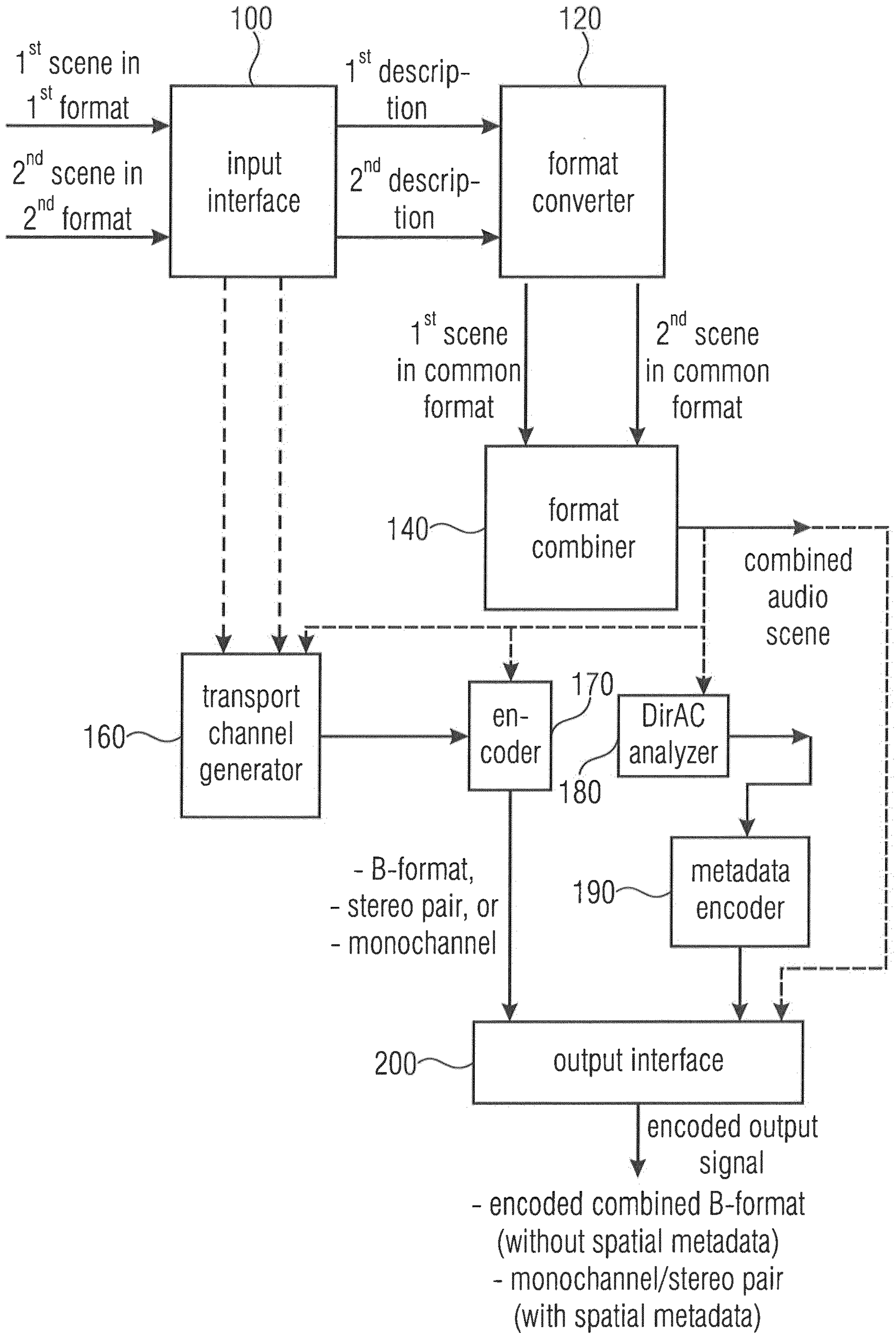
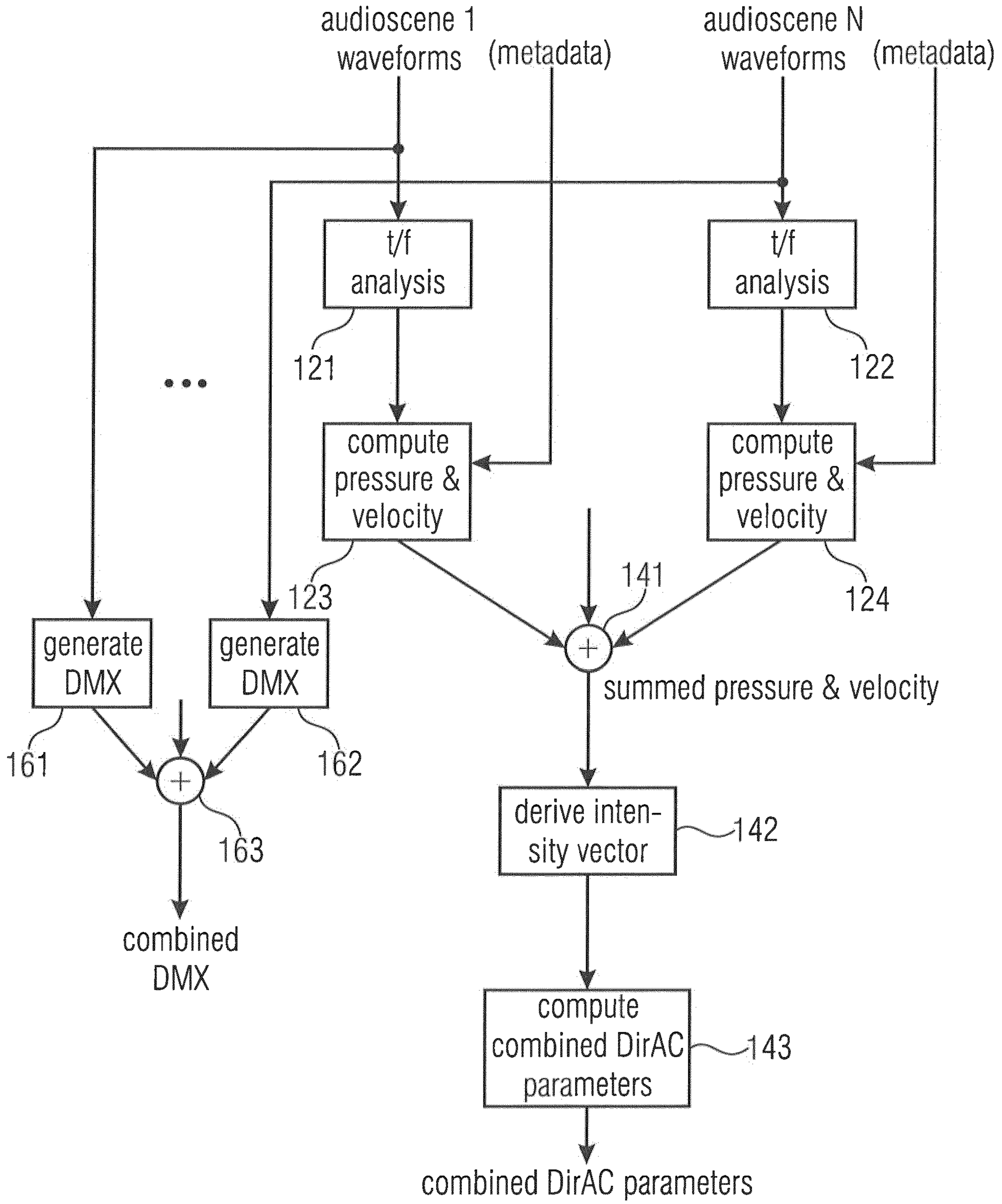
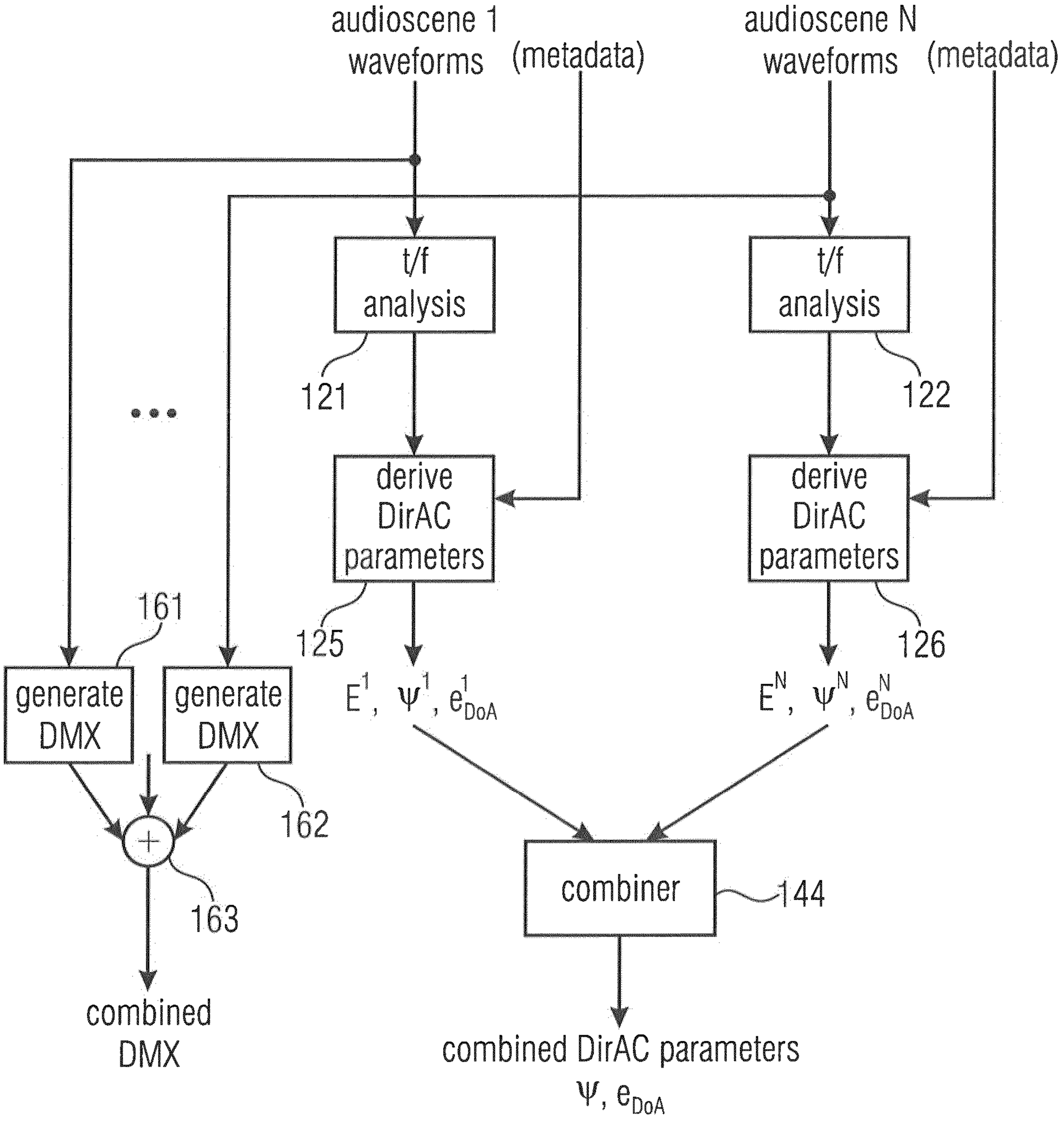

100
120
1st scene in 1st format
2nd scene in 2nd format
input interface
1st description
2nd description
format converter
1st scene in common format
2nd scene in common format
140
format combiner
combined audio scene
transport channel generator
160
en-coder
170
DirAC analyzer
180
- B-format,
- stereo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

Fig. 1a



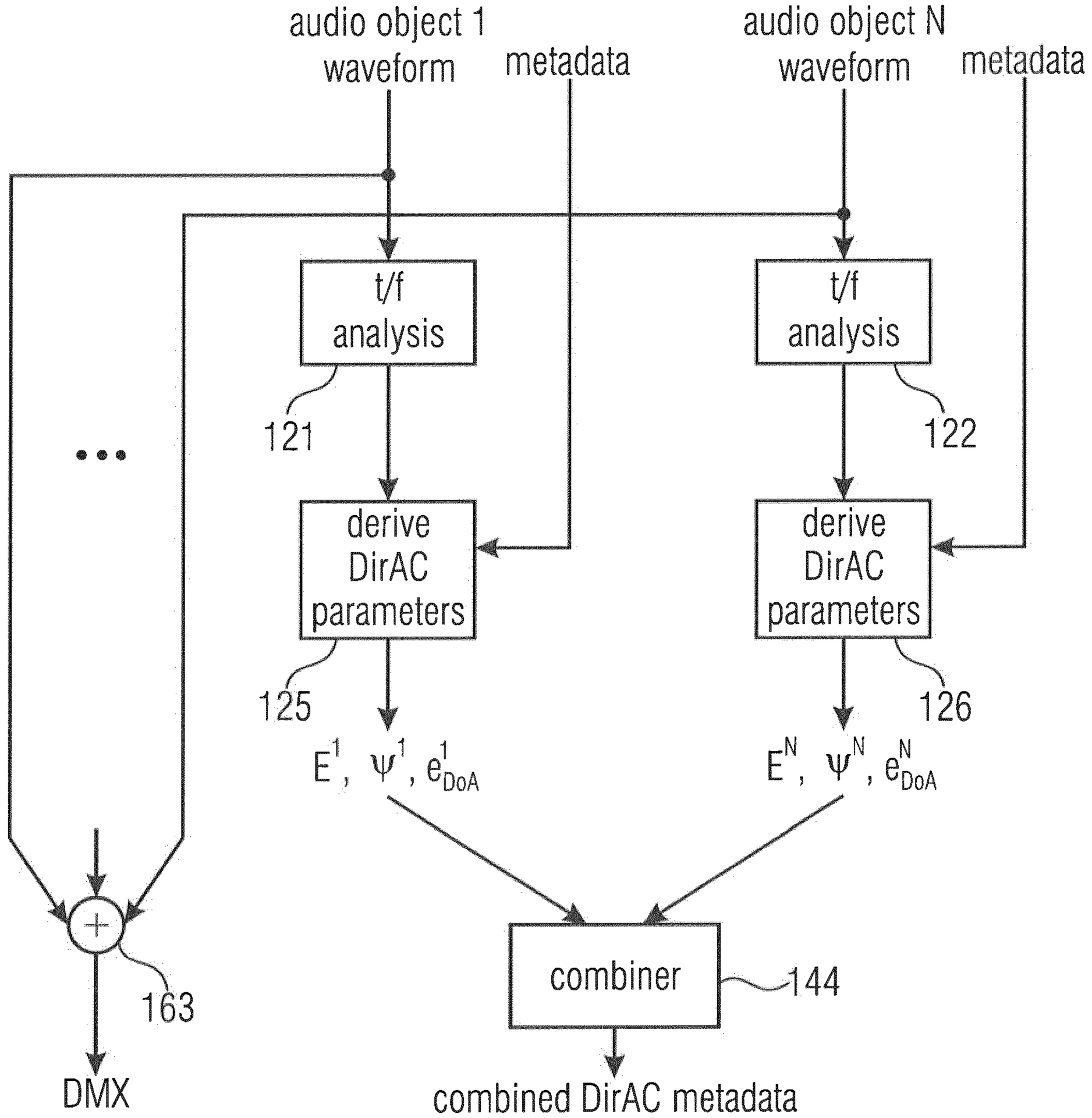
- block diagram for combining different audio scenes in different formats in the pressure/velocity domain

Fig. 1b



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

Fig. 1c



combiner could be:

alt. ①

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

alt. ②

$$\begin{cases} \psi = 0 \quad \text{(since audio objects have no diffuseness usually)} \\ e_{DoA} = e_{DoA}^{\operatorname{argmax}(E^i)} \end{cases}$$

