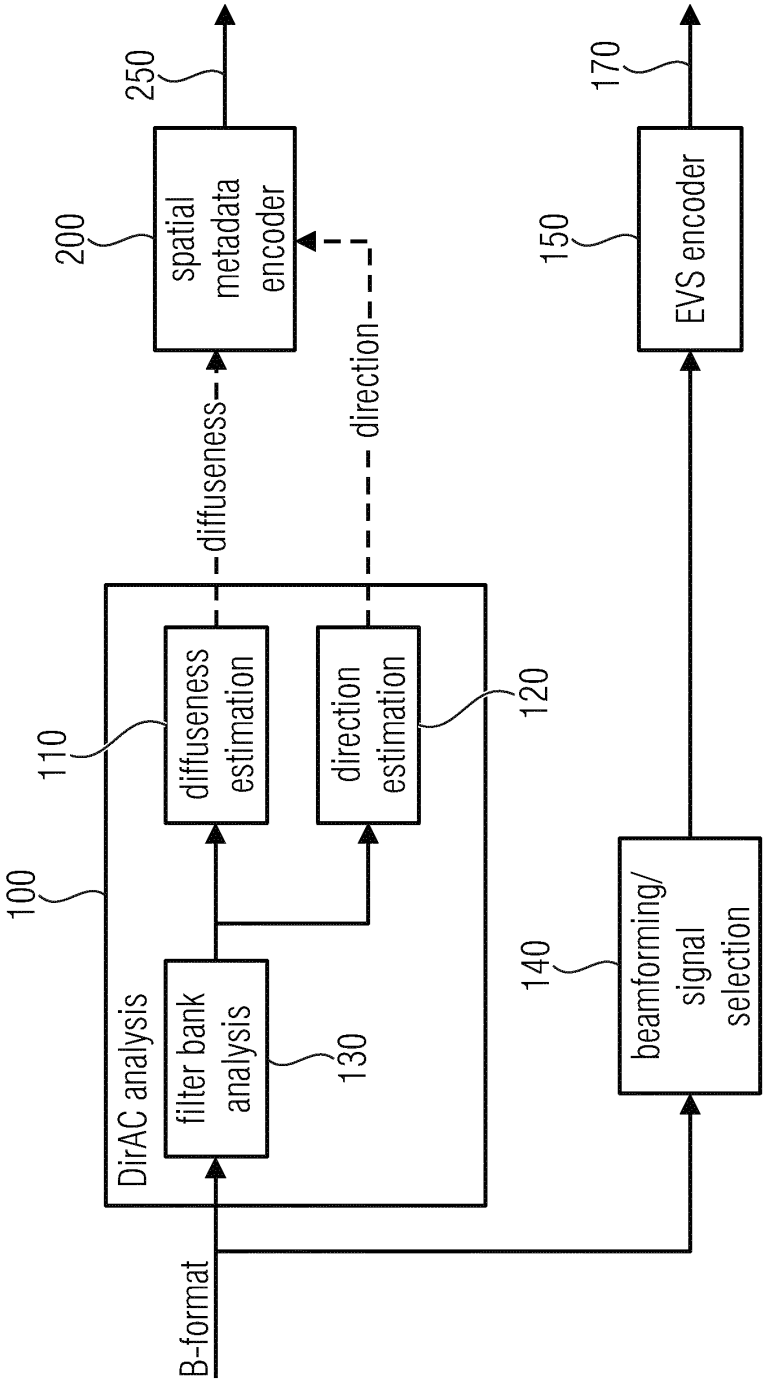

250
spatial metadata encoder
200
diffuseness
direction
170
EVS encoder
150
diffuseness estimation
110
direction estimation
120
100
DirAC analysis
filter bank analysis
130
beamforming/ signal selection
140
B-format

Fig. 1a

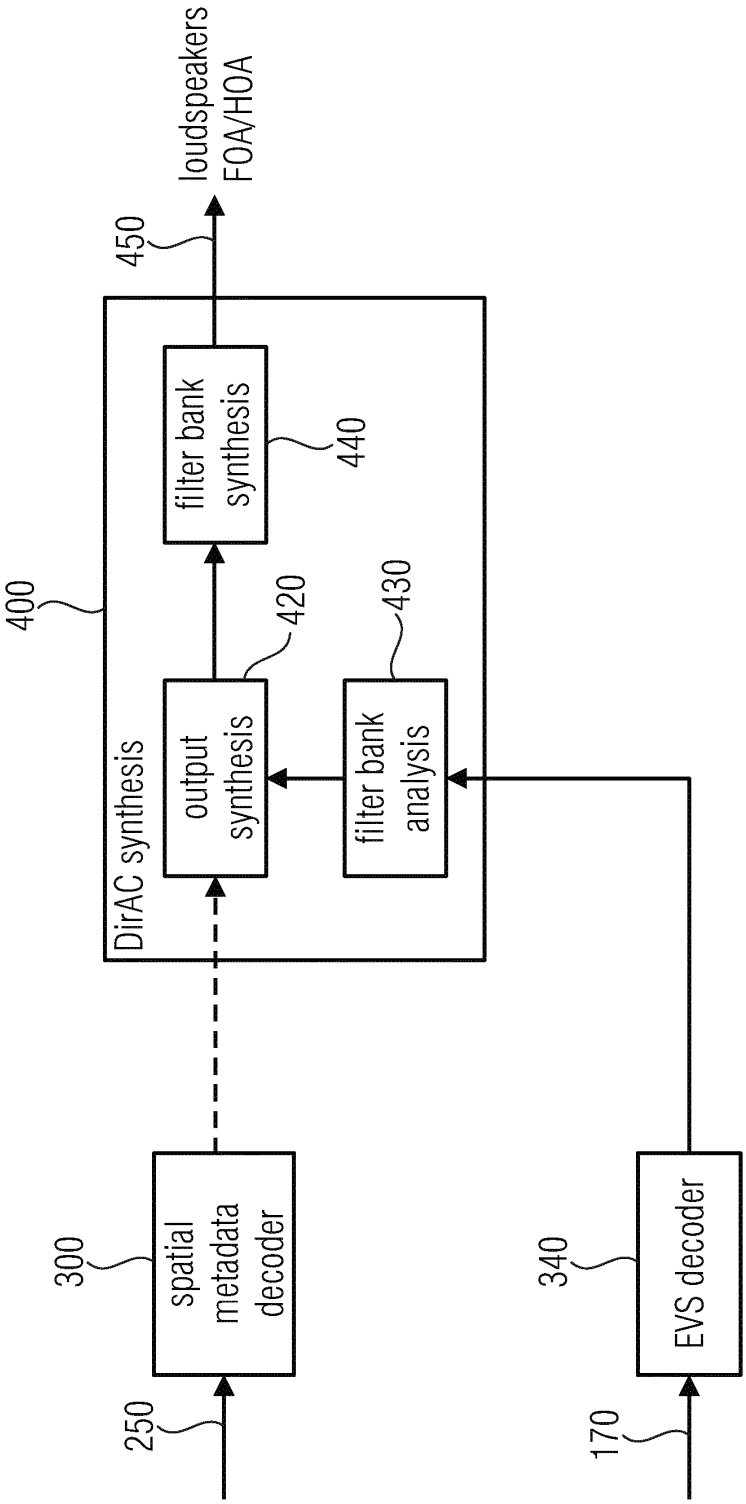

loudspeakers
FOA/HOA
450
filter bank
synthesis
440
400
420
430
DirAC synthesis
output
synthesis
filter bank
analysis
300
spatial
metadata
decoder
340
EVS decoder
250
170

Fig. 1b

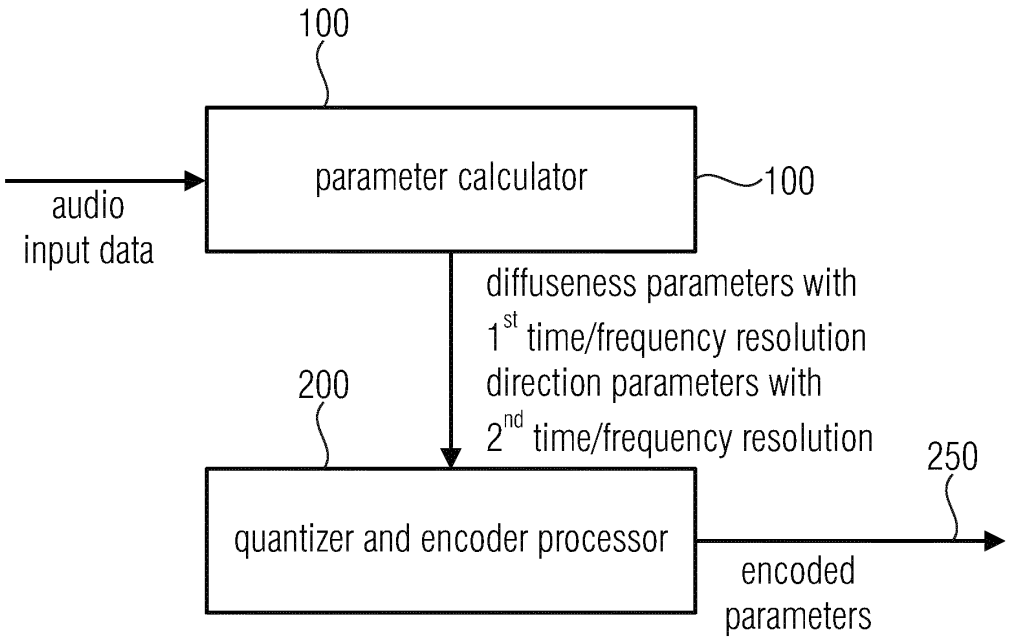

100
parameter calculator
100
audio
input data
diffuseness parameters with
1st time/frequency resolution
direction parameters with
2nd time/frequency resolution
200
250
quantizer and encoder processor
encoded
parameters

Fig. 2a

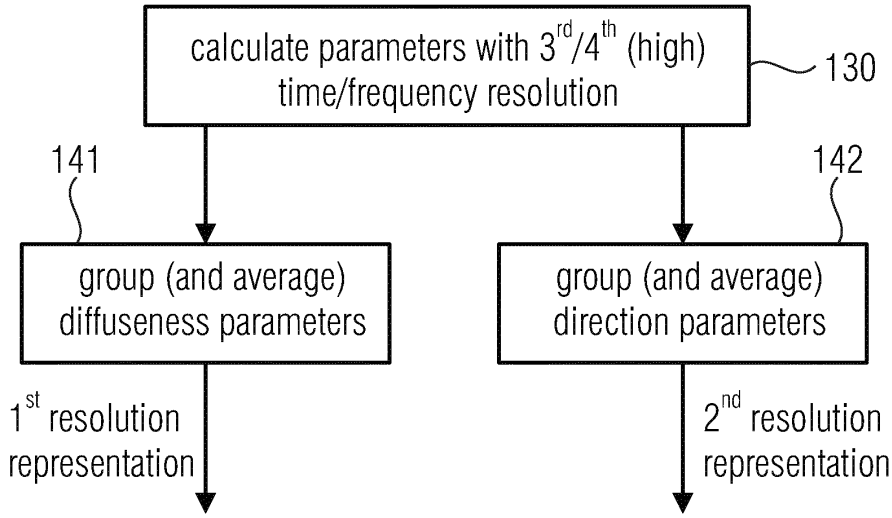

calculate parameters with 3rd/4th (high)
time/frequency resolution
130
141
142
group (and average)
diffuseness parameters
group (and average)
direction parameters
1st resolution
representation
2nd resolution
representation

Fig. 2b

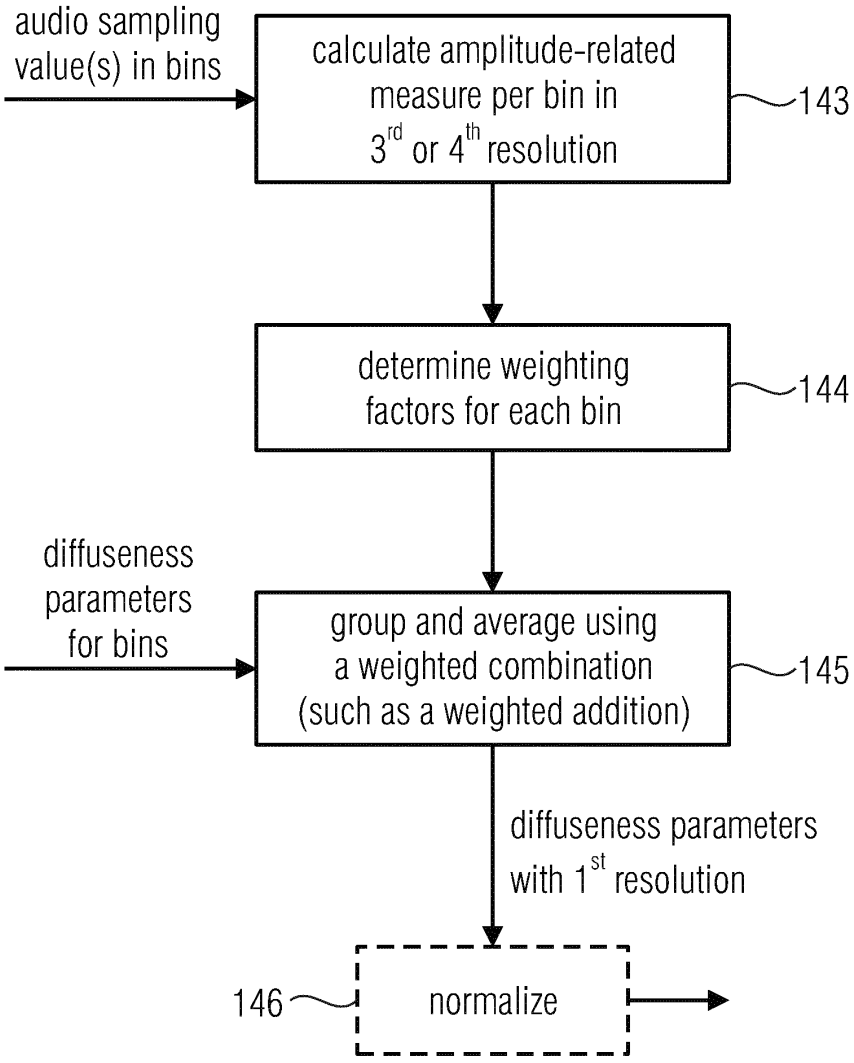

audio sampling value(s) in bins
calculate amplitude-related measure per bin in 3rd or 4th resolution
143
determine weighting factors for each bin
144
diffuseness parameters for bins
group and average using a weighted combination (such as a weighted addition)
145
diffuseness parameters with 1st resolution
146
normalize

