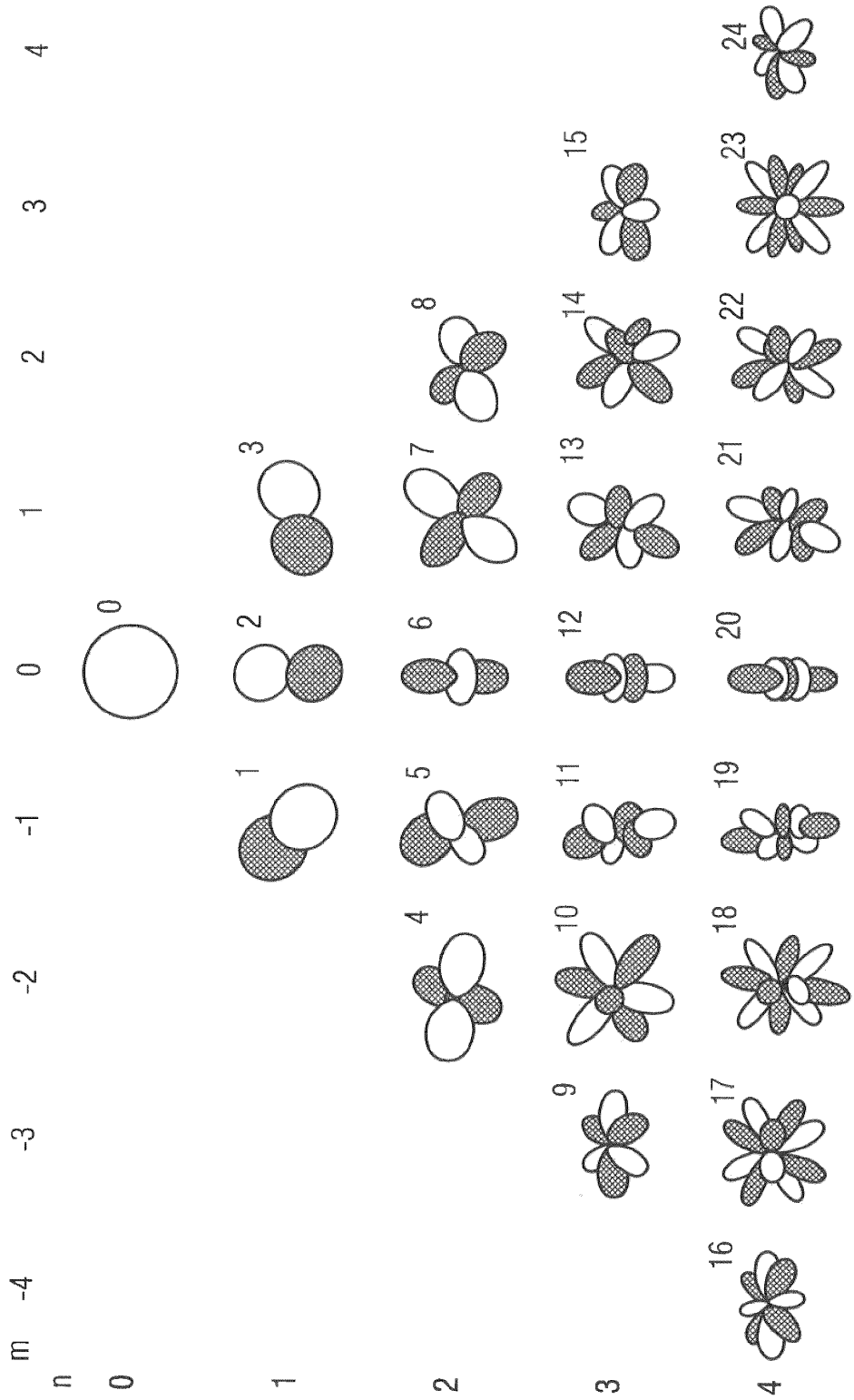

m
4
3
2
1
0
-1
-2
-3
-4
n
0
1
2
3
4
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Fig. 1a

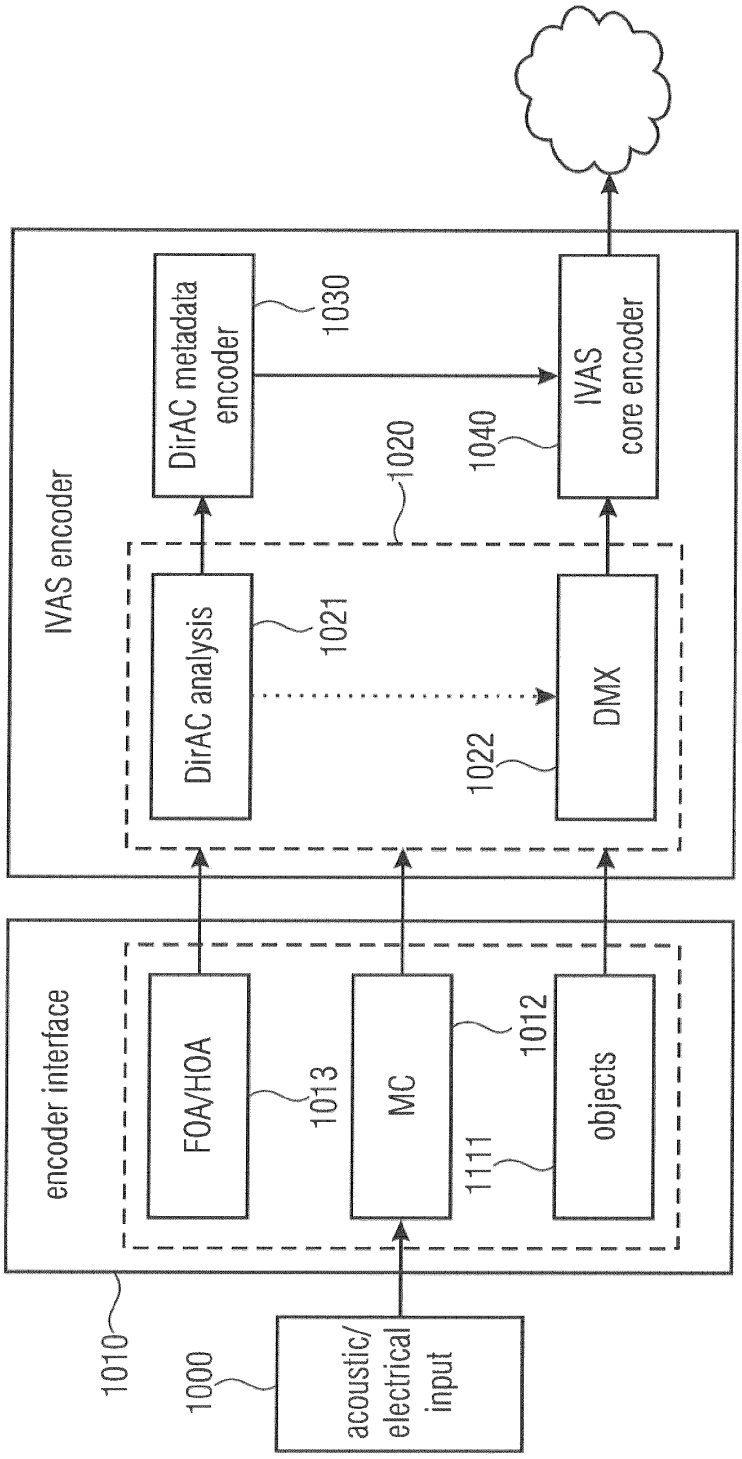

IVAS encoder
DirAC metadata encoder
1030
IVAS core encoder
1040
1020
DirAC analysis
1021
DMX
1022
encoder interface
FOA/HOA
1013
MC
1012
objects
1111
1010
1000
acoustic/ electrical input