Fig. 1d

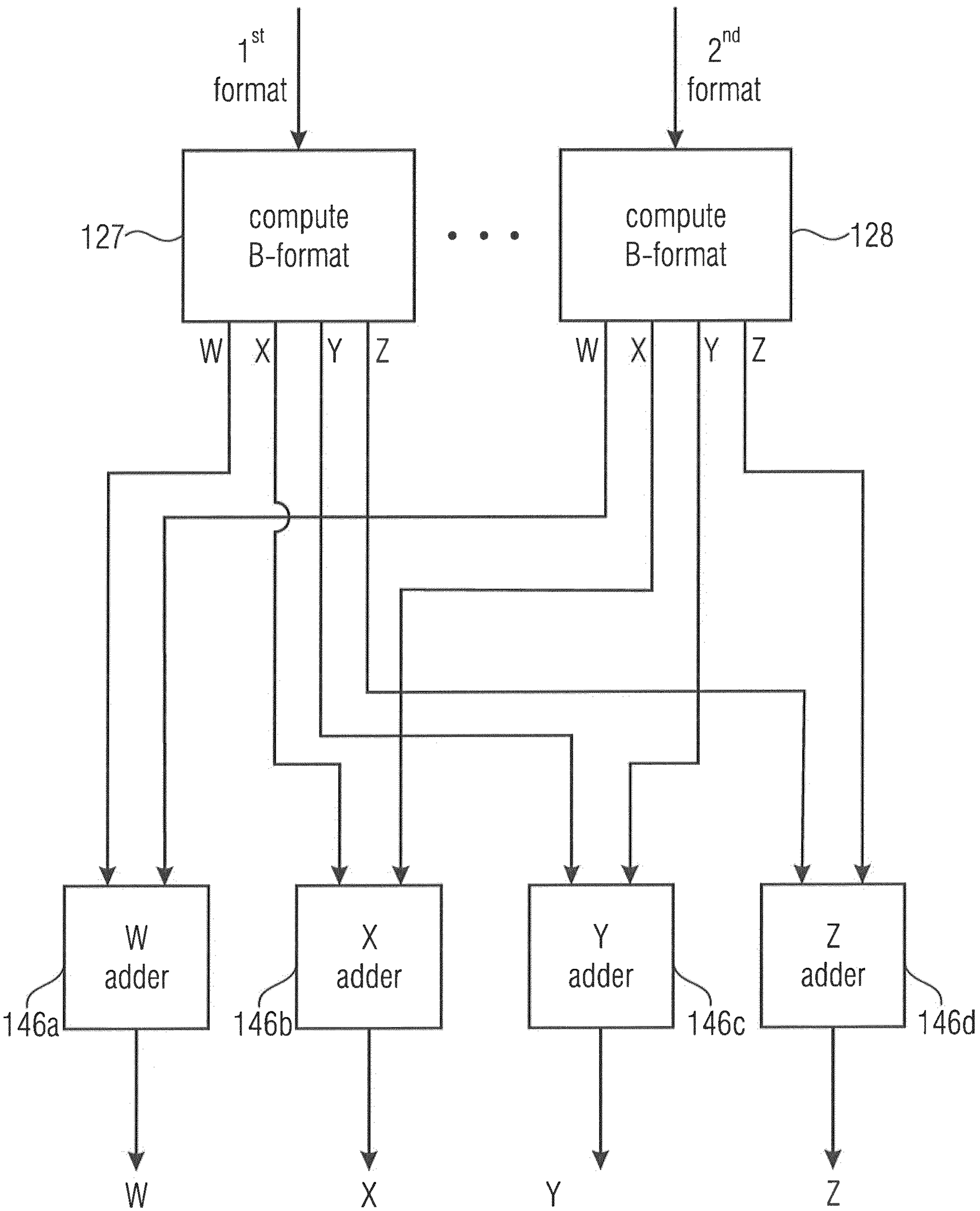
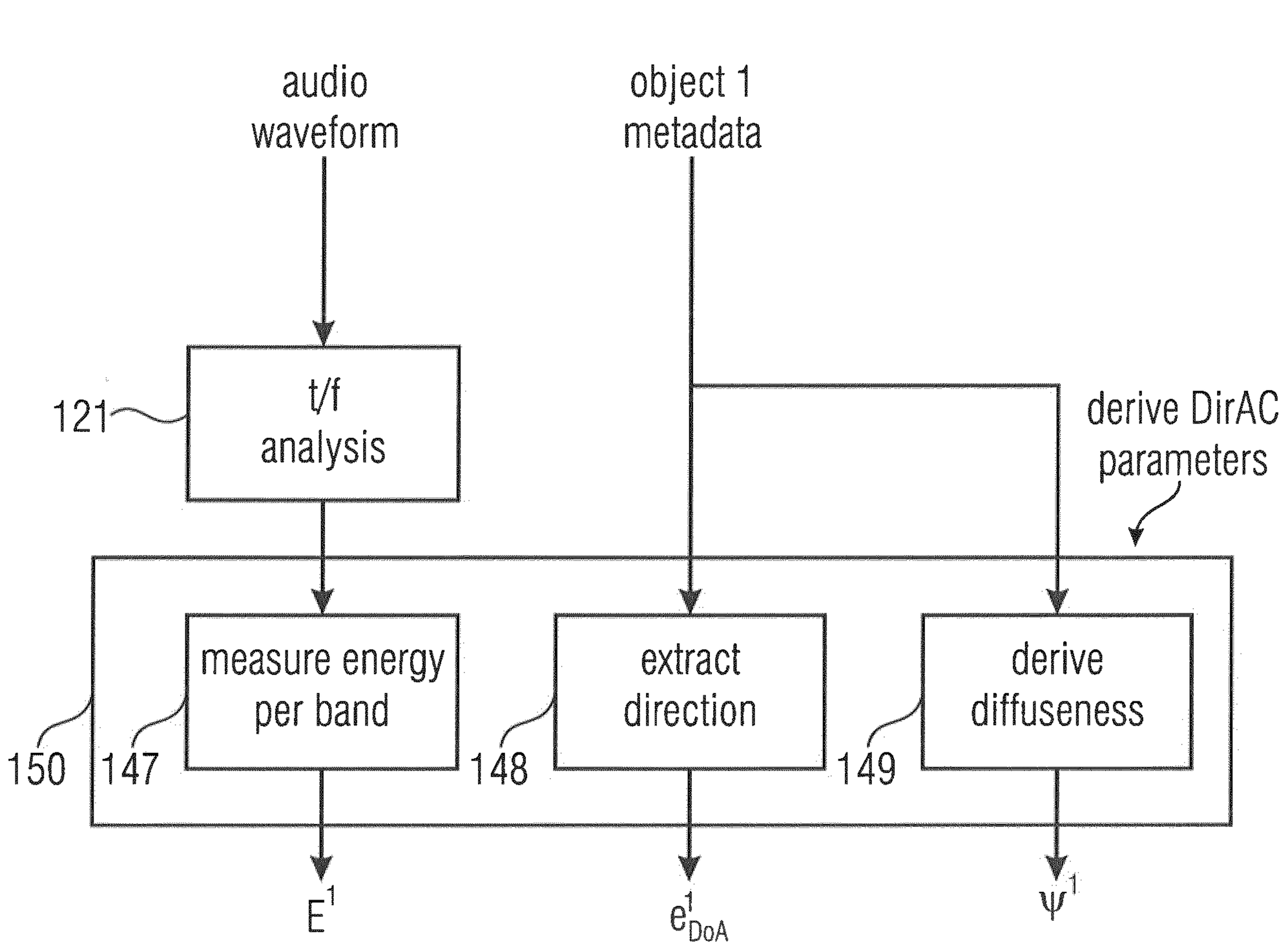

1st format
2nd format
compute B-format
127
compute B-format
128
W
X
Y
Z
W
X
Y
Z
W adder
X adder
Y adder
Z adder
146a
146b
146c
146d
W
X
Y
Z

Fig. 1e



- block diagram for deriving DirAC parameters from an audio object

Fig. 1f

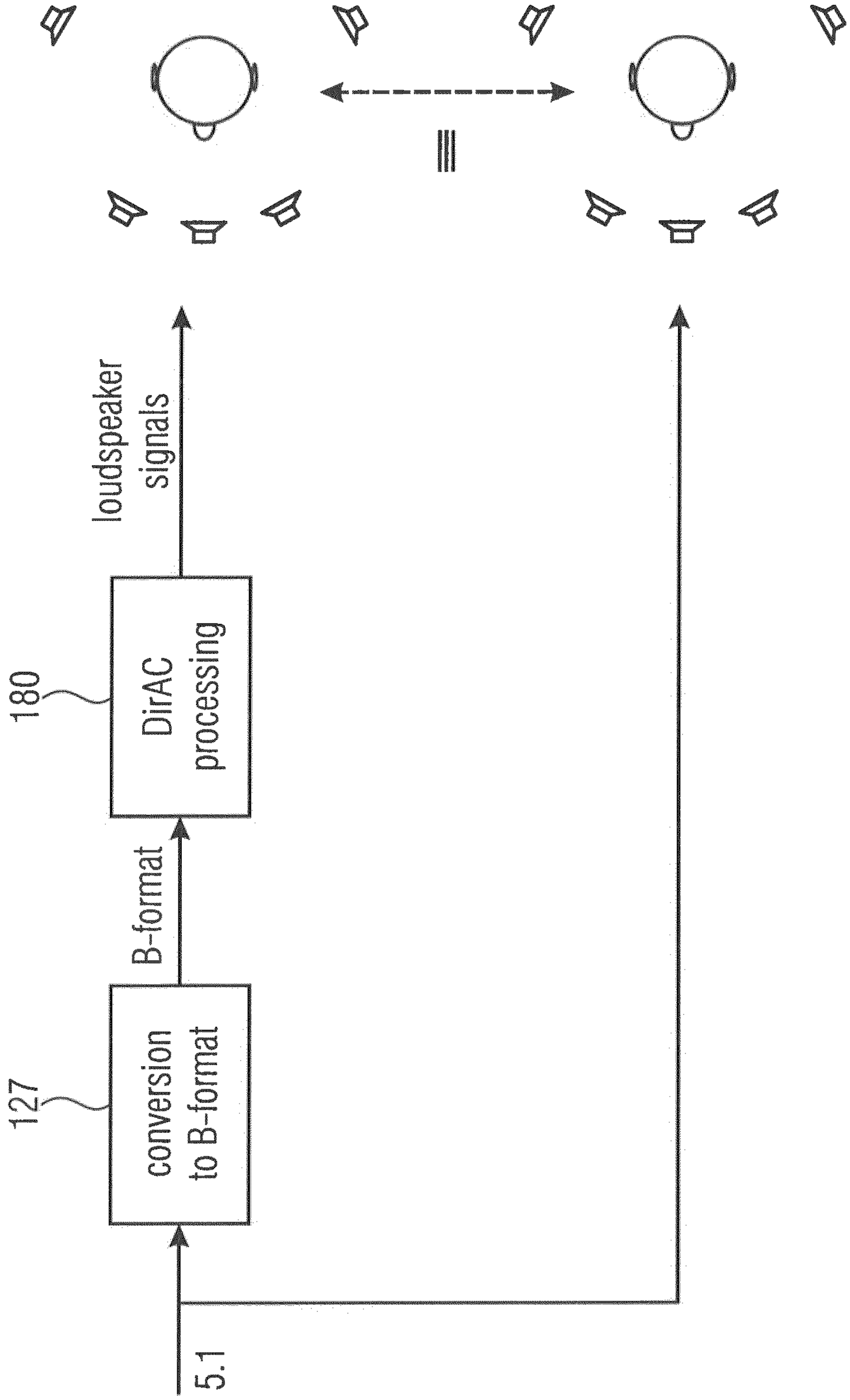

loudspeaker signals
180
DirAC processing
B-format
127
conversion to B-format
5.1

Fig. 1g

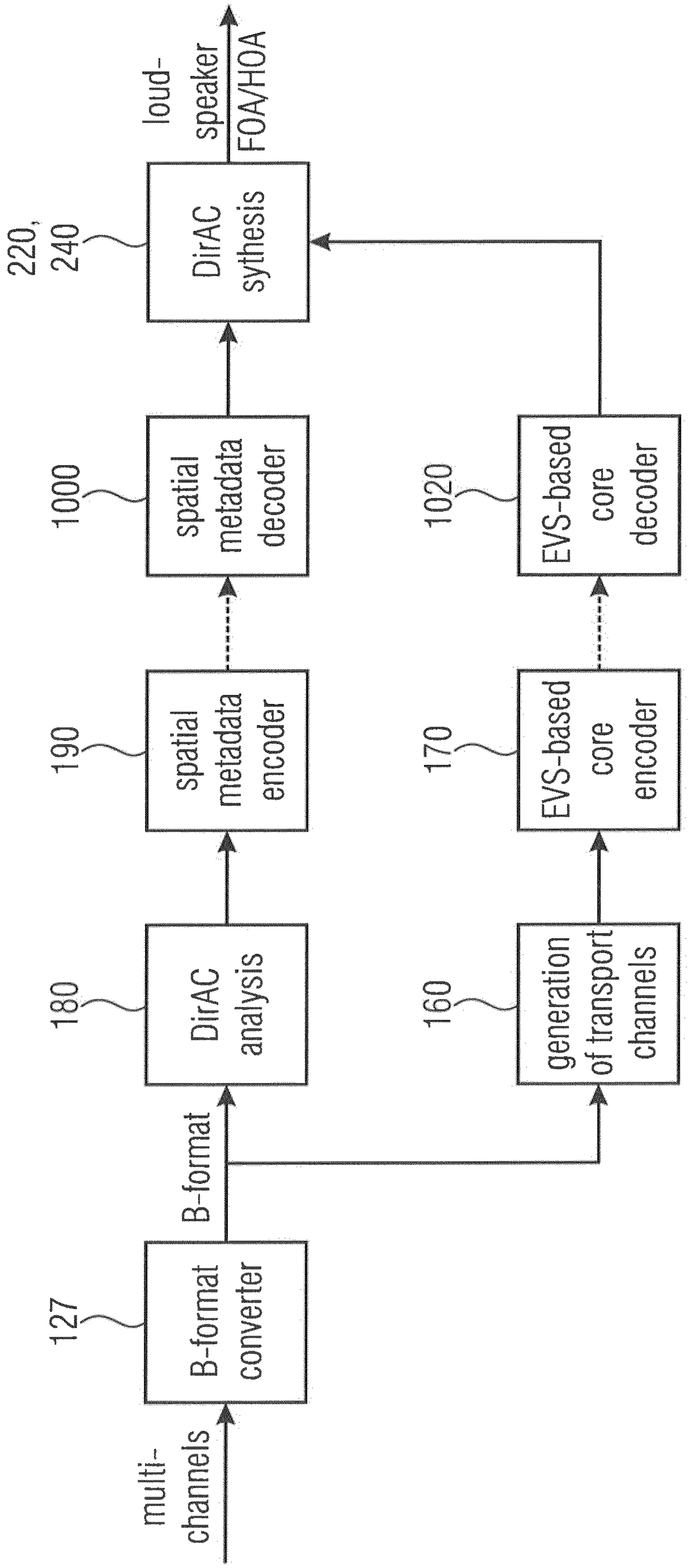

multi-channels
B-format
B-format converter
127
DirAC analysis
180
spatial metadata encoder
190
generation of transport channels
160
EVS-based core encoder
170
loud-speaker
FOA/HOA
DirAC sythesis
220, 240
spatial metadata decoder
1000
EVS-based core decoder
1020