Fig. 2c

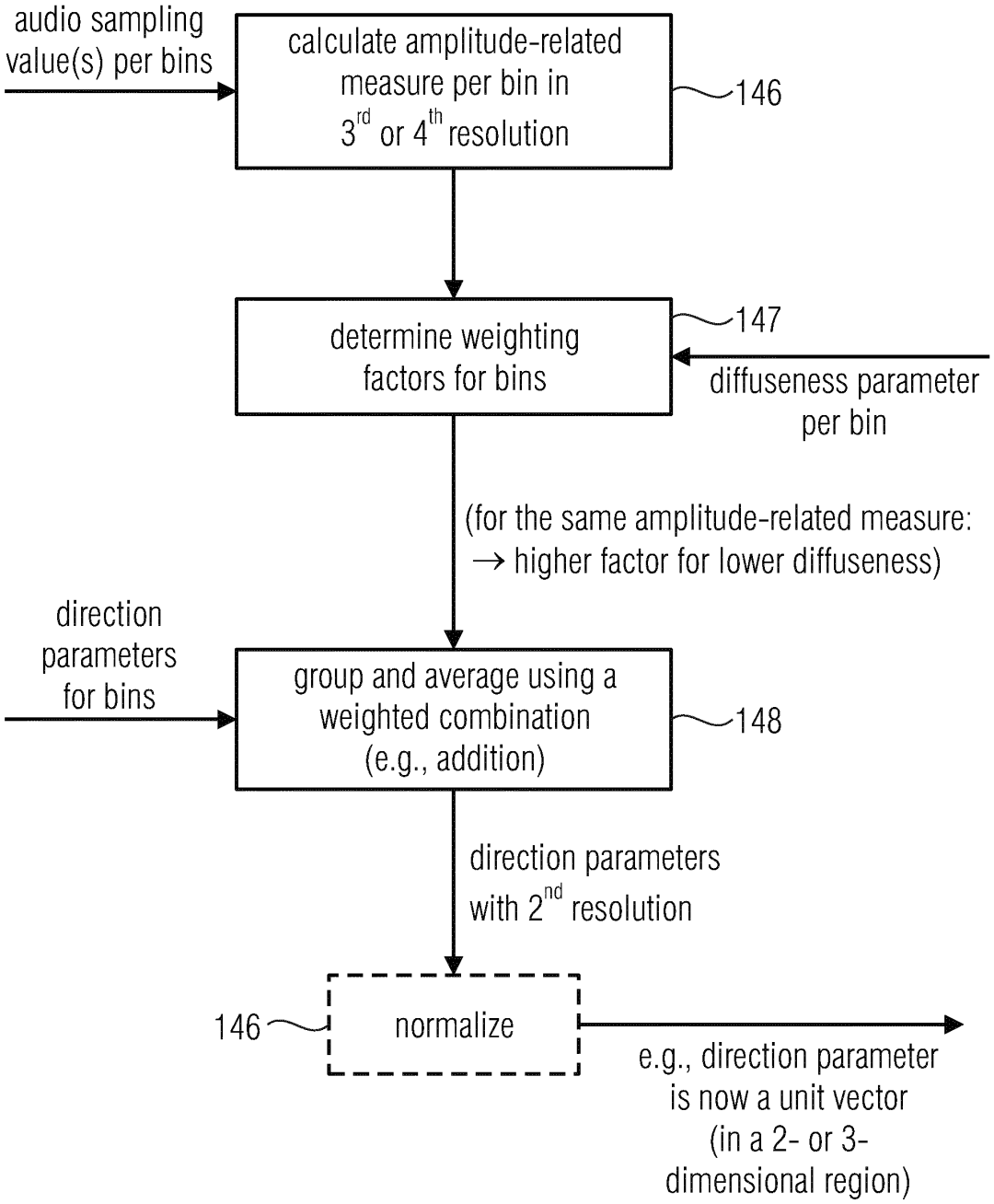

audio sampling value(s) per bins
calculate amplitude-related measure per bin in 3rd or 4th resolution
146
determine weighting factors for bins
147
diffuseness parameter per bin
(for the same amplitude-related measure: → higher factor for lower diffuseness)
direction parameters for bins
group and average using a weighted combination (e.g., addition)
148
direction parameters with 2nd resolution
146
normalize
e.g., direction parameter is now a unit vector (in a 2- or 3-dimensional region)

Fig. 2d

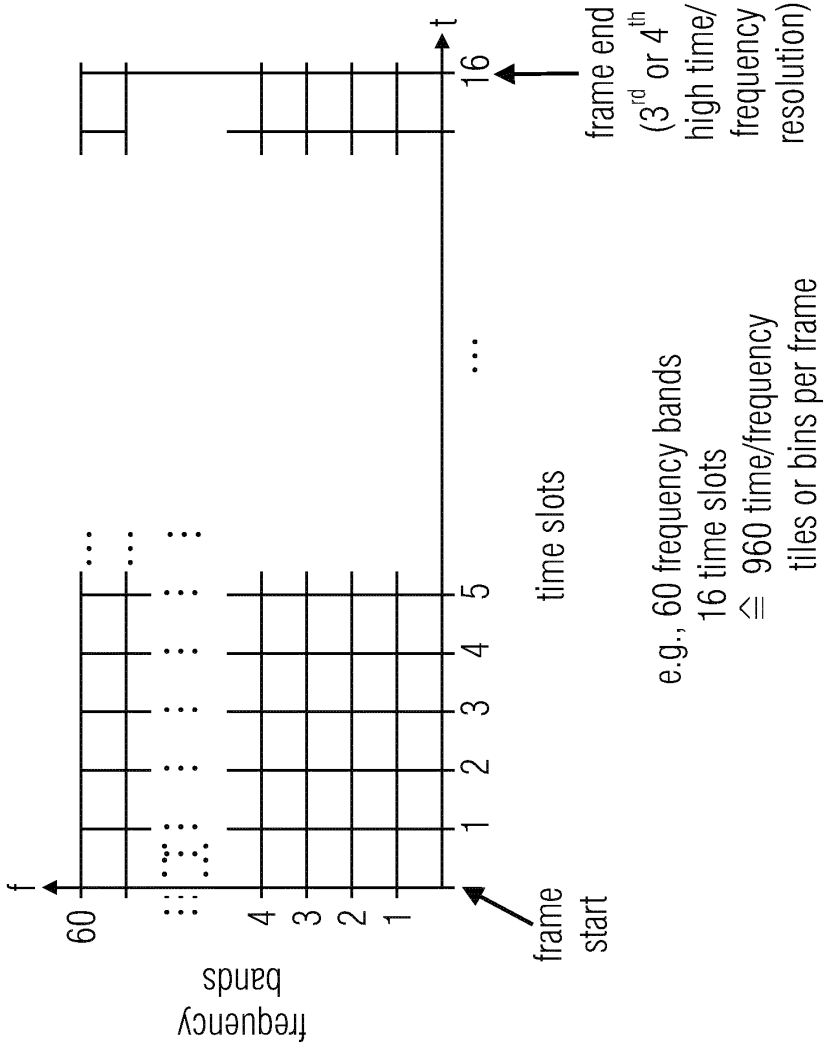
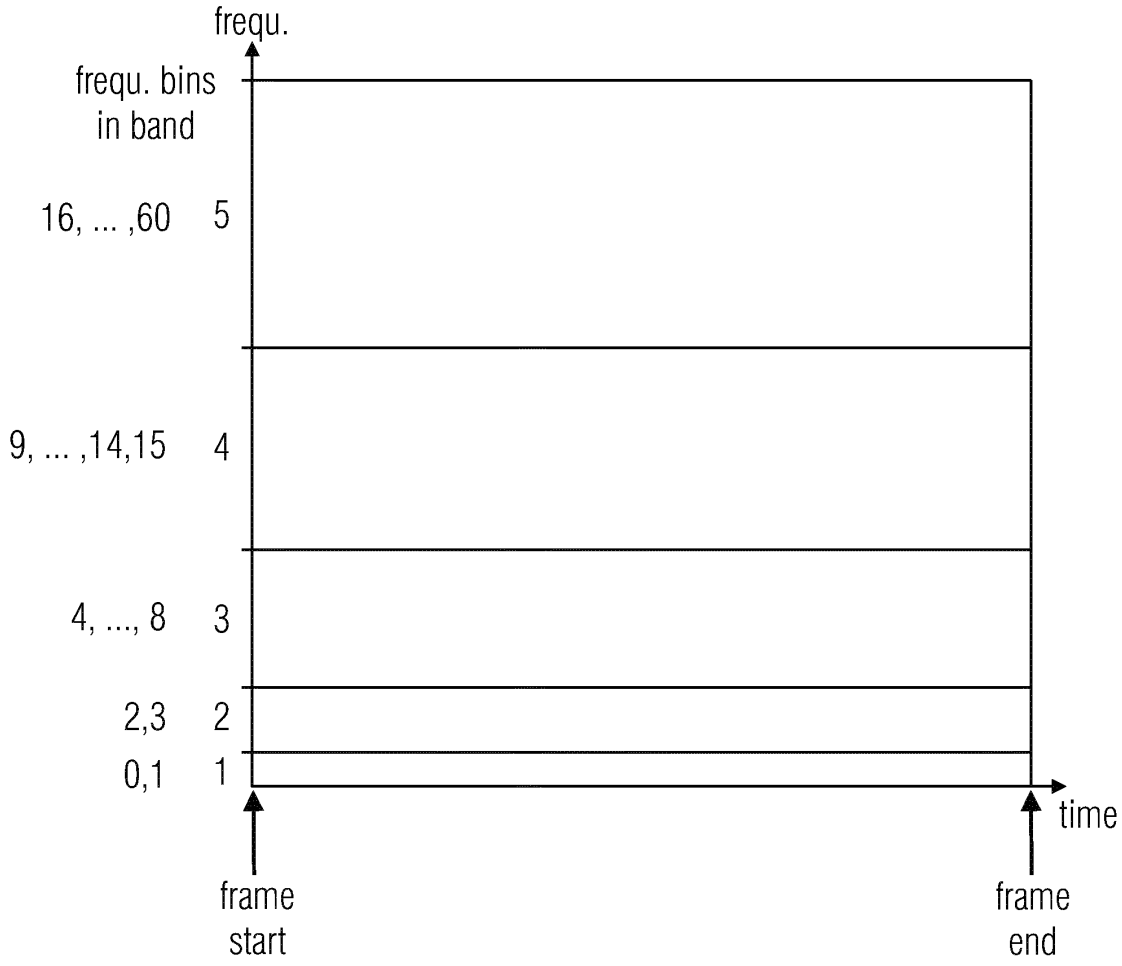


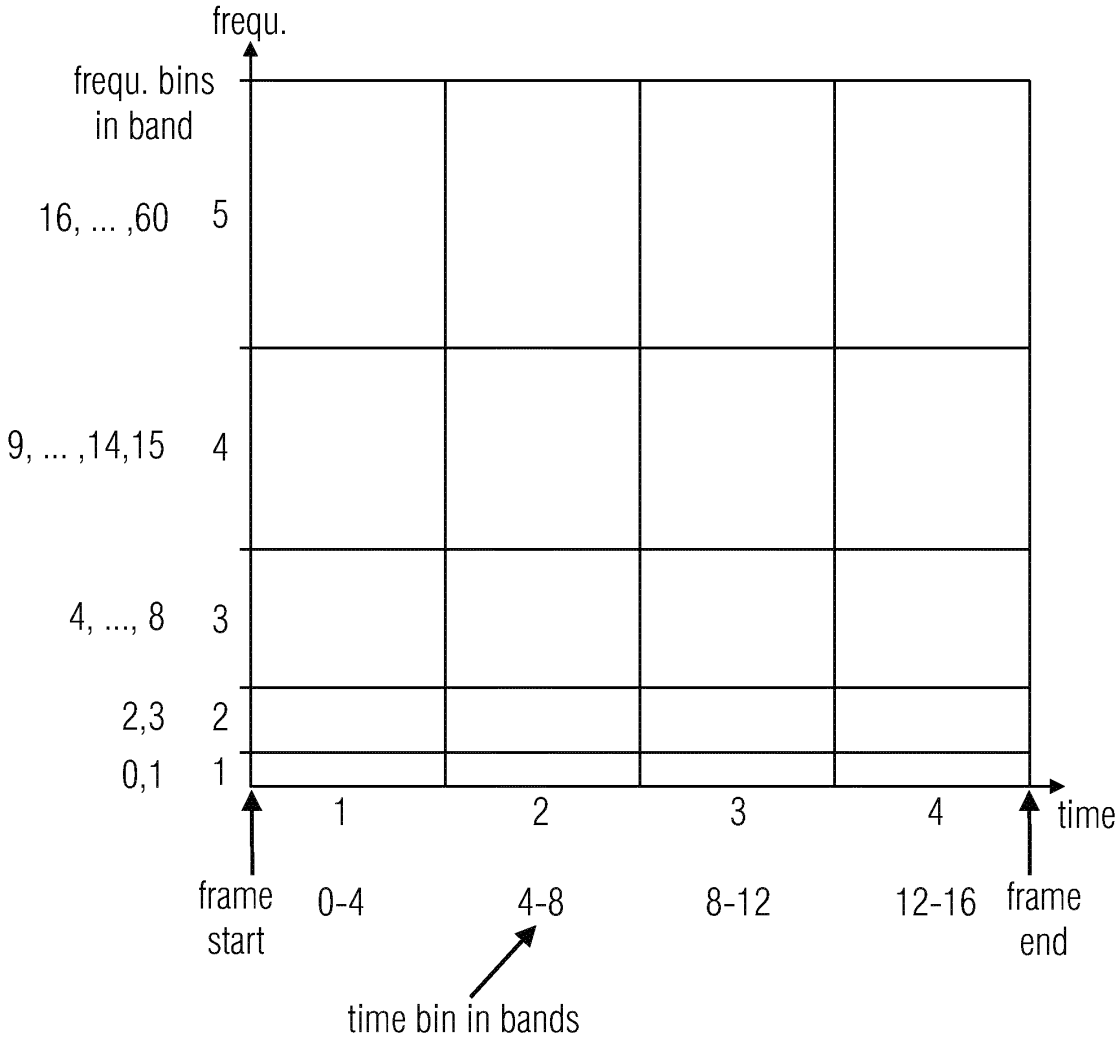
Fig. 3a
(high resolution)



diffuseness grouping
(1^{st} time/frequency resolution)

e.g., 1 time slot
5 bands (bandwidth increases from low high frequency)
$\hat{=}$ 5 grouped bins per frame

Fig. 3b
(low resolution)

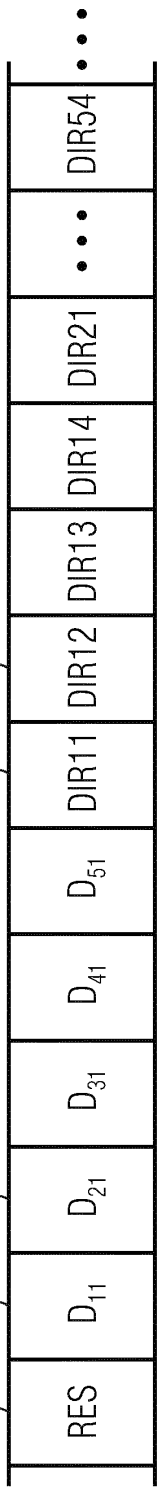


direction parameter grouping
(2nd time/frequency resolution)

e.g., 4 time slots (each having four time bins)
5 bands (bandwidth increases from low to high)

≙ 20 grouped bins per frame

Fig. 3c
(medium resolution)



D_{11} : encoded diffuseness parameter for 1st band and 1st time slot

D_{21} : encoded diffuseness parameter for 2nd band and 1st time slot

• • •

DIR11 : encoded dir. parameter for for 1st band and 1st time slot

• • •

DIR54 : encoded dir. parameter for 5th band and 4th time slot

RES : resolution indication for 1st and/or 2nd resolution (not necessary, when fixedly set by encoder and decoder)

Fig. 3d

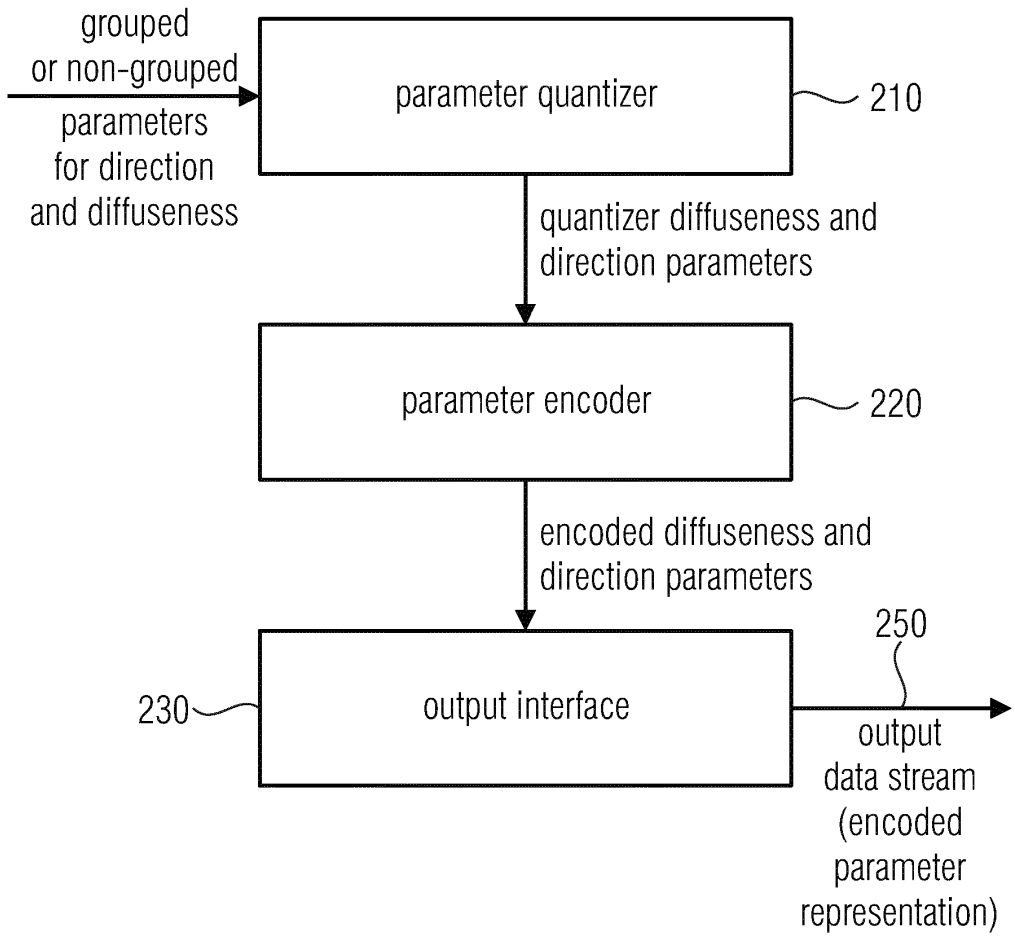

grouped
or non-grouped
parameters
for direction
and diffuseness
parameter quantizer
210
quantizer diffuseness and
direction parameters
parameter encoder
220
encoded diffuseness and
direction parameters
250
230
output interface
output
data stream
(encoded
parameter
representation)

