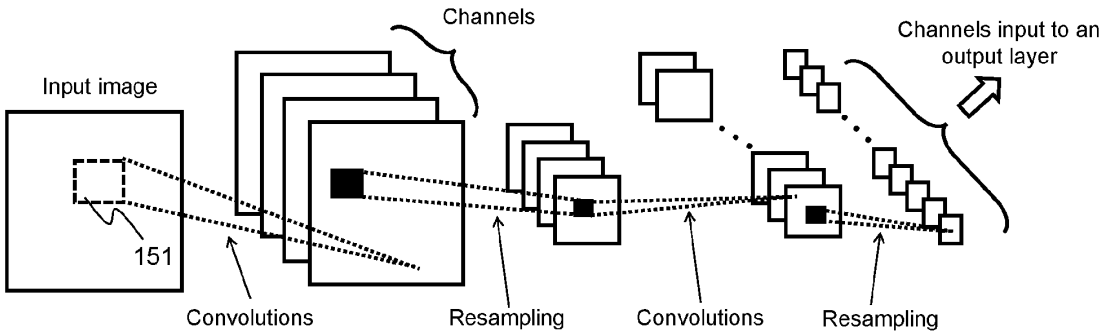

Channels
Channels input to an output layer
Input image
151
Convolutions
Resampling
Convolutions
Resampling

FIG. 1

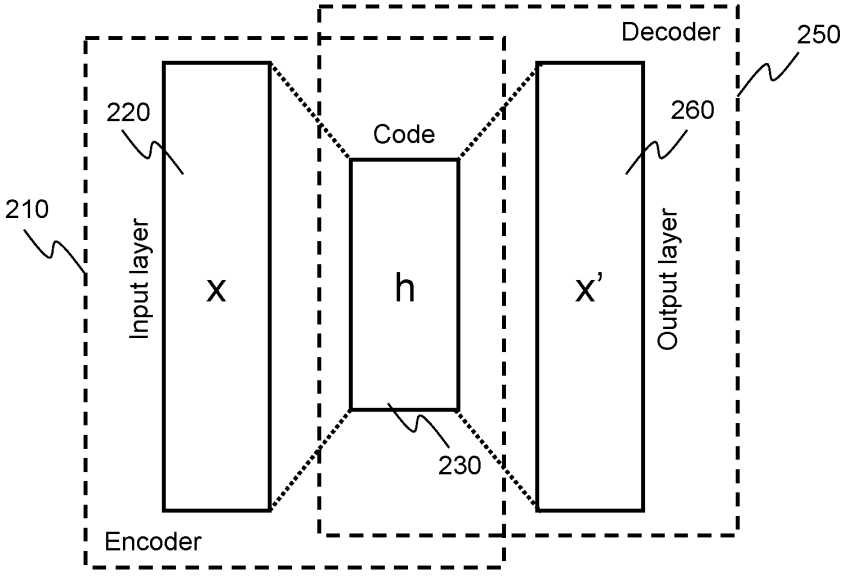

Decoder
250
220
Code
260
210
Input layer
x
h
x'
Output layer
230
Encoder

FIG. 2

FIG. 3A

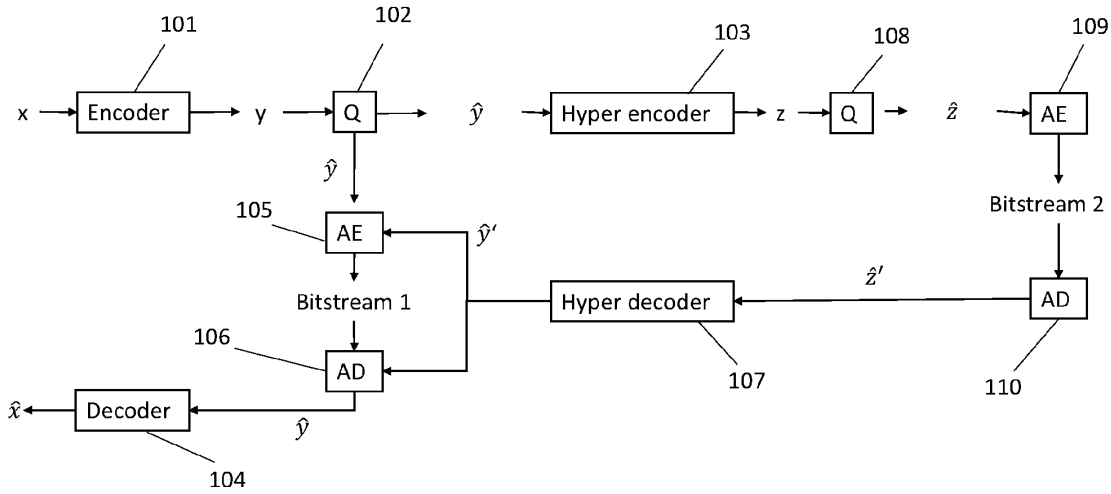

101
102
103
108
109
x
Encoder
y
Q
$\hat{y}$
Hyper encoder
z
Q
$\hat{z}$
AE
$\hat{y}$
Bitstream 2
105
AE
$\hat{y}'$
$\hat{z}'$
Bitstream 1
Hyper decoder
AD
106
AD
107
110
$\hat{x}$
Decoder
$\hat{y}$
104

FIG. 3B

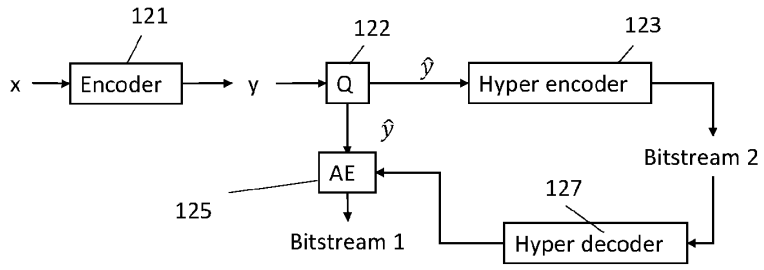

121
122
123
x
Encoder
y
Q
$\hat{y}$
Hyper encoder
$\hat{y}$
Bitstream 2
AE
125
127
Bitstream 1
Hyper decoder