Fig. 1h

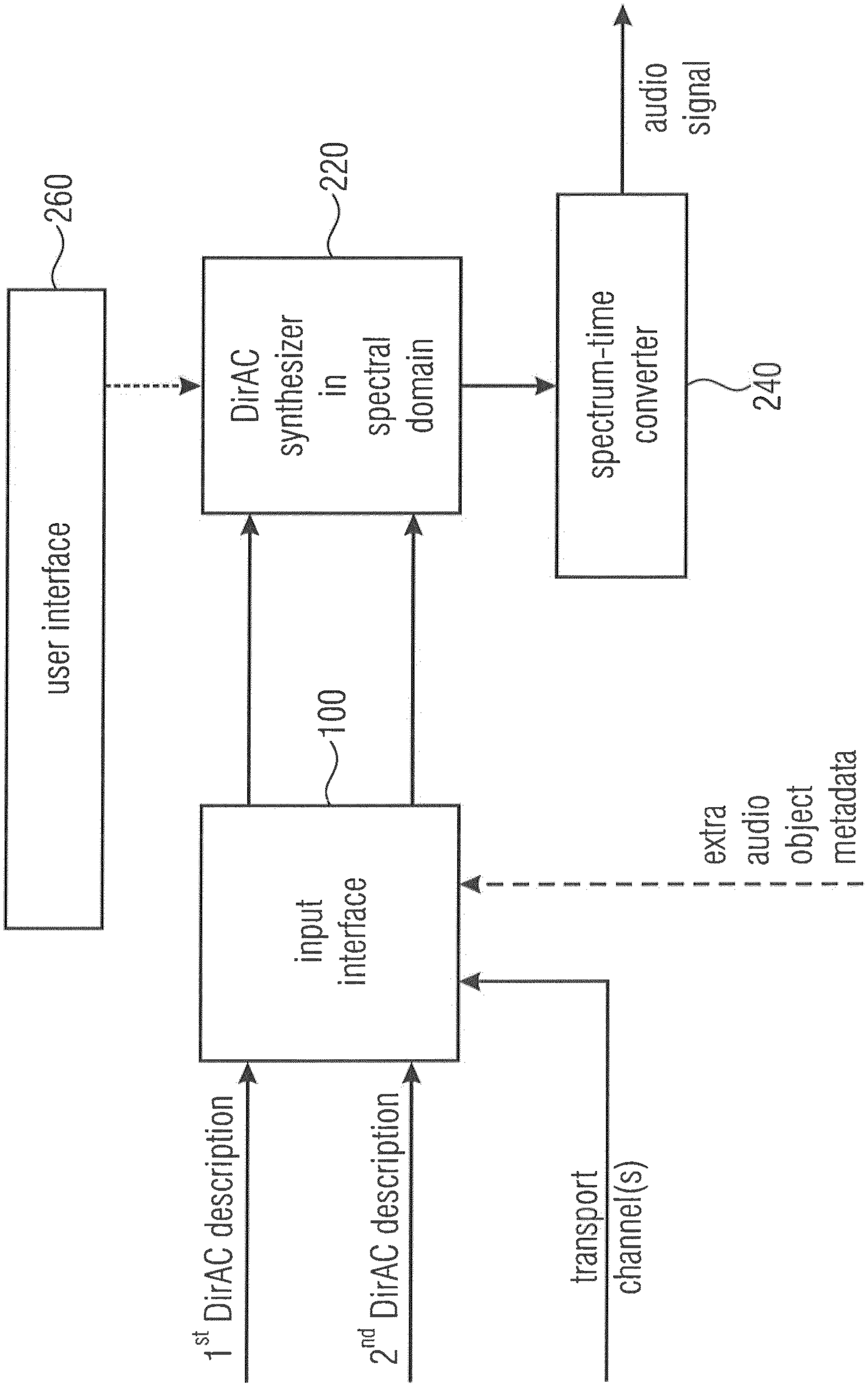
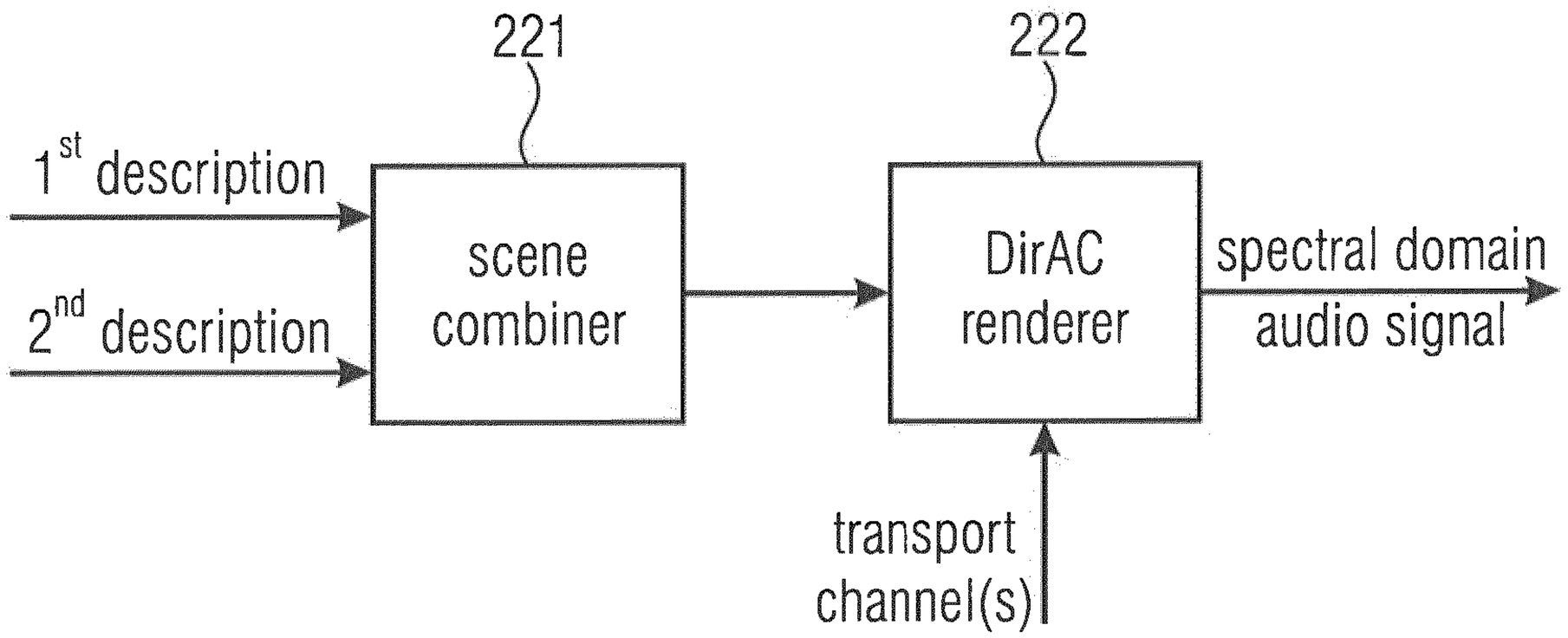
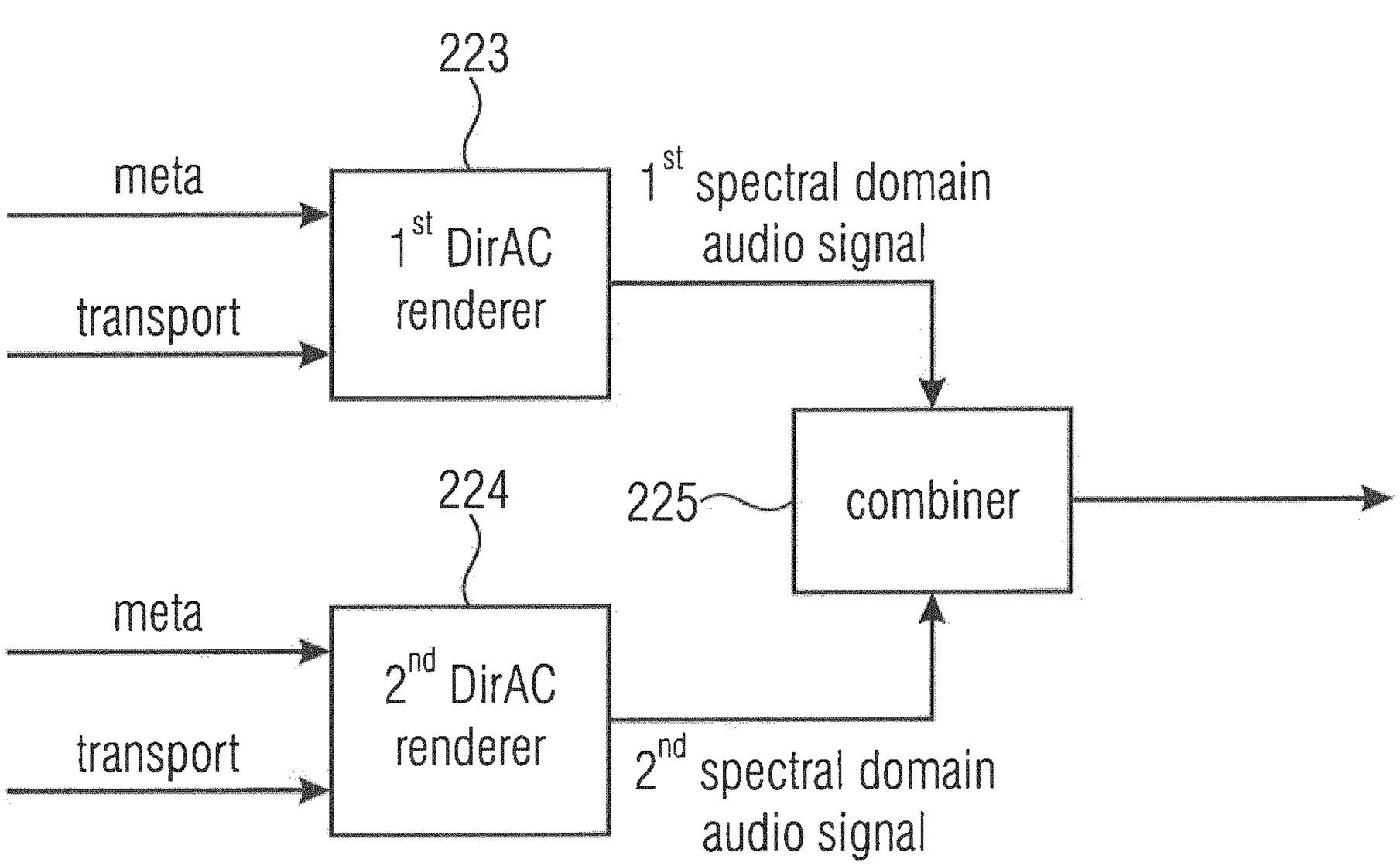

260
user interface
220
DirAC synthesizer in spectral domain
audio signal
spectrum-time converter
240
100
input interface
extra audio object metadata
1st DirAC description
2nd DirAC description
transport channel(s)

Fig. 2a



DirAC synthesizer

Fig. 2b



DirAC synthesizer

Fig. 2c

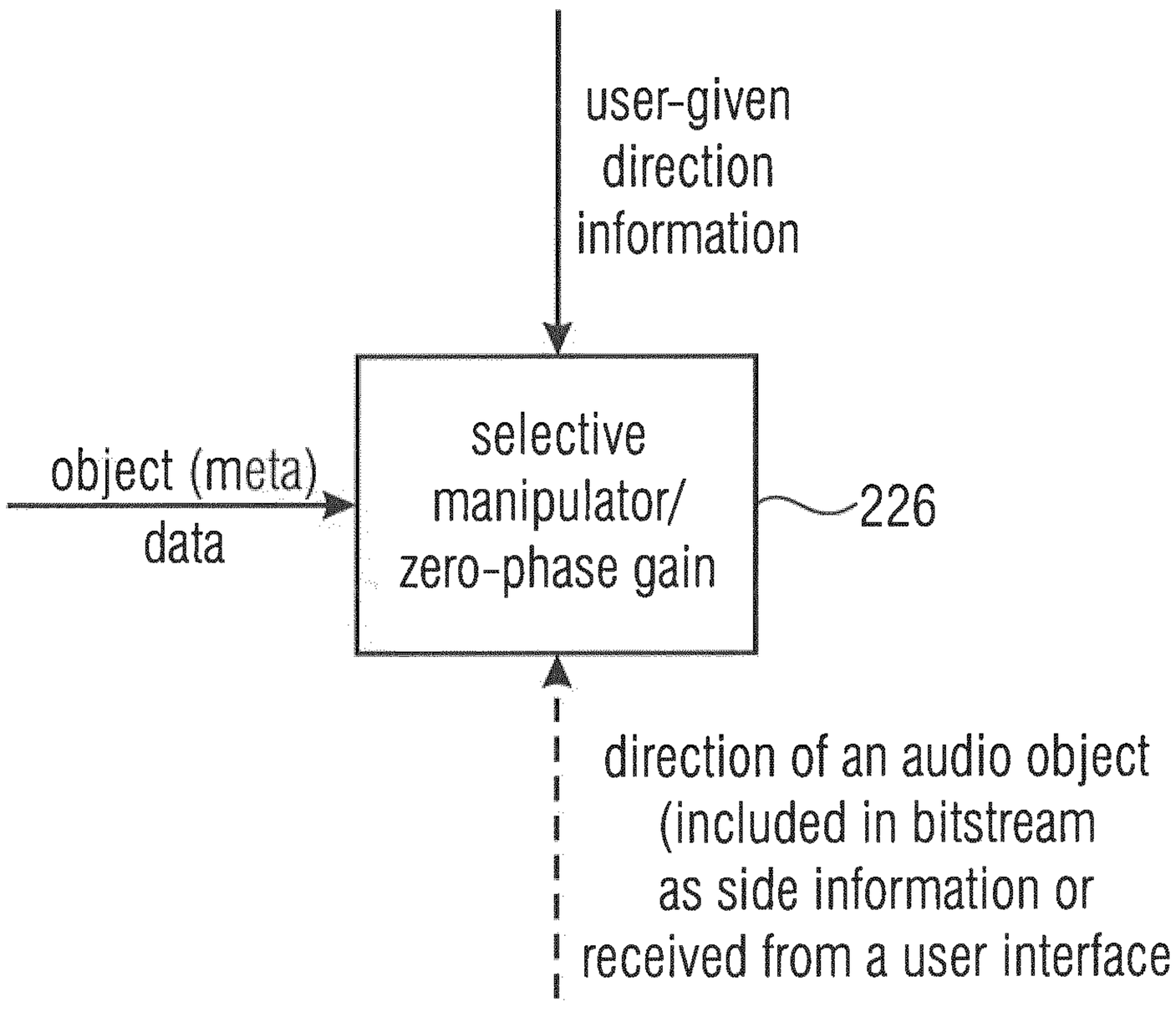

user-given
direction
information
object (meta)
data
selective
manipulator/
zero-phase gain
226
direction of an audio object
(included in bitstream
as side information or
received from a user interface

Fig. 2d

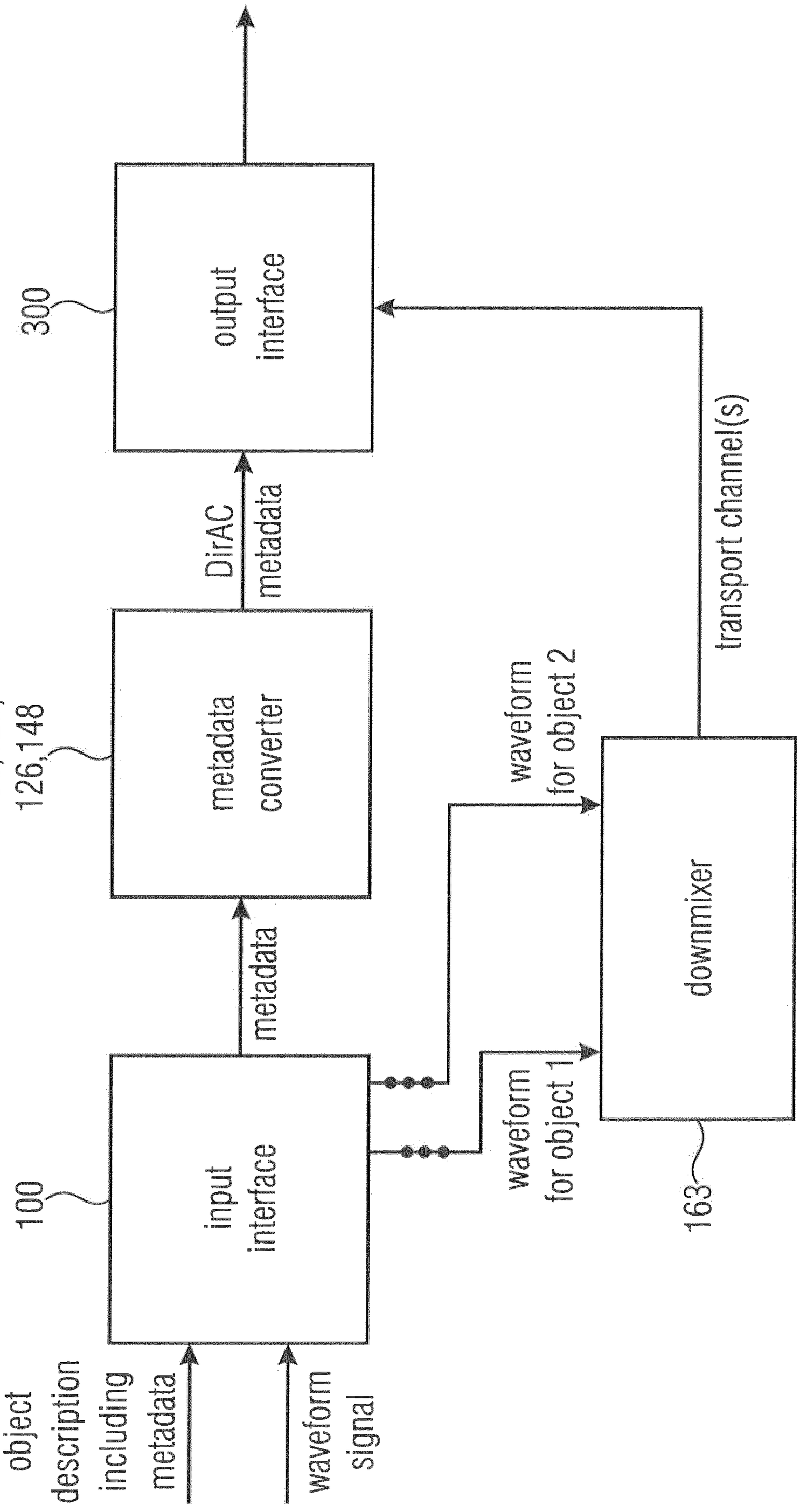

output interface
300
DirAC metadata
metadata converter
150,125, 126,148
input interface
100
waveform for object 2
waveform for object 1
downmixer
163
transport channel(s)
object description including metadata
waveform signal

Fig. 3a

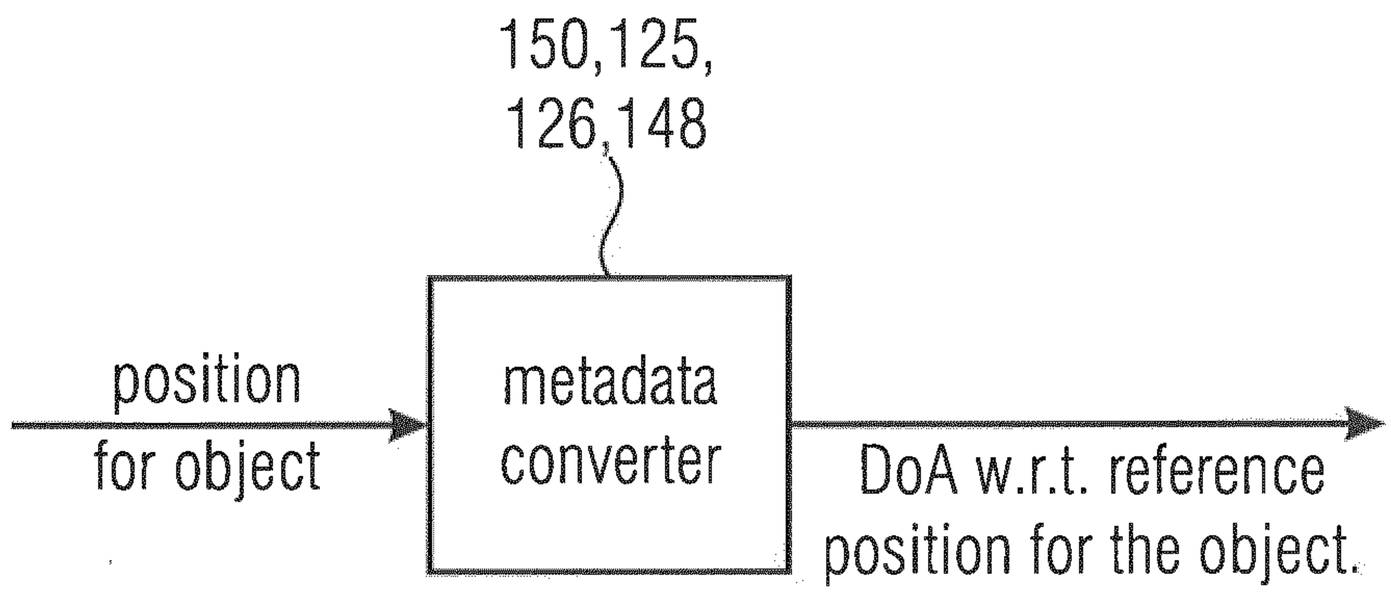

150,125,
126,148
position
for object
metadata
converter
DoA w.r.t. reference
position for the object.