Fig. 4a

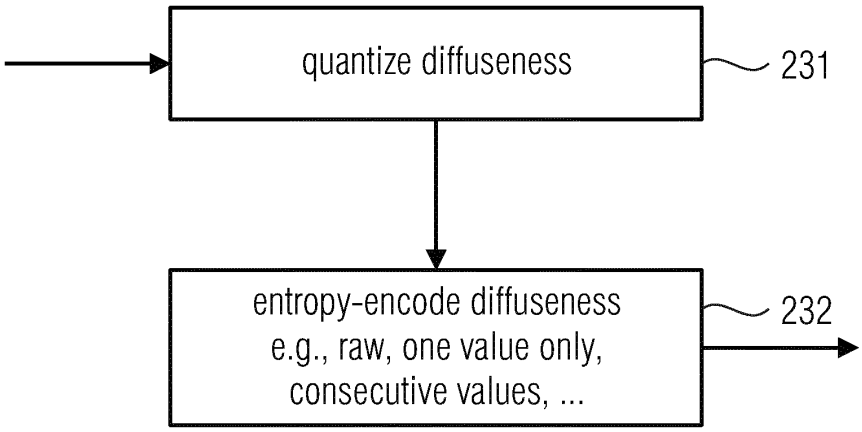

quantize diffuseness
231
entropy-encode diffuseness
e.g., raw, one value only,
consecutive values, ...
232

Fig. 4b

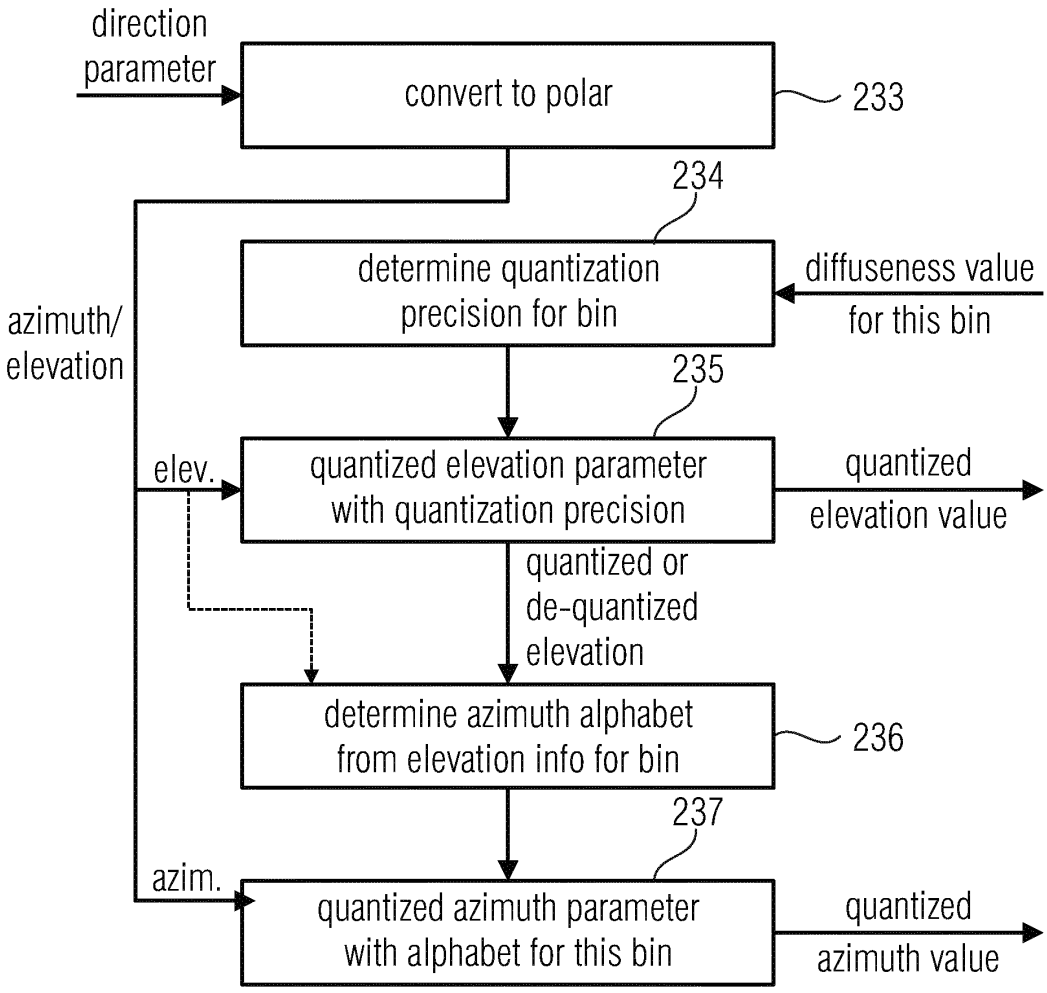

direction
parameter
convert to polar
233
234
determine quantization
precision for bin
diffuseness value
for this bin
azimuth/
elevation
235
elev.
quantized elevation parameter
with quantization precision
quantized
elevation value
quantized or
de-quantized
elevation
determine azimuth alphabet
from elevation info for bin
236
237
azim.
quantized azimuth parameter
with alphabet for this bin
quantized
azimuth value

Fig. 4c

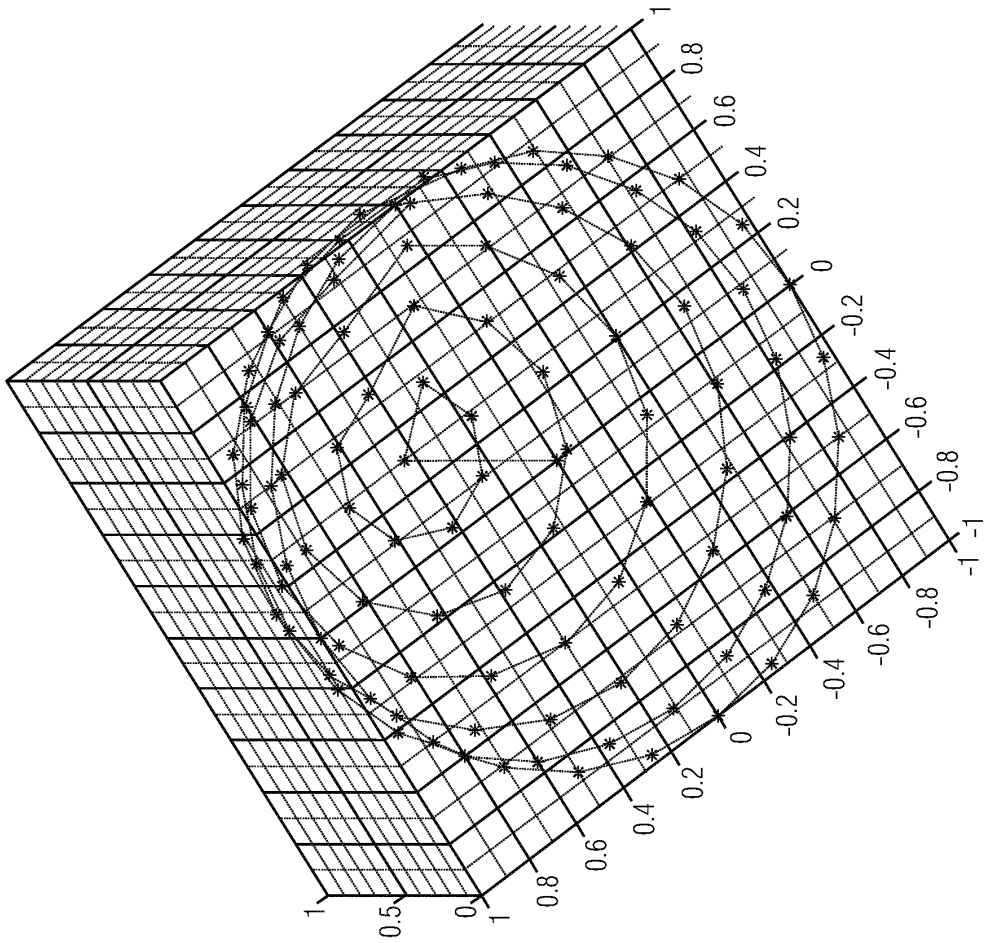


Fig. 4d
quasi-uniform coverage of the unit sphere

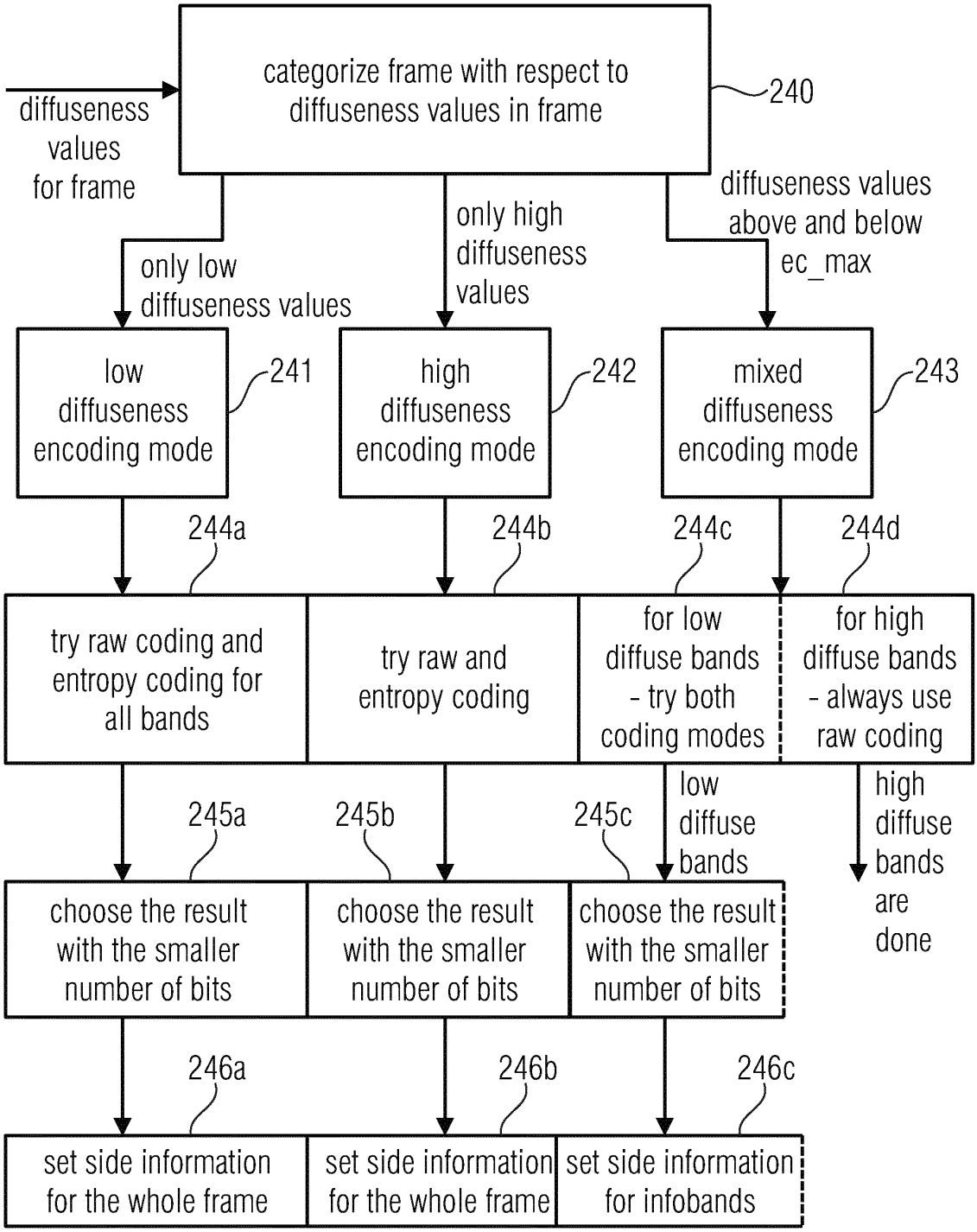

diffuseness values for frame
categorize frame with respect to diffuseness values in frame
240
only low diffuseness values
only high diffuseness values
diffuseness values above and below ec_max
low diffuseness encoding mode
241
high diffuseness encoding mode
242
mixed diffuseness encoding mode
243
244a
244b
244c
244d
try raw coding and entropy coding for all bands
try raw and entropy coding
for low diffuse bands - try both coding modes
for high diffuse bands - always use raw coding
245a
245b
245c
low diffuse bands
high diffuse bands are done
choose the result with the smaller number of bits
choose the result with the smaller number of bits
choose the result with the smaller number of bits
246a
246b
246c
set side information for the whole frame
set side information for the whole frame
set side information for infobands

Fig. 5a

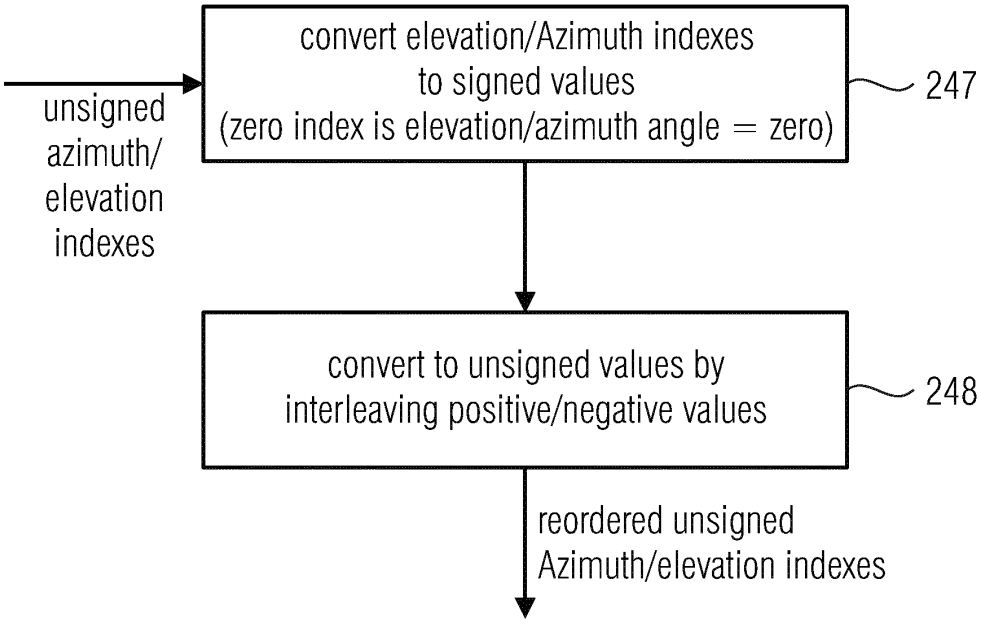


Fig. 5b
(optional pre-processing
of direction indexes for
both modes)