FIG. 3C

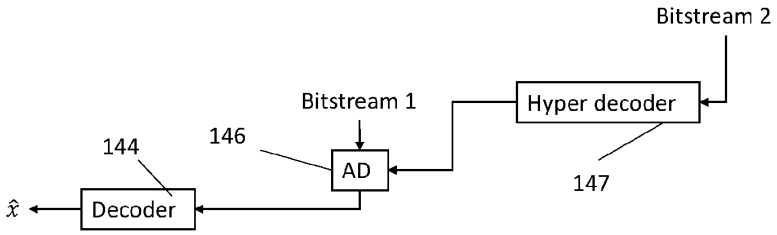
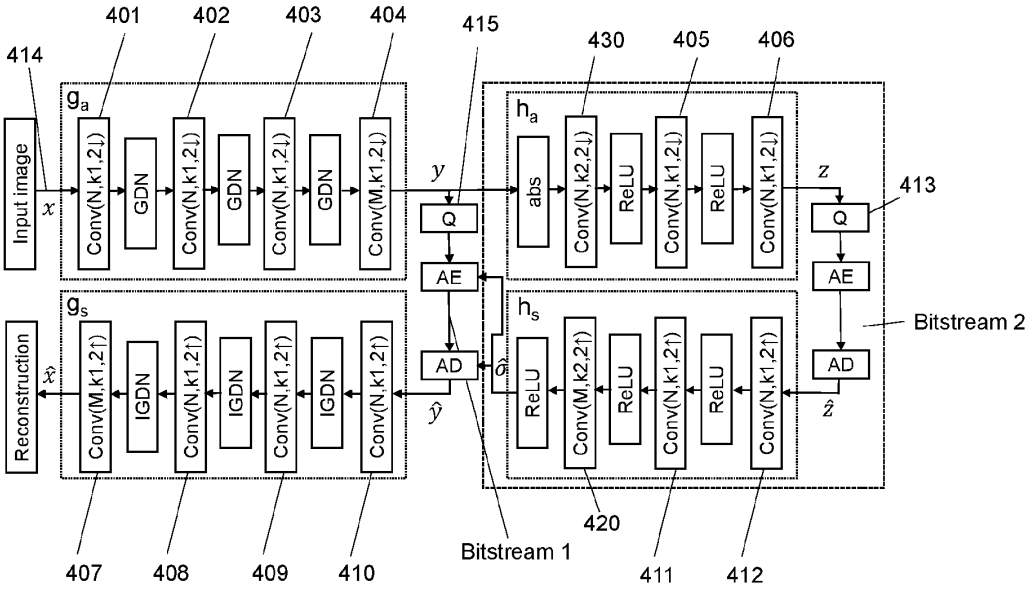

Bitstream 2
Bitstream 1
Hyper decoder
146
144
AD
147
$\hat{x}$
Decoder

FIG. 4


401
402
403
404
415
430
405
406
414
g_a
h_a
Input image
x
Conv(N,k1,2↓)
GDN
Conv(M,k1,2↓)
y
Q
AE
AD
abs
Conv(N,k2,2↓)
ReLU
z
413
Bitstream 2
g_s
h_s
Reconstruction
$\hat{x}$
Conv(M,k1,2↑)
IGDN
Conv(N,k1,2↑)
$\hat{y}$
Conv(M,k2,2↑)
$\hat{z}$
420
Bitstream 1
407
408
409
410
411
412

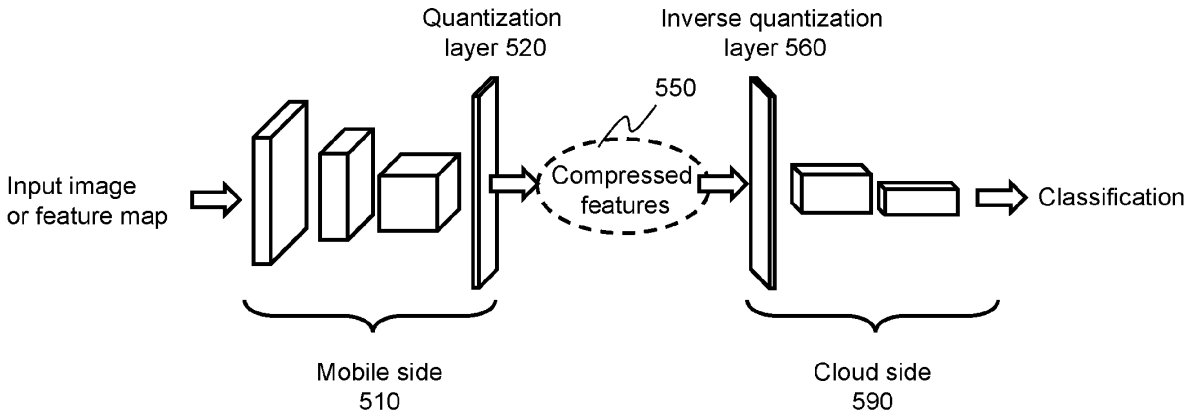

Quantization
layer 520
Inverse quantization
layer 560
550
Input image
or feature map
Compressed
features
Classification
Mobile side
510
Cloud side
590

FIG. 5

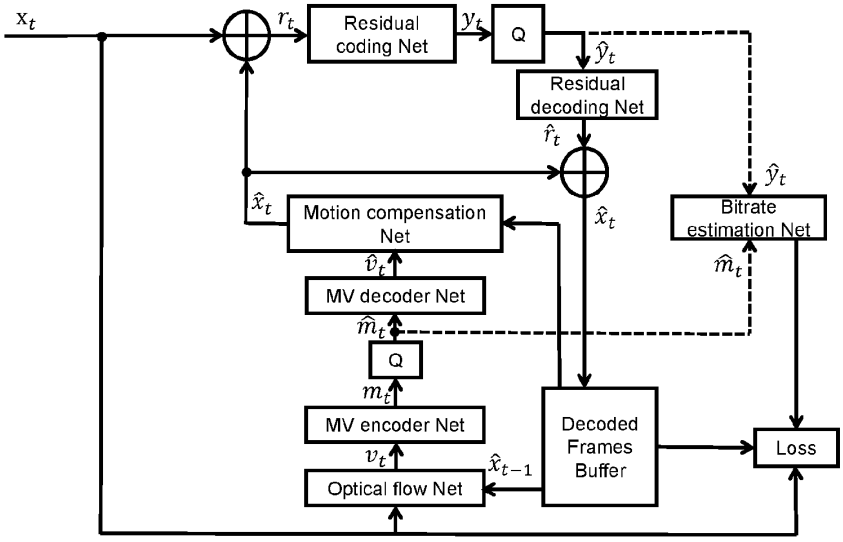

x_t
r_t
Residual coding Net
y_t
Q
$\hat{y}_t$
Residual decoding Net
$\hat{r}_t$
$\hat{x}_t$
Motion compensation Net
$\hat{v}_t$
MV decoder Net
$\hat{m}_t$
Q
m_t
MV encoder Net
v_t
Optical flow Net
$\hat{x}_{t-1}$
Decoded Frames Buffer
$\hat{x}_t$
$\hat{y}_t$
Bitrate estimation Net
$\hat{m}_t$
Loss

