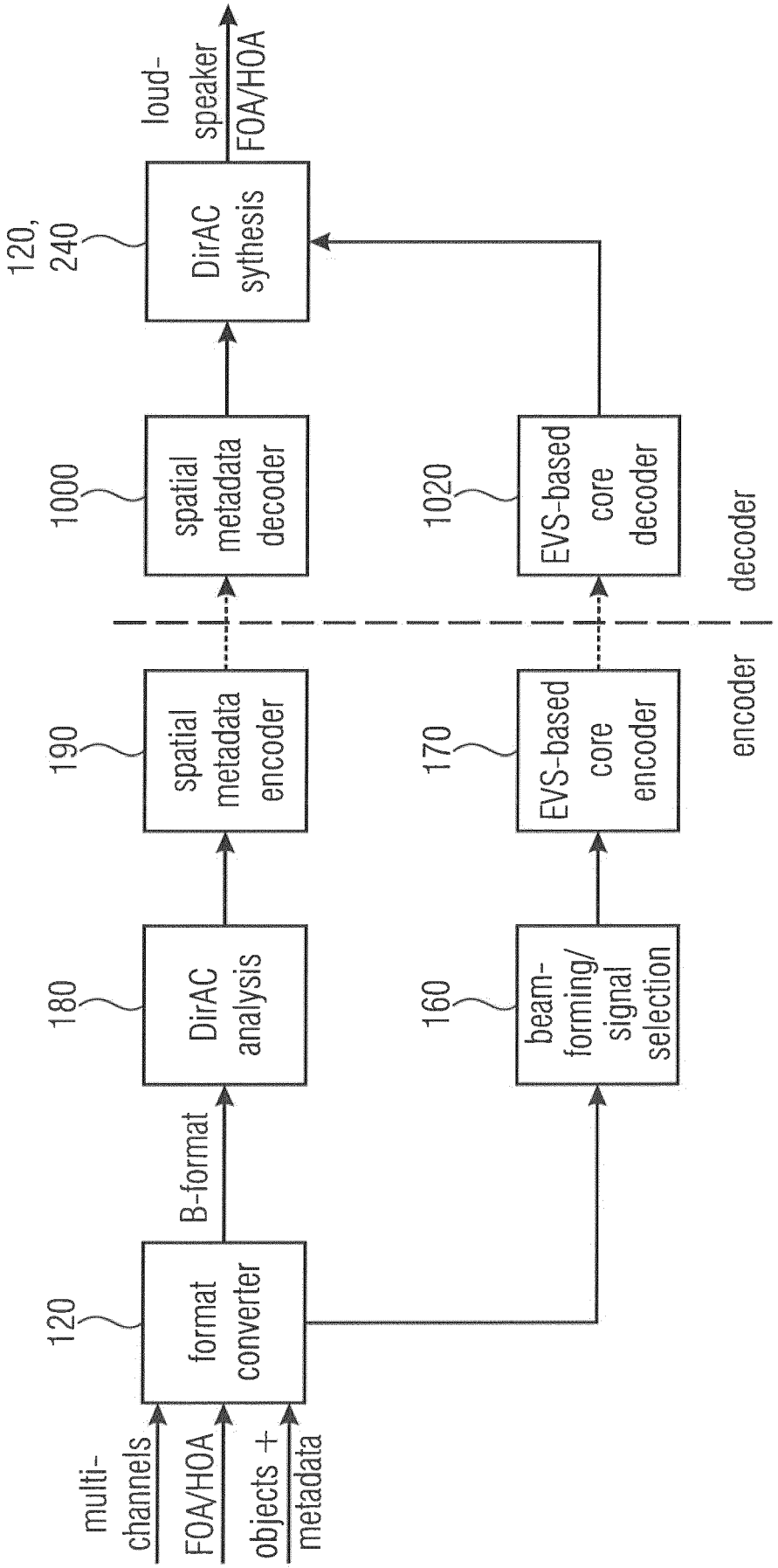

multi-channels
FOA/HOA
objects + metadata
120
format converter
B-format
180
DirAC analysis
190
spatial metadata encoder
160
beam-forming/ signal selection
170
EVS-based core encoder
encoder
1000
spatial metadata decoder
1020
EVS-based core decoder
decoder
120, 240
DirAC sythesis
loud-speaker FOA/HOA