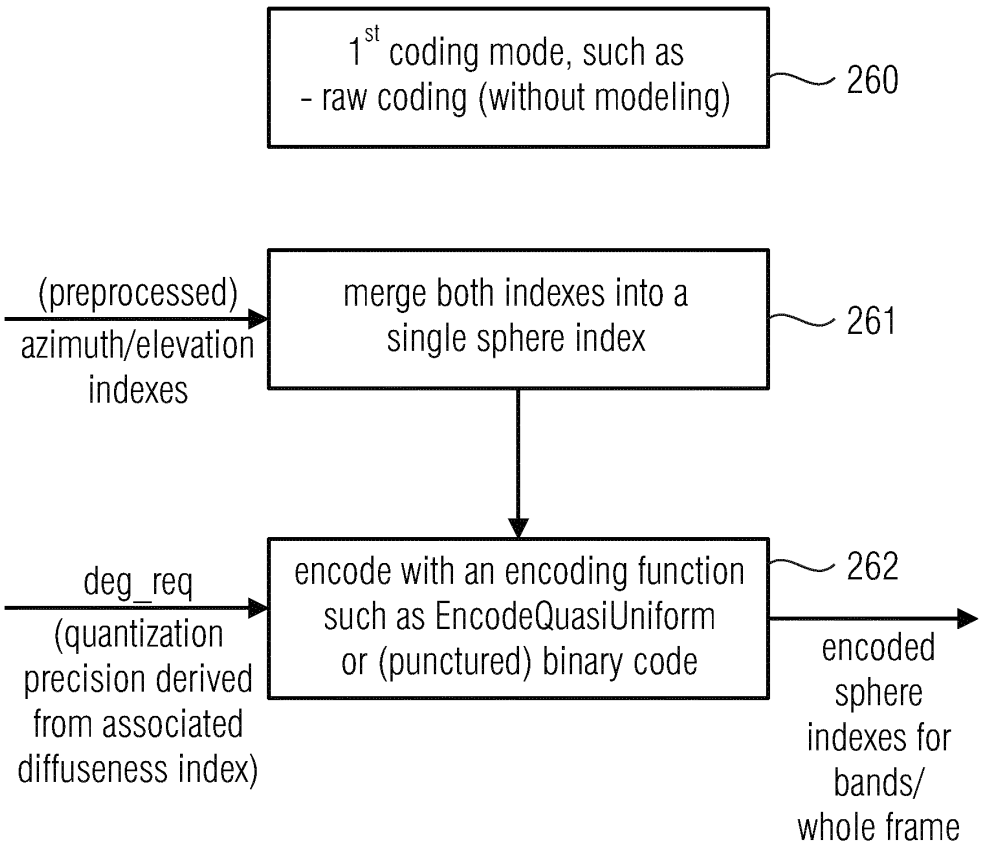

1st coding mode, such as
- raw coding (without modeling)
260
(preprocessed)
azimuth/elevation
indexes
merge both indexes into a
single sphere index
261
deg_req
(quantization
precision derived
from associated
diffuseness index)
encode with an encoding function
such as EncodeQuasiUniform
or (punctured) binary code
262
encoded
sphere
indexes for
bands/
whole frame

Fig. 5c

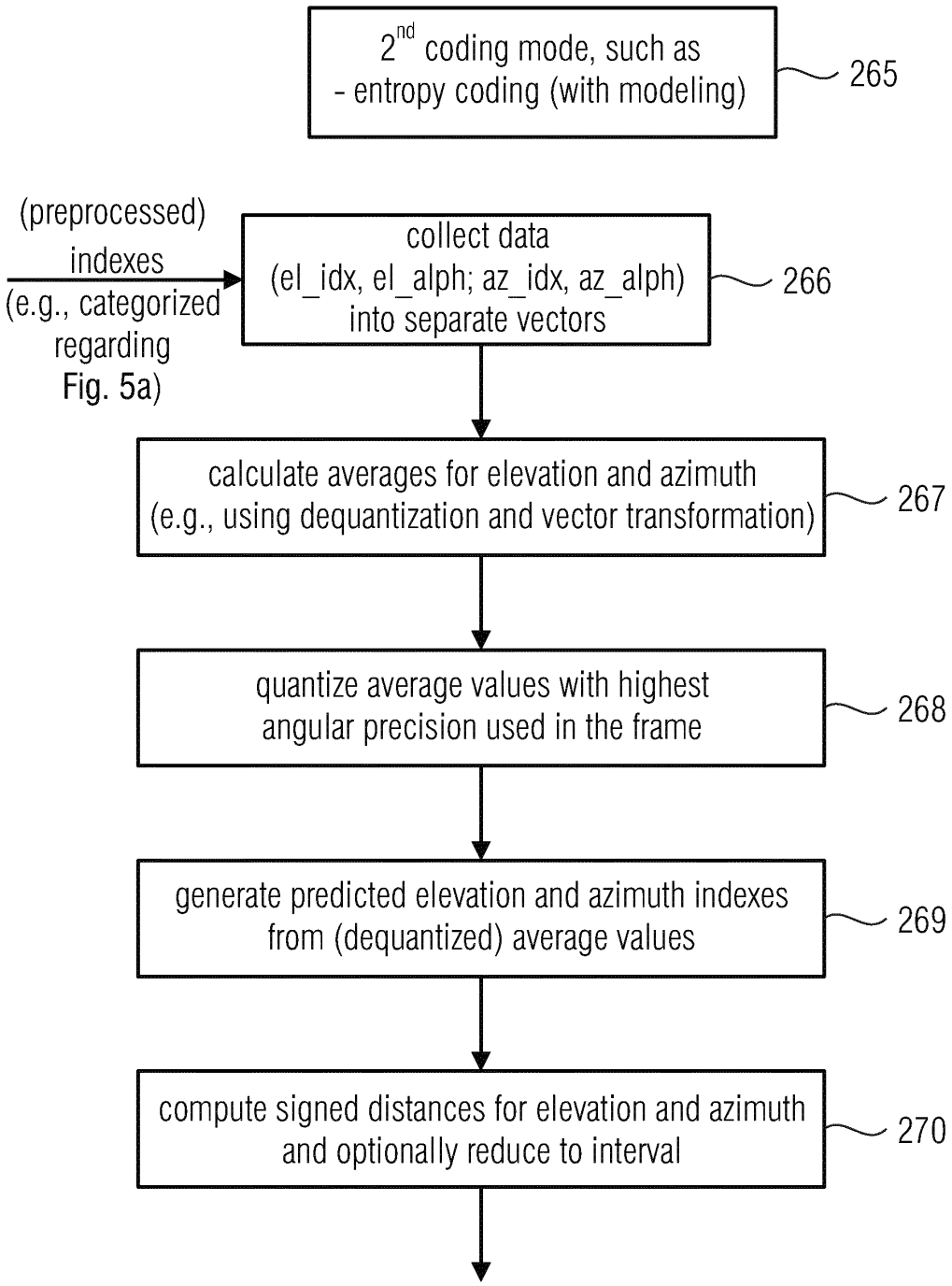


Fig. 5d
(modeling with projection for prediction)

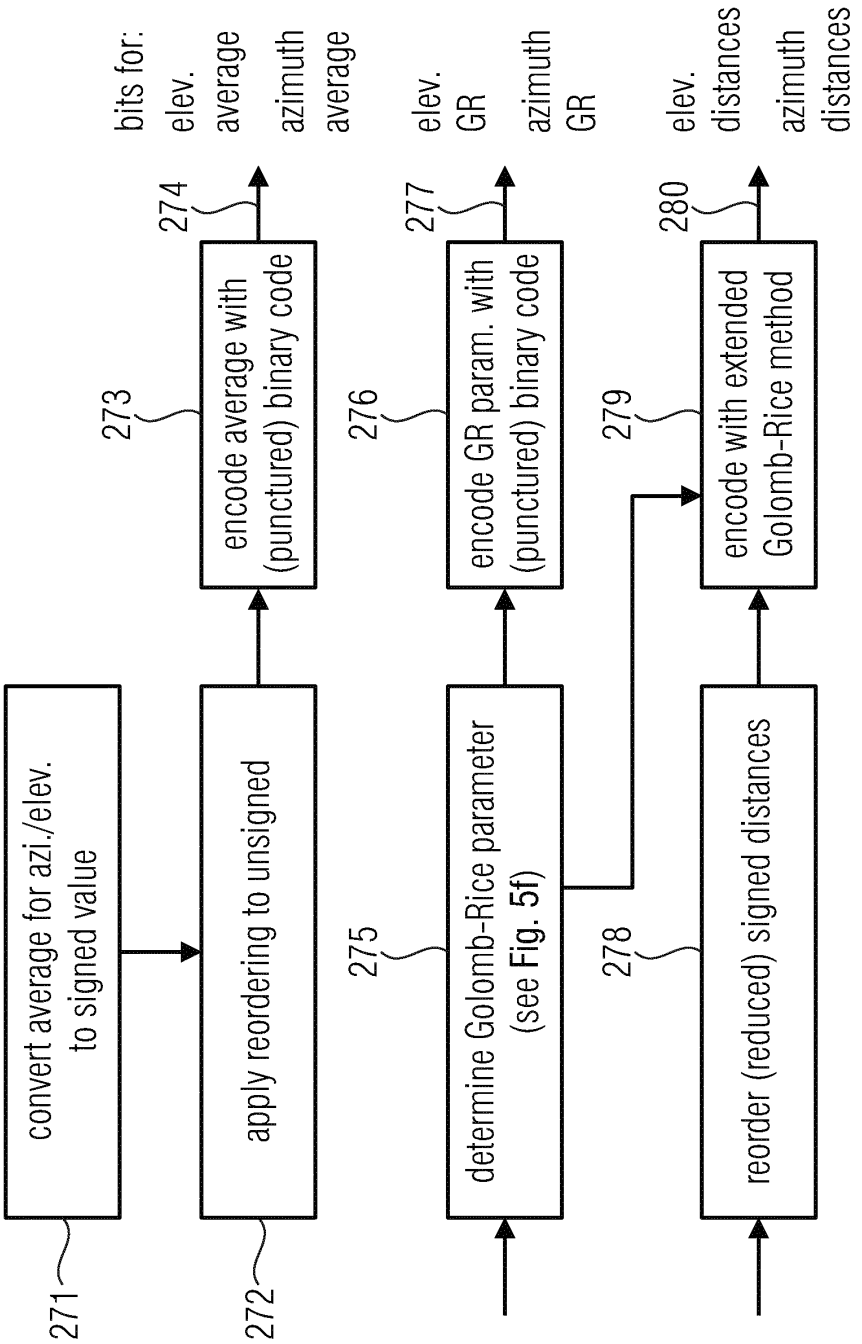


Fig. 5e
(encoding, i.e., conversion into bits for transmission storage)

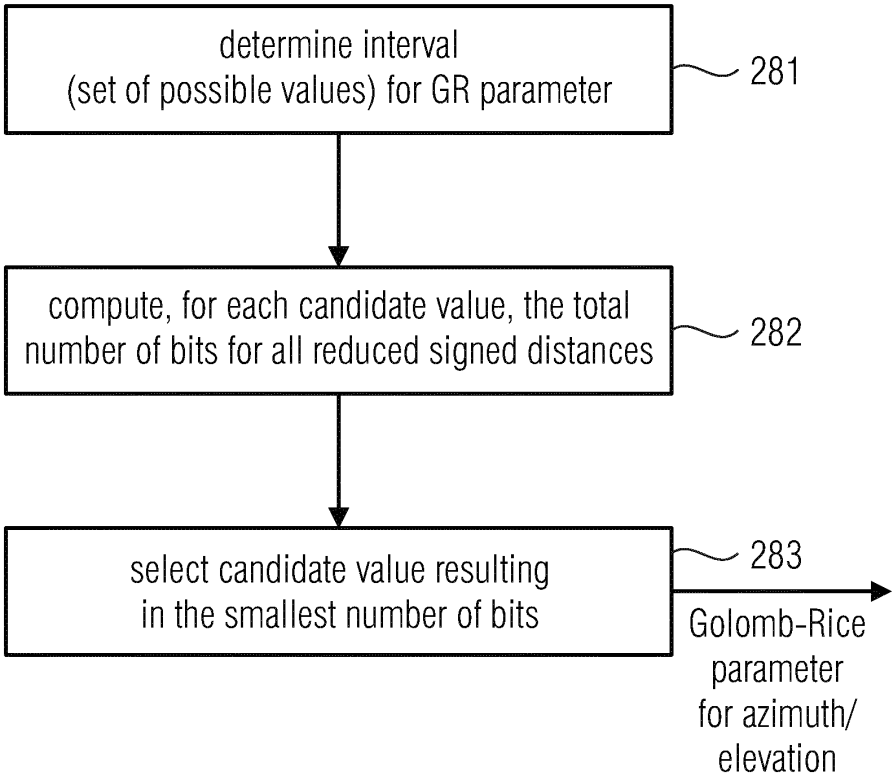

determine interval
(set of possible values) for GR parameter
281
compute, for each candidate value, the total
number of bits for all reduced signed distances
282
select candidate value resulting
in the smallest number of bits
283
Golomb-Rice
parameter
for azimuth/
elevation

Fig. 5f

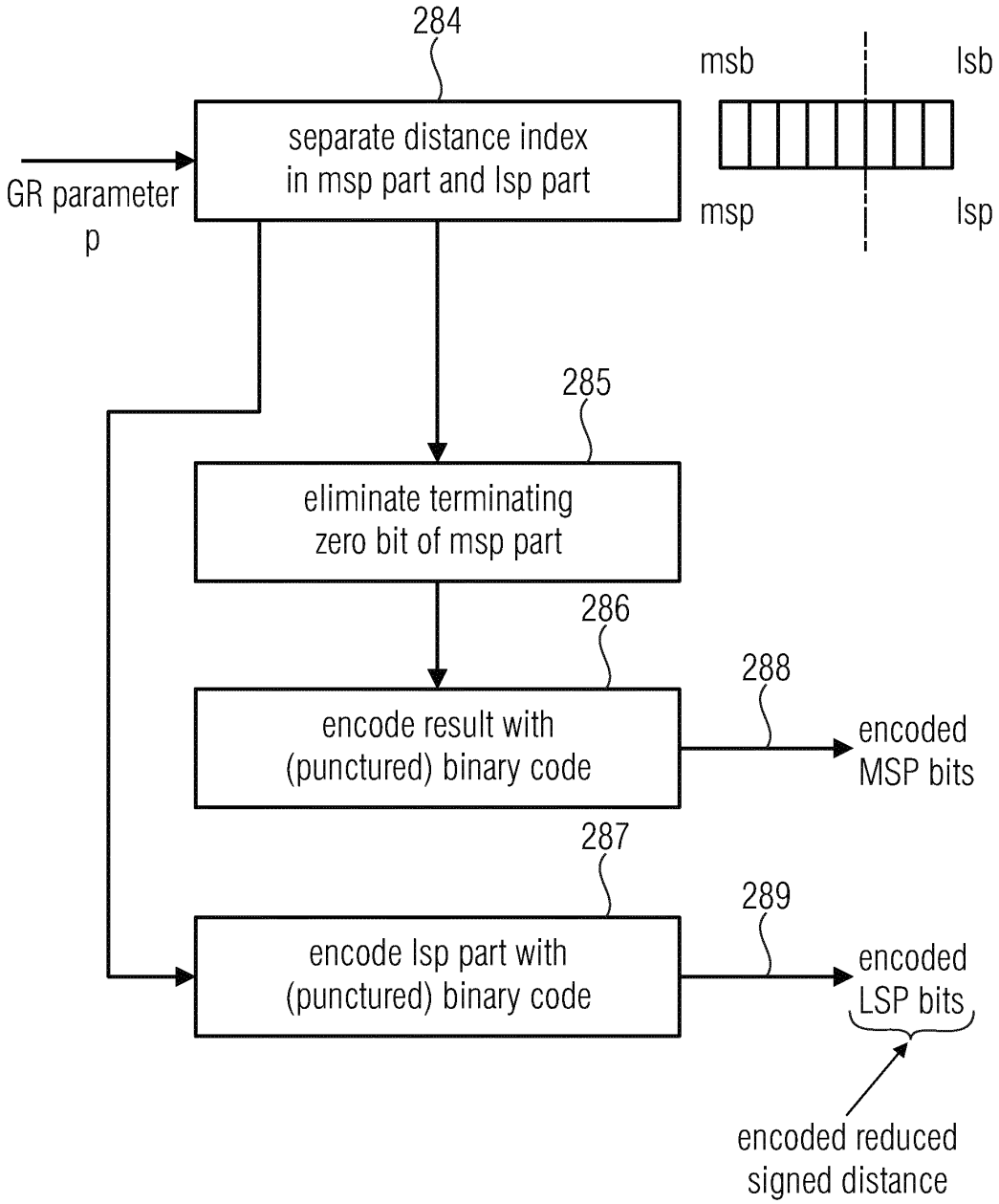

284
msb
lsb
separate distance index
in msp part and lsp part
GR parameter
p
msp
lsp
285
eliminate terminating
zero bit of msp part
286
288
encode result with
(punctured) binary code
encoded
MSP bits
287
289
encode lsp part with
(punctured) binary code
encoded
LSP bits
encoded reduced
signed distance