Fig. 3b

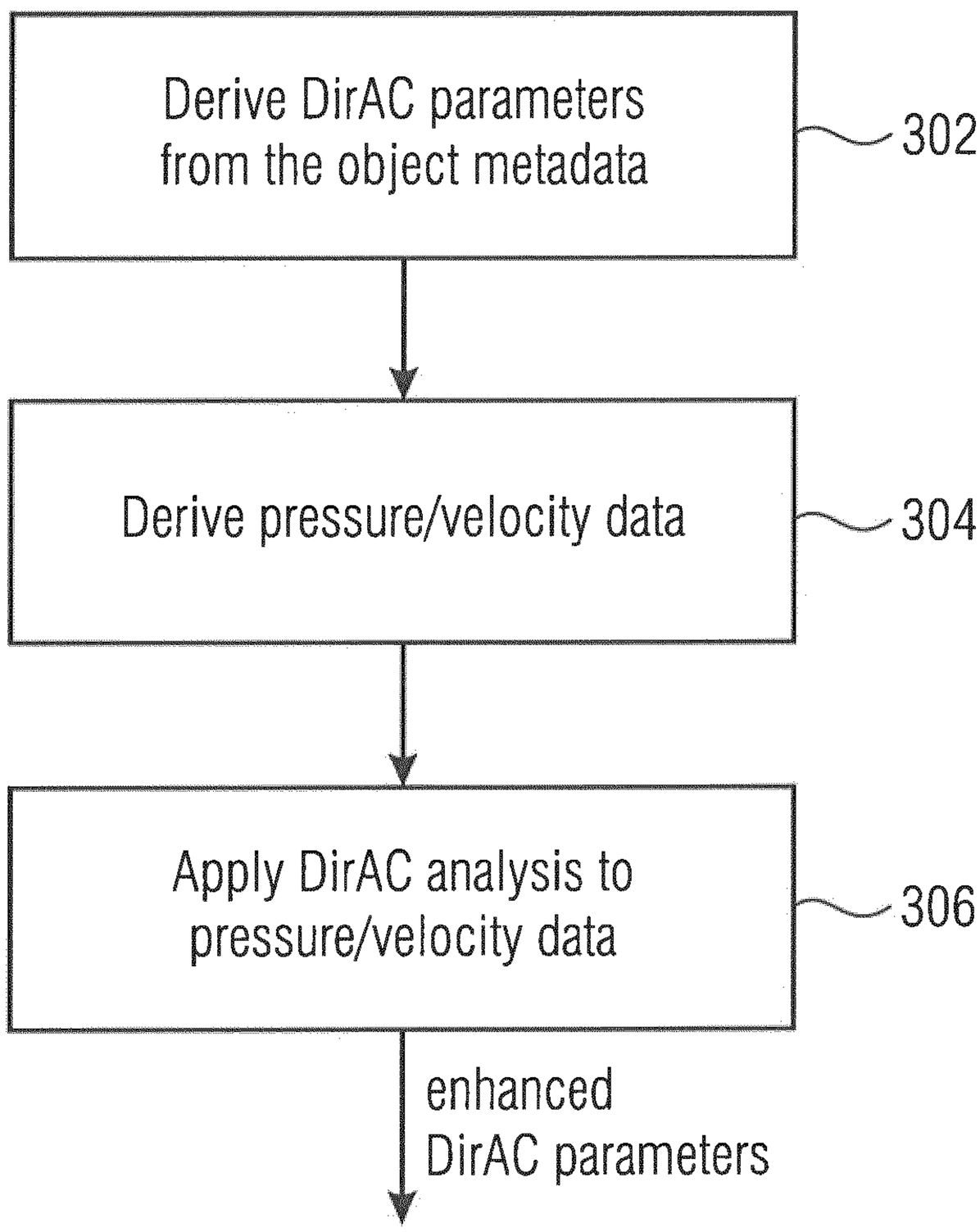

Derive DirAC parameters
from the object metadata
302
Derive pressure/velocity data
304
Apply DirAC analysis to
pressure/velocity data
306
enhanced
DirAC parameters

Fig. 3c

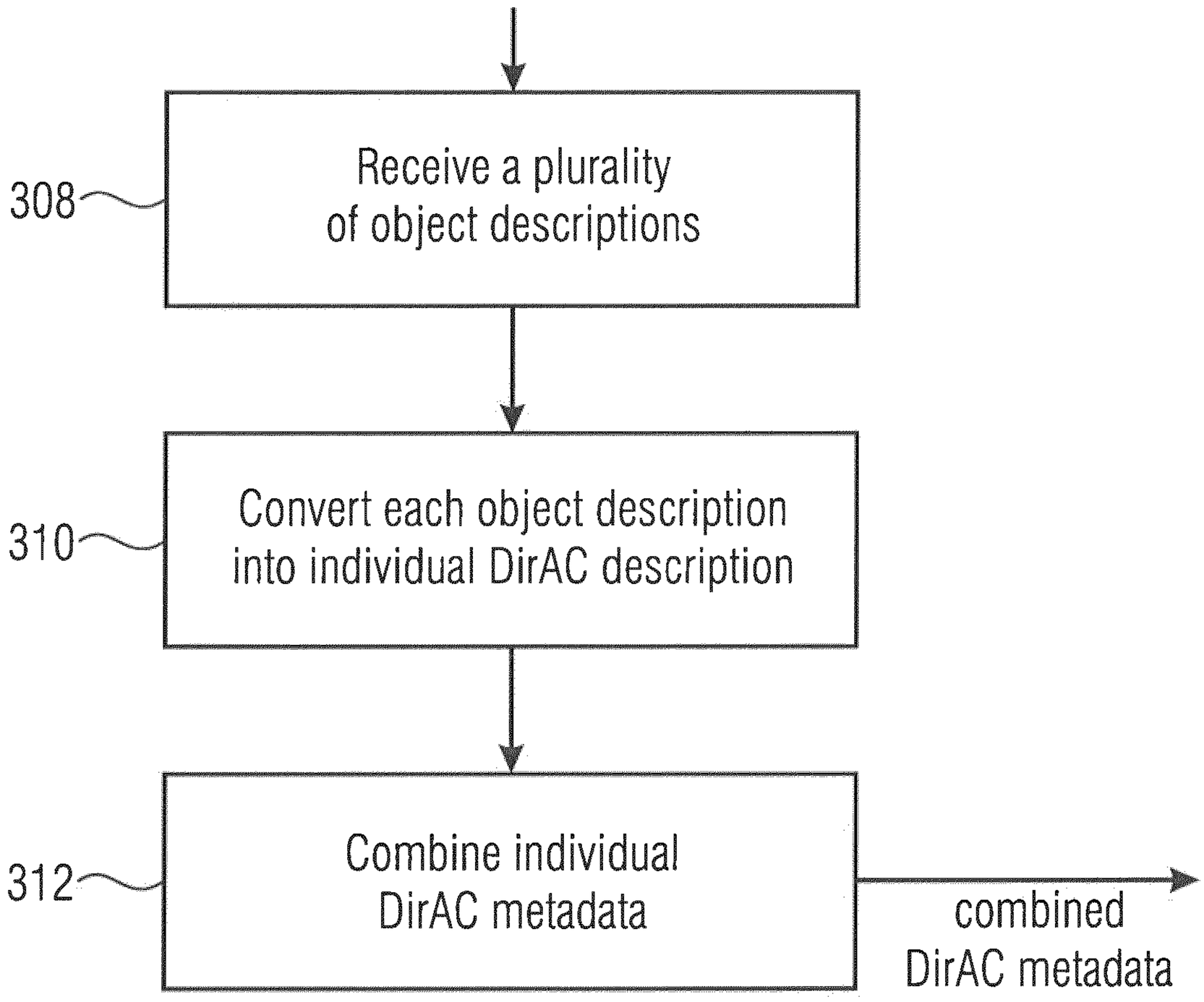

Receive a plurality of object descriptions
308
Convert each object description into individual DirAC description
310
Combine individual DirAC metadata
312
combined DirAC metadata

Fig. 3d

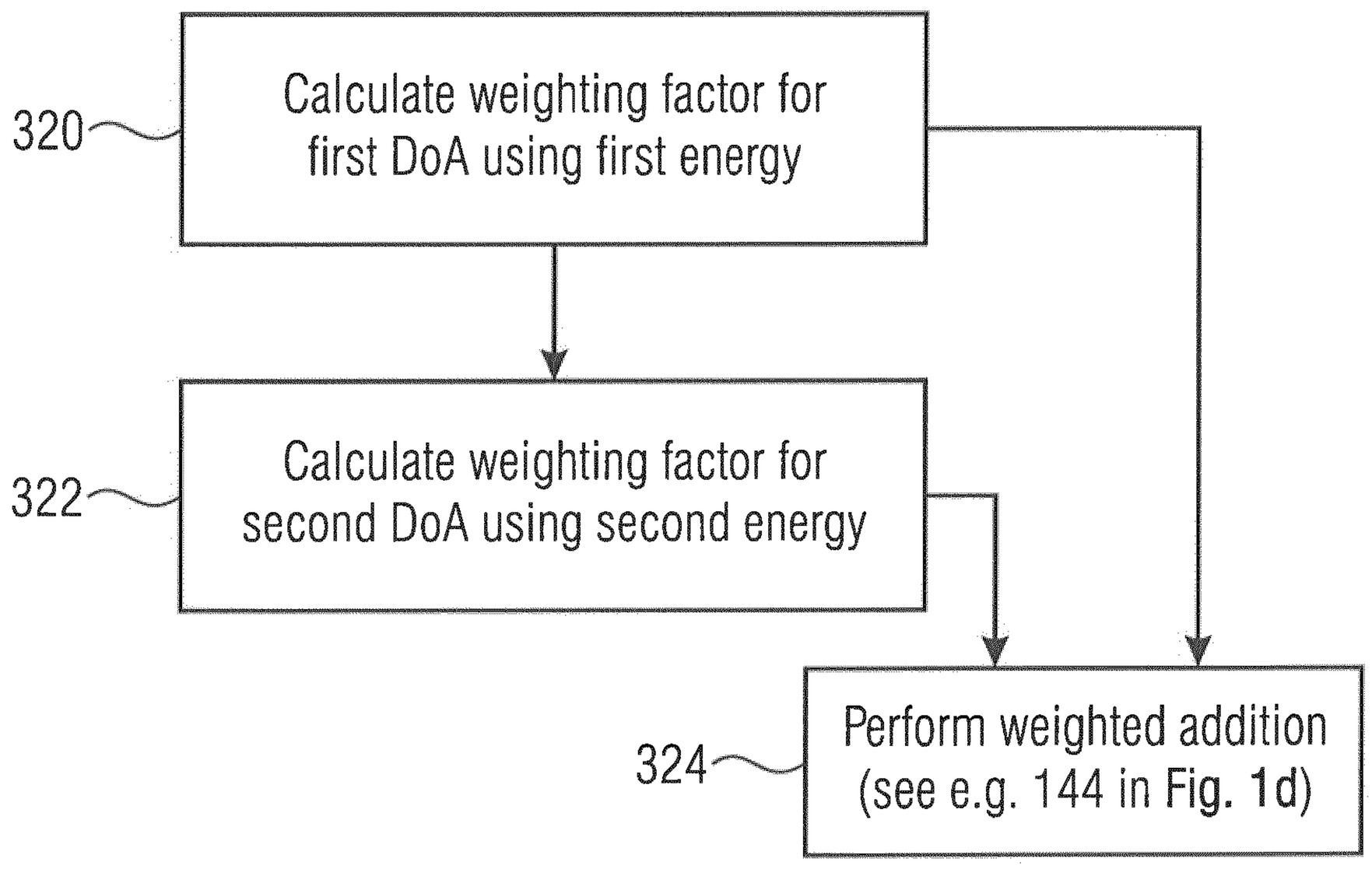
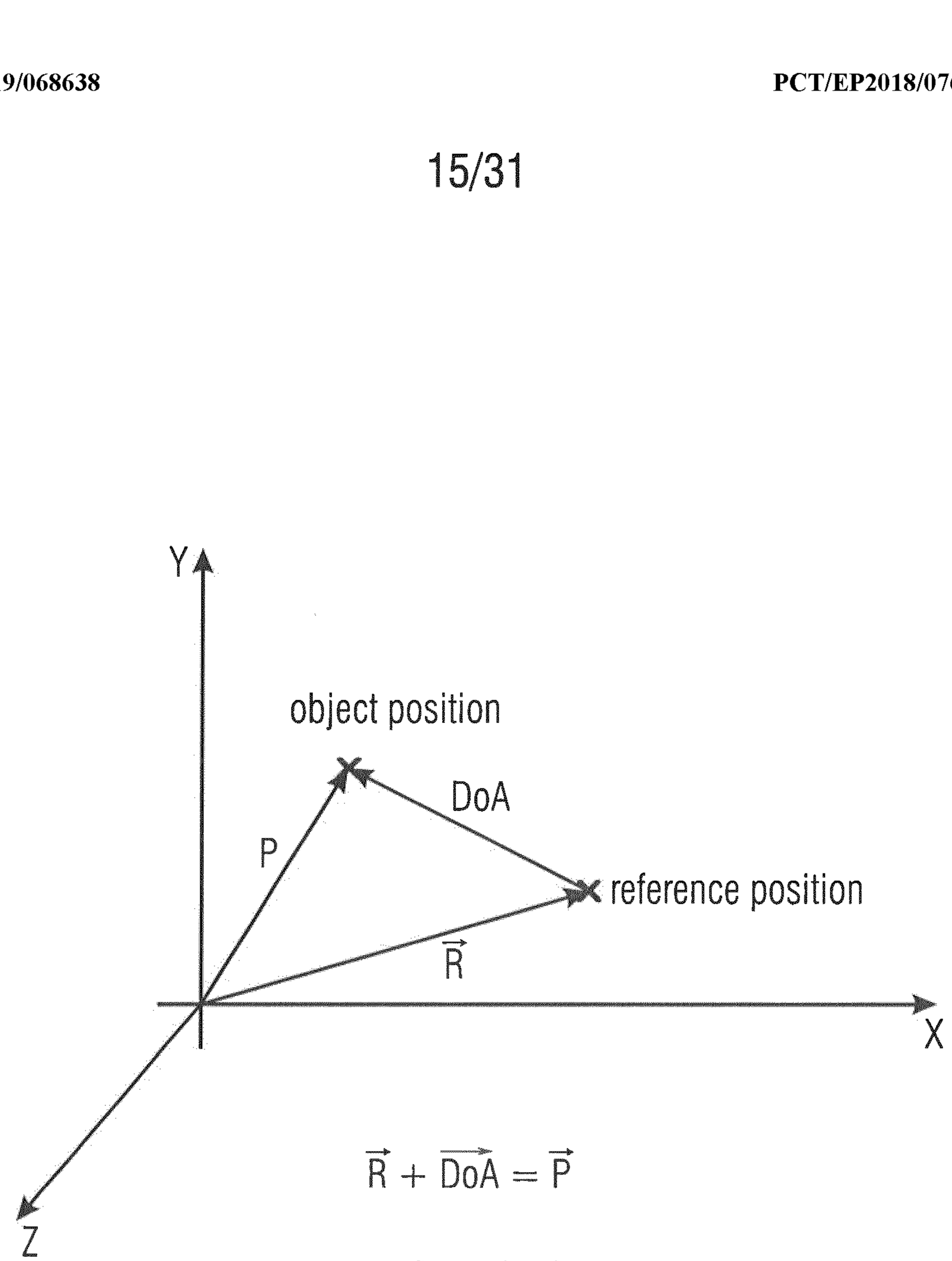