Fig. 11

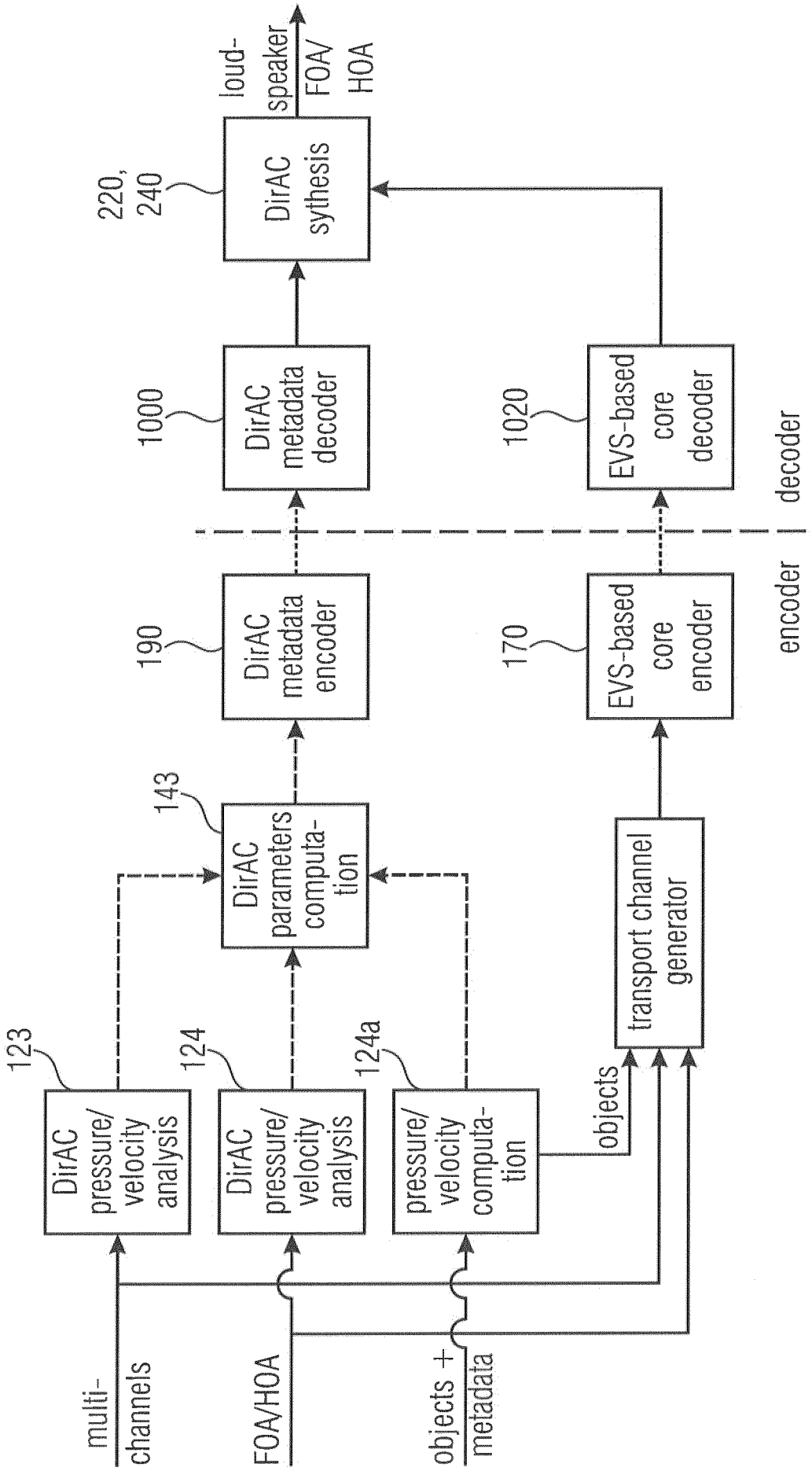

multi-channels
FOA/HOA
objects + metadata
DirAC pressure/ velocity analysis
123
DirAC pressure/ velocity analysis
124
pressure/ velocity computa-tion
124a
DirAC parameters computa-tion
143
objects
transport channel generator
DirAC metadata encoder
190
EVS-based core encoder
170
encoder
DirAC metadata decoder
1000
EVS-based core decoder
1020
decoder
DirAC sythesis
220, 240
loud-speaker FOA/ HOA