Fig. 5g

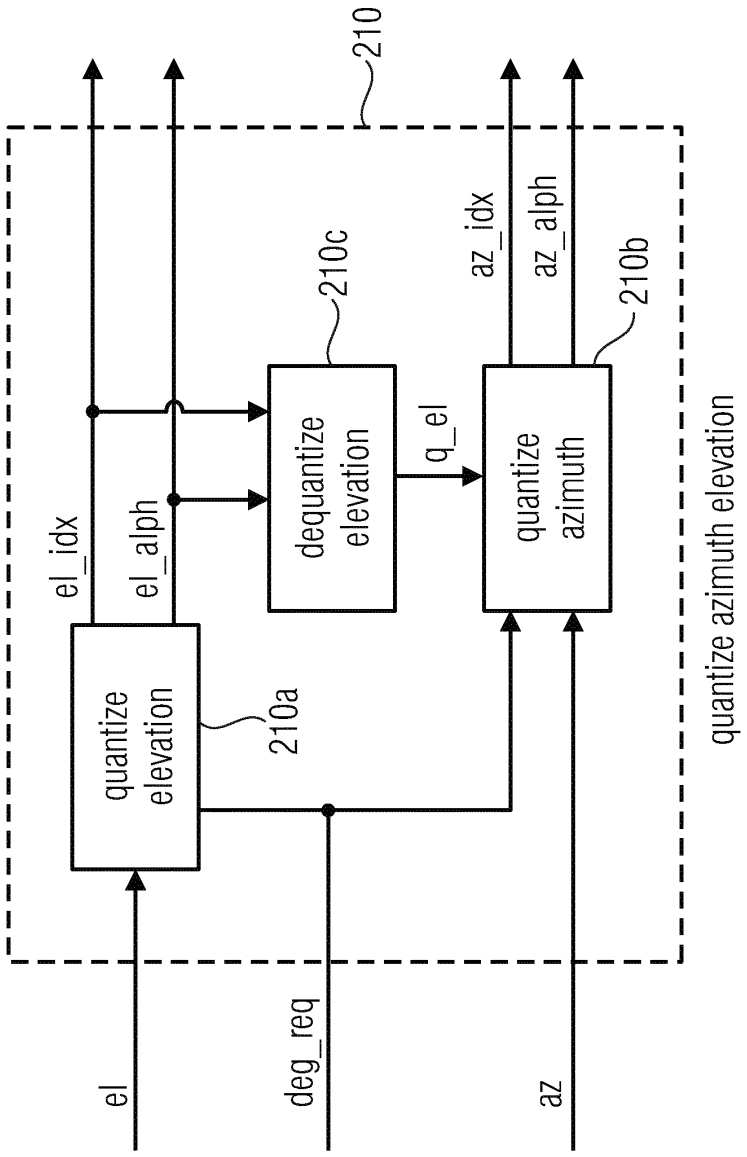


Fig. 6a
Encoder

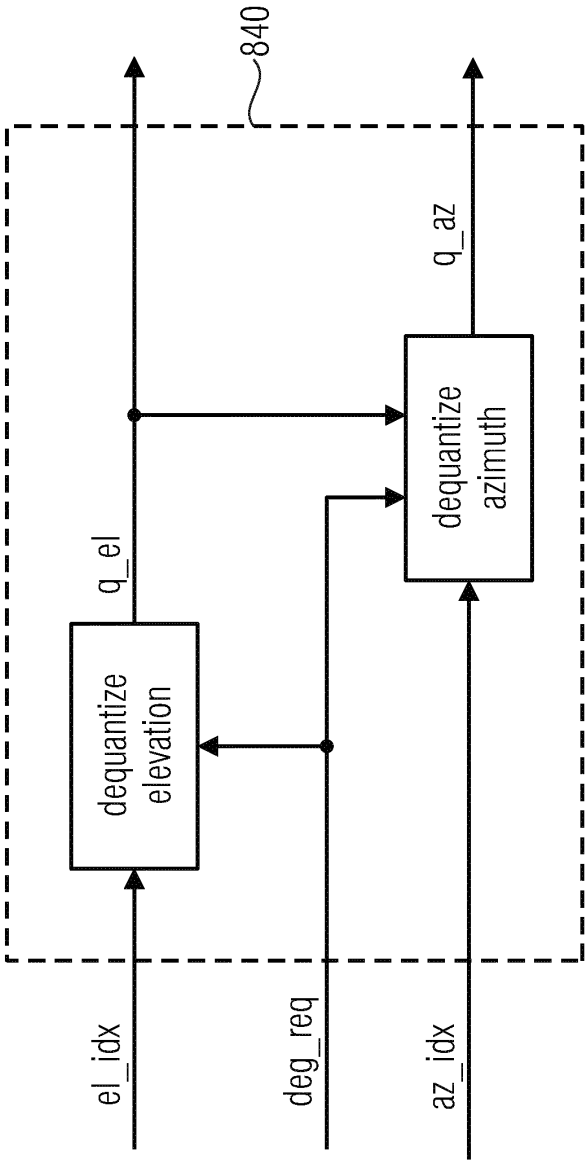


Fig. 6b
Encoder/Decoder

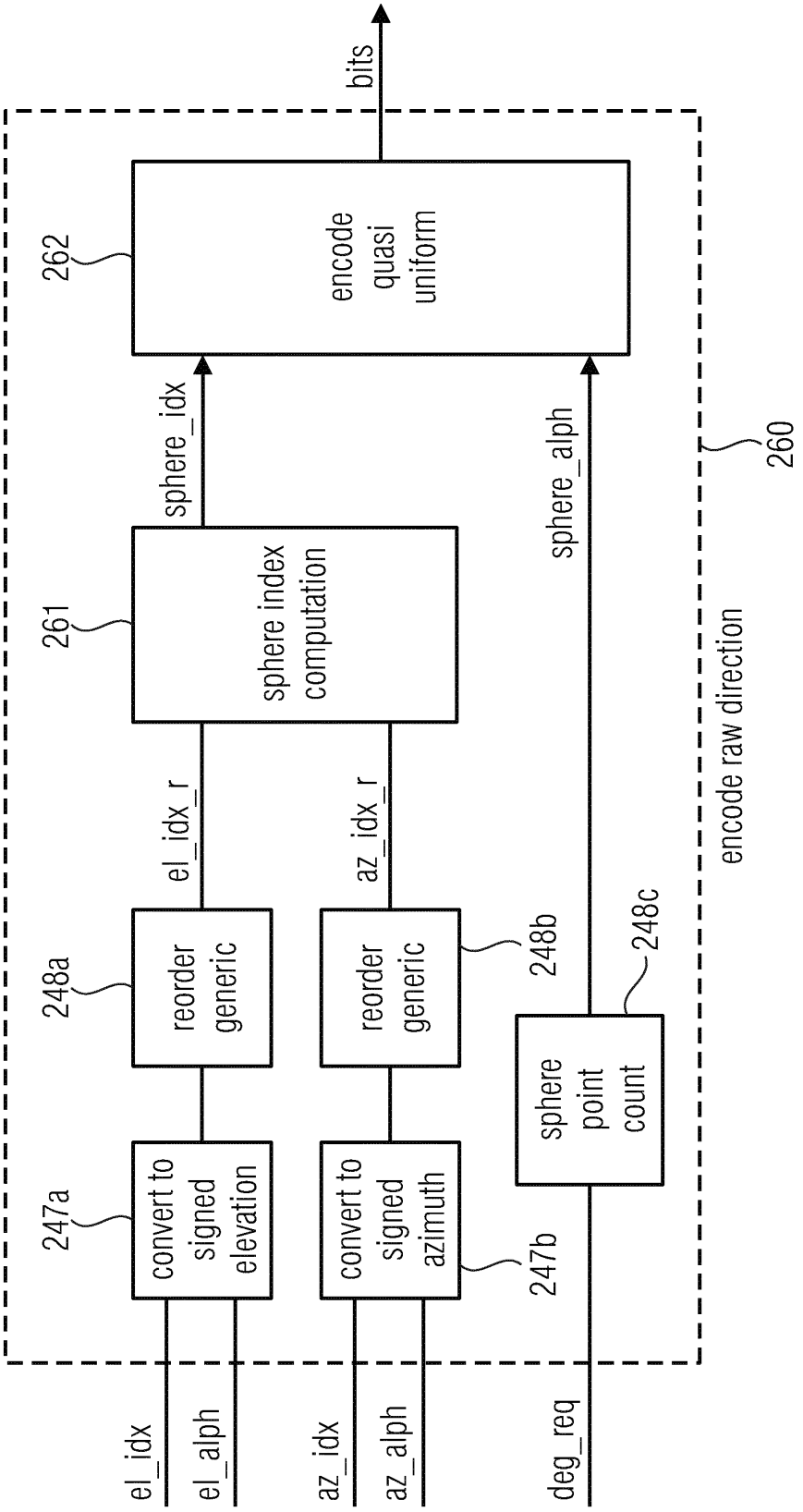


Fig. 6c
Encoder

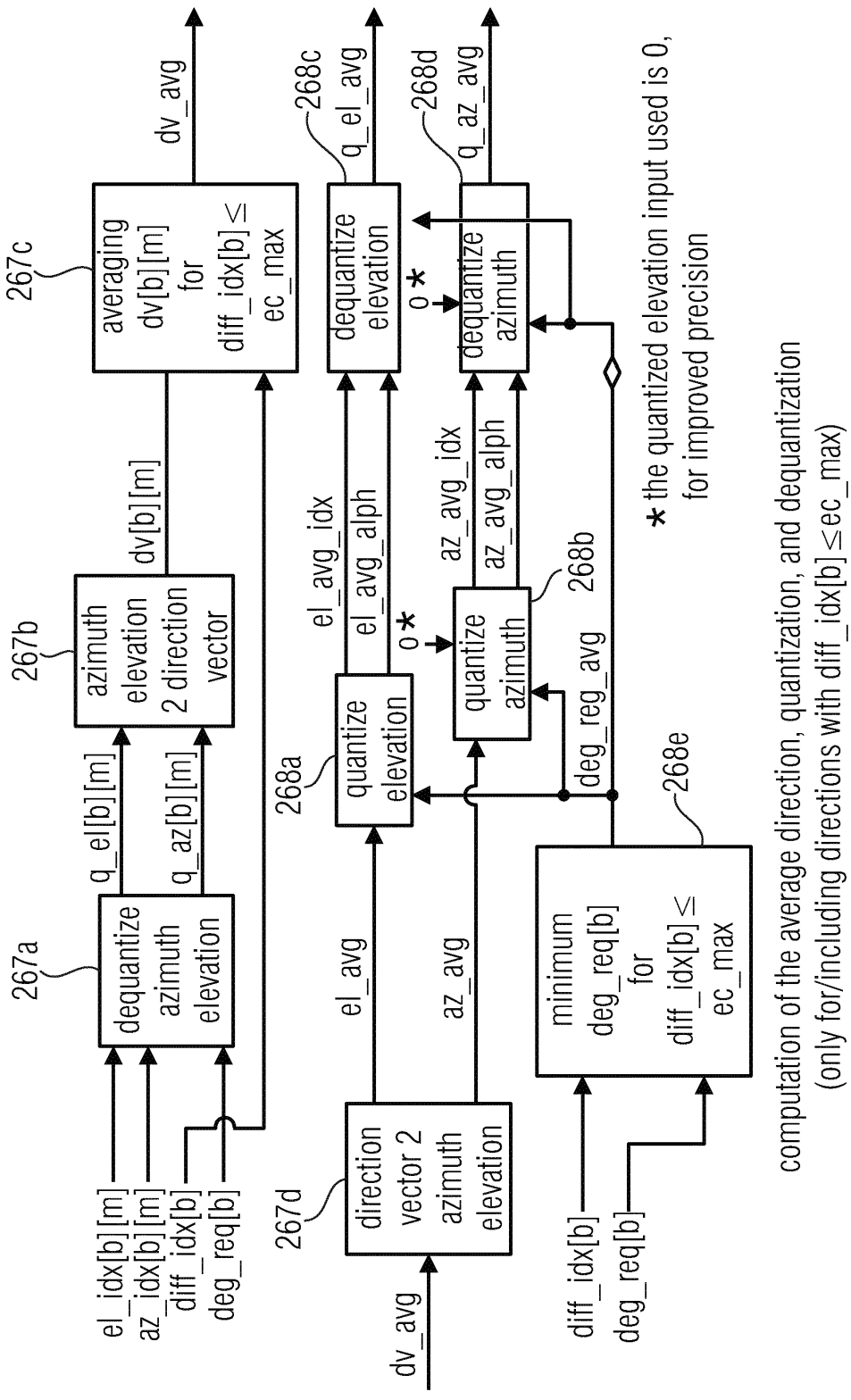


Fig. 6d
Encoder

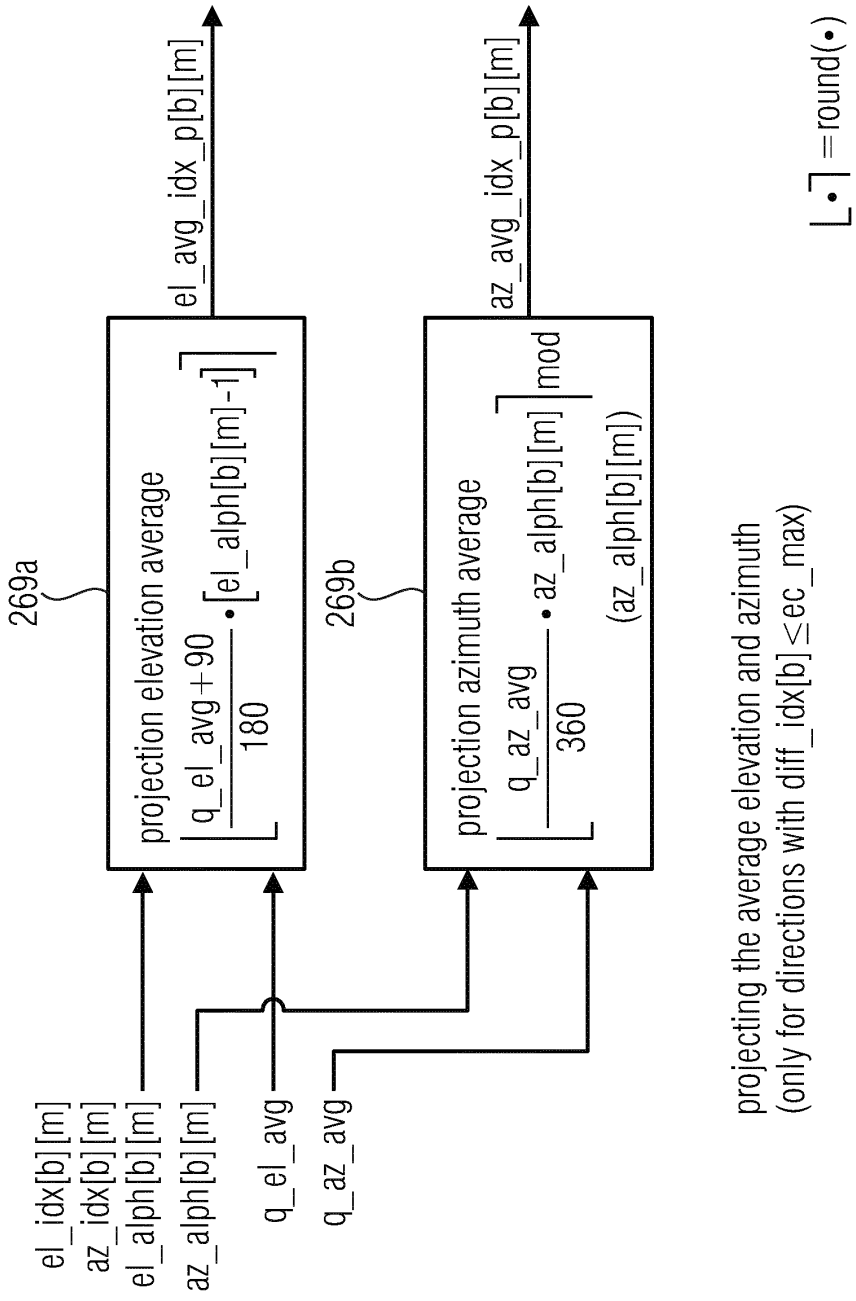
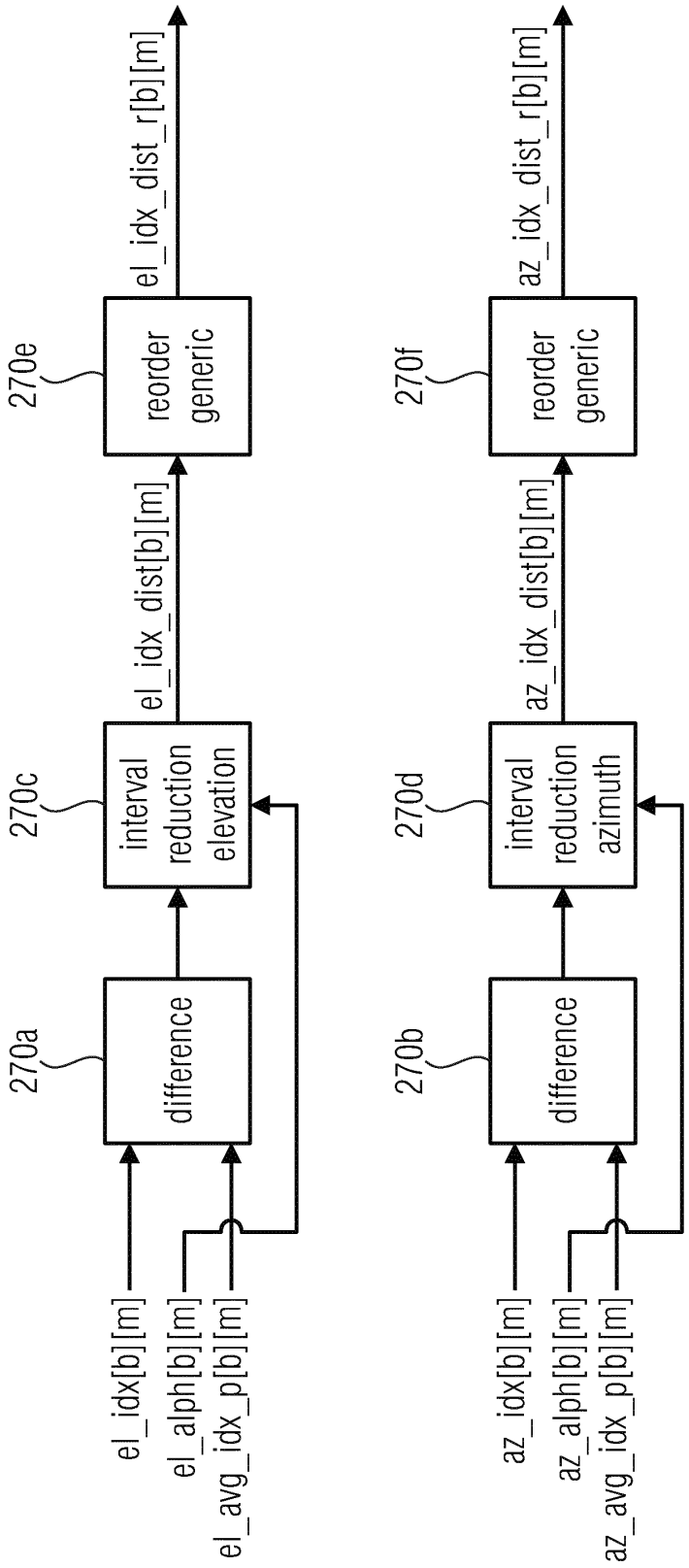