Calculate weighting factor for first DoA using first energy
320
Calculate weighting factor for second DoA using second energy
322
Perform weighted addition (see e.g. 144 in Fig. 1d)
324

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

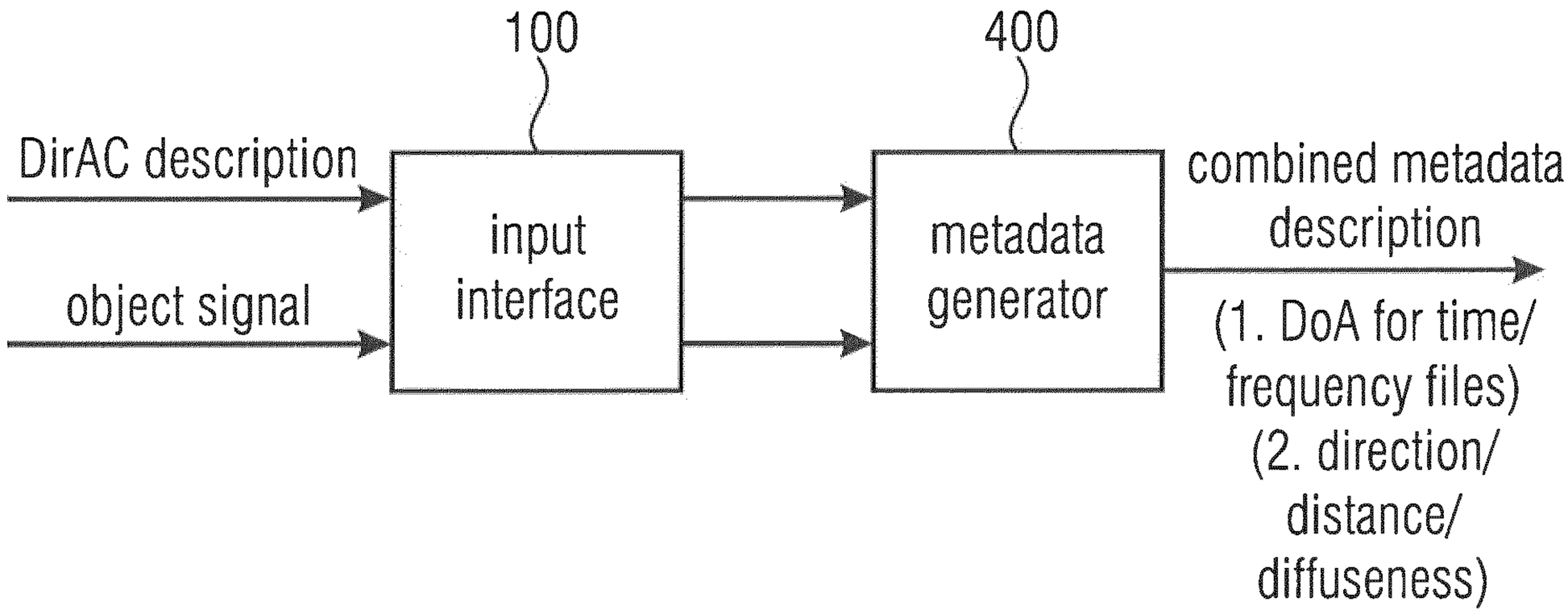

100
400
DirAC description
object signal
input interface
metadata generator
combined metadata description
(1. DoA for time/ frequency files)
(2. direction/ distance/ diffuseness)

Fig. 4a

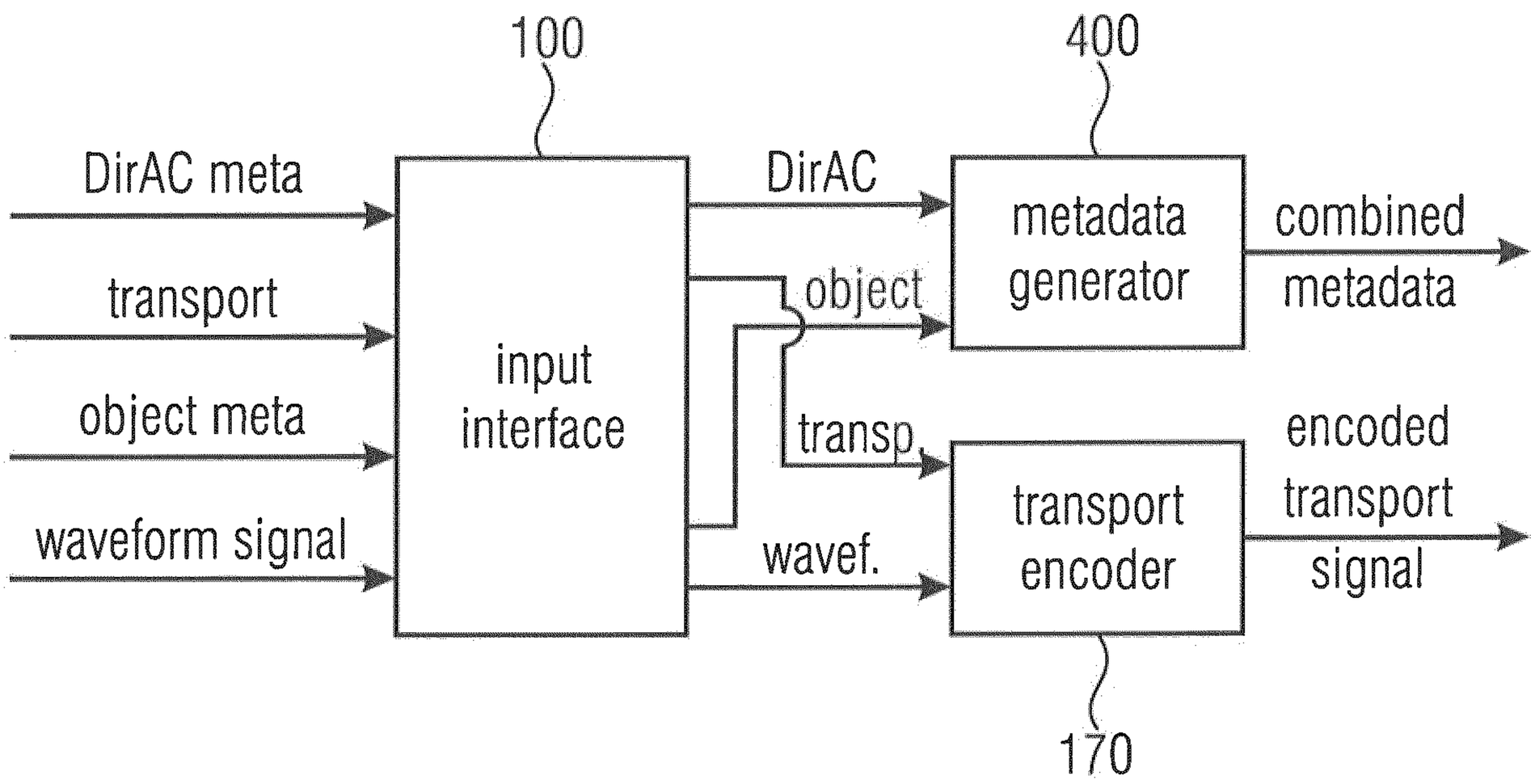

100
400
DirAC meta
transport
object meta
waveform signal
input interface
DirAC
object
transp.
wavef.
metadata generator
combined metadata
transport encoder
encoded transport signal
170

Fig. 4b

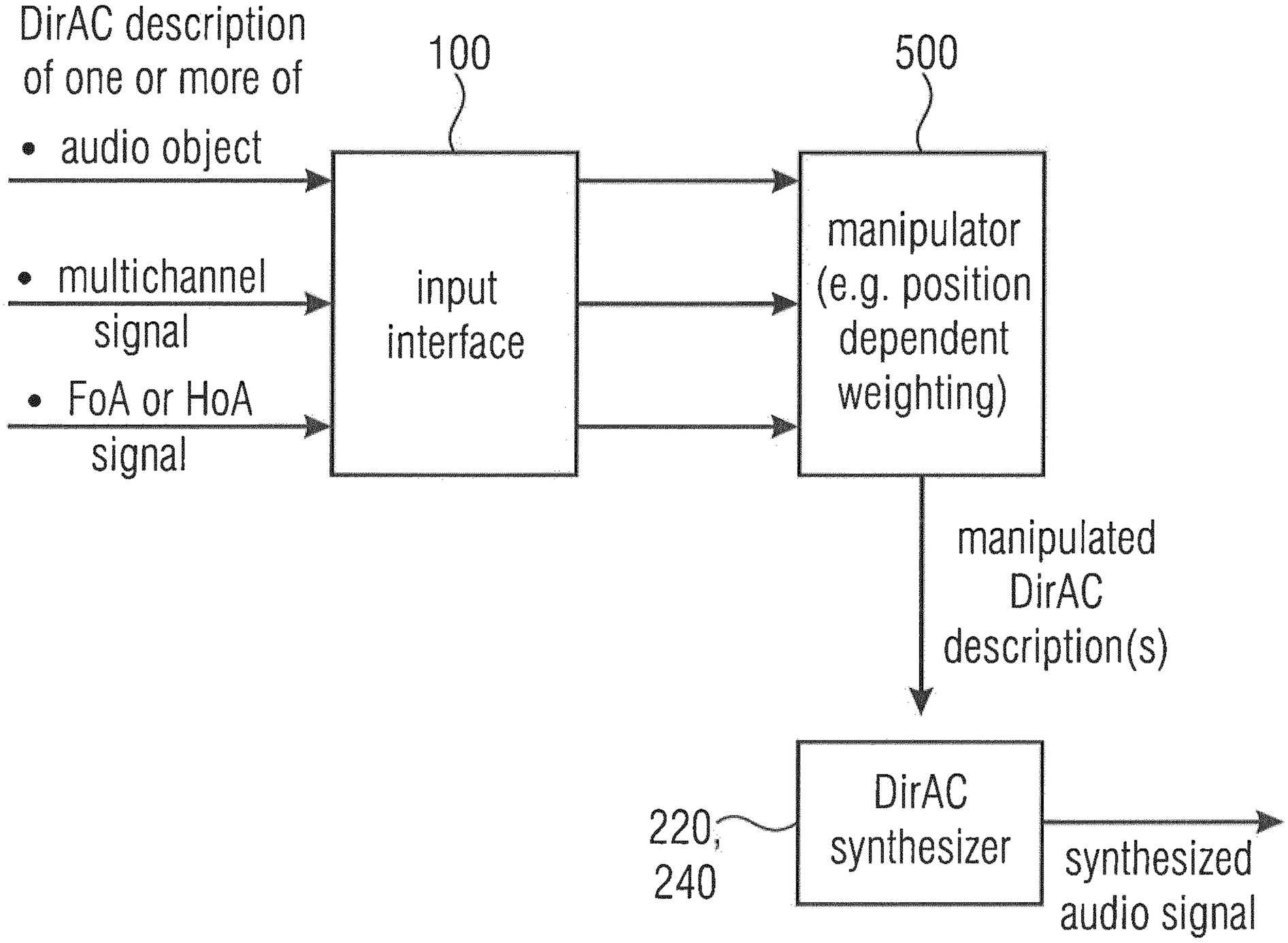

DirAC description of one or more of
• audio object
• multichannel signal
• FoA or HoA signal
100
input interface
500
manipulator (e.g. position dependent weighting)
manipulated DirAC description(s)
220; 240
DirAC synthesizer
synthesized audio signal