Fig. 1b

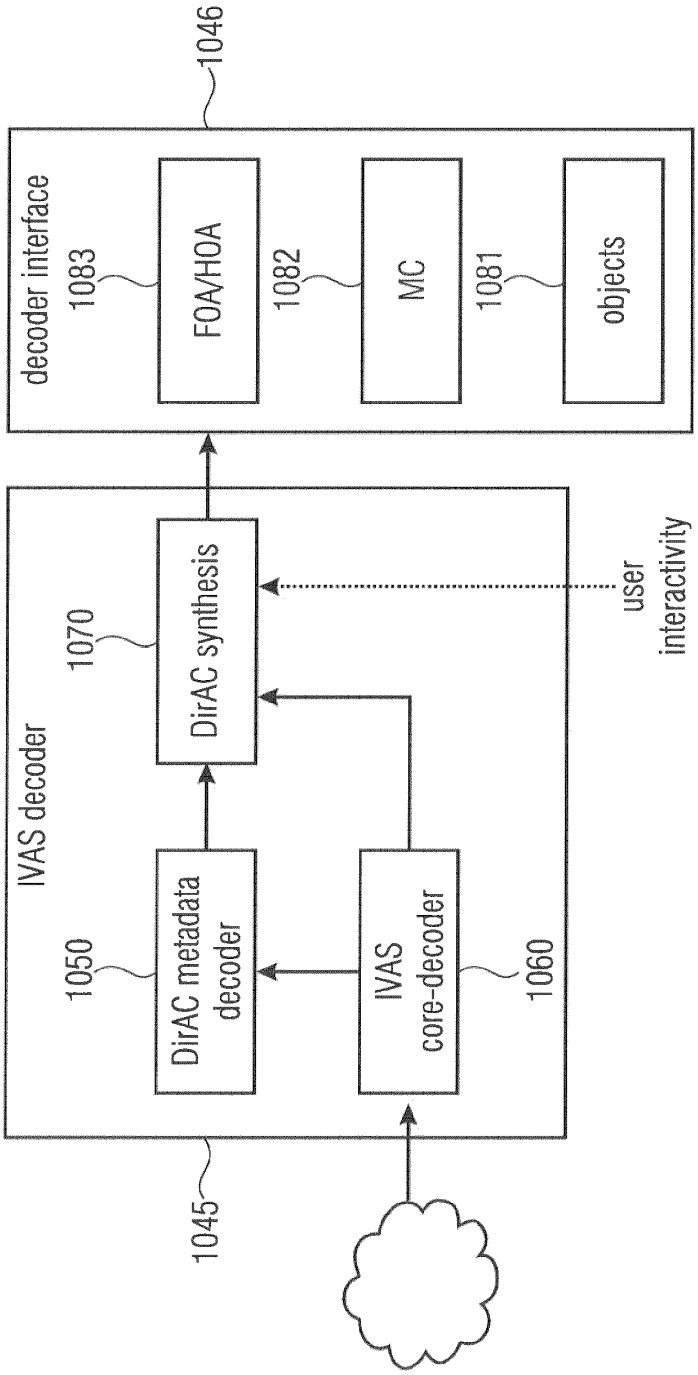

1046
decoder interface
1083
FOA/HOA
1082
MC
1081
objects
IVAS decoder
1070
DirAC synthesis
user interactivity
1050
DirAC metadata decoder
IVAS core-decoder
1060
1045

Fig. 2a

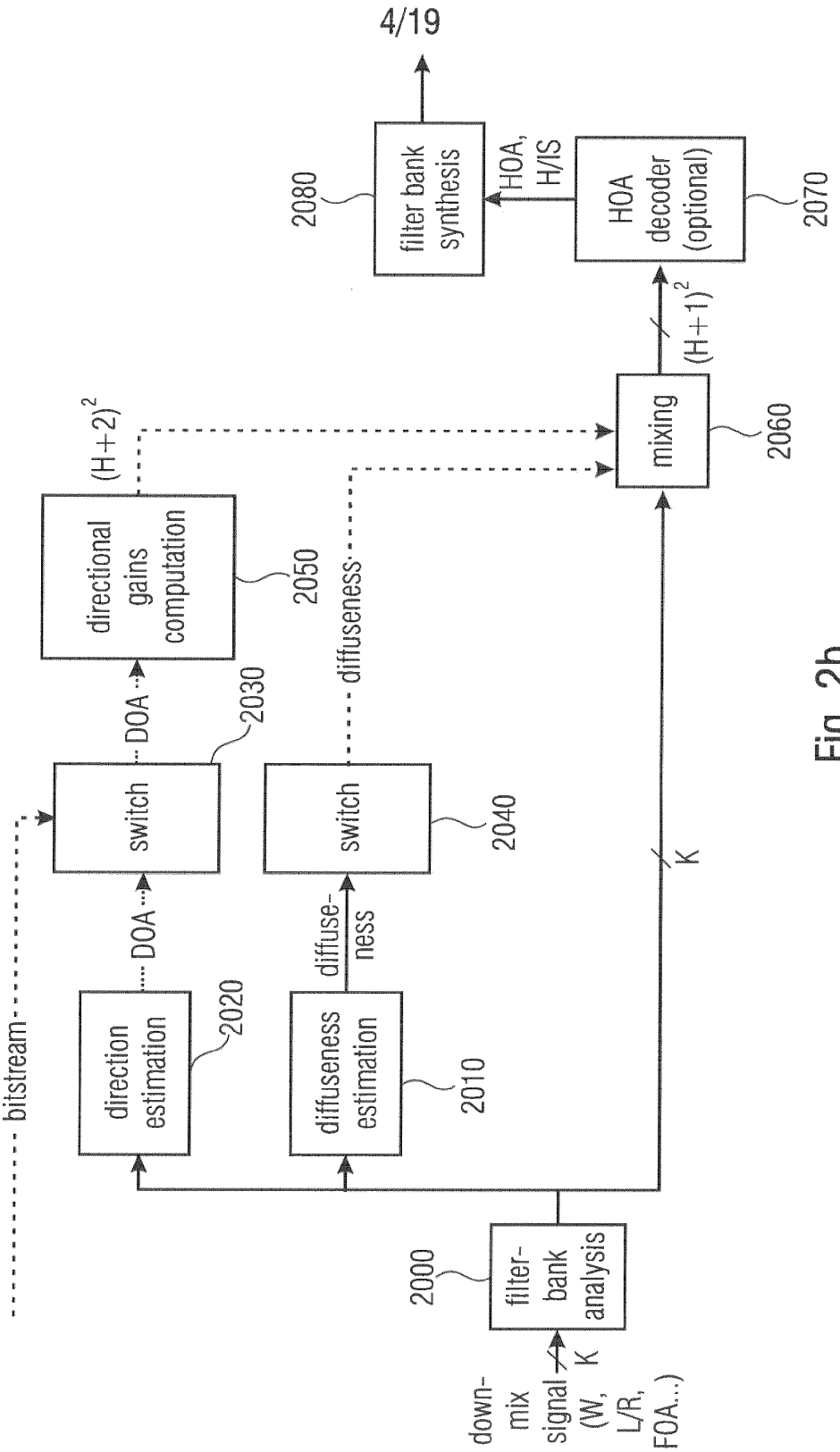

down-mix signal (W, L/R, FOA...)
K
2000
filter-bank analysis
bitstream
direction estimation
2020
DOA
diffuseness estimation
2010
diffuse-ness
switch
2030
DOA
switch
2040
diffuseness
directional gains computation
2050
$(H+2)^2$
mixing
2060
K
$(H+1)^2$
HOA decoder (optional)
2070
HOA, H/IS
filter bank synthesis
2080