Fig. 6e
Encoder



computing the values which will be encoded with extended Golomb-Rice, separately for elevation and azimuth (only for directions with diff_idx[b] ≤ ec_max)

Extended Golomb-Rice: see e.g., Figs. 5e to 5g

Fig. 6f
Encoder

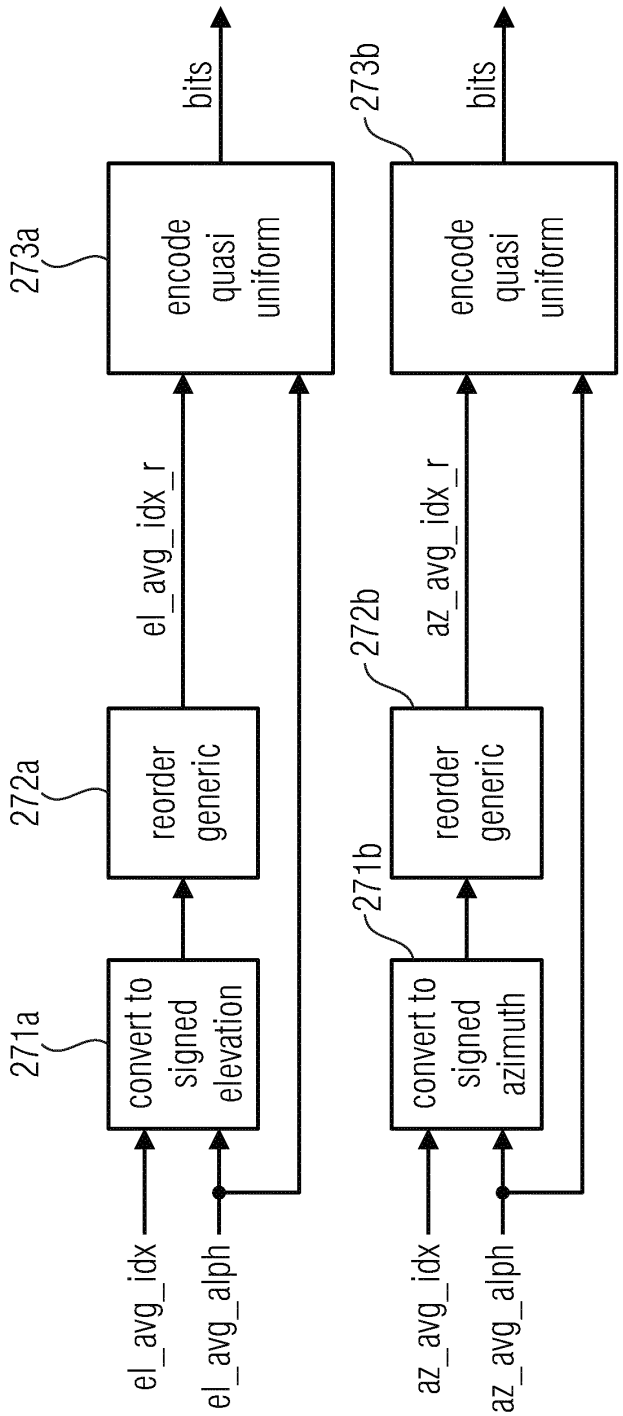


Fig. 6g
Encoder

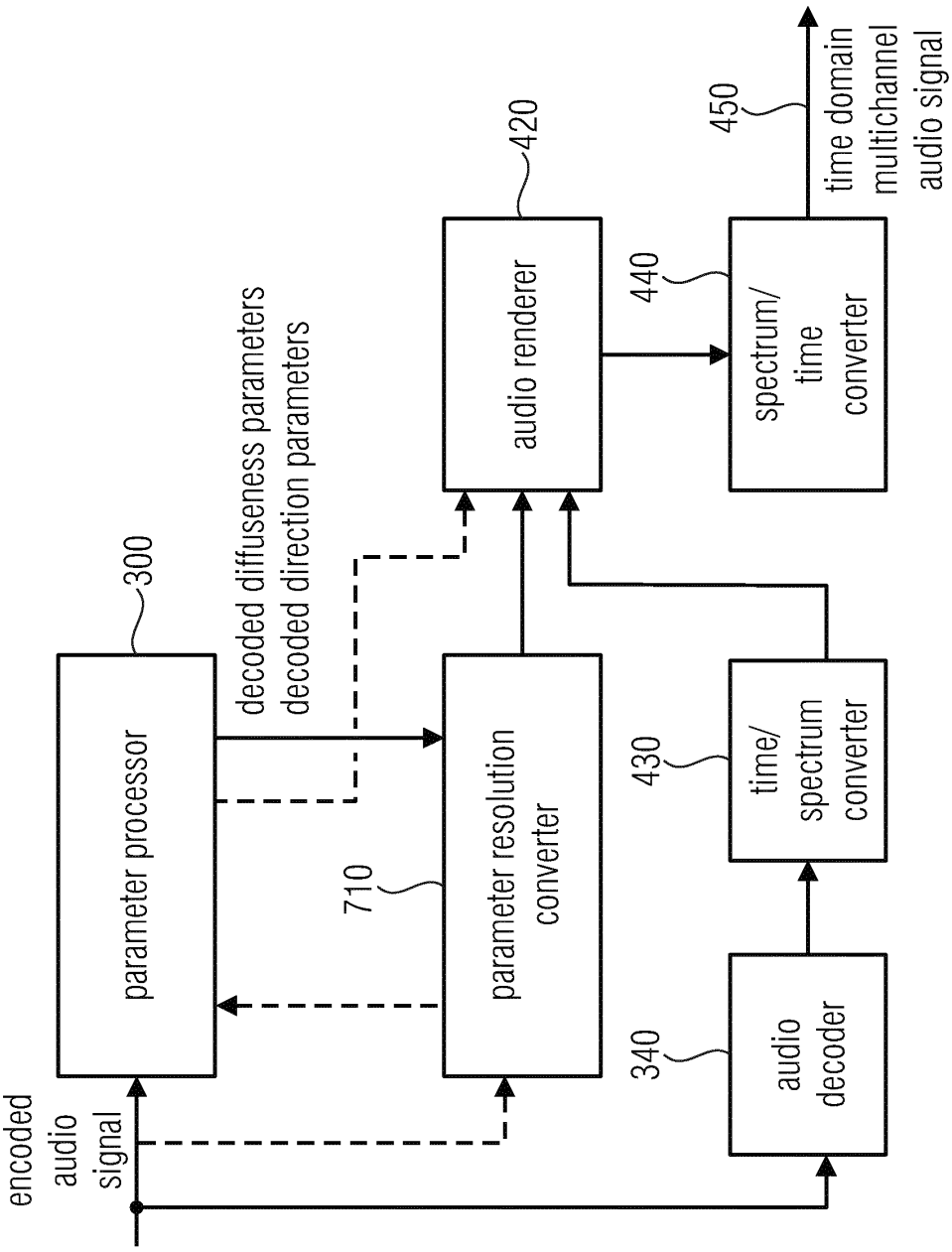

encoded audio signal
parameter processor
300
decoded diffuseness parameters
decoded direction parameters
parameter resolution converter
710
audio renderer
420
audio decoder
340
time/ spectrum converter
430
spectrum/ time converter
440
450
time domain multichannel audio signal

Fig. 7a

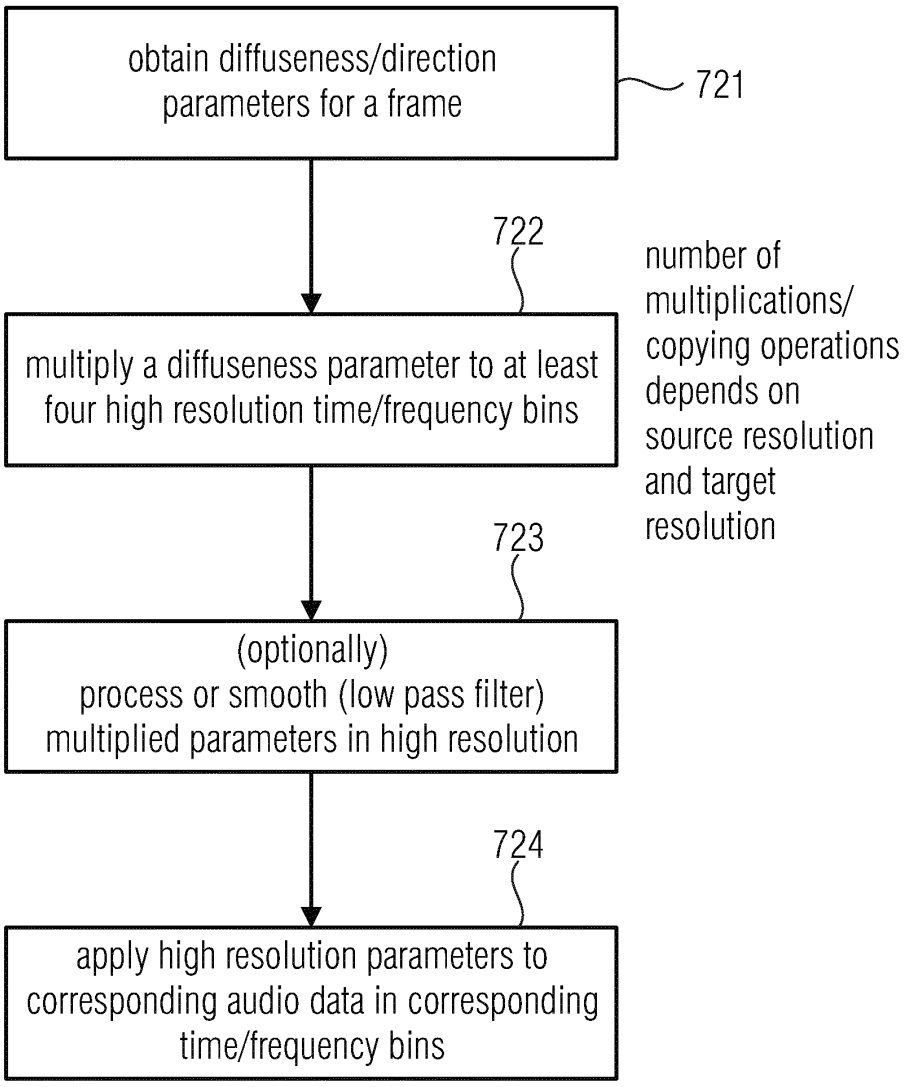

obtain diffuseness/direction parameters for a frame
721
722
multiply a diffuseness parameter to at least four high resolution time/frequency bins
number of multiplications/ copying operations depends on source resolution and target resolution
723
(optionally) process or smooth (low pass filter) multiplied parameters in high resolution
724
apply high resolution parameters to corresponding audio data in corresponding time/frequency bins

Fig. 7b

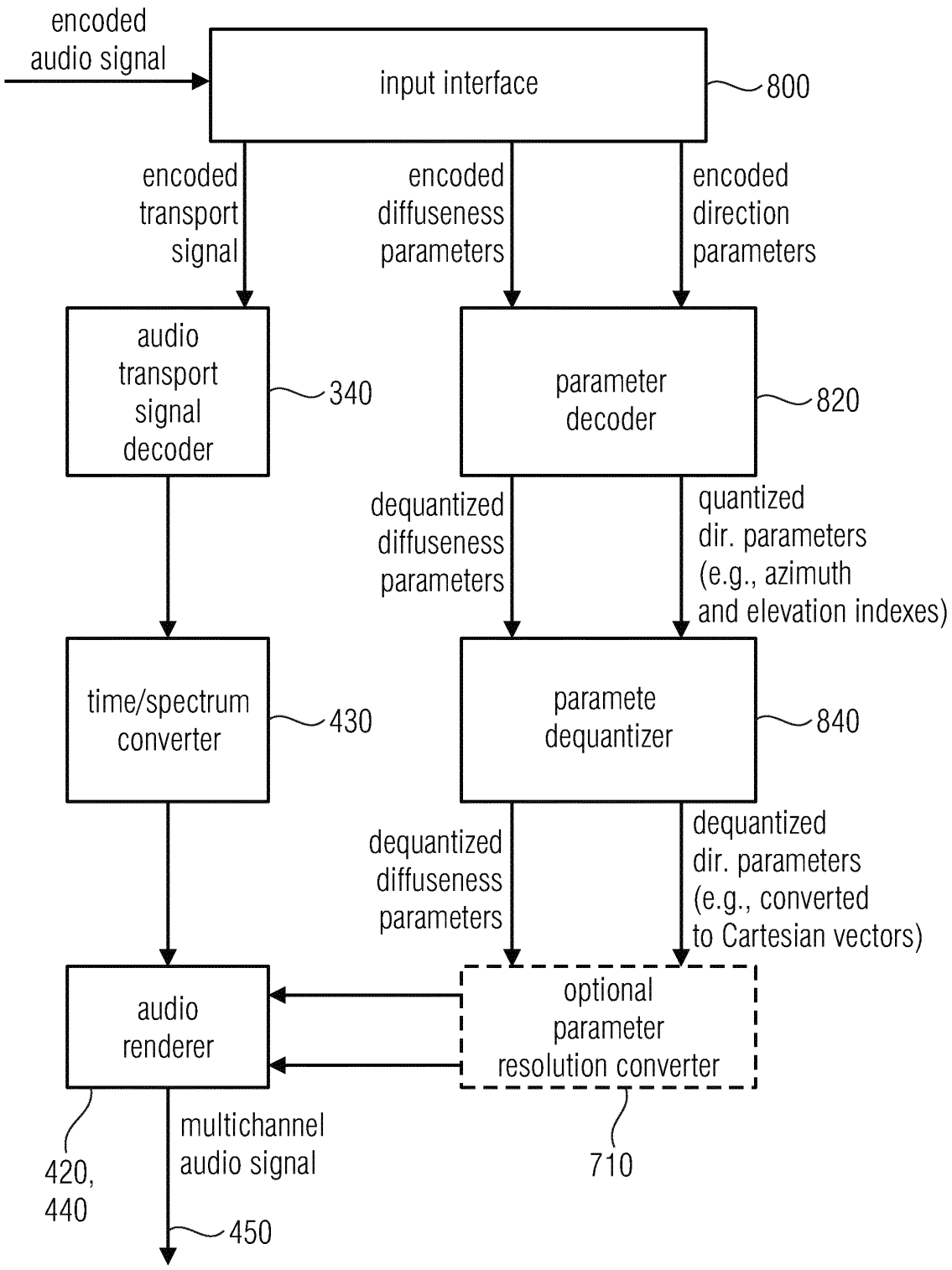

encoded audio signal
input interface
800
encoded transport signal
encoded diffuseness parameters
encoded direction parameters
audio transport signal decoder
340
parameter decoder
820
dequantized diffuseness parameters
quantized dir. parameters (e.g., azimuth and elevation indexes)
time/spectrum converter
430
paramete dequantizer
840
dequantized diffuseness parameters
dequantized dir. parameters (e.g., converted to Cartesian vectors)
audio renderer
optional parameter resolution converter
420, 440
multichannel audio signal
450
710