Fig. 12

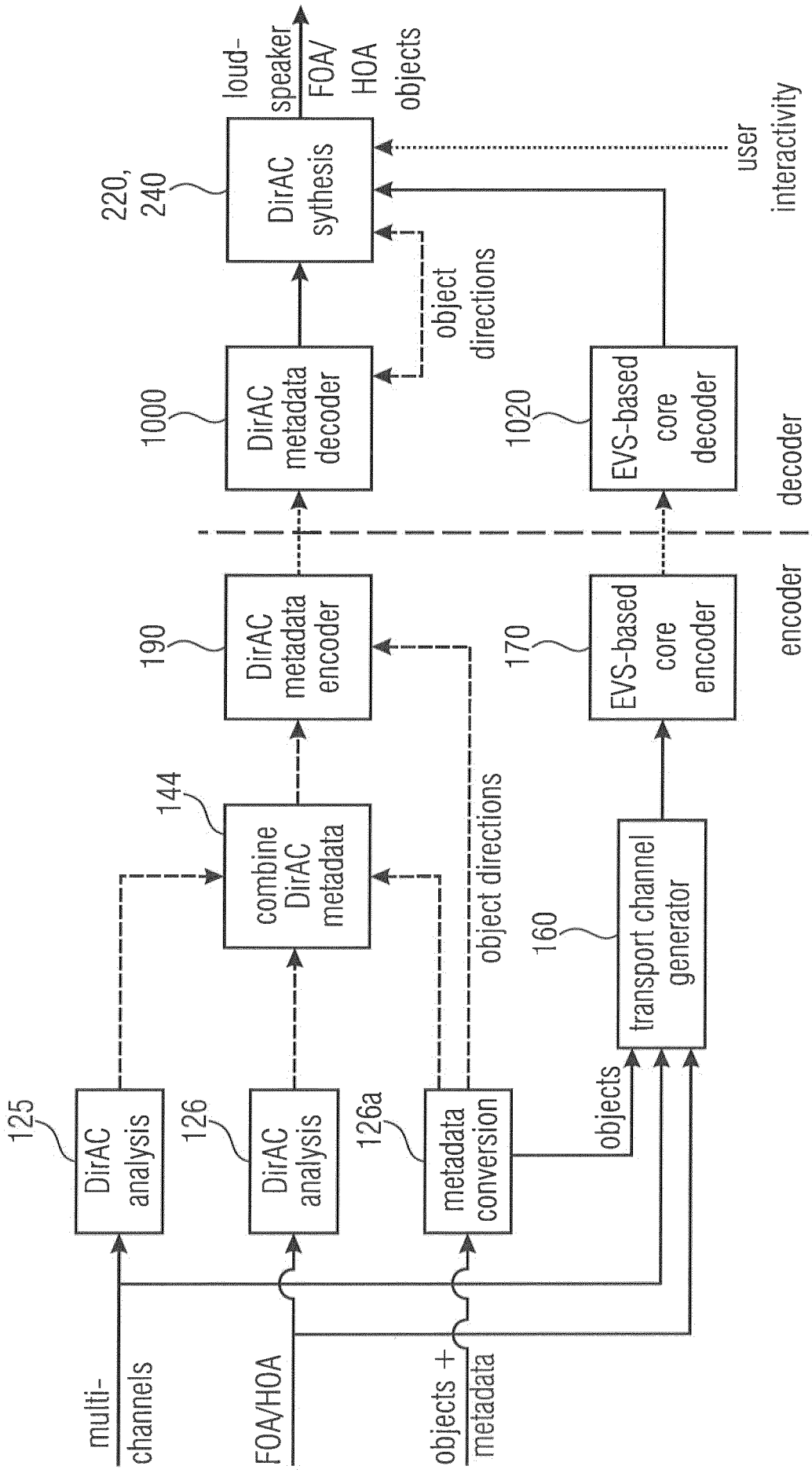

multi-channels
FOA/HOA
objects + metadata
125
DirAC analysis
126
DirAC analysis
126a
metadata conversion
144
combine DirAC metadata
190
DirAC metadata encoder
160
transport channel generator
objects
170
EVS-based core encoder
encoder
object directions
1000
DirAC metadata decoder
1020
EVS-based core decoder
decoder
220, 240
DirAC sythesis
loud-speaker
FOA/ HOA
objects
user interactivity