FIG. 6A

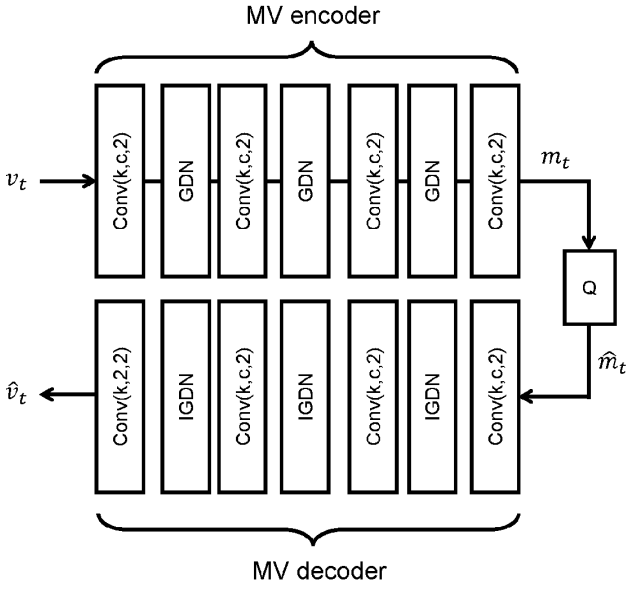

MV encoder
v_t
Conv(k,c,2)
GDN
Conv(k,c,2)
GDN
Conv(k,c,2)
GDN
Conv(k,c,2)
m_t
Q
$\hat{m}_t$
$\hat{v}_t$
Conv(k,2,2)
IGDN
Conv(k,c,2)
IGDN
Conv(k,c,2)
IGDN
Conv(k,c,2)
MV decoder

FIG. 6B

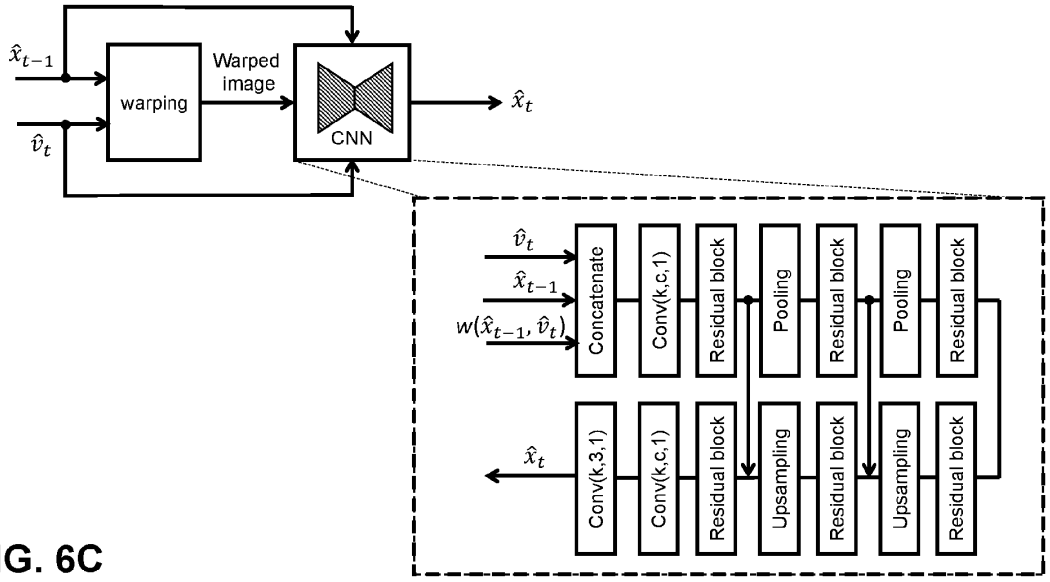

$\hat{x}_{t-1}$
$\hat{v}_t$
warping
Warped image
CNN
$\hat{x}_t$
$\hat{v}_t$
$\hat{x}_{t-1}$
$w(\hat{x}_{t-1}, \hat{v}_t)$
Concatenate
Conv(k,c,1)
Residual block
Pooling
Residual block
Pooling
Residual block
$\hat{x}_t$
Conv(k,3,1)
Conv(k,c,1)
Residual block
Upsampling
Residual block
Upsampling
Residual block

FIG. 6C

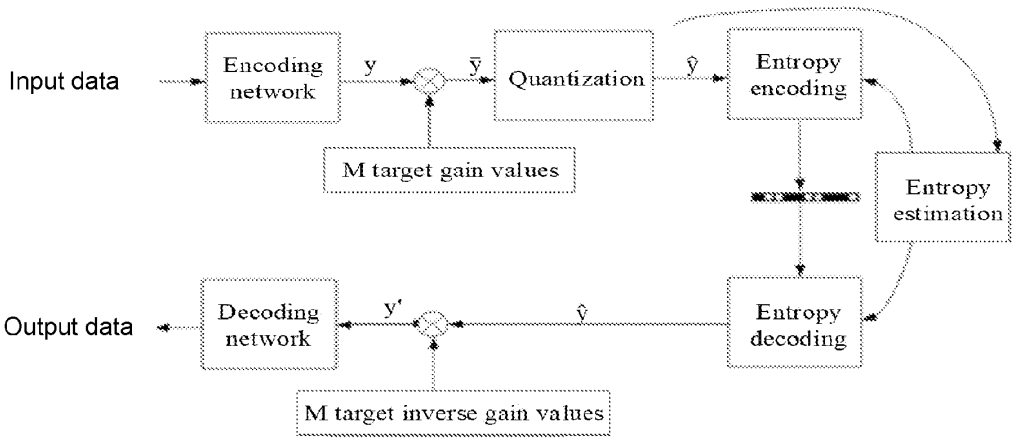

Input data
Encoding network
y
$\bar{y}$
Quantization
$\hat{y}$
Entropy encoding
M target gain values
Entropy estimation
Output data
Decoding network
y'
$\hat{y}$
Entropy decoding
M target inverse gain values

FIG. 7

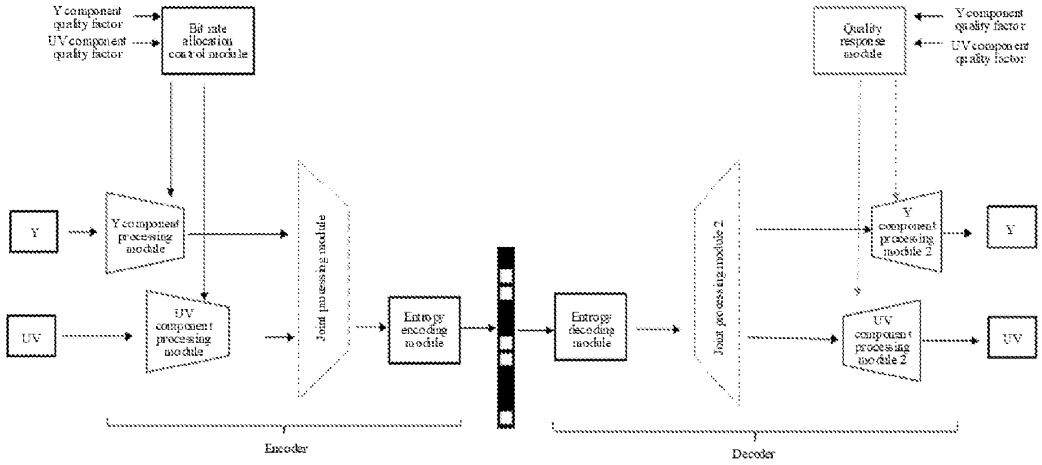

Y component quality factor
UV component quality factor
Bit rate allocation control module
Y component processing module
UV component processing module
Joint processing module
Entropy encoding module
Entropy decoding module
Joint processing module 2
Quality response module
Y component quality factor
UV component quality factor
Y component processing module 2
UV component processing module 2
Y
UV
Y
UV
Encoder
Decoder