Fig. 8a

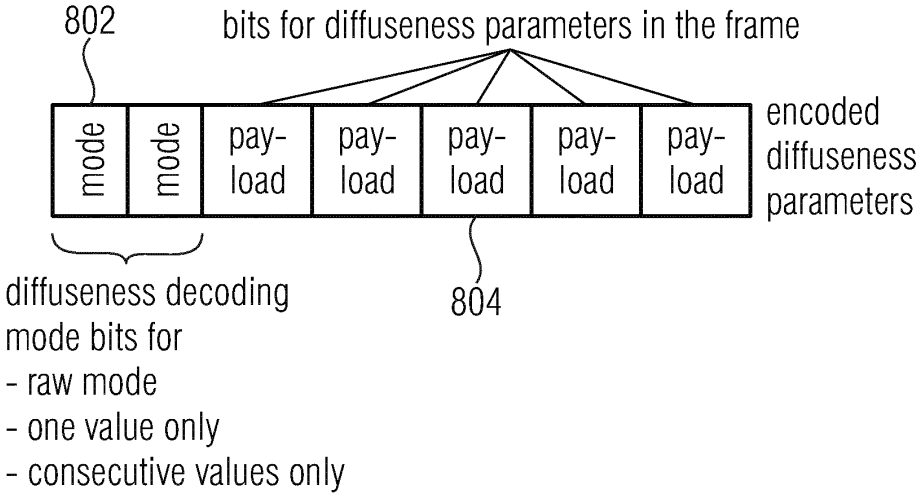


Fig. 8b

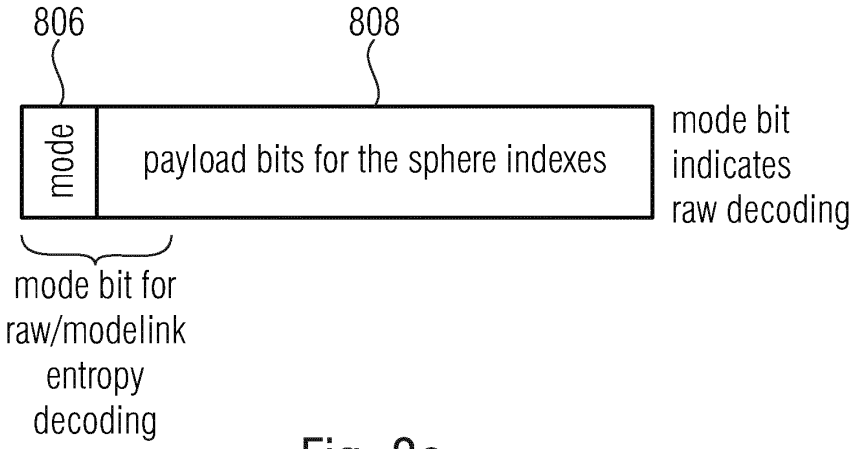


Fig. 8c

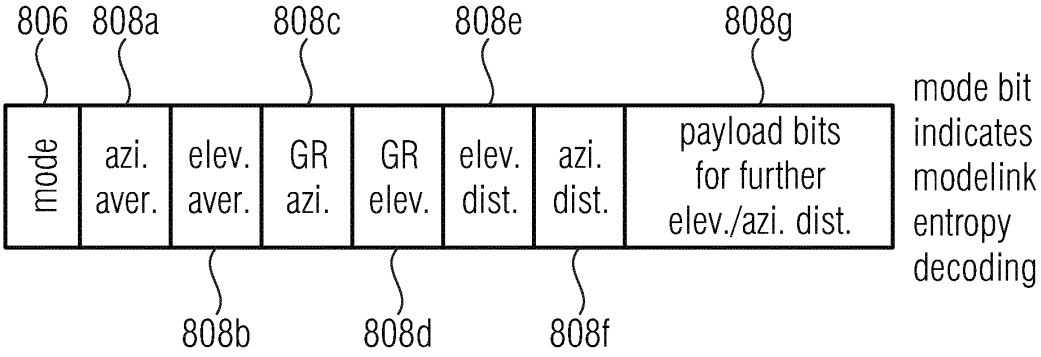


Fig. 8d

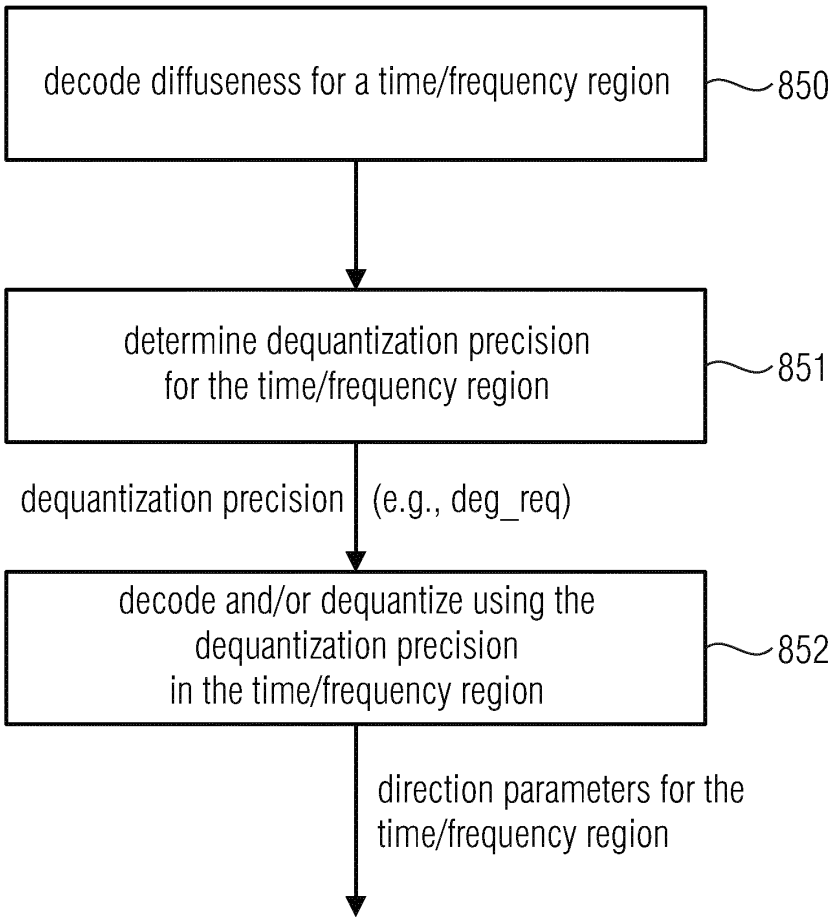

decode diffuseness for a time/frequency region
850
determine dequantization precision for the time/frequency region
851
dequantization precision (e.g., deg_req)
decode and/or dequantize using the dequantization precision in the time/frequency region
852
direction parameters for the time/frequency region

Fig. 8e

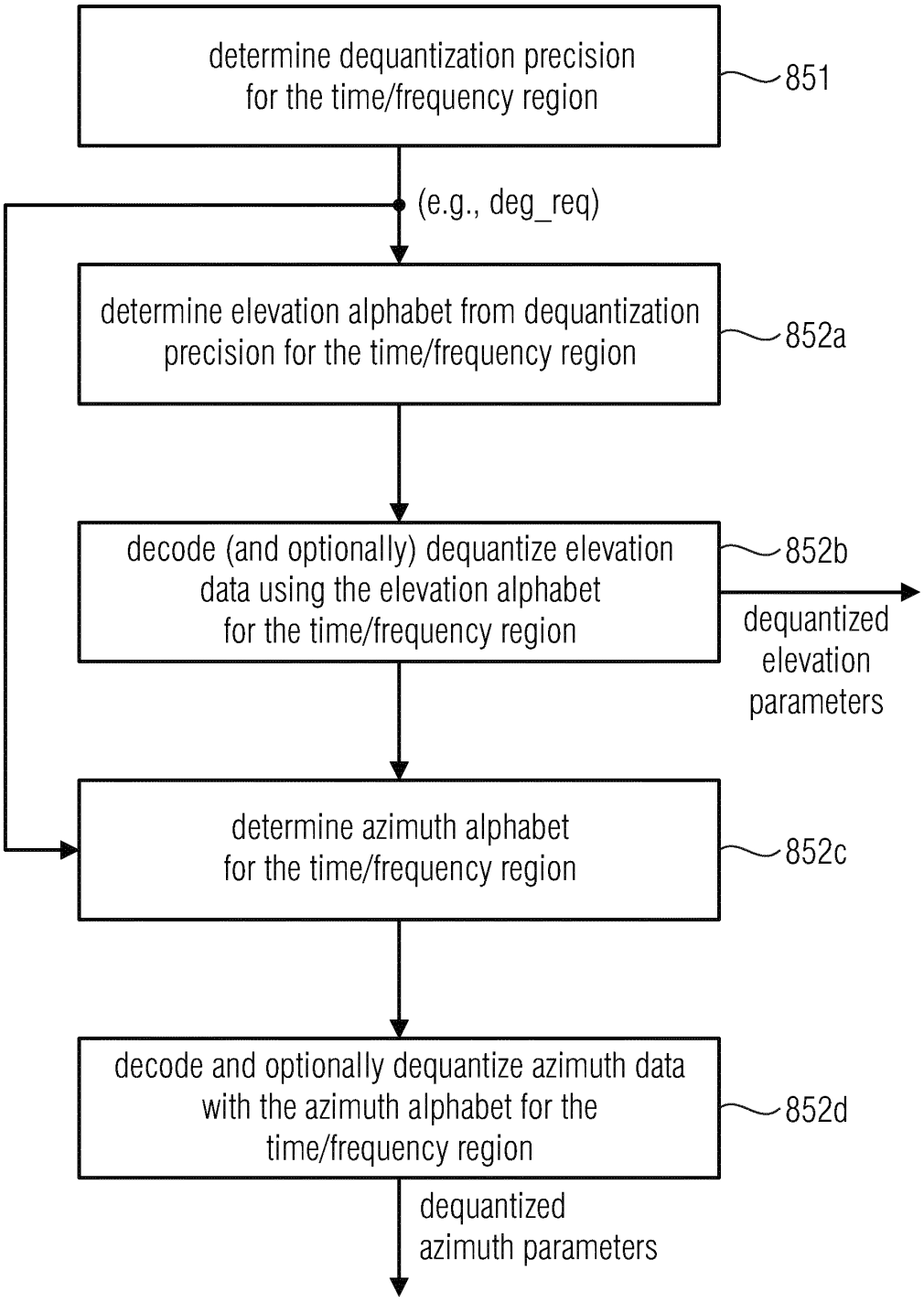

determine dequantization precision for the time/frequency region
851
(e.g., deg_req)
determine elevation alphabet from dequantization precision for the time/frequency region
852a
decode (and optionally) dequantize elevation data using the elevation alphabet for the time/frequency region
852b
dequantized elevation parameters
determine azimuth alphabet for the time/frequency region
852c
decode and optionally dequantize azimuth data with the azimuth alphabet for the time/frequency region
852d
dequantized azimuth parameters

Fig. 8f

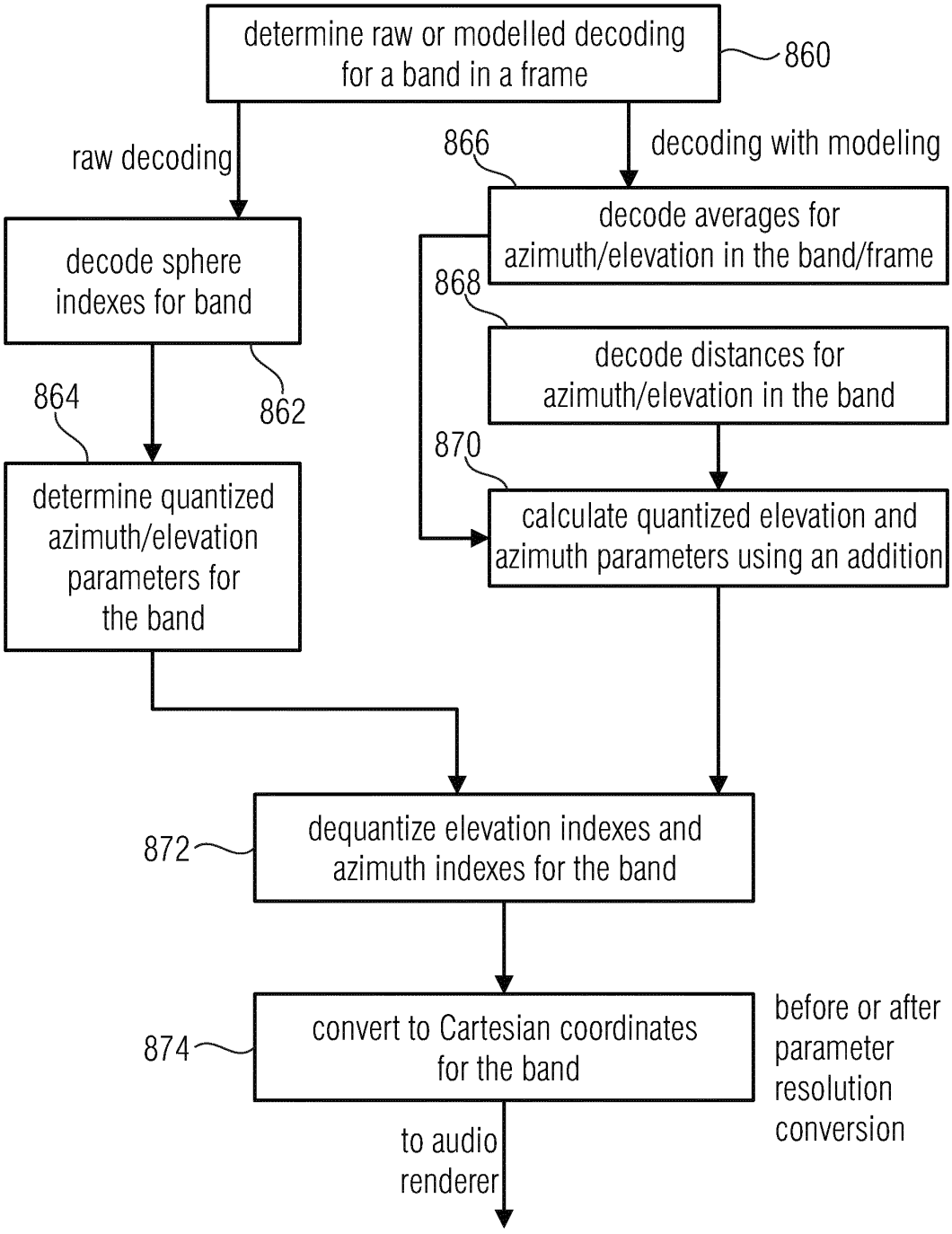
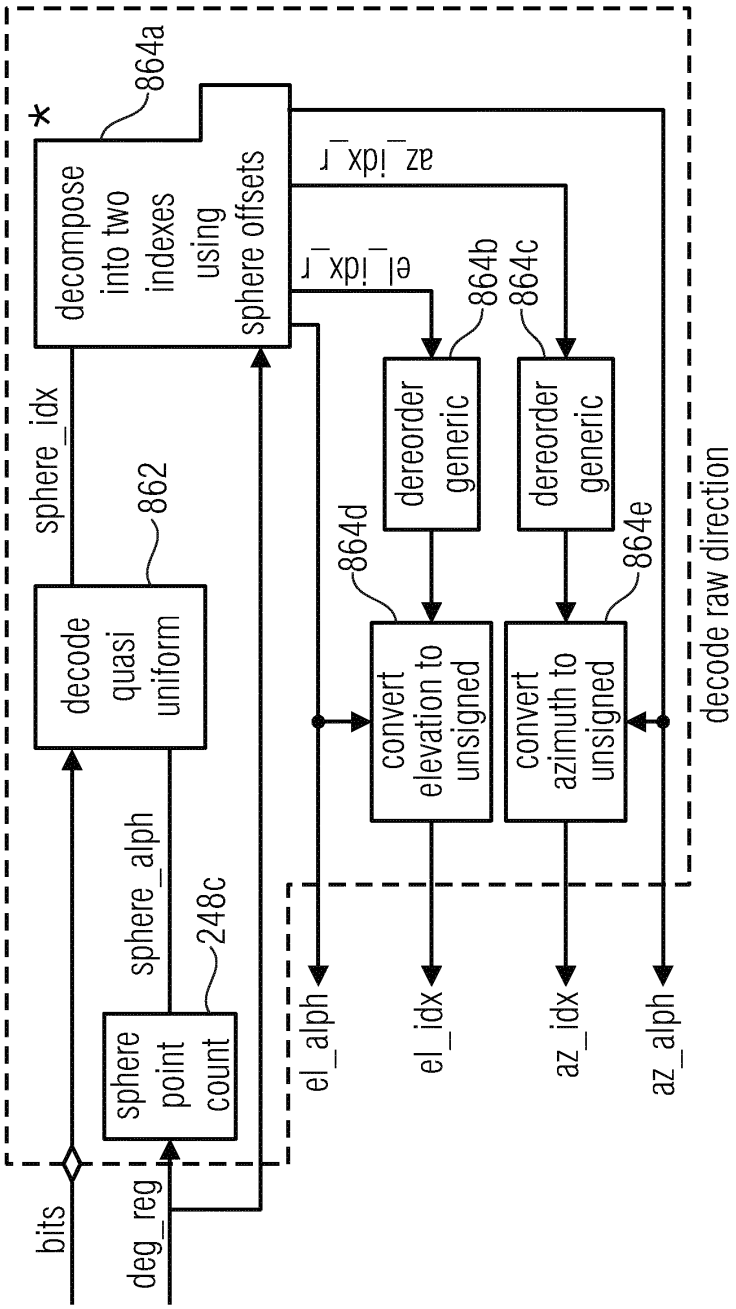

determine raw or modelled decoding for a band in a frame
860
raw decoding
866
decoding with modeling
decode sphere indexes for band
decode averages for azimuth/elevation in the band/frame
868
decode distances for azimuth/elevation in the band
862
864
870
determine quantized azimuth/elevation parameters for the band
calculate quantized elevation and azimuth parameters using an addition
872
dequantize elevation indexes and azimuth indexes for the band
874
convert to Cartesian coordinates for the band
before or after parameter resolution conversion
to audio renderer