Fig. 2b

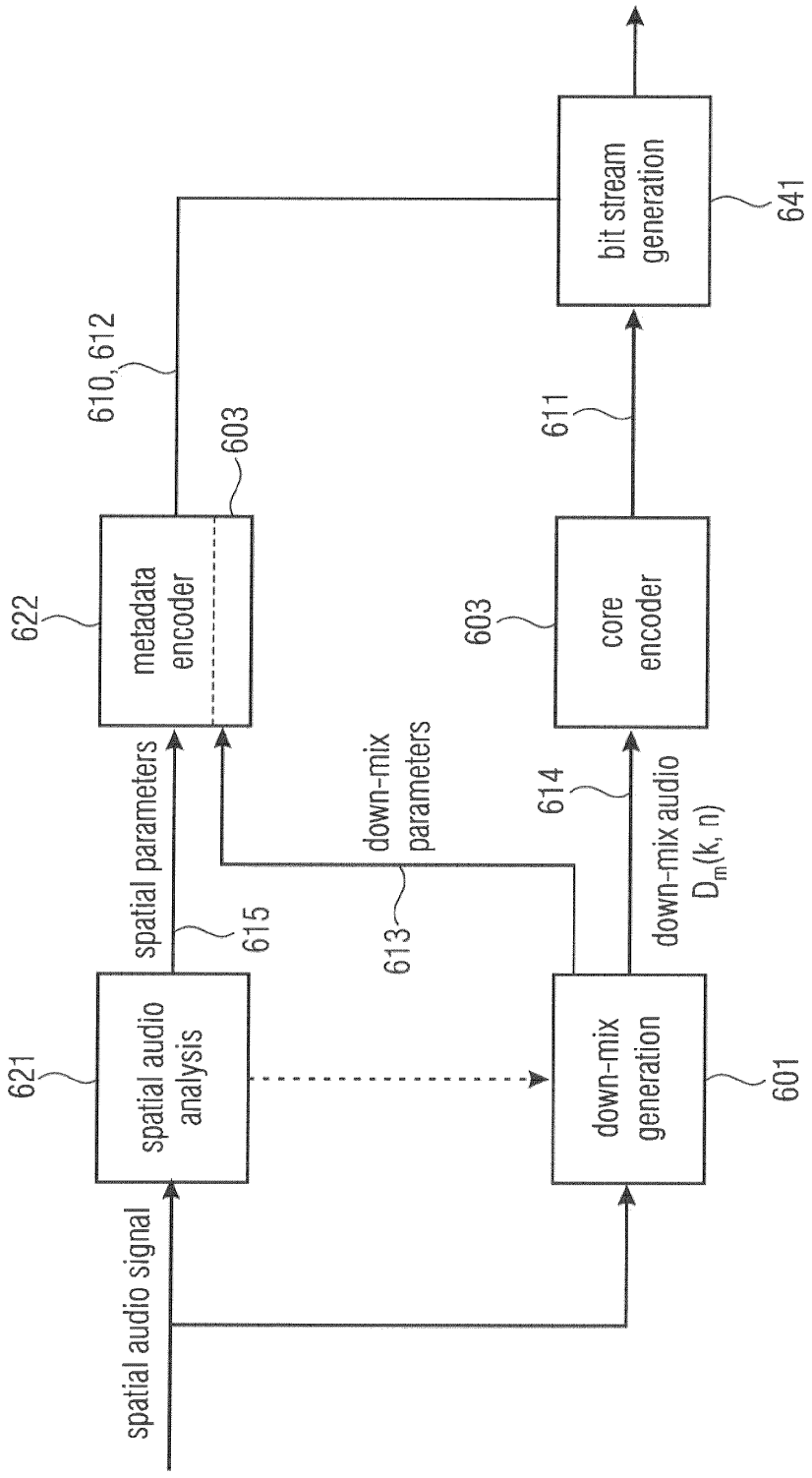

spatial audio signal
spatial audio analysis
621
spatial parameters
615
metadata encoder
622
610, 612
603
down-mix parameters
613
down-mix generation
601
down-mix audio $D_m(k, n)$
614
core encoder
603
611
bit stream generation
641

Fig. 3

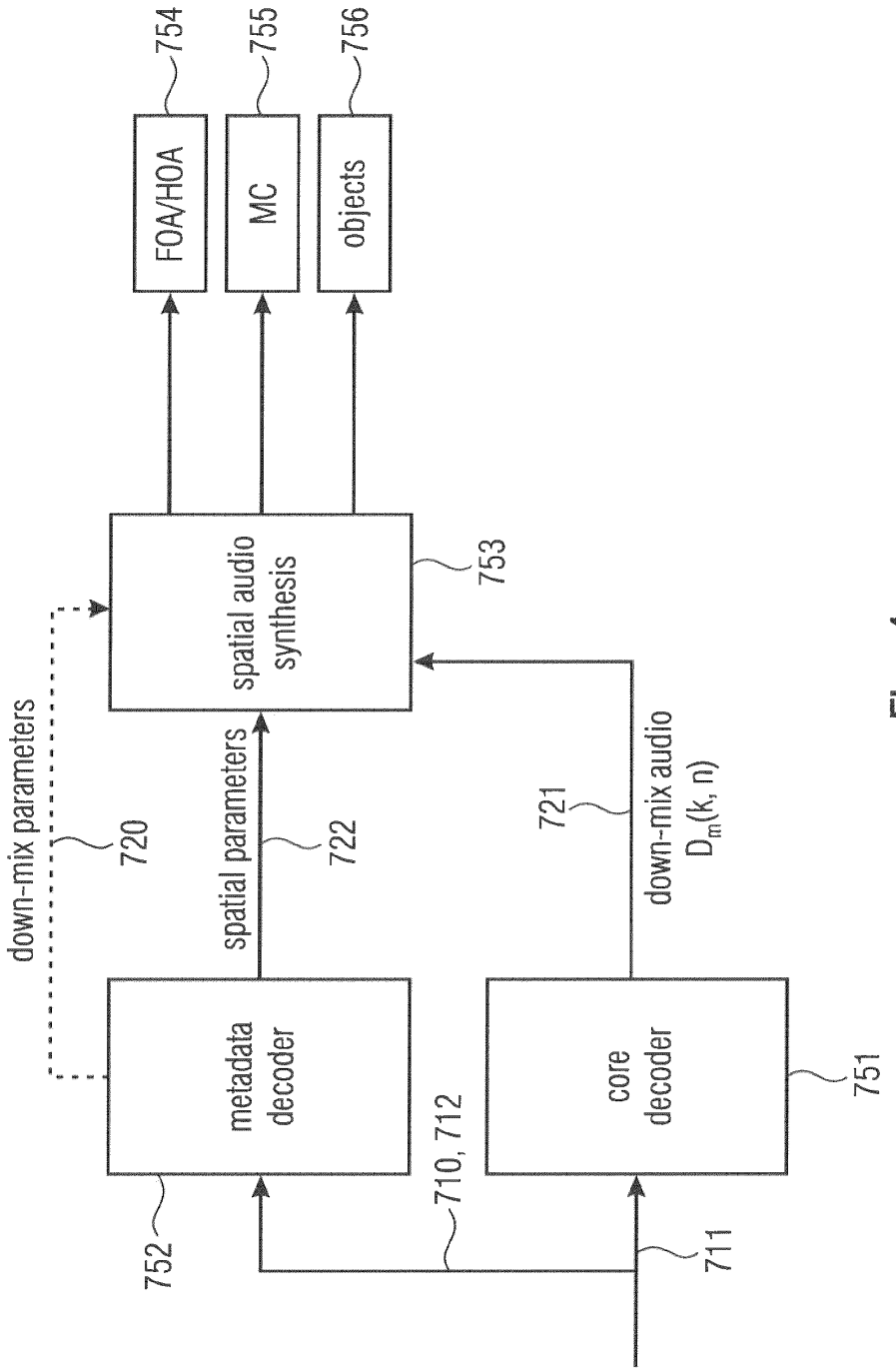

754
755
756
FOA/HOA
MC
objects
spatial audio synthesis
753
down-mix parameters
720
spatial parameters
722
721
down-mix audio
$D_m(k, n)$
metadata decoder
core decoder
751
710, 712
752
711