FIG. 8

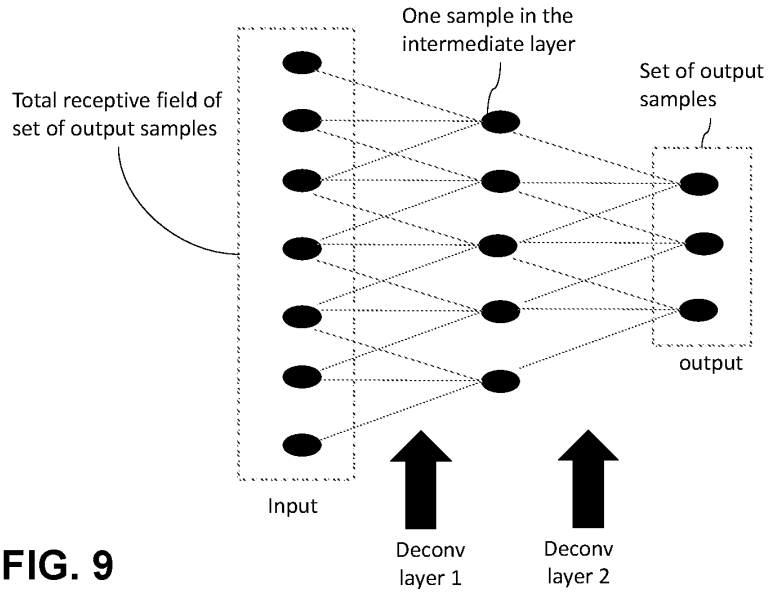

One sample in the intermediate layer
Set of output samples
Total receptive field of set of output samples
output
Input
Deconv layer 1
Deconv layer 2

FIG. 9

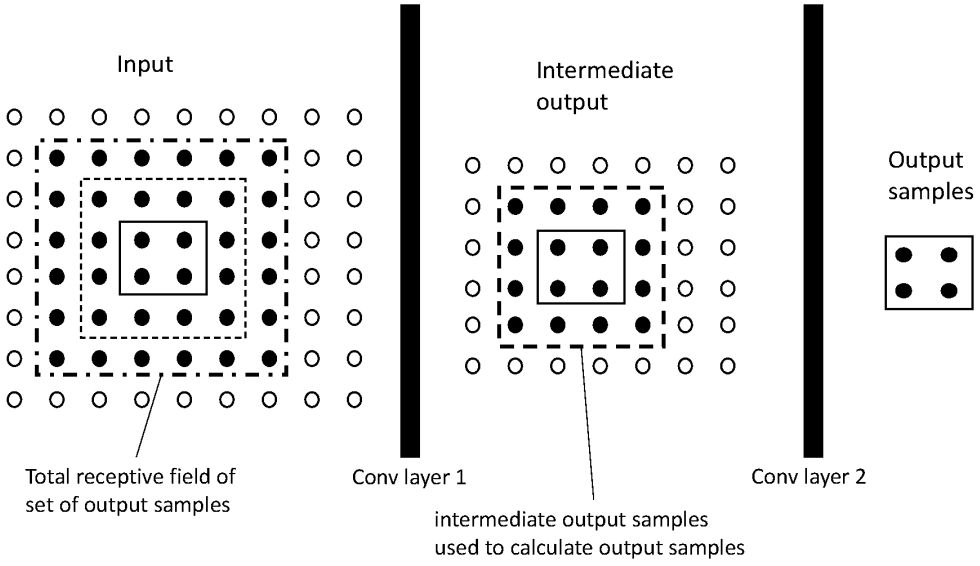

Input
Intermediate output
Output samples
Total receptive field of set of output samples
Conv layer 1
Conv layer 2
intermediate output samples used to calculate output samples

FIG. 10

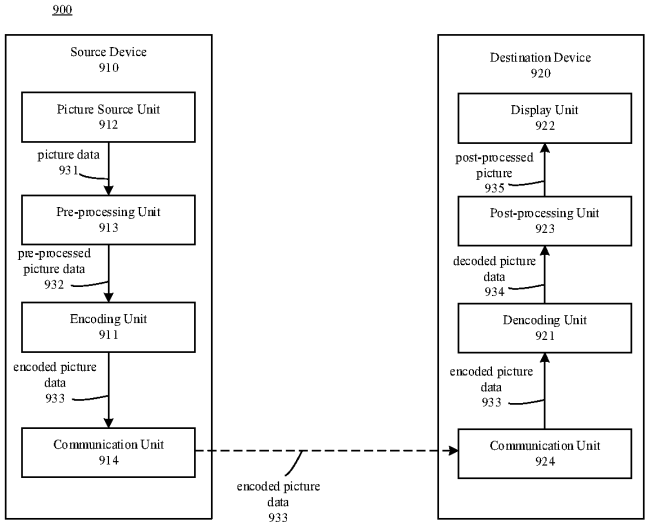

900
Source Device
910
Picture Source Unit
912
picture data
931
Pre-processing Unit
913
pre-processed
picture data
932
Encoding Unit
911
encoded picture
data
933
Communication Unit
914
encoded picture
data
933
Destination Device
920
Display Unit
922
post-processed
picture
935
Post-processing Unit
923
decoded picture
data
934
Dencoding Unit
921
encoded picture
data
933
Communication Unit
924