Fig. 5a

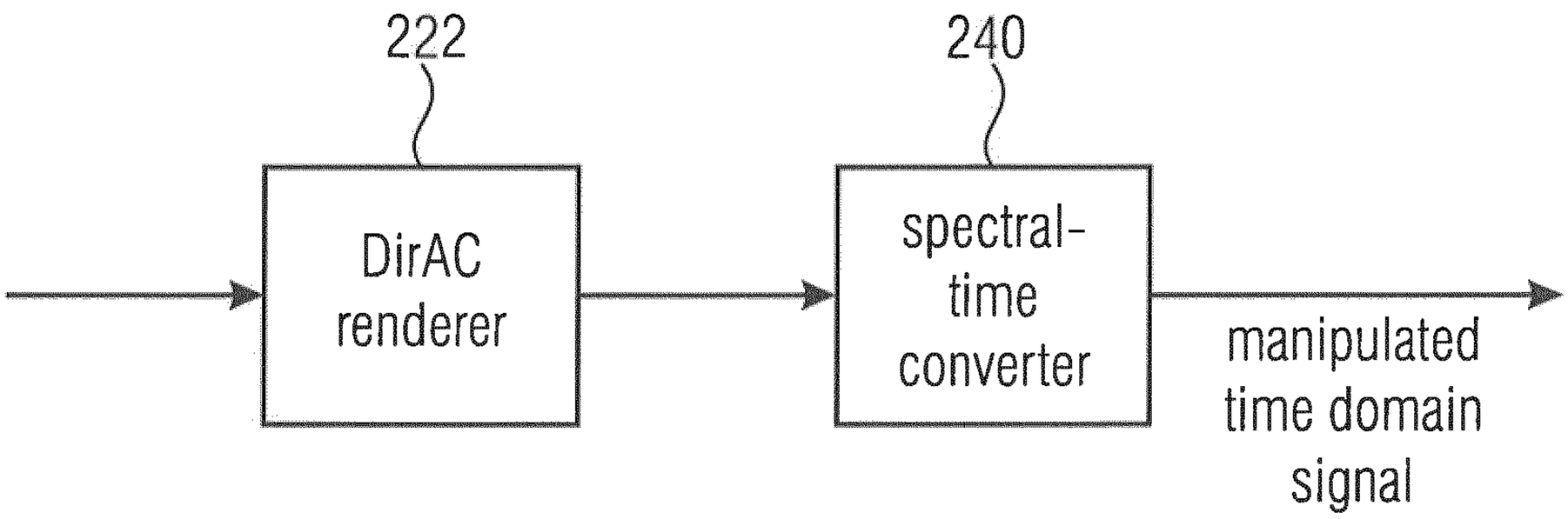

222
DirAC renderer
240
spectral-time converter
manipulated time domain signal

Fig. 5b

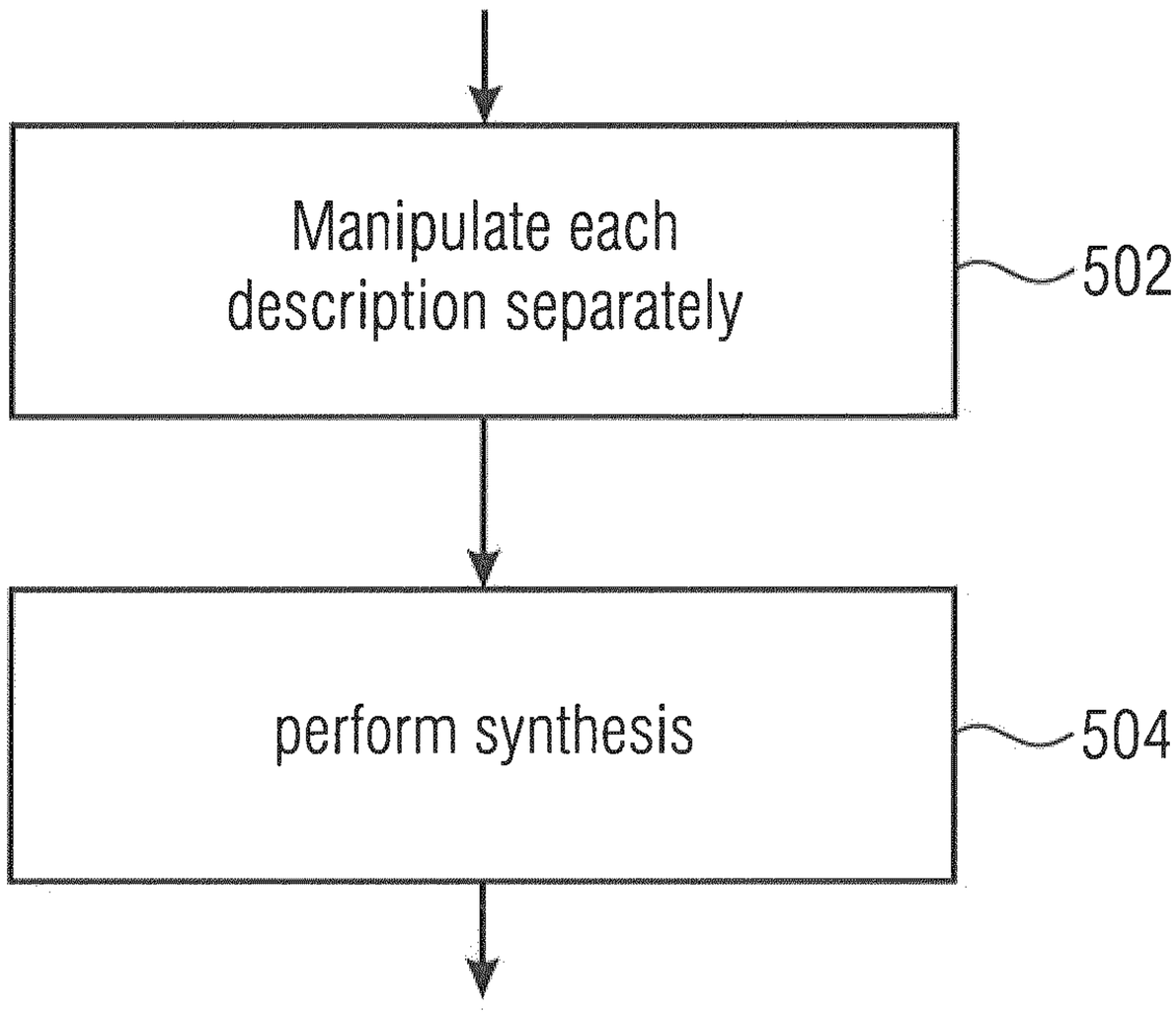

Manipulate each description separately
502
perform synthesis
504

Fig. 5c

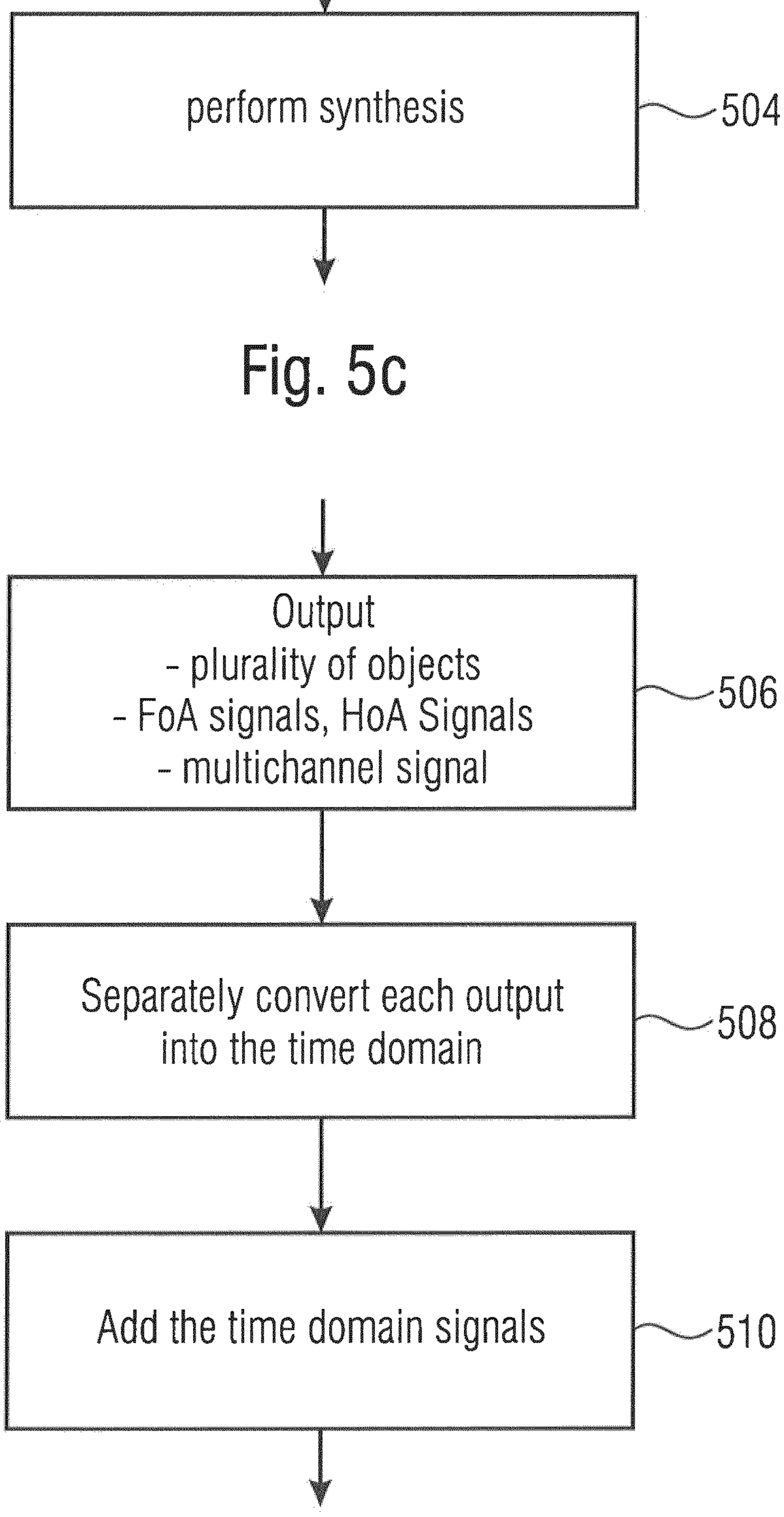
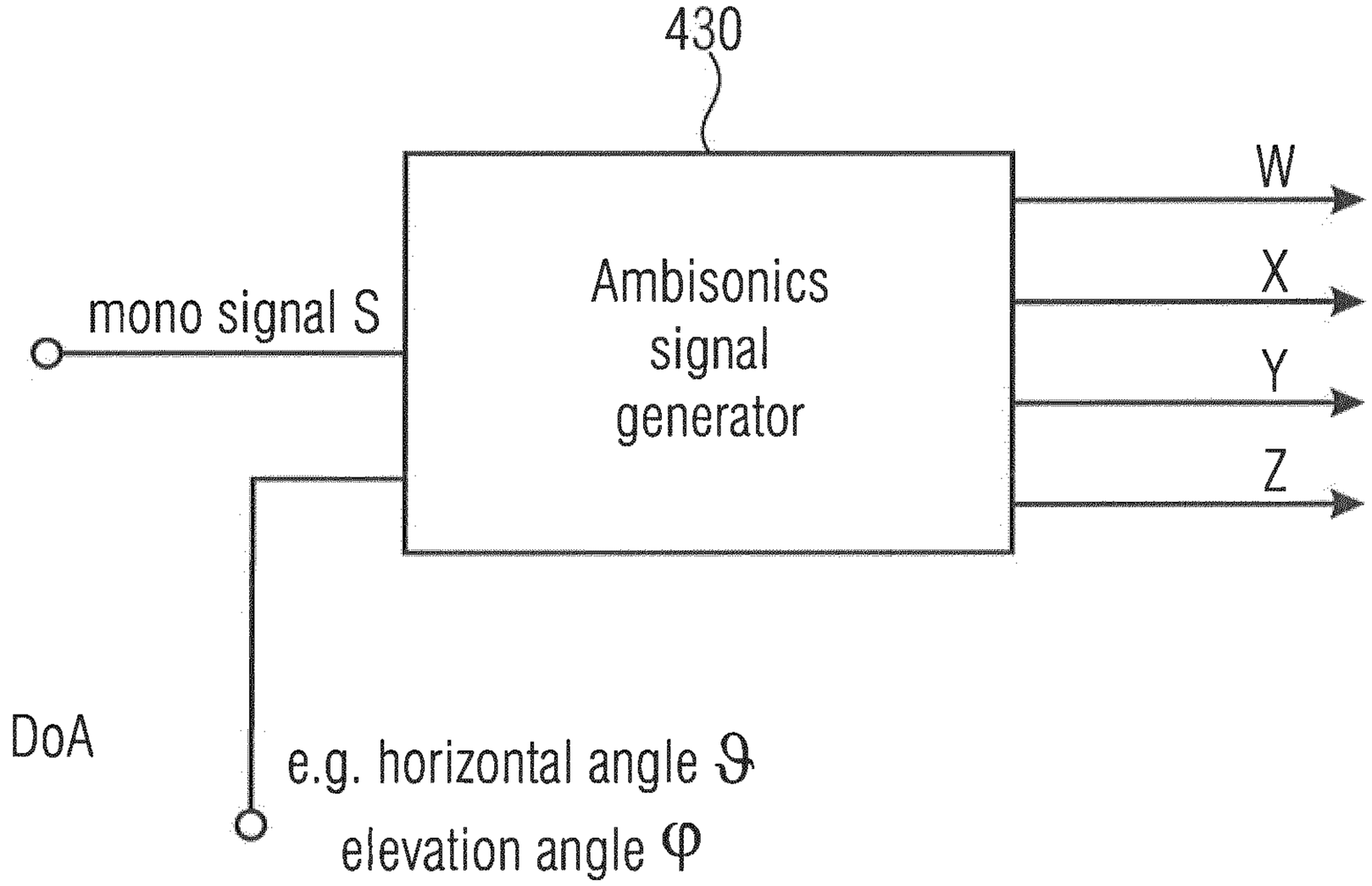

Output
- plurality of objects
- FoA signals, HoA Signals
- multichannel signal
506
Separately convert each output into the time domain
508
Add the time domain signals
510

Fig. 5d



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

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

Fig. 6

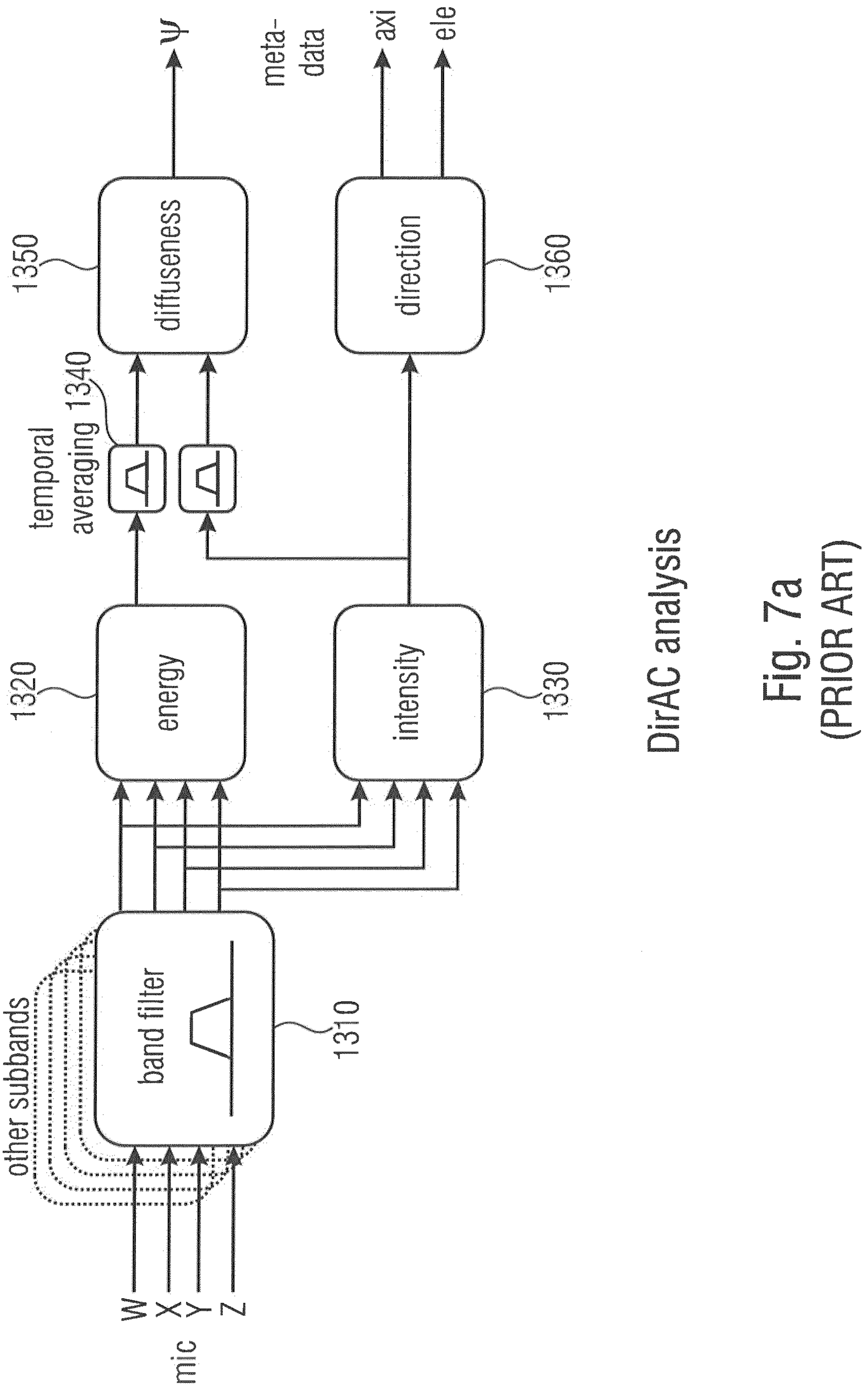
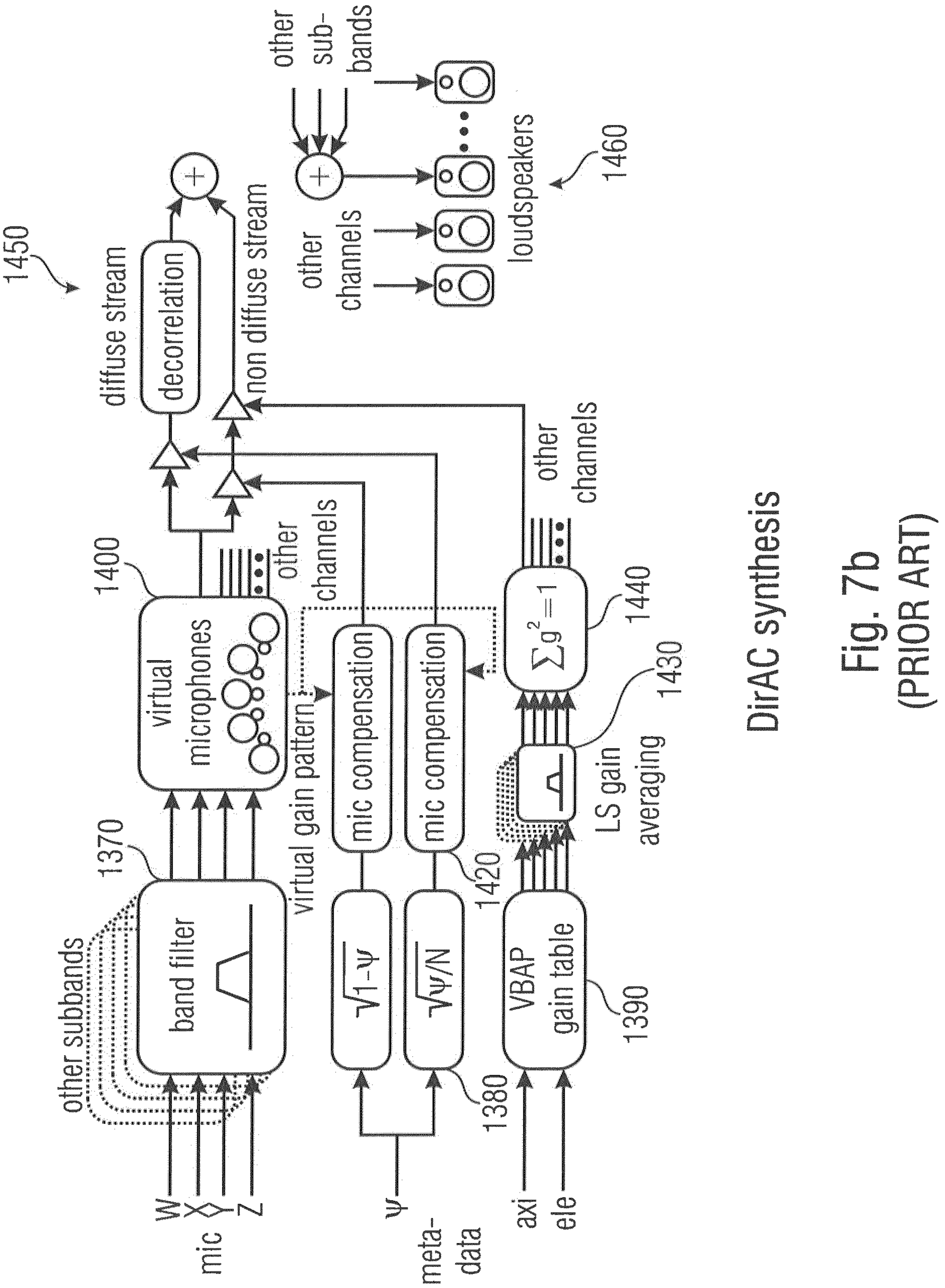


Fig. 7a
(PRIOR ART)



DirAC synthesis

Fig. 7b
(PRIOR ART)