Fig. 13

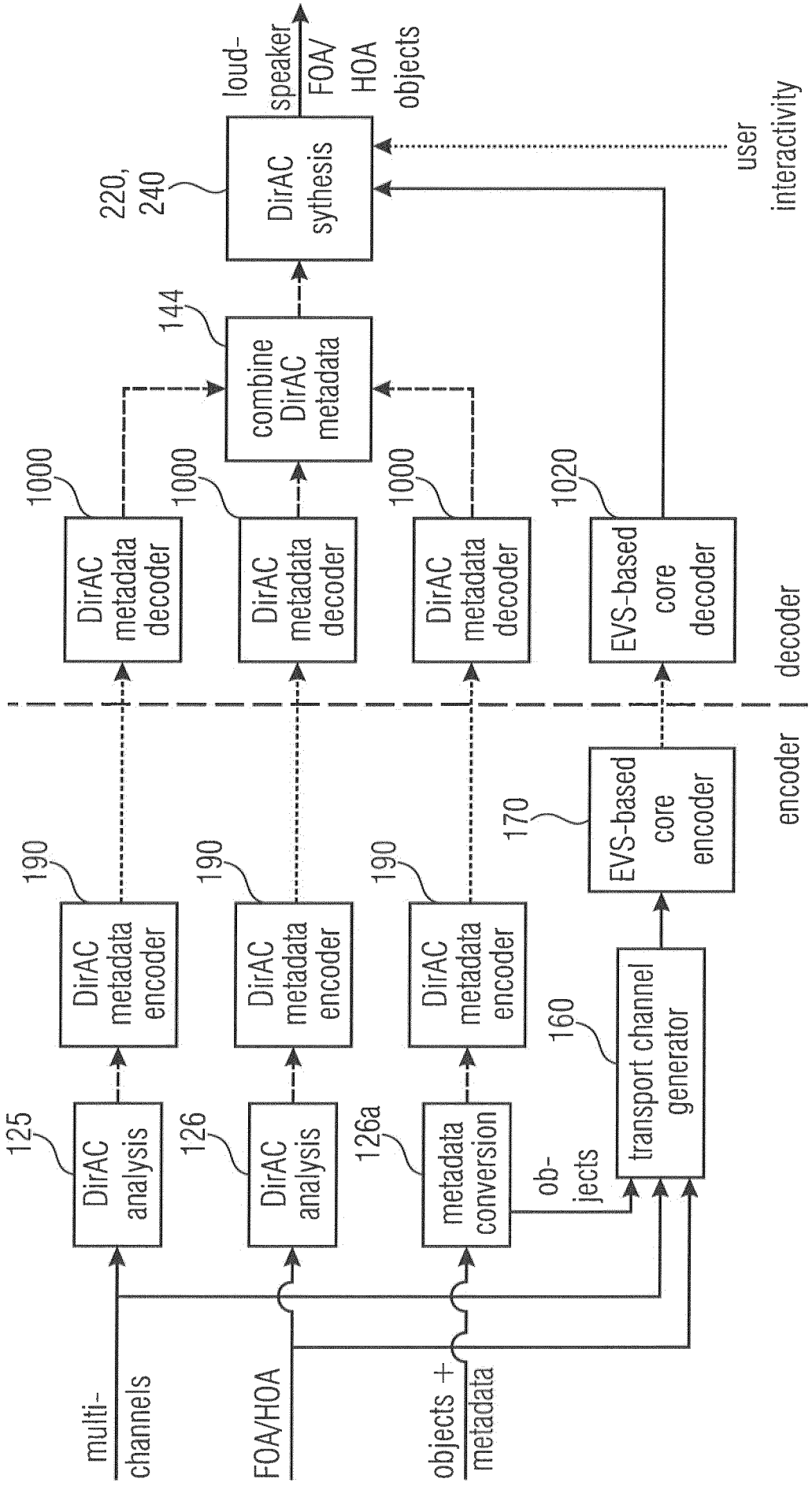

multi-channels
FOA/HOA
objects + metadata
125
DirAC analysis
126
DirAC analysis
126a
metadata conversion
ob-jects
160
transport channel generator
190
DirAC metadata encoder
190
DirAC metadata encoder
190
DirAC metadata encoder
170
EVS-based core encoder
encoder
1000
DirAC metadata decoder
1000
DirAC metadata decoder
1000
DirAC metadata decoder
1020
EVS-based core decoder
decoder
144
combine DirAC metadata
220, 240
DirAC sythesis
loud-speaker
FOA/ HOA objects
user interactivity