Fig. 4

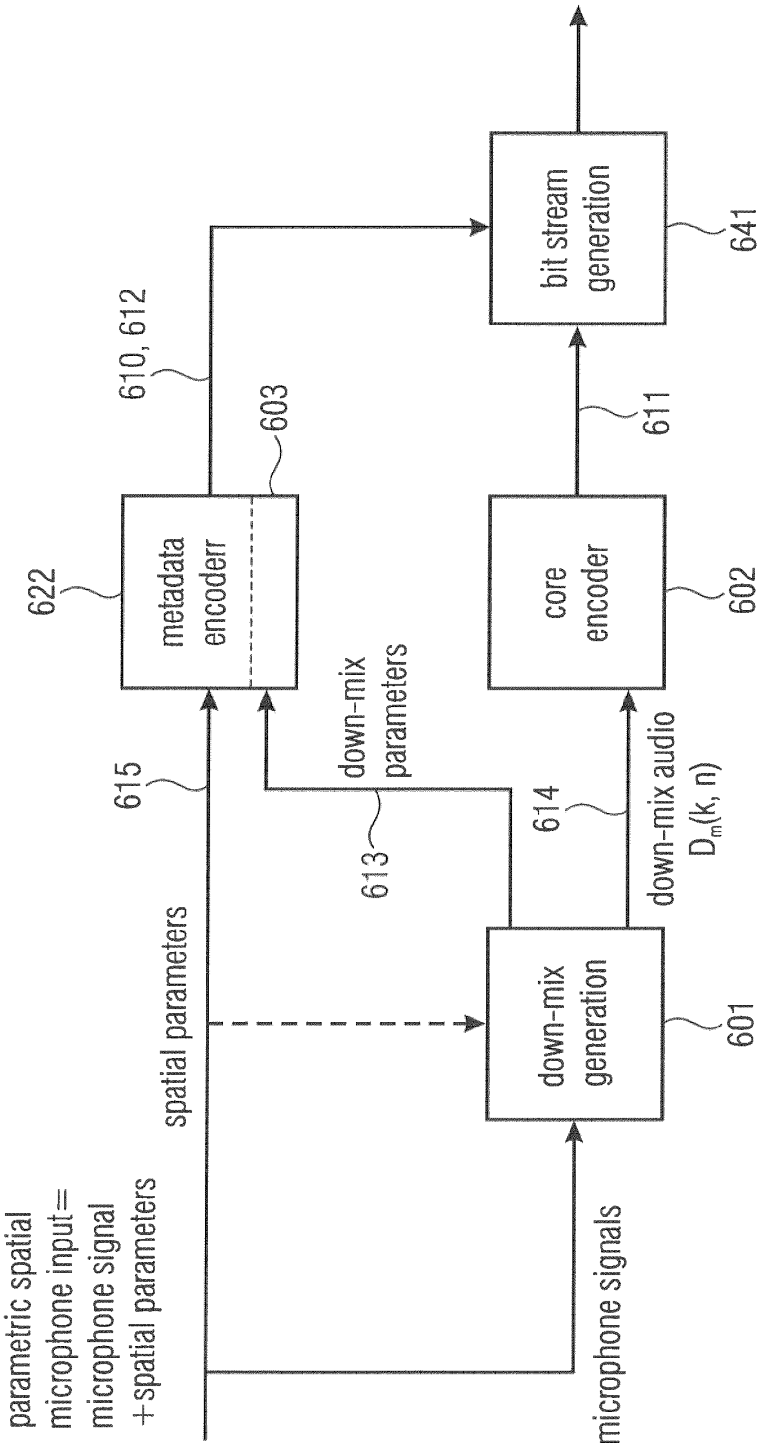

parametric spatial microphone input = microphone signal + spatial parameters
spatial parameters
microphone signals
down-mix generation
601
down-mix audio $D_m(k, n)$
614
down-mix parameters
613
615
core encoder
602
611
metadata encoderr
622
603
610, 612
bit stream generation
641

Fig. 5

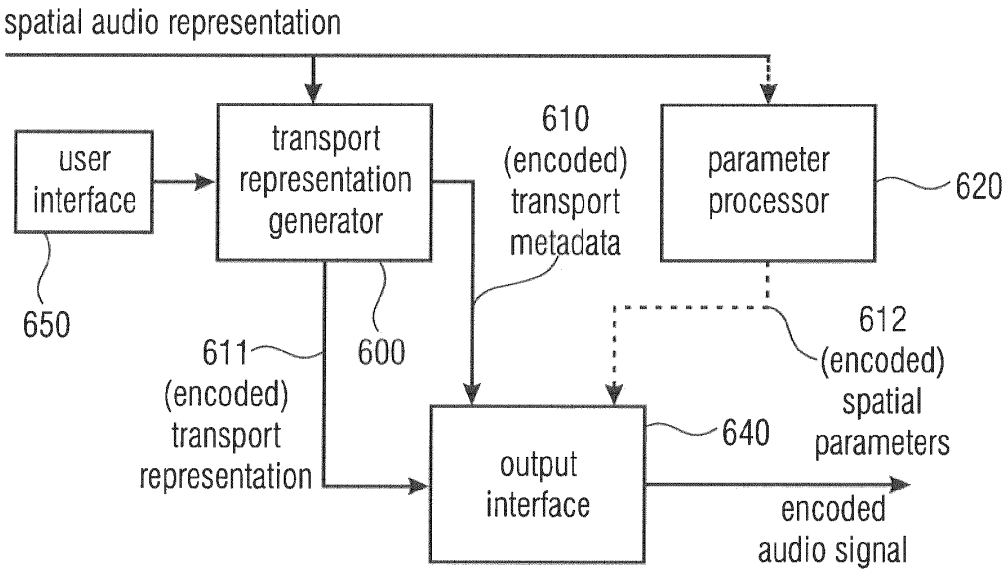

spatial audio representation
user interface
transport representation generator
610 (encoded) transport metadata
parameter processor
620
650
611 (encoded) transport representation
600
612 (encoded) spatial parameters
640
output interface
encoded audio signal

Fig. 6

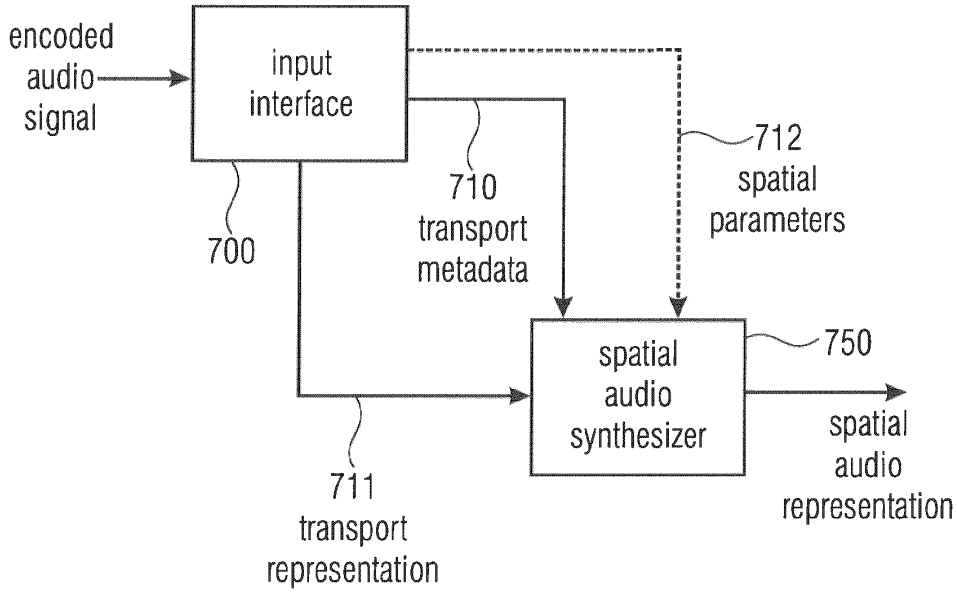

encoded audio signal
input interface
712 spatial parameters
710 transport metadata
700
750
spatial audio synthesizer
spatial audio representation
711 transport representation

Fig. 7

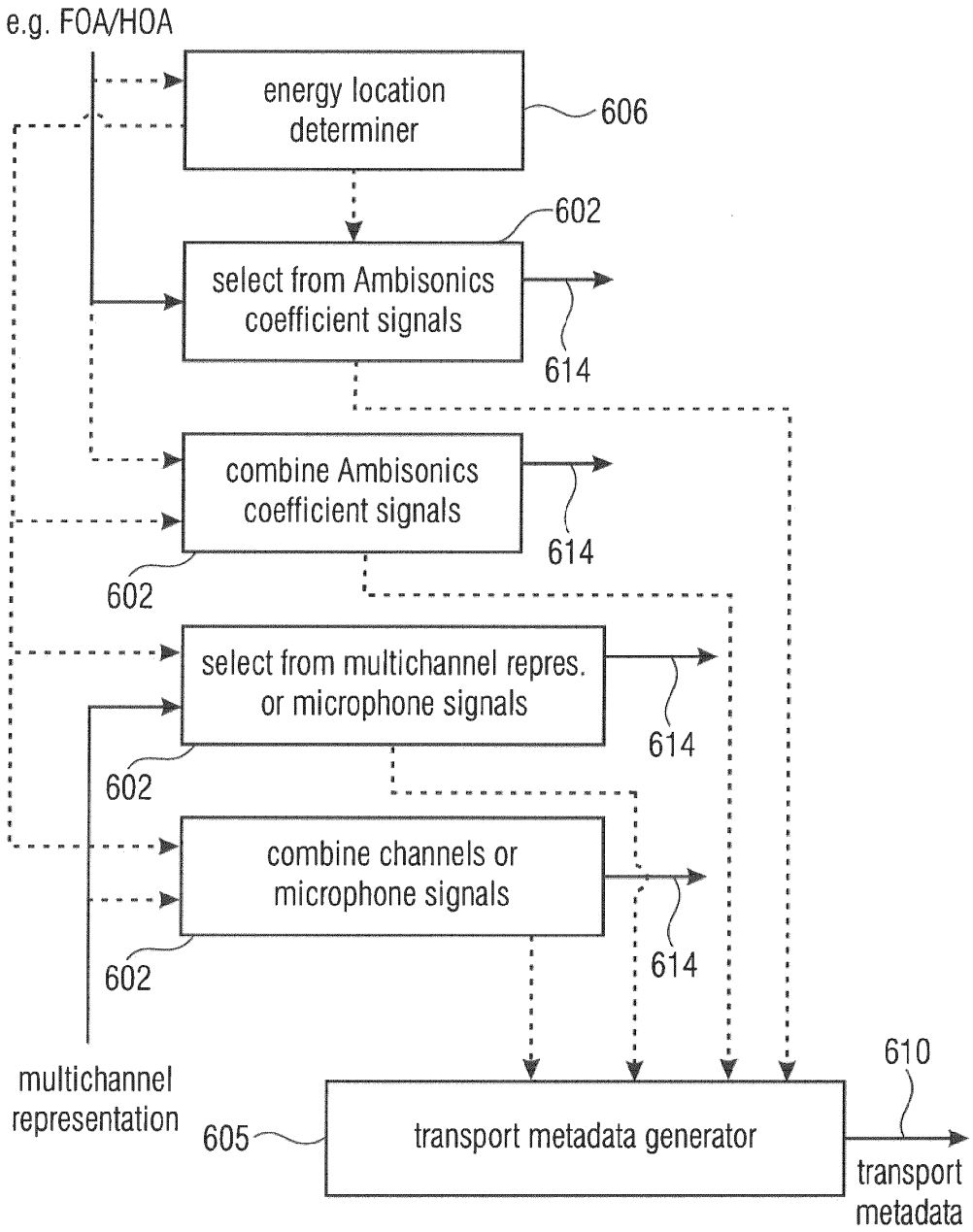

e.g. FOA/HOA
energy location determiner
606
602
select from Ambisonics coefficient signals
614
combine Ambisonics coefficient signals
614
602
select from multichannel repres. or microphone signals
614
602
combine channels or microphone signals
614
602
multichannel representation
605
transport metadata generator
610
transport metadata