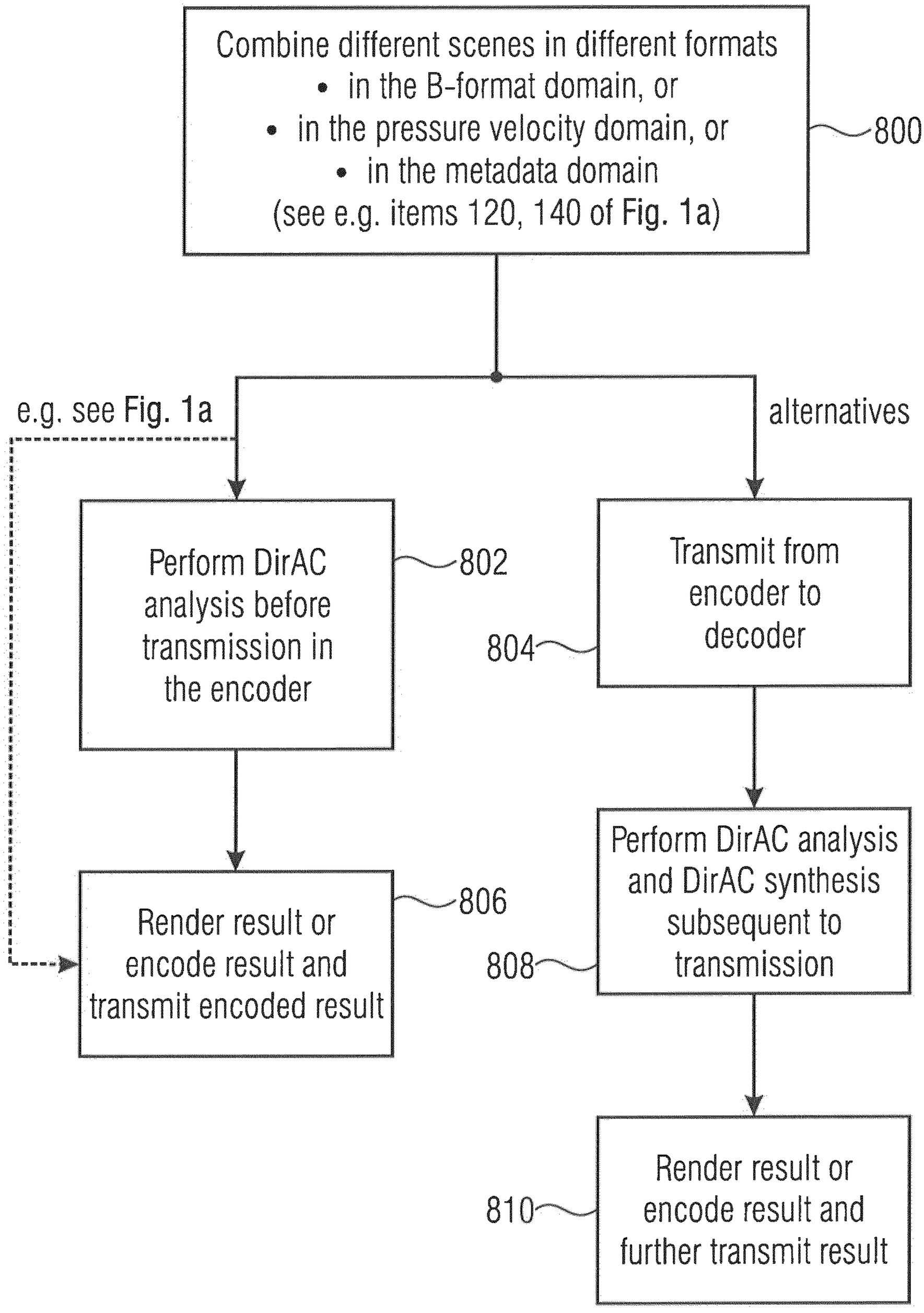

Combine different scenes in different formats
• in the B-format domain, or
• in the pressure velocity domain, or
• in the metadata domain
(see e.g. items 120, 140 of Fig. 1a)
800
e.g. see Fig. 1a
alternatives
Perform DirAC analysis before transmission in the encoder
802
Transmit from encoder to decoder
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

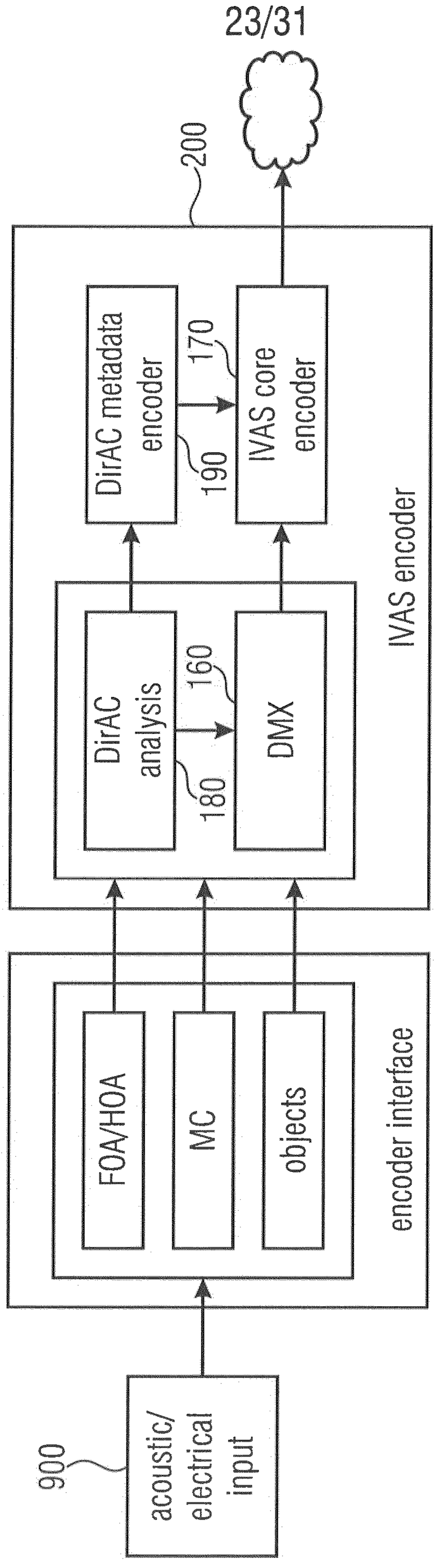

23/31
200
IVAS encoder
DirAC analysis
180
160
DMX
DirAC metadata encoder
190
170
IVAS core encoder
encoder interface
FOA/HOA
MC
objects
900
acoustic/ electrical input

Fig. 9

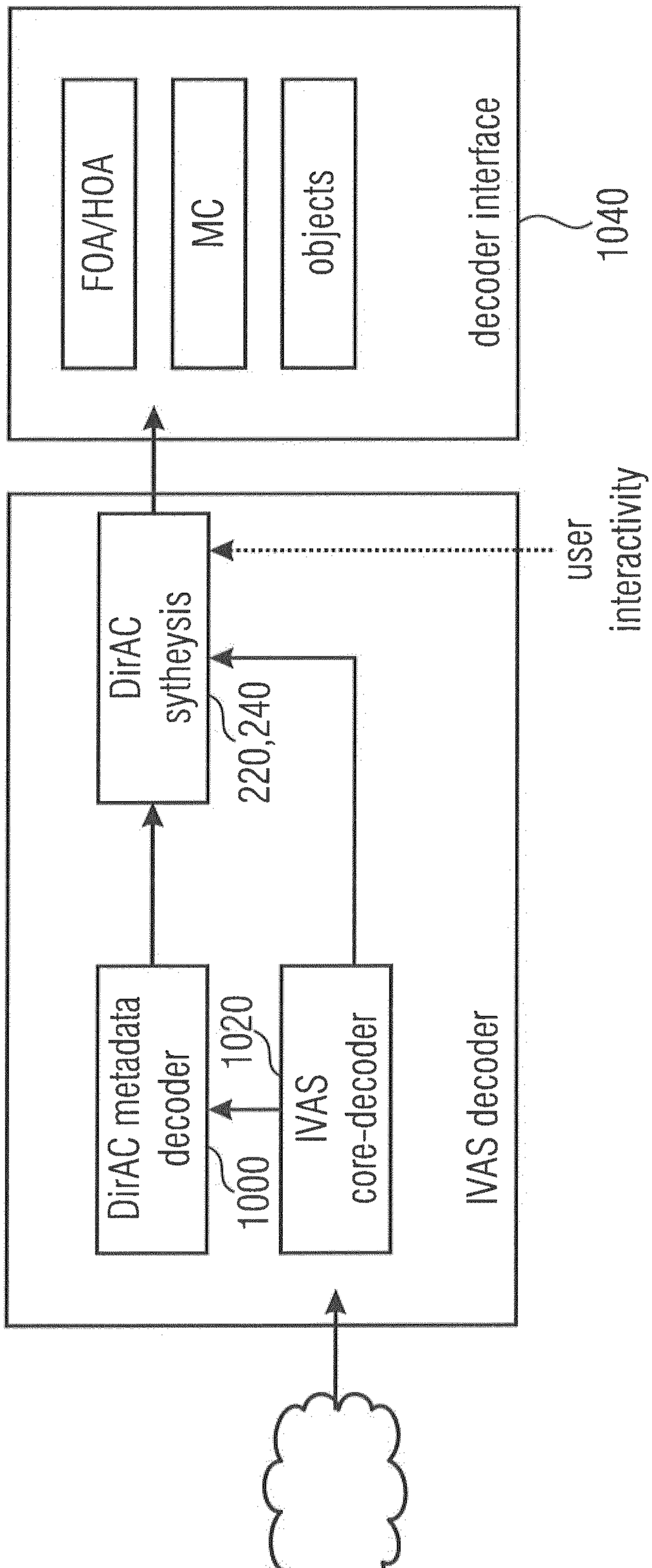

FOA/HOA
MC
objects
decoder interface
1040
user interactivity
DirAC sythesis
220,240
DirAC metadata decoder
1000
IVAS core-decoder
1020
IVAS decoder

Fig. 10

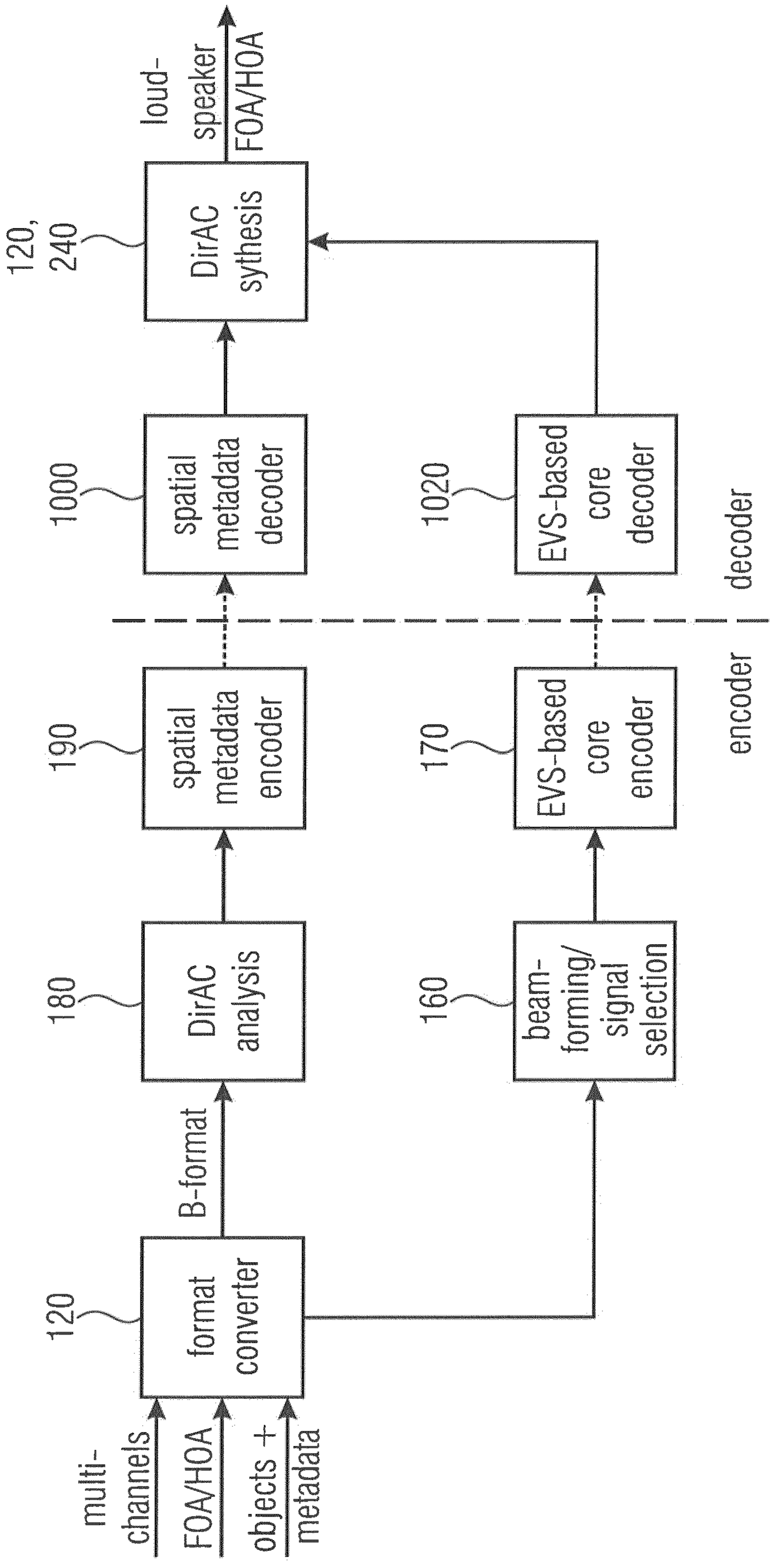

loud-speaker
FOA/HOA
120, 240
DirAC sythesis
1000
spatial metadata decoder
1020
EVS-based core decoder
decoder
190
spatial metadata encoder
170
EVS-based core encoder
encoder
180
DirAC analysis
160
beam-forming/ signal selection
B-format
120
format converter
multi-channels
FOA/HOA
objects + metadata

Fig. 11

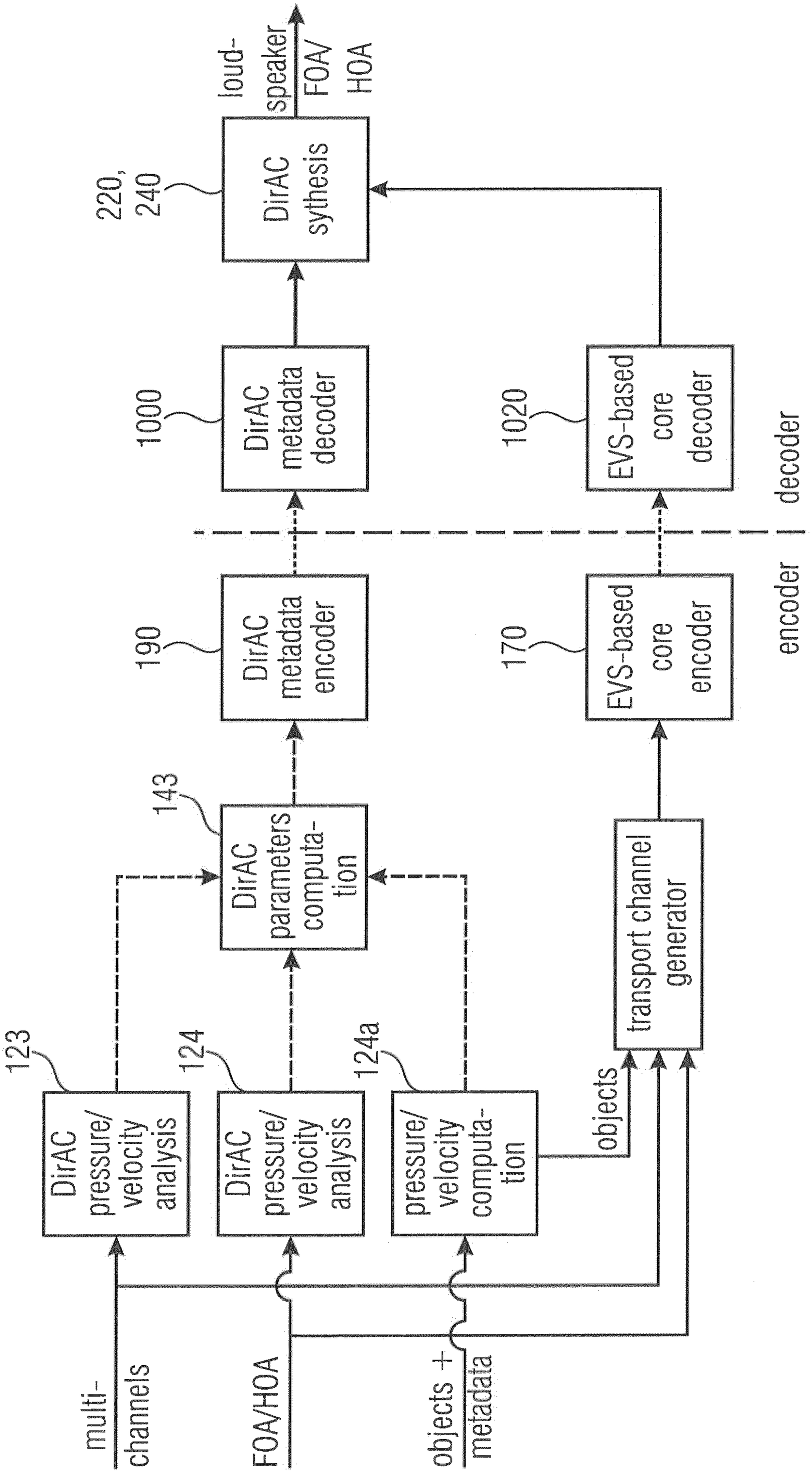

loud-speaker
FOA/HOA
220, 240
DirAC sythesis
1000
DirAC metadata decoder
1020
EVS-based core decoder
decoder
encoder
190
DirAC metadata encoder
170
EVS-based core encoder
143
DirAC parameters computa-tion
123
124
124a
DirAC pressure/ velocity analysis
DirAC pressure/ velocity analysis
pressure/ velocity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