Fig. 14

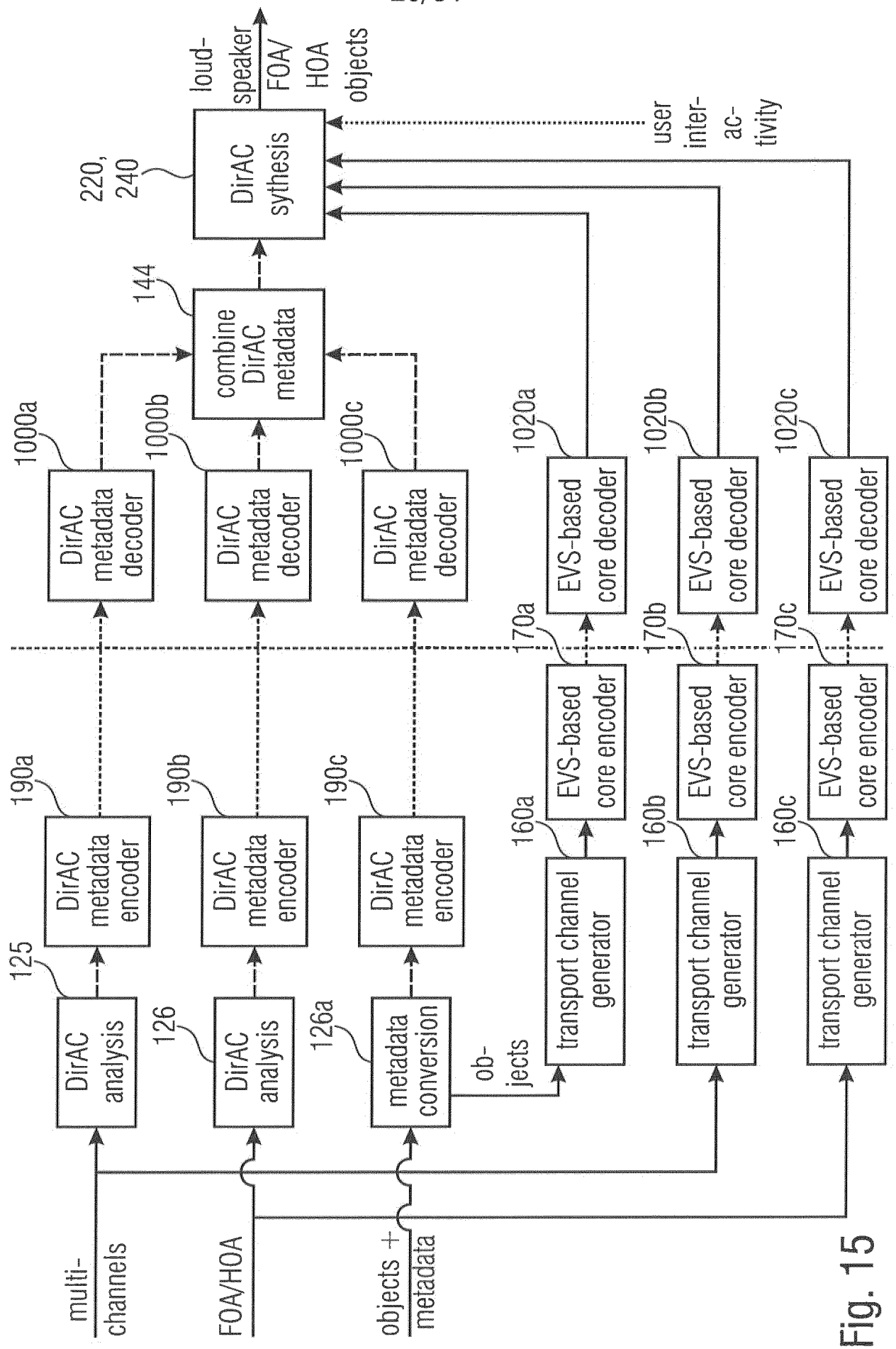

multi-channels
FOA/HOA
objects + metadata
DirAC analysis
DirAC analysis
metadata conversion
125
126
126a
DirAC metadata encoder
DirAC metadata encoder
DirAC metadata encoder
190a
190b
190c
transport channel generator
transport channel generator
transport channel generator
ob-jects
160a
160b
160c
EVS-based core encoder
EVS-based core encoder
EVS-based core encoder
170a
170b
170c
EVS-based core decoder
EVS-based core decoder
EVS-based core decoder
1020a
1020b
1020c
DirAC metadata decoder
DirAC metadata decoder
DirAC metadata decoder
1000a
1000b
1000c
combine DirAC metadata
144
DirAC sythesis
220, 240
loud-speaker
FOA/ HOA objects
user inter-ac-tivity