Fig. 8a

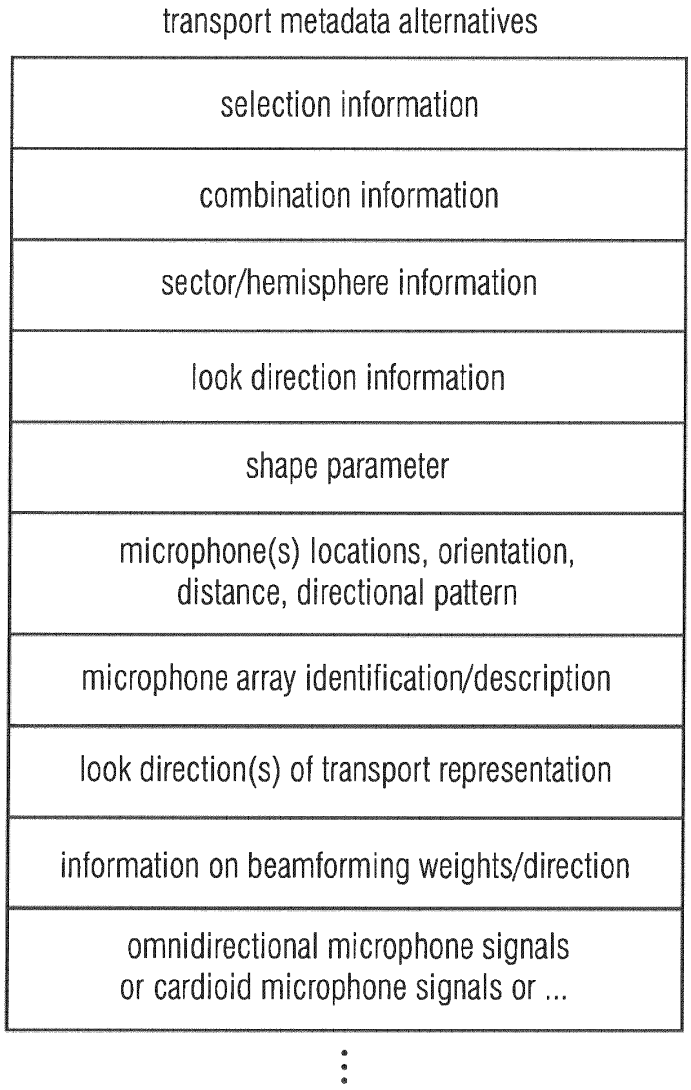

transport metadata alternatives
selection information
combination information
sector/hemisphere information
look direction information
shape parameter
microphone(s) locations, orientation, distance, directional pattern
microphone array identification/description
look direction(s) of transport representation
information on beamforming weights/direction
omnidirectional microphone signals or cardioid microphone signals or ...

Fig. 8b

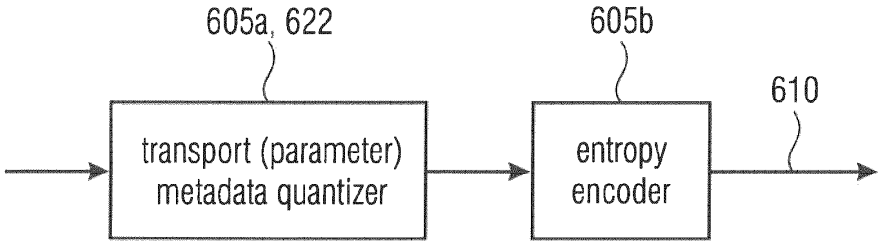

605a, 622
605b
610
transport (parameter) metadata quantizer
entropy encoder

Fig. 8c

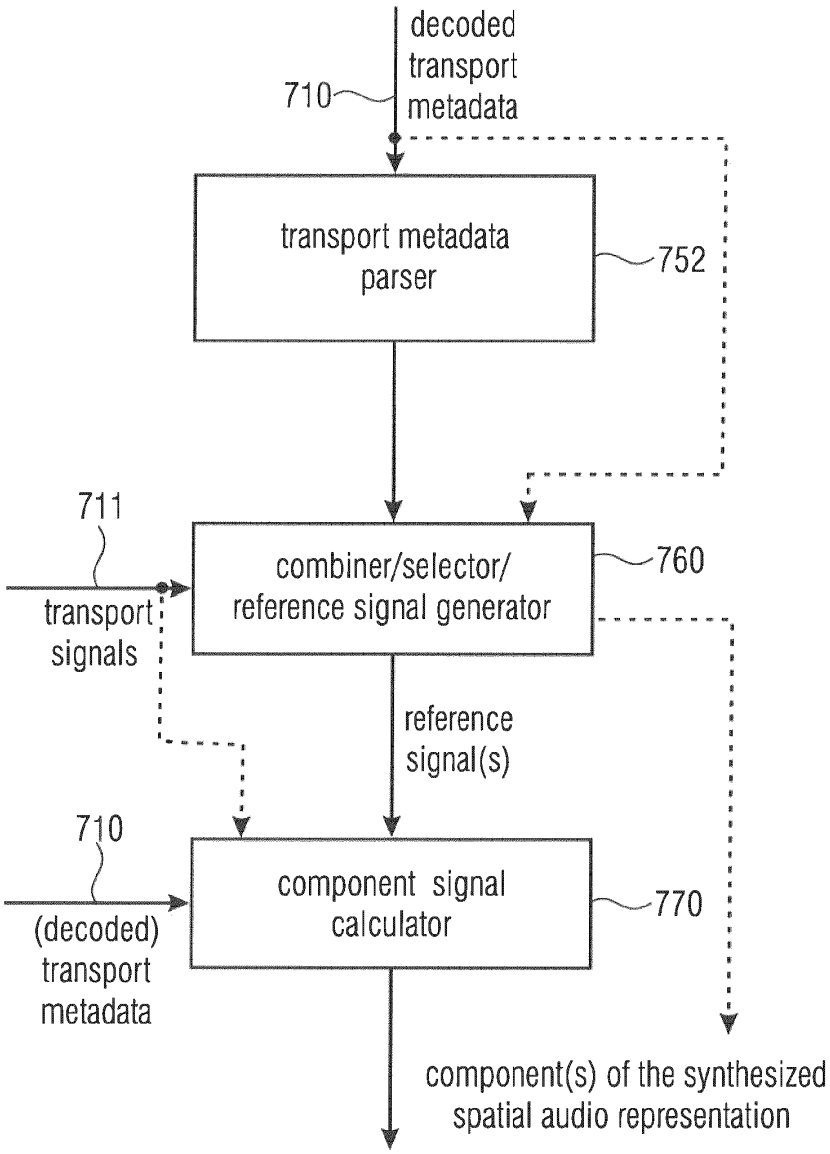

decoded transport metadata
710
transport metadata parser
752
711
transport signals
combiner/selector/ reference signal generator
760
reference signal(s)
710
(decoded) transport metadata
component signal calculator
770
component(s) of the synthesized spatial audio representation

Fig. 9a

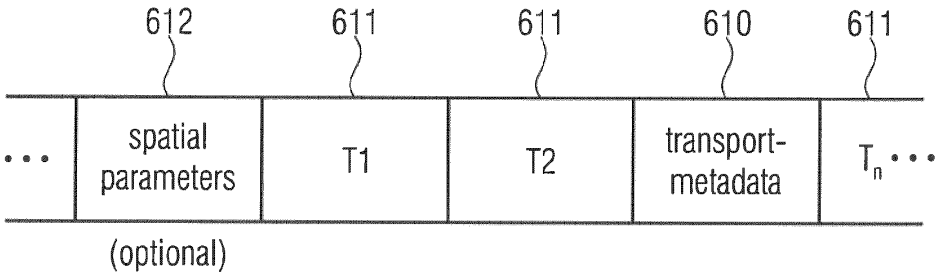


Fig. 9b

771

speaker	ref. signal sel./generator	T-metadata
L	T1/T1	T1: left/front (omni or cardioid)
R	T2/T1	
L_s	T1/T2	T2: right/back (omni or cardioid)
R_s	T2/T2	
C	T1+T2/T1	