Fig. 12

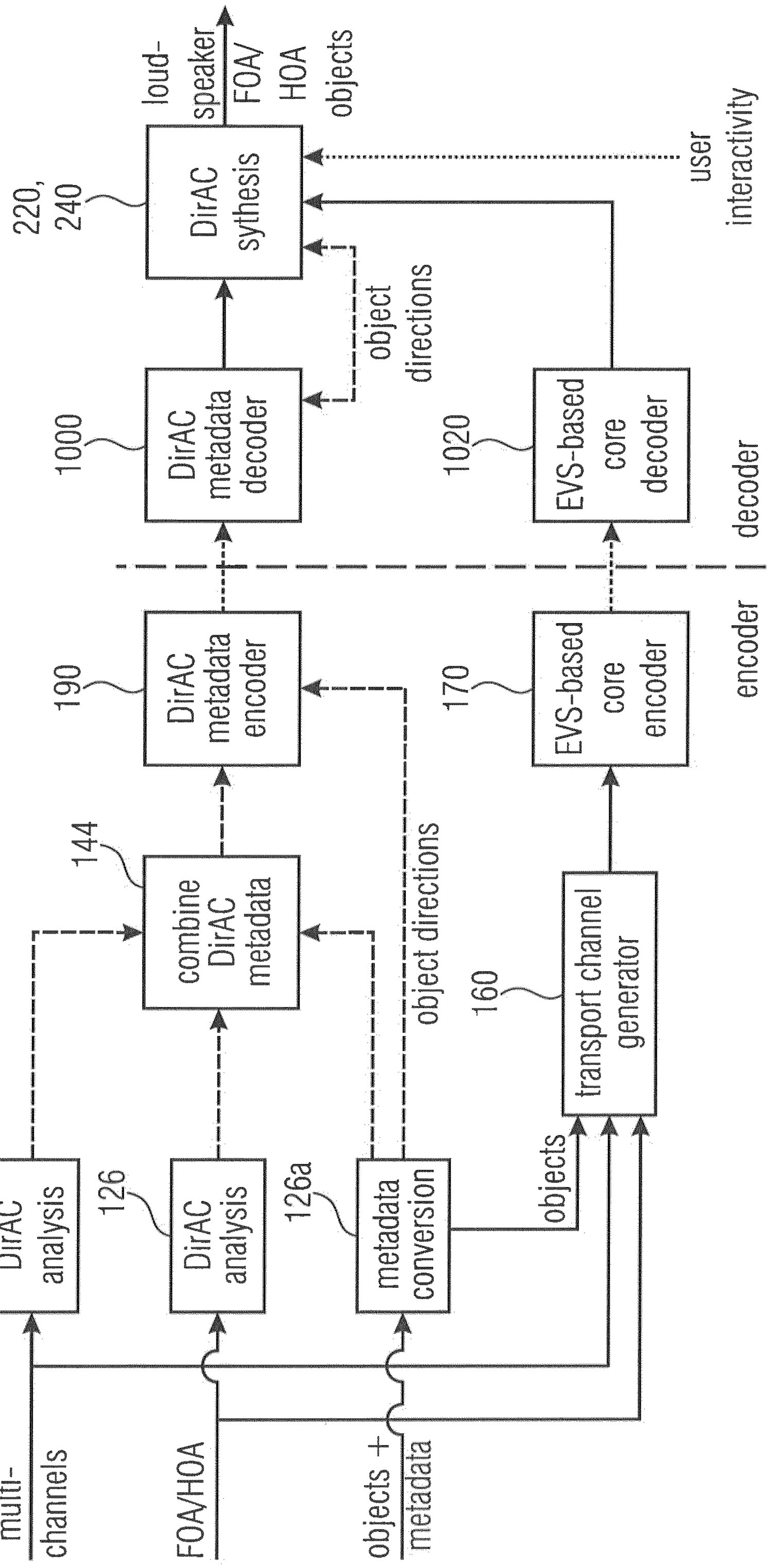

multi-channels
FOA/HOA
objects + metadata
125
DirAC analysis
126
DirAC analysis
126a
metadata conversion
160
transport channel generator
objects
144
combine DirAC metadata
object directions
190
DirAC metadata encoder
1000
DirAC metadata decoder
170
EVS-based core encoder
1020
EVS-based core decoder
encoder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

Fig. 13

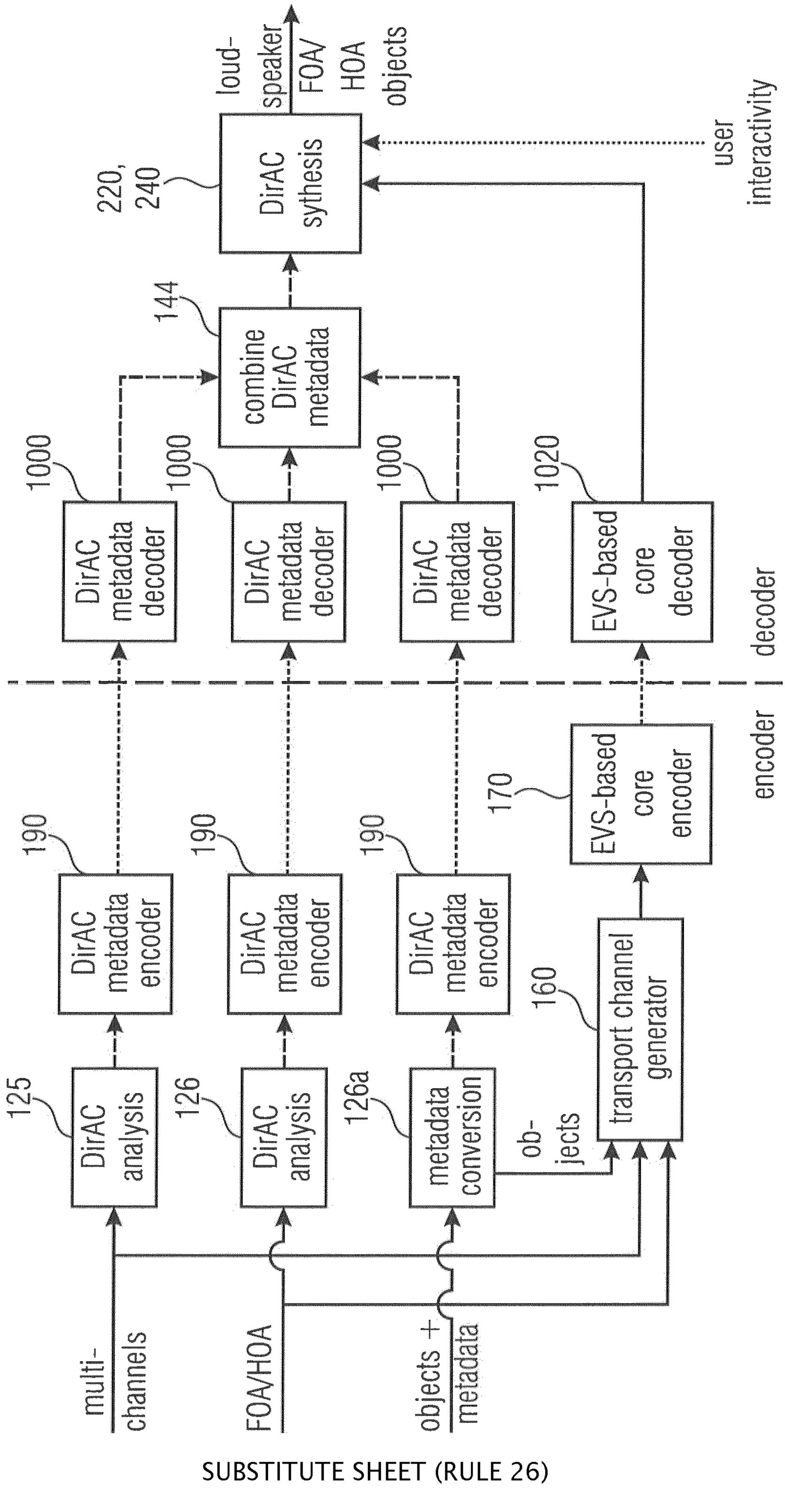

loud-speaker
FOA/HOA
objects
user interactivity
220, 240
DirAC sythesis
144
combine DirAC metadata
1000
DirAC metadata decoder
1020
EVS-based core decoder
decoder
190
DirAC metadata encoder
170
EVS-based core encoder
encoder
160
transport channel generator
125
DirAC analysis
126
126a
metadata conversion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

Fig. 14

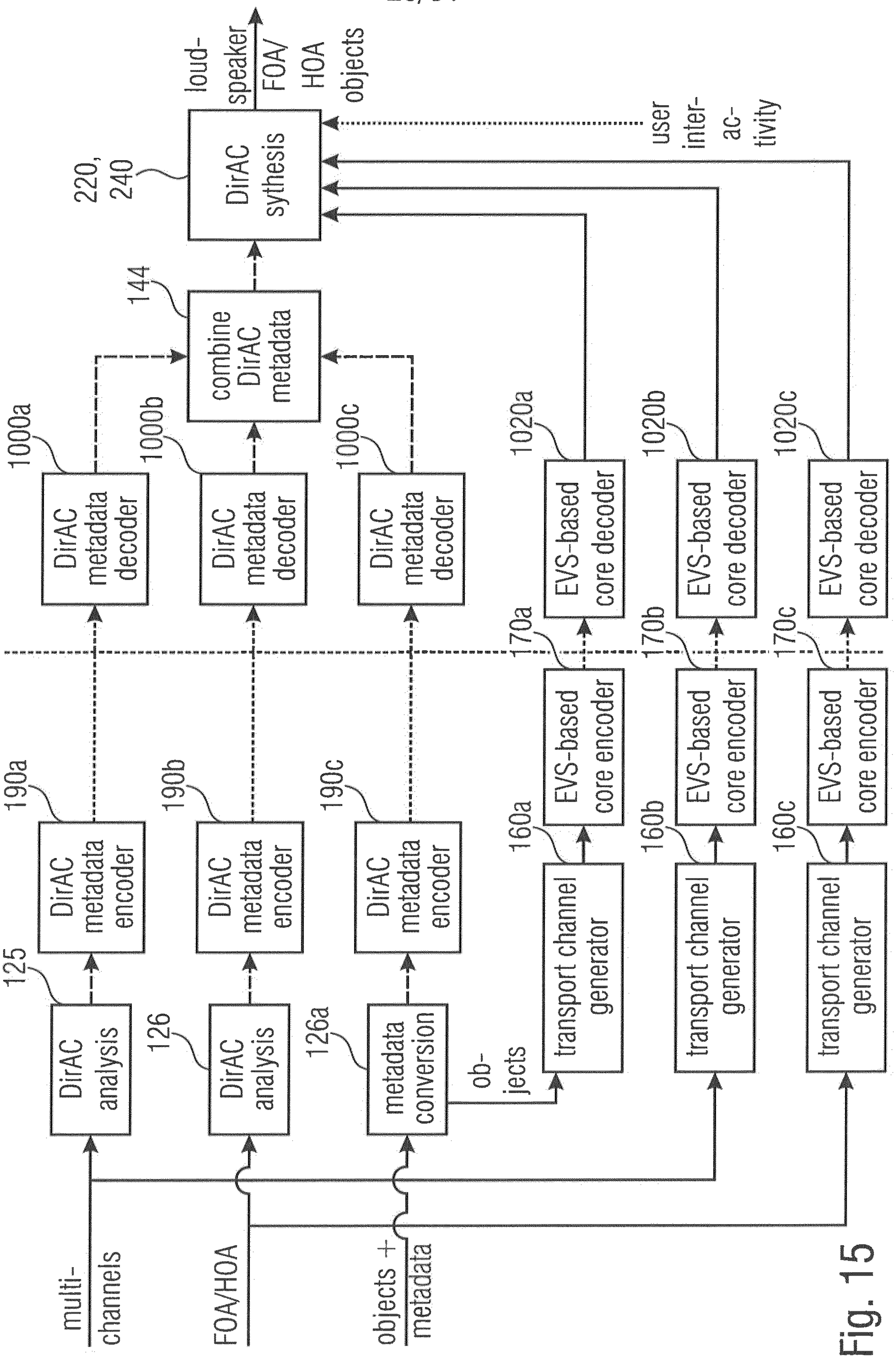

loud-speaker
FOA/HOA
objects
user inter-ac-tivity
220, 240
DirAC sythesis
144
combine DirAC metadata
1000a
DirAC metadata decoder
1000b
DirAC metadata decoder
1000c
DirAC metadata decoder
1020a
EVS-based core decoder
1020b
EVS-based core decoder
1020c
EVS-based core decoder
170a
170b
170c
190a
DirAC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

Fig. 15

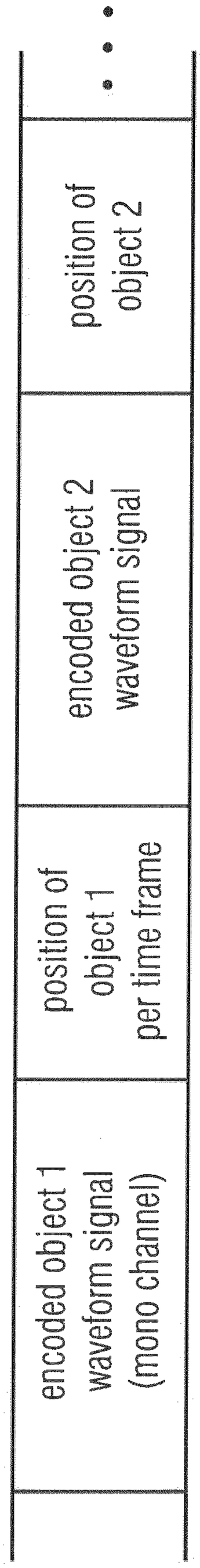


Fig. 16a

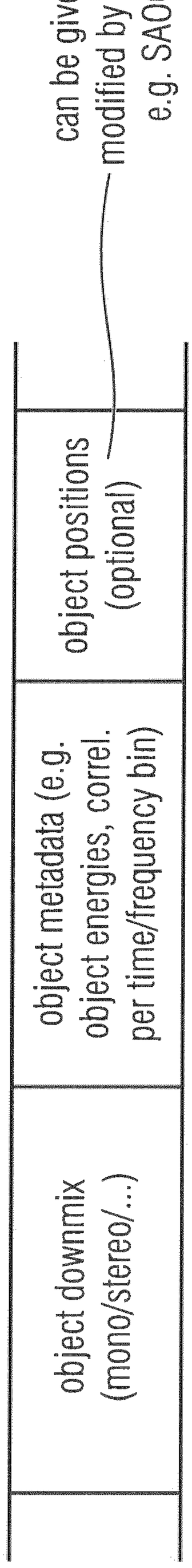


Fig. 16b

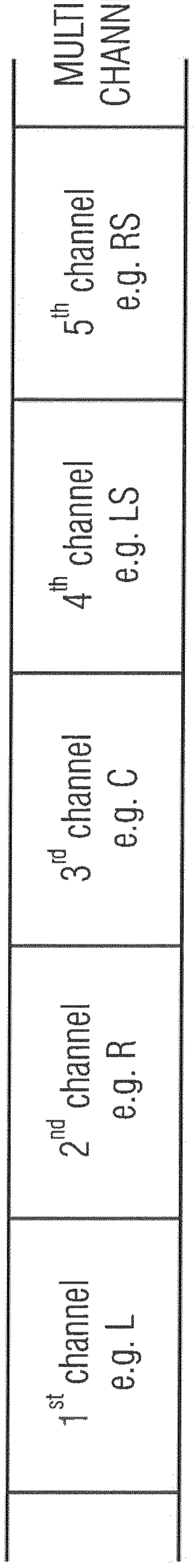


Fig. 16c

e.g. MPEG SURROUND	FoA HoA	DirAC
channel downmix (mono/stereo/...)	W	DirAC downmix (mono or stereo ...)
parametric side info as channel metadata for time/frequency bin	X	parametric side info (direction of arrival, (optional) diffuseness per time/frequency bin)
	Y	
	Z	
	optional higher components	

Fig. 16d Fig. 16e Fig. 16f