Fig. 15

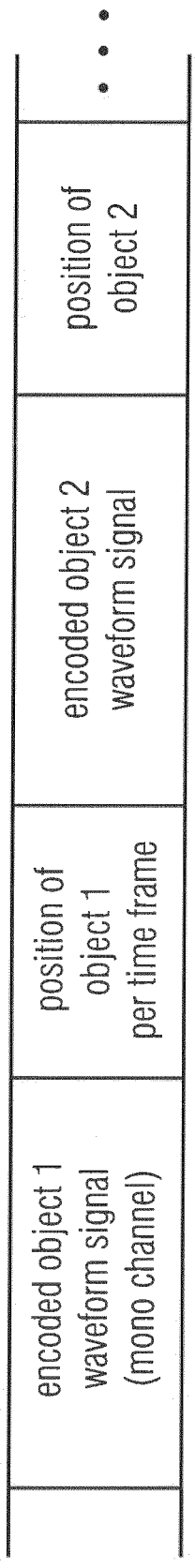


Fig. 16a

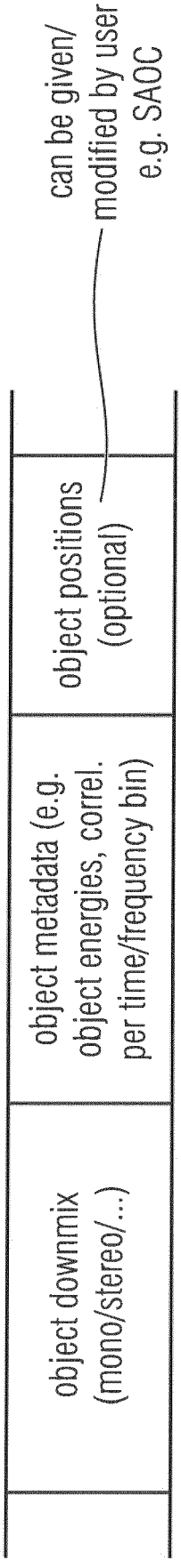


Fig. 16b

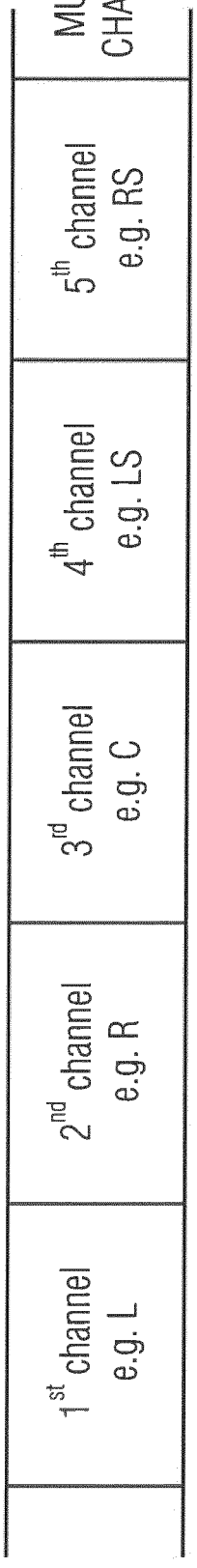


Fig. 16c

e.g. MPEG SURROUND