FIG. 11

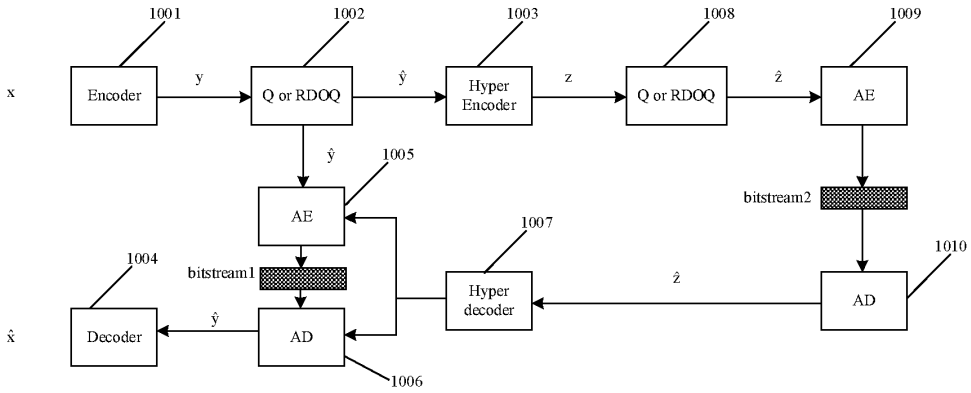

1001
1002
1003
1008
1009
x
Encoder
y
Q or RDOQ
ŷ
Hyper Encoder
z
Q or RDOQ
ẑ
AE
ŷ
1005
AE
bitstream2
1007
1010
1004
bitstream1
Hyper decoder
ẑ
AD
x̂
Decoder
ŷ
AD
1006

FIG. 12

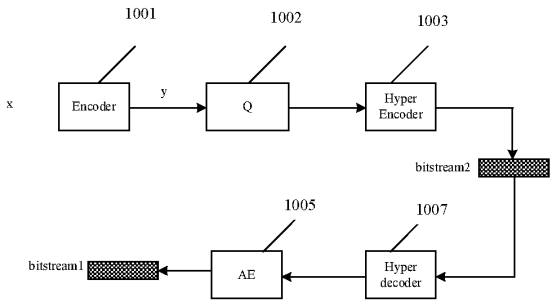

1001
1002
1003
x
Encoder
y
Q
Hyper Encoder
bitstream2
1005
1007
bitstream1
AE
Hyper decoder

FIG. 13

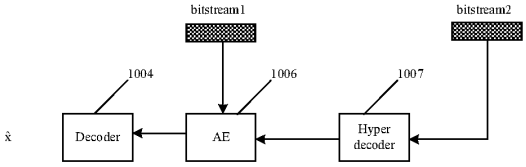

bitstream1
bitstream2
1004
1006
1007
x̂
Decoder
AE
Hyper decoder

FIG. 14

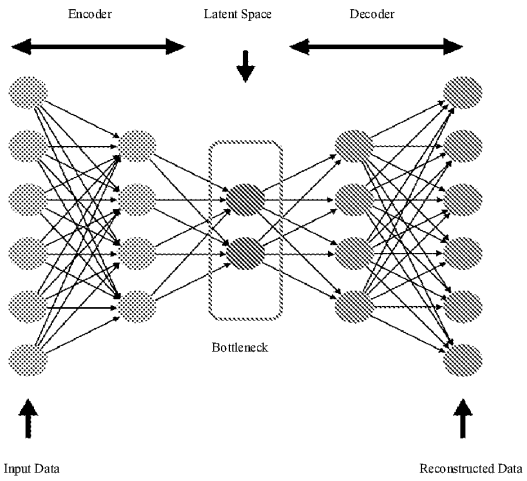

Encoder
Latent Space
Decoder
Bottleneck
Input Data
Reconstructed Data

FIG. 15

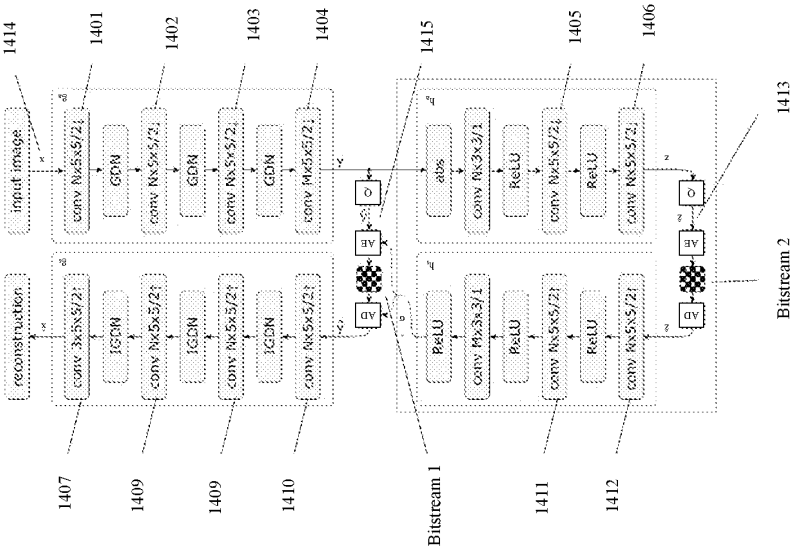

1414
1401
1402
1403
1404
1415
1405
1406
1413
input image
reconstruction
conv Nx5x5/2↓
GDN
conv Nx5x5/2↓
GDN
conv Nx5x5/2↓
GDN
conv Mx5x5/2↓
abs
conv Nx3x3/1
ReLU
conv Nx5x5/2↓
ReLU
conv Nx5x5/2↓
Q
AE
AD
conv 3x5x5/2↑
IGDN
conv Nx5x5/2↑
IGDN
conv Nx5x5/2↑
IGDN
conv Nx5x5/2↑
ReLU
conv Mx3x3/1
ReLU
conv Nx5x5/2↑
ReLU
conv Nx5x5/2↑
Bitstream 2
1407
1409
1409
1410
Bitstream 1
1411
1412

FIG. 16

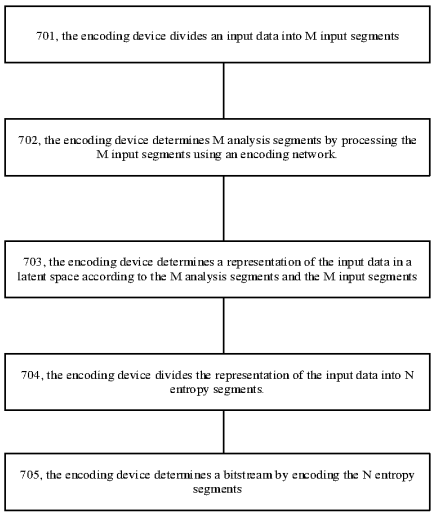

701, the encoding device divides an input data into M input segments
702, the encoding device determines M analysis segments by processing the M input segments using an encoding network.
703, the encoding device determines a representation of the input data in a latent space according to the M analysis segments and the M input segments
704, the encoding device divides the representation of the input data into N entropy segments.
705, the encoding device determines a bitstream by encoding the N entropy segments

FIG. 17

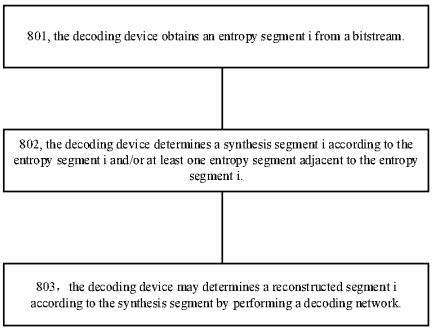

801, the decoding device obtains an entropy segment i from a bitstream.
802, the decoding device determines a synthesis segment i according to the entropy segment i and/or at least one entropy segment adjacent to the entropy segment i.
803，the decoding device may determines a reconstructed segment i according to the synthesis segment by performing a decoding network.