Fig. 9c

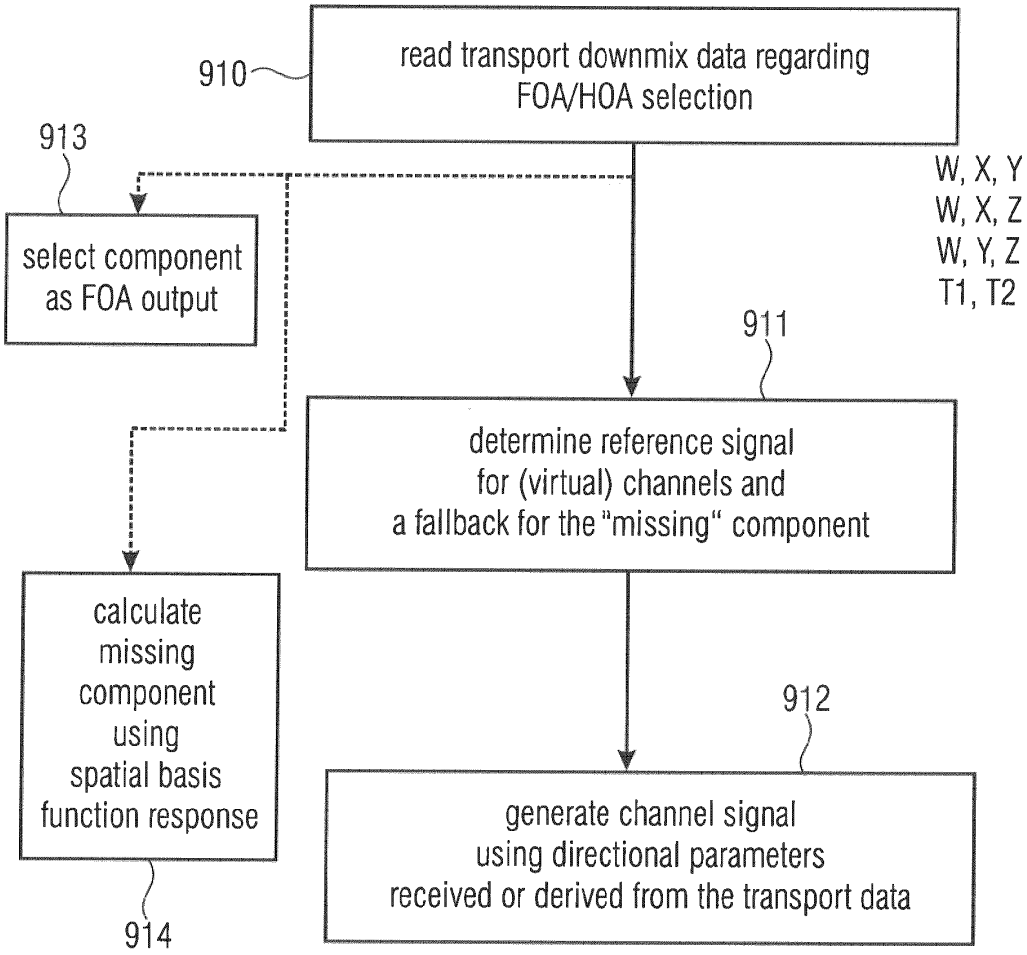

910
read transport downmix data regarding FOA/HOA selection
913
select component as FOA output
W, X, Y
W, X, Z
W, Y, Z
T1, T2
911
determine reference signal for (virtual) channels and a fallback for the "missing" component
calculate missing component using spatial basis function response
914
912
generate channel signal using directional parameters received or derived from the transport data

Fig. 9d

	shape parameter	look direction
$C_M = 1$	omnidirectional	left
		right
$C_M = 0.5$	cardioid	front
		back
$C_M = 0$	dipole	up
		down
		(quantized) ϕ, ϑ
		left/right pair
		front/back pair

Fig. 9e

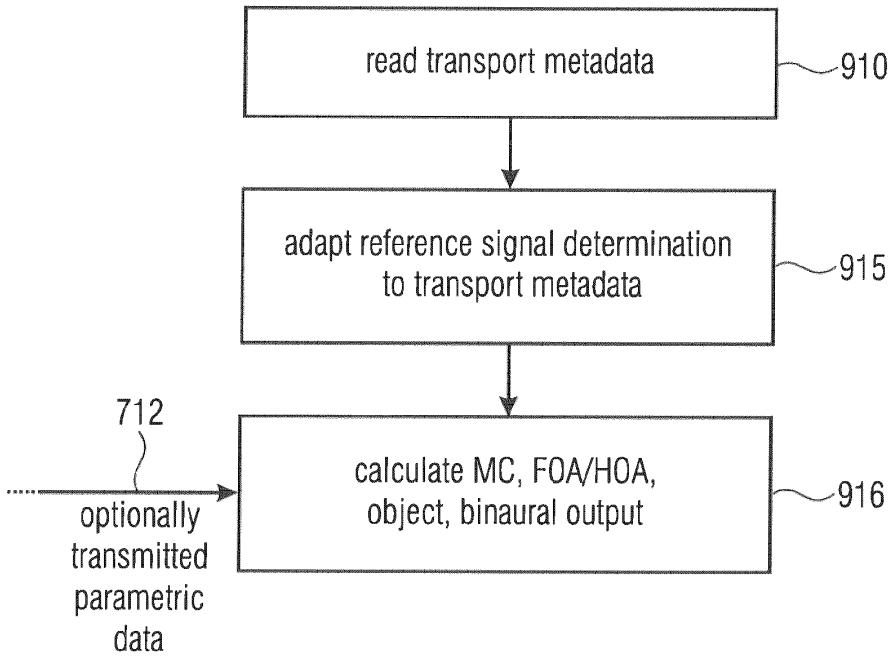


Fig. 9f

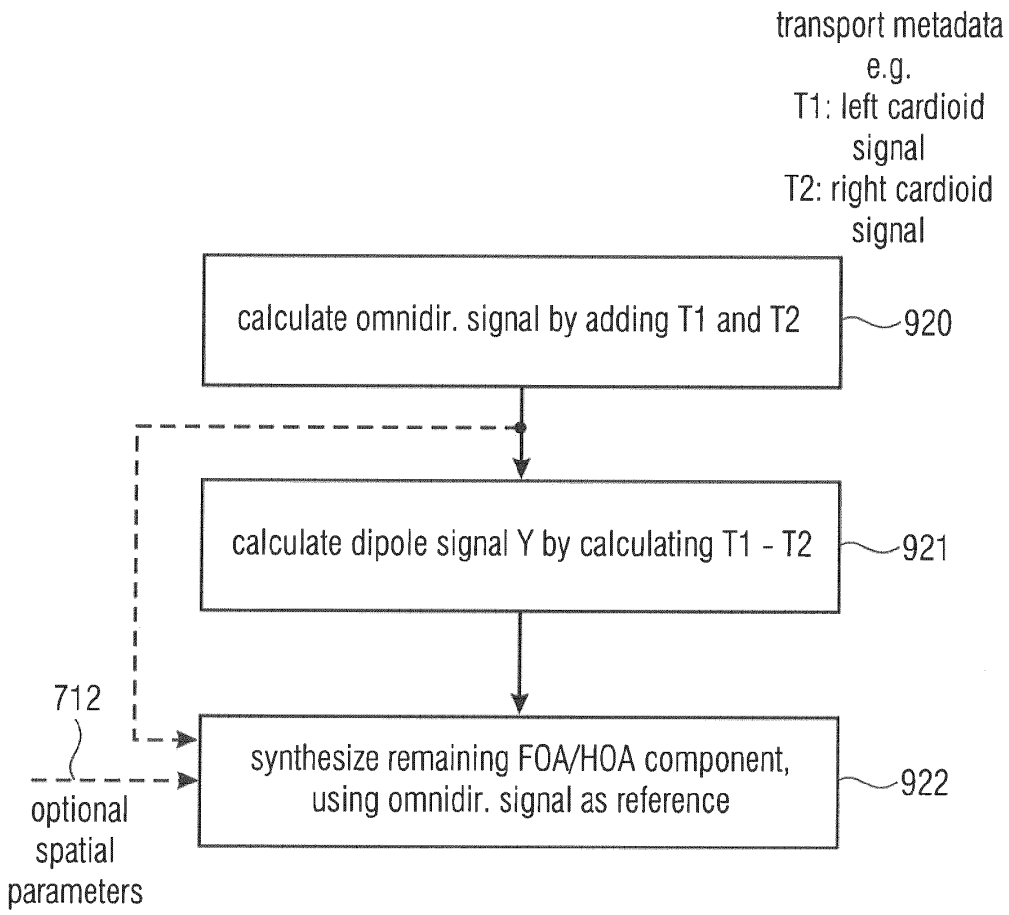

transport metadata
e.g.
T1: left cardioid signal
T2: right cardioid signal
calculate omnidir. signal by adding T1 and T2
920
calculate dipole signal Y by calculating T1 - T2
921
712
synthesize remaining FOA/HOA component, using omnidir. signal as reference
922
optional spatial parameters