Fig. 8g



* n_points=90 : ⌈90 : deg_req⌉
el_alph =2 · n_points+1
el_idx_r is found that satisfies the relation
A=sphere_offsets(deg_req, el_idx_r)≤shere_idx<sphere_offsets(deg_req, el_idx_r+1)=B
az_alph=B-A
az_idx_r=sphere_idx-A

Fig. 9a
Decoder

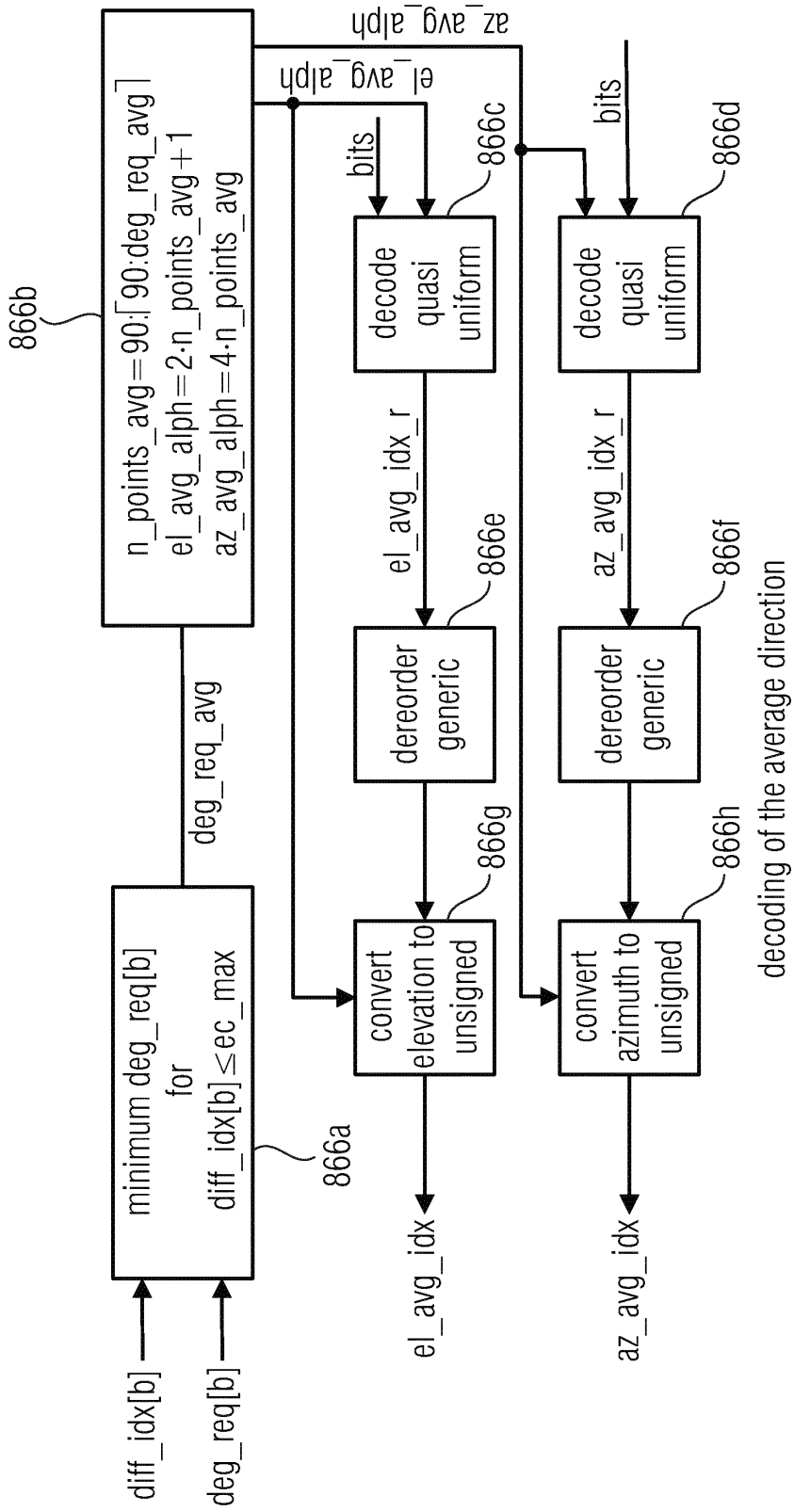


Fig. 9b
Decoder

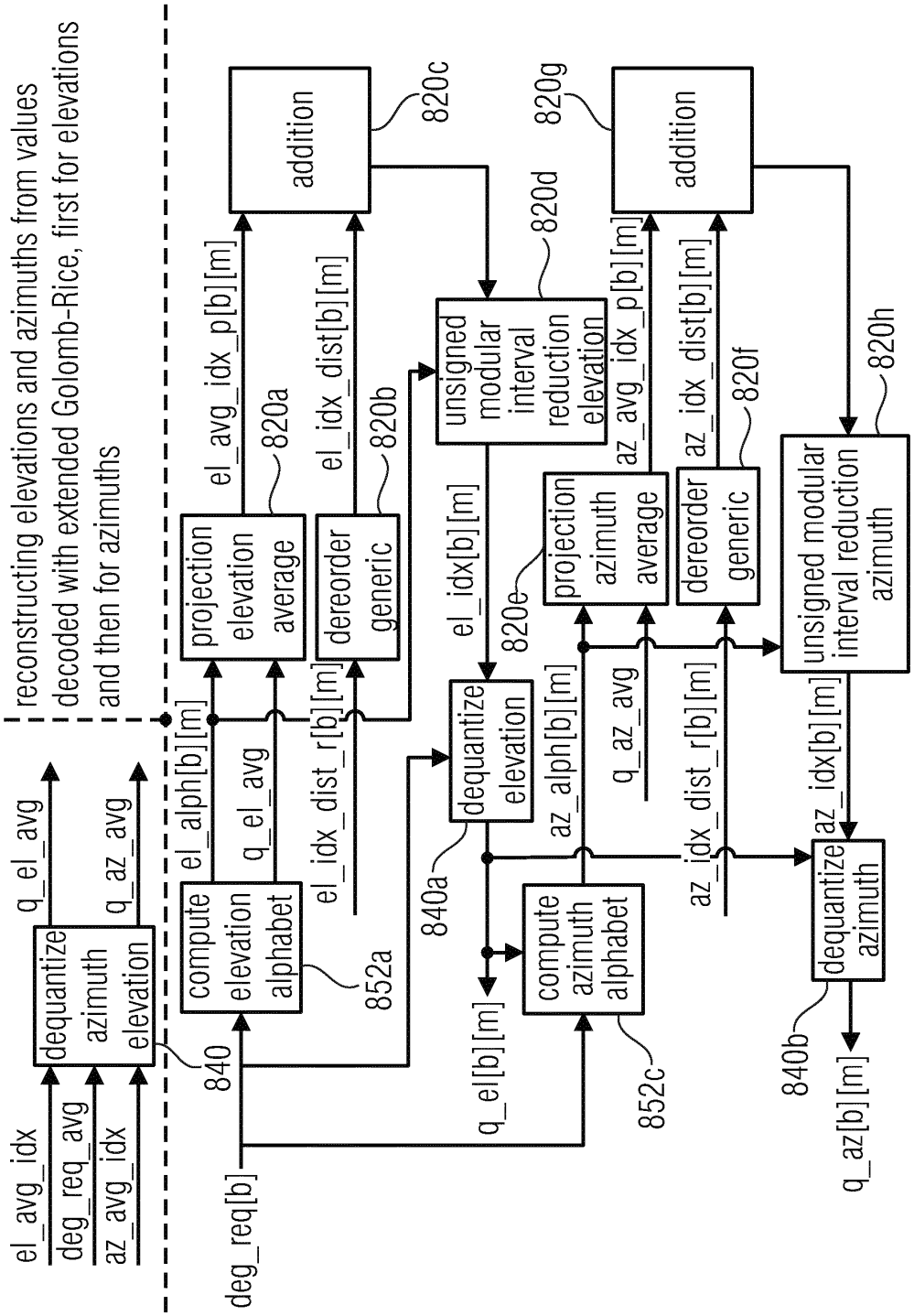


Fig. 9c
Decoder

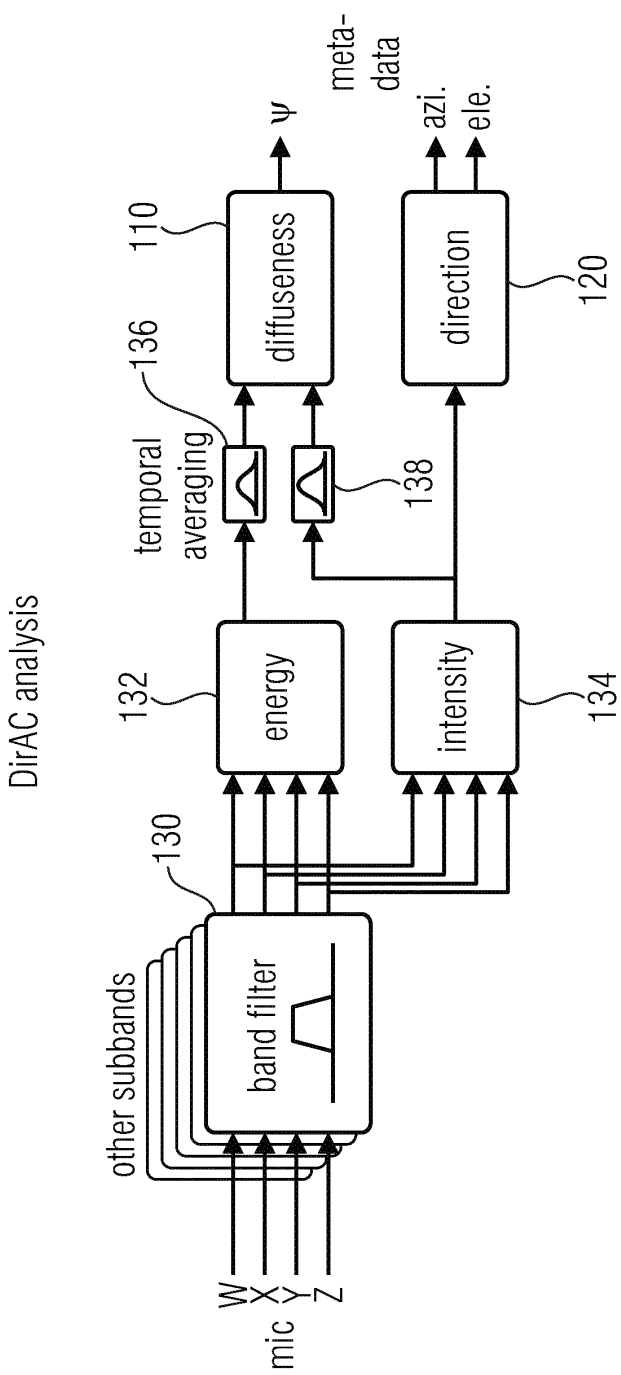


Fig. 10a
(PRIOR ART)

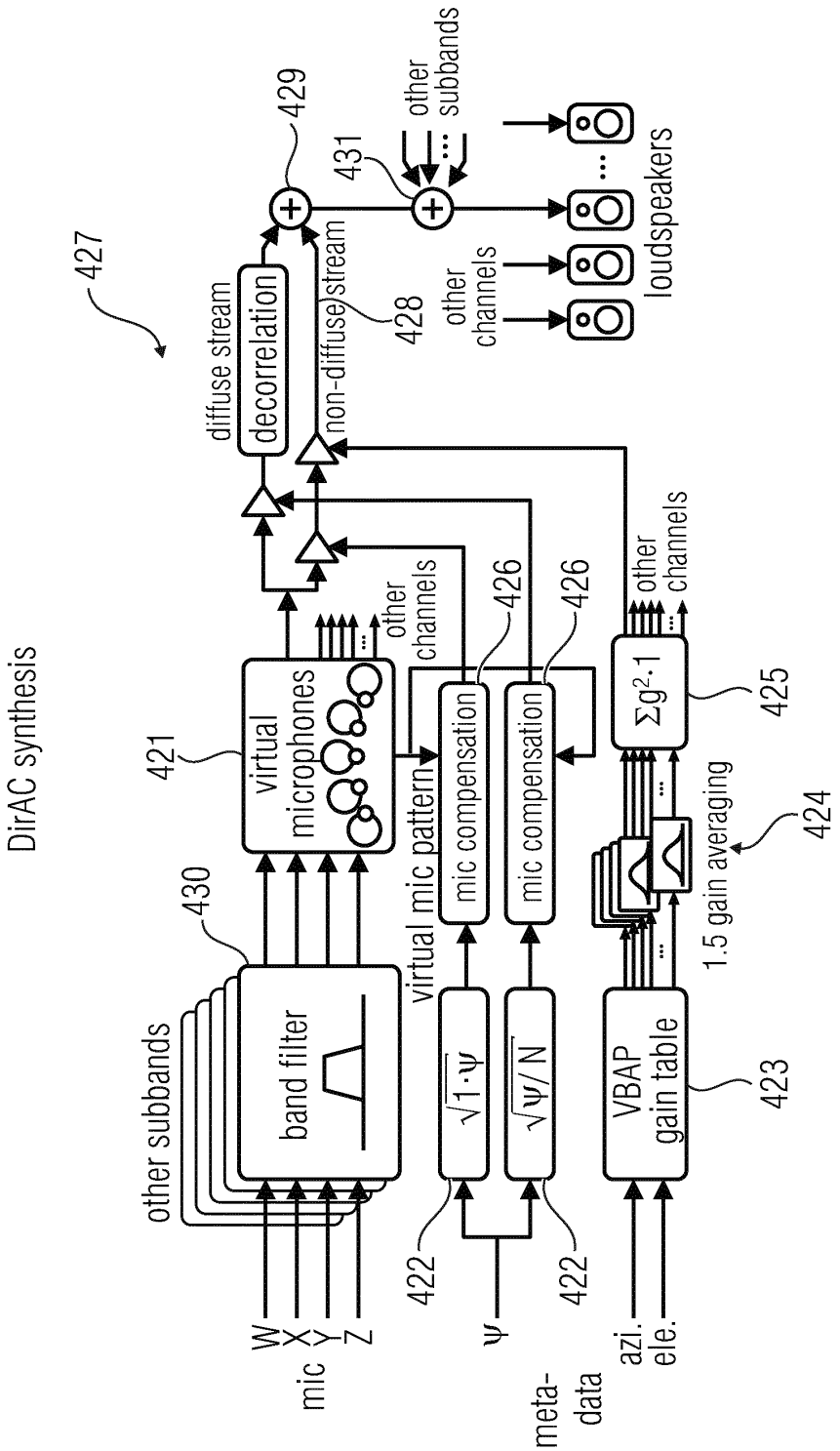


Fig. 10b
(PRIOR ART)