FIG. 18

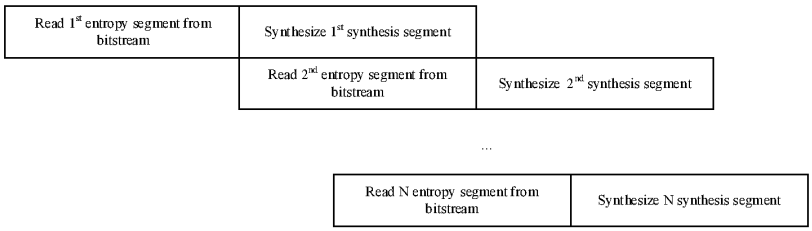

Read 1st entropy segment from bitstream
Synthesize 1st synthesis segment
Read 2nd entropy segment from bitstream
Synthesize 2nd synthesis segment
...
Read N entropy segment from bitstream
Synthesize N synthesis segment

FIG. 19

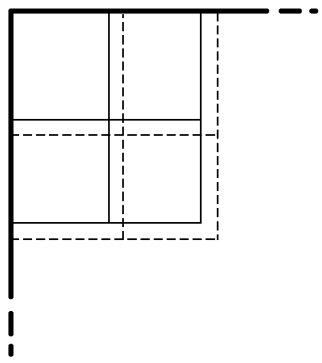


FIG. 20

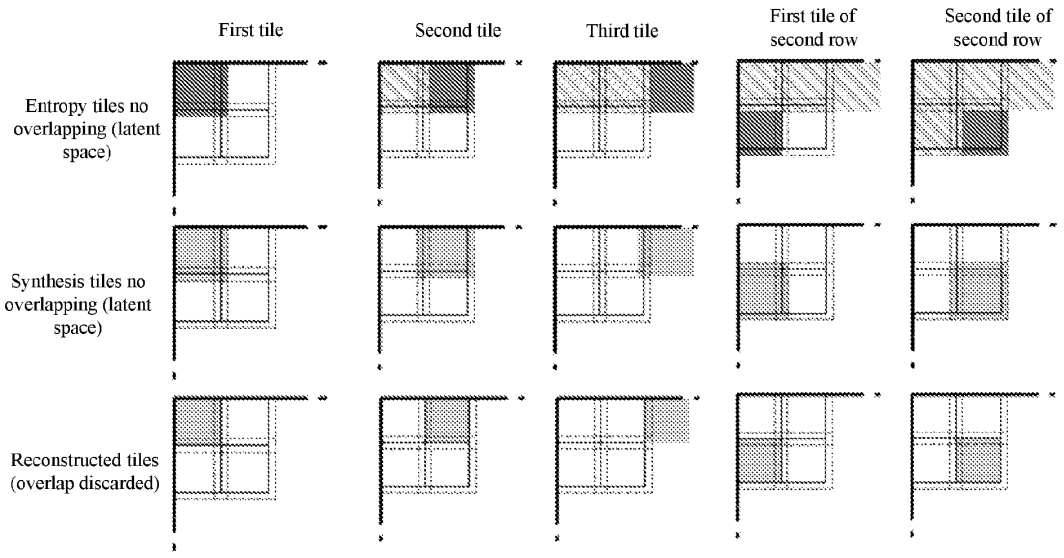

First tile
Second tile
Third tile
First tile of second row
Second tile of second row
Entropy tiles no overlapping (latent space)
Synthesis tiles no overlapping (latent space)
Reconstructed tiles (overlap discarded)

FIG. 21

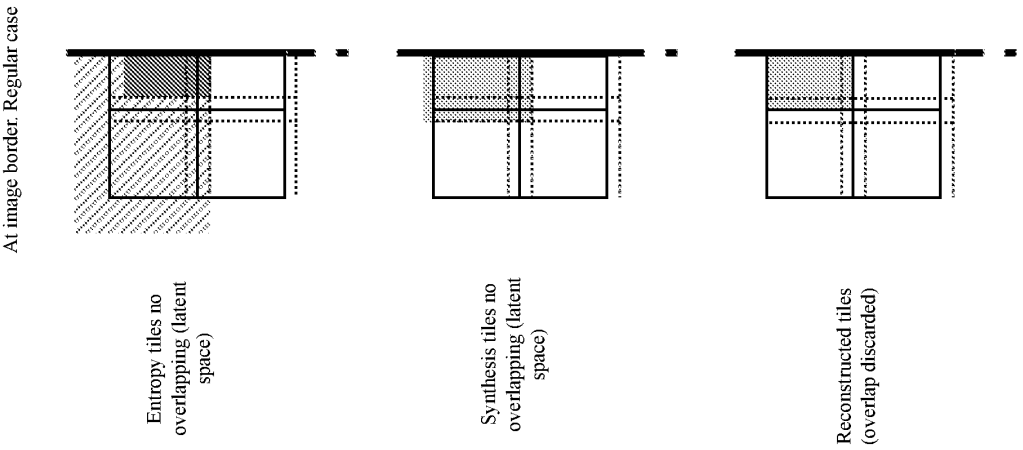

At image border. Regular case
Entropy tiles no overlapping (latent space)
Synthesis tiles no overlapping (latent space)
Reconstructed tiles (overlap discarded)

FIG. 22

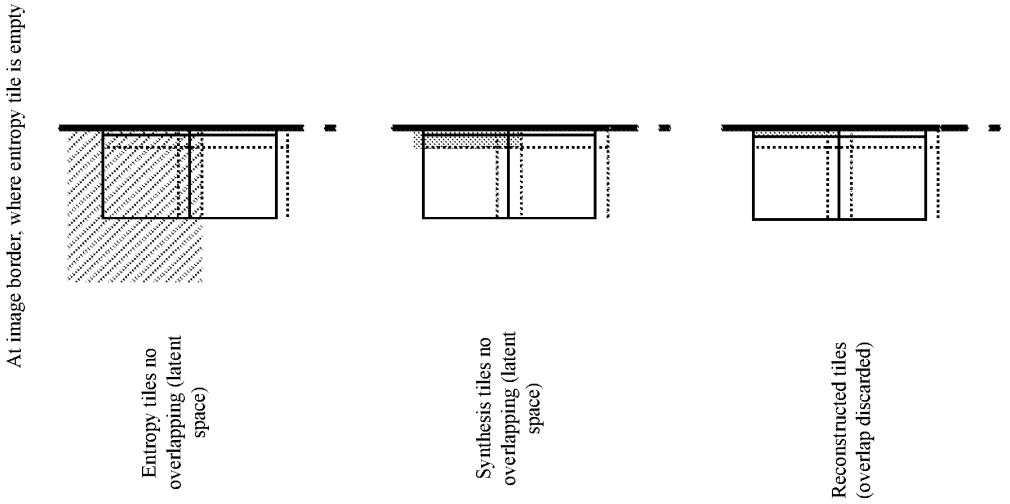

At image border, where entropy tile is empty
Entropy tiles no overlapping (latent space)
Synthesis tiles no overlapping (latent space)
Reconstructed tiles (overlap discarded)