Fig. 9g

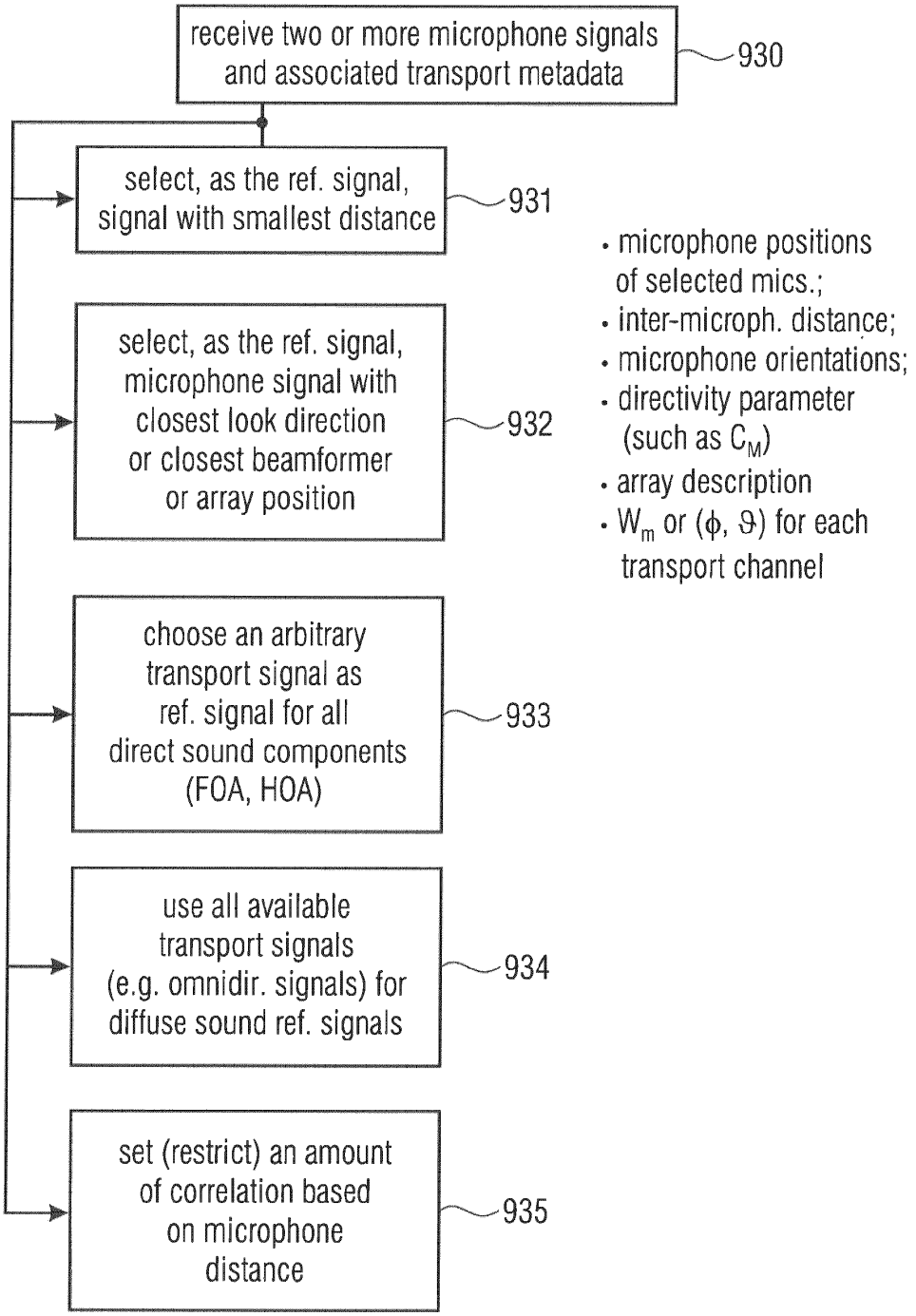

receive two or more microphone signals and associated transport metadata
930
select, as the ref. signal, signal with smallest distance
931
select, as the ref. signal, microphone signal with closest look direction or closest beamformer or array position
932
choose an arbitrary transport signal as ref. signal for all direct sound components (FOA, HOA)
933
use all available transport signals (e.g. omnidir. signals) for diffuse sound ref. signals
934
set (restrict) an amount of correlation based on microphone distance
935
• microphone positions of selected mics.;
• inter-microph. distance;
• microphone orientations;
• directivity parameter (such as C_M)
• array description
• W_m or (ϕ, ϑ) for each transport channel

Fig. 9h

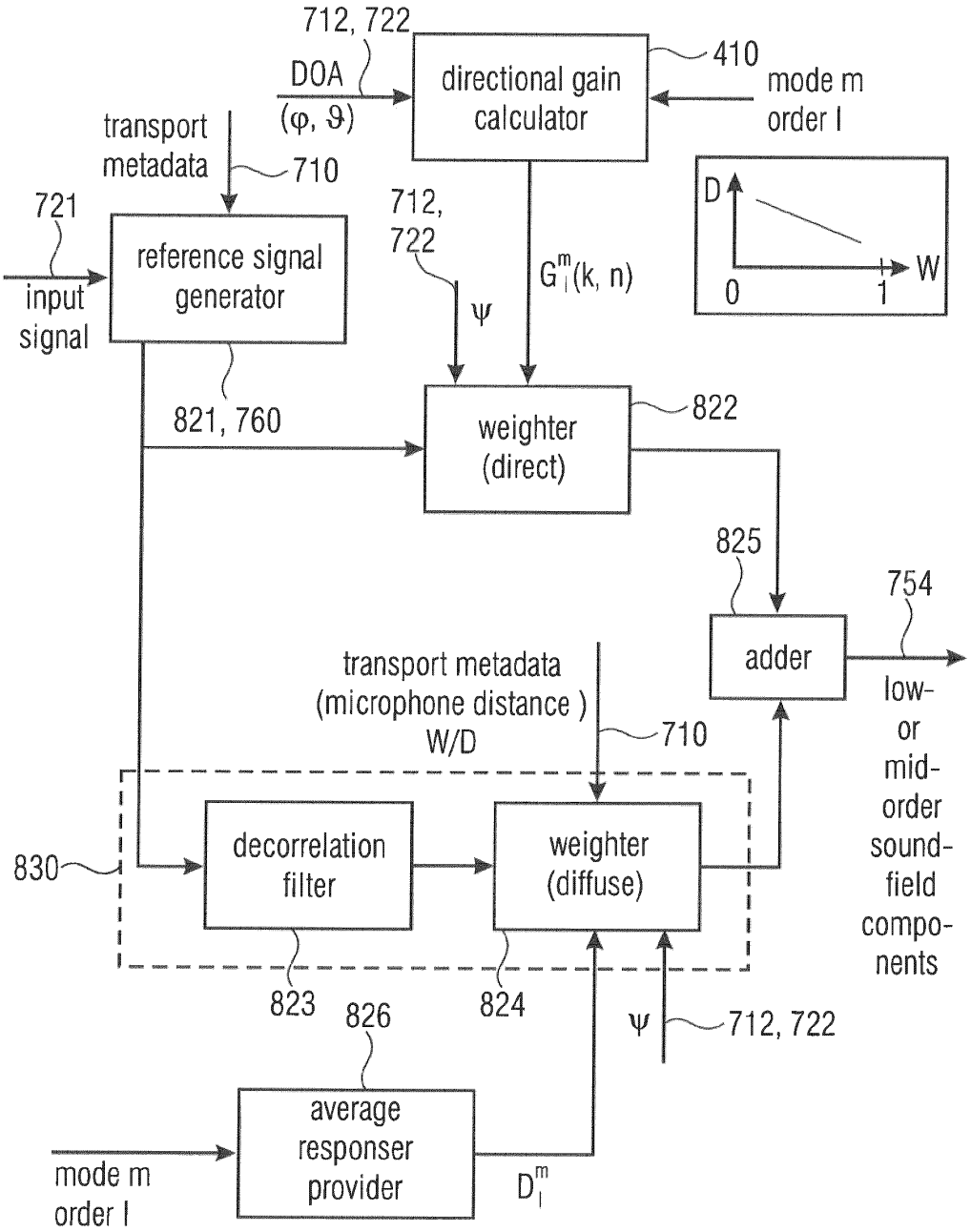

712, 722
DOA
(φ, ϑ)
directional gain calculator
410
mode m
order l
transport metadata
710
721
input signal
reference signal generator
712, 722
ψ
$G^m_l(k, n)$
D
W
0
1
821, 760
weighter (direct)
822
825
754
adder
transport metadata (microphone distance) W/D
710
low- or mid-order sound-field compo-nents
830
decorrelation filter
weighter (diffuse)
823
824
826
ψ
712, 722
average responser provider
mode m
order l
D^m_l