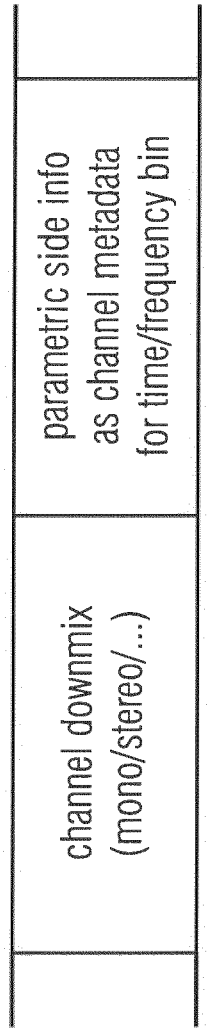


Fig. 16d

FoA HoA

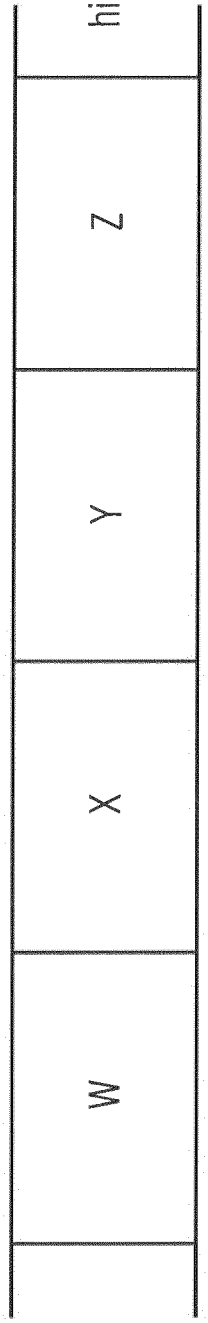


Fig. 16e

DirAC

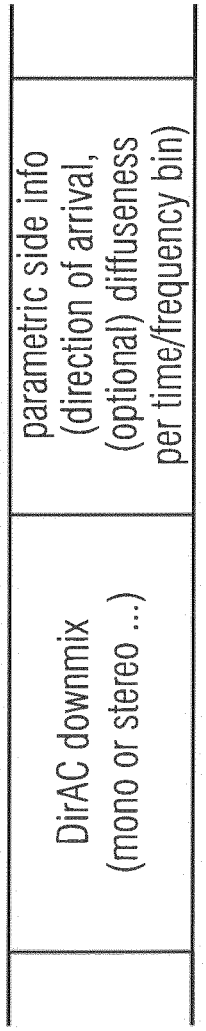


Fig. 16f