FIG. 23

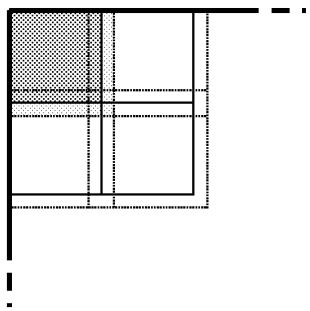


FIG. 24a

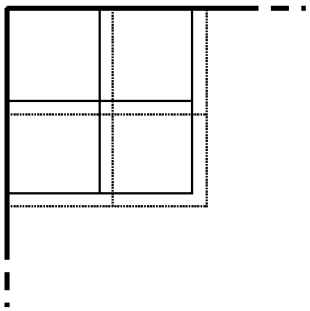


FIG. 24b

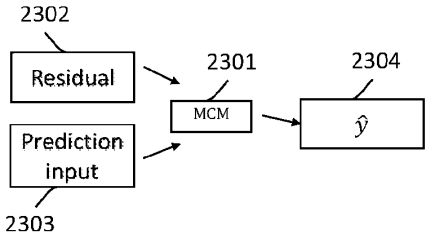

2302
Residual
Prediction input
2303
2301
MCM
2304
$\hat{y}$

FIG. 25

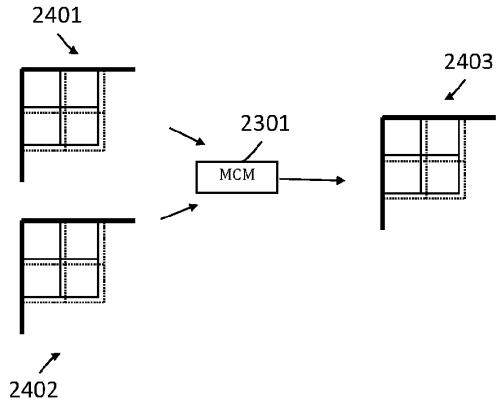

2401
2402
2301
MCM
2403

FIG. 26

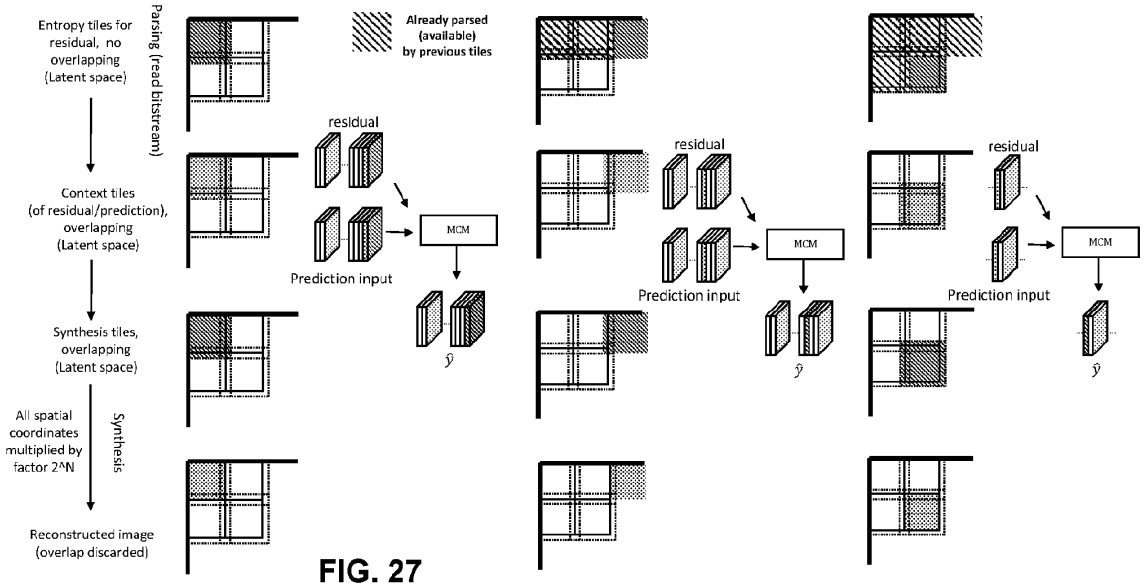

Entropy tiles for residual, no overlapping (Latent space)
Parsing (read bitstream)
Context tiles (of residual/prediction), overlapping (Latent space)
Synthesis tiles, overlapping (Latent space)
All spatial coordinates multiplied by factor 2^N
Synthesis
Reconstructed image (overlap discarded)
Already parsed (available) by previous tiles
residual
Prediction input
MCM
$\hat{y}$

FIG. 27

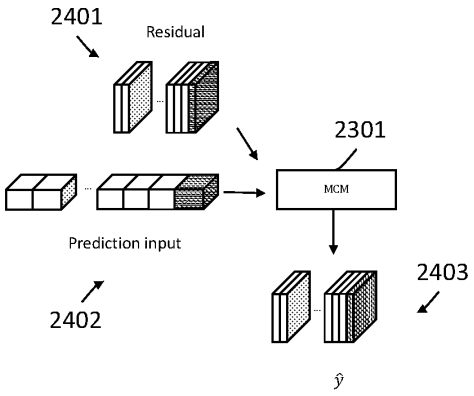


FIG. 28

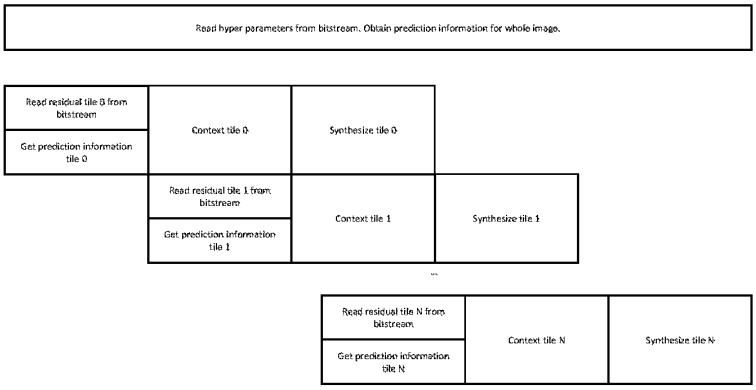

Read hyper parameters from bitstream. Obtain prediction information for whole image.
Read residual tile 0 from bitstream
Get prediction information tile 0
Context tile 0
Synthesize tile 0
Read residual tile 1 from bitstream
Get prediction information tile 1
Context tile 1
Synthesize tile 1
...
Read residual tile N from bitstream
Get prediction information tile N
Context tile N
Synthesize tile N