Fig. 10

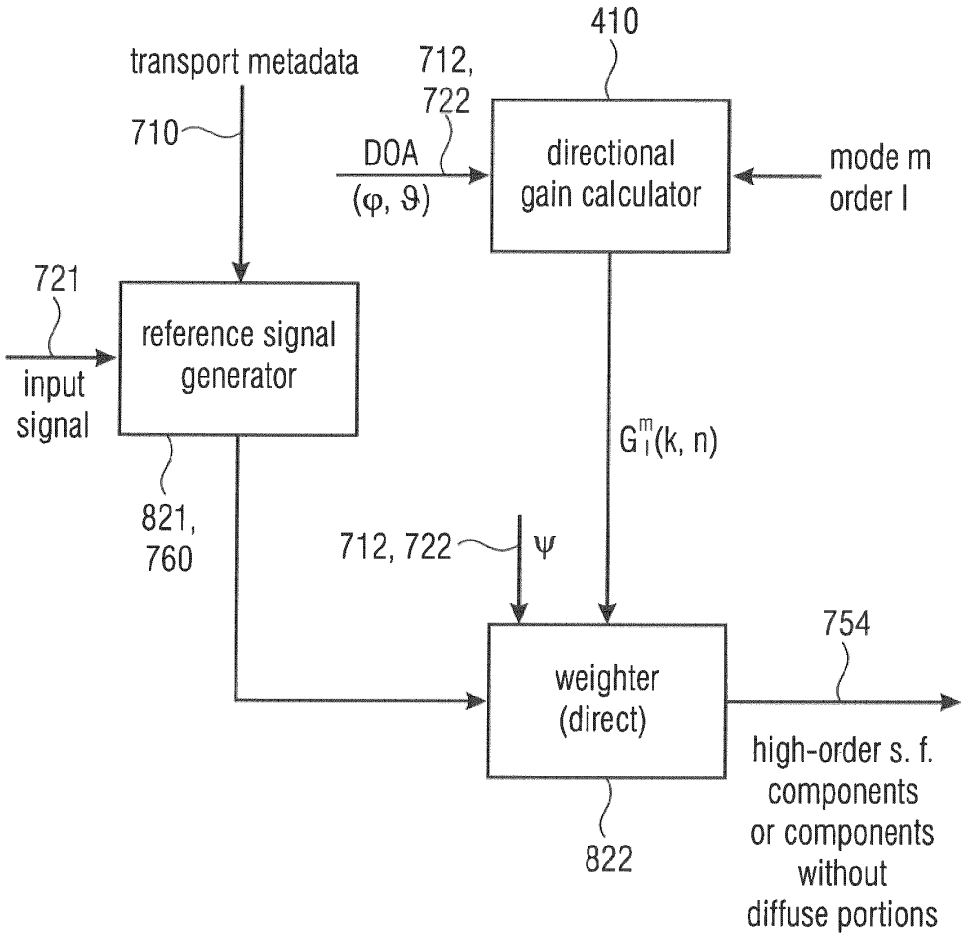

transport metadata
710
712,
722
410
DOA
(φ, ϑ)
directional
gain calculator
mode m
order l
721
input
signal
reference signal
generator
821,
760
$G_l^m(k, n)$
712, 722
Ψ
754
weighter
(direct)
822
high-order s. f.
components
or components
without
diffuse portions

Fig. 11

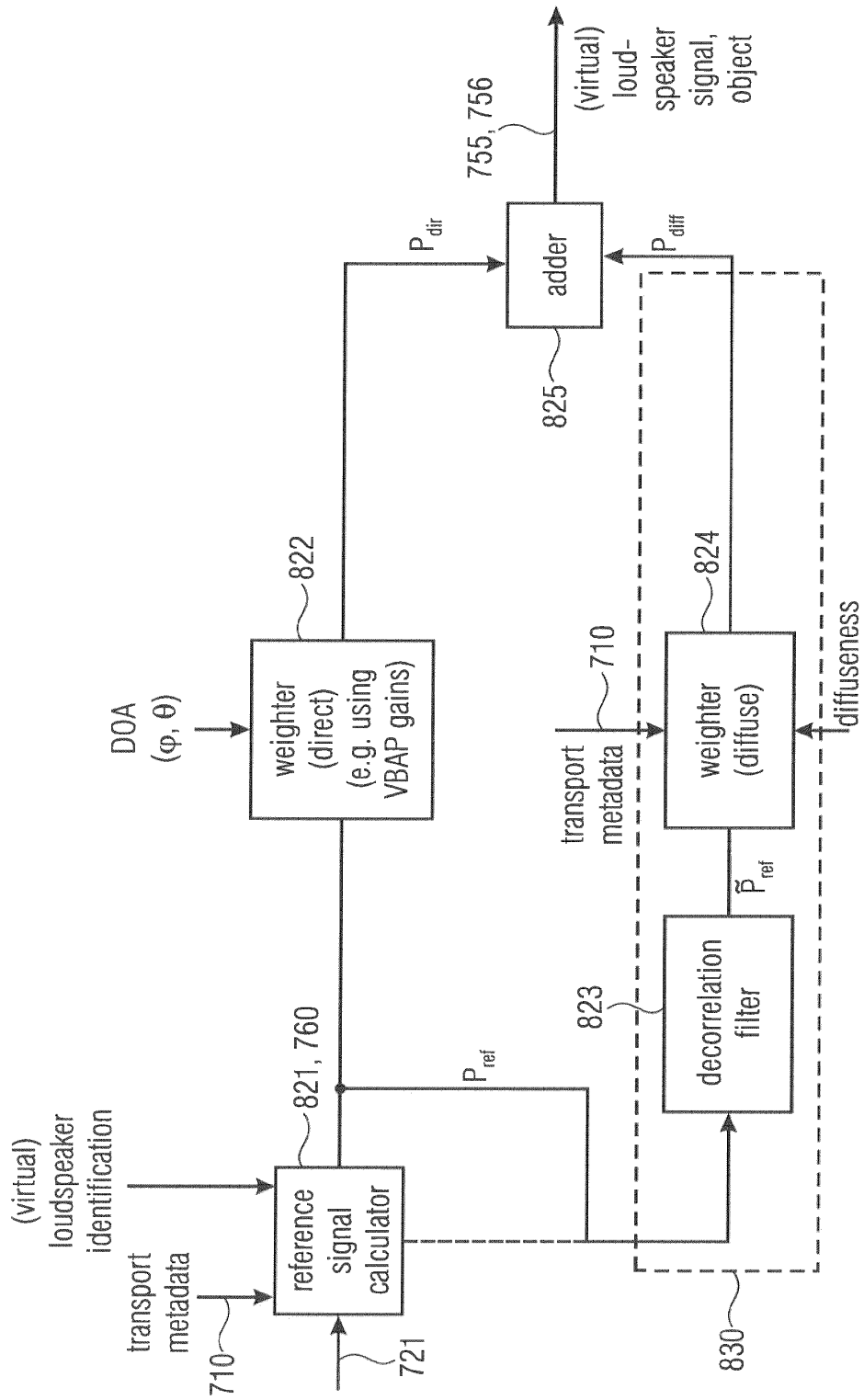

(virtual) loudspeaker identification
transport metadata
710
reference signal calculator
721
821, 760
P_{ref}
823
decorrelation filter
$\tilde{P}_{ref}$
824
weighter (diffuse)
diffuseness
710
transport metadata
DOA (φ, θ)
822
weighter (direct) (e.g. using VBAP gains)
P_{dir}
P_{diff}
825
adder
755, 756
(virtual) loud-speaker signal, object
830

Fig. 12