FIG. 29

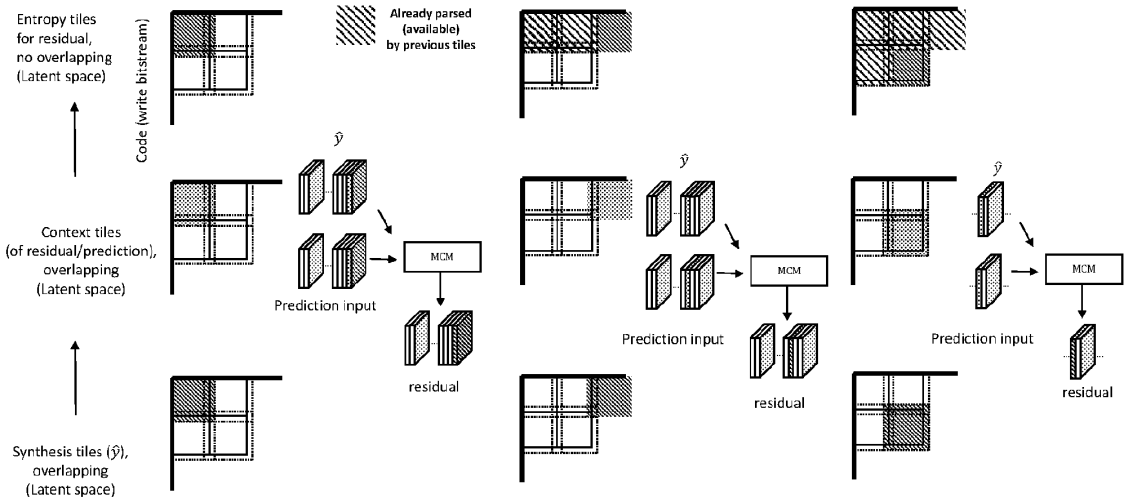

Entropy tiles
for residual,
no overlapping
(Latent space)
Code (write bitstream)
Already parsed
(available)
by previous tiles
Context tiles
(of residual/prediction),
overlapping
(Latent space)
Synthesis tiles ($\hat{y}$),
overlapping
(Latent space)
$\hat{y}$
MCM
Prediction input
residual
$\hat{y}$
MCM
Prediction input
residual
$\hat{y}$
MCM
Prediction input
residual

FIG. 30

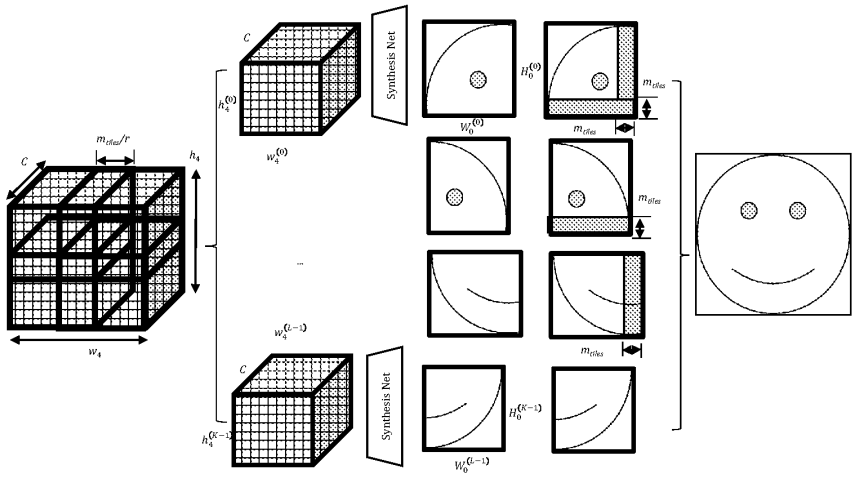


FIG. 31

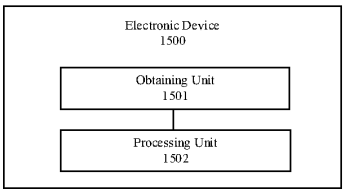


FIG. 32

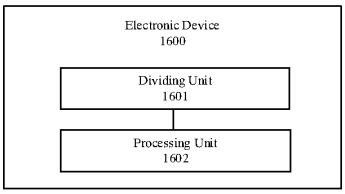


FIG. 33

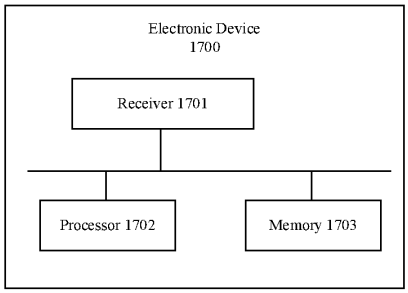


FIG. 34

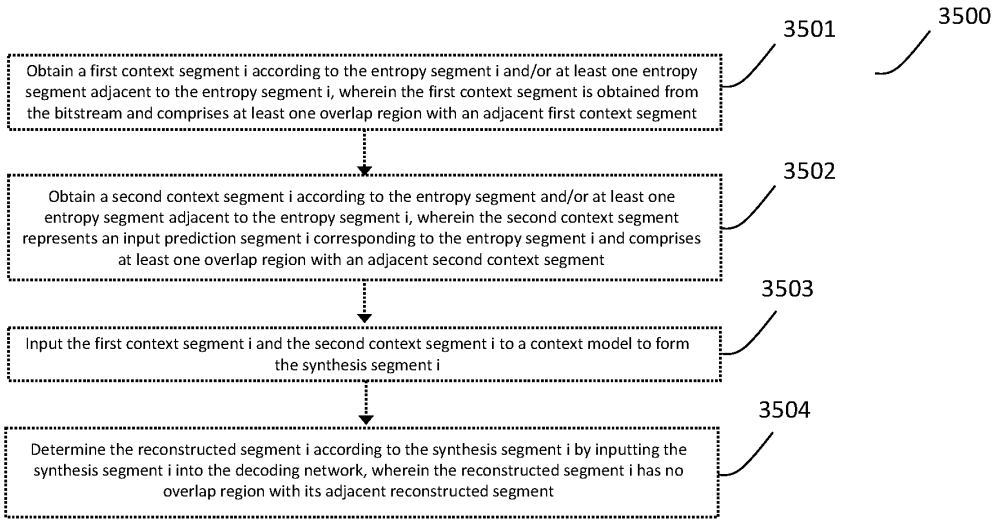

3500
3501
Obtain a first context segment i according to the entropy segment i and/or at least one entropy segment adjacent to the entropy segment i, wherein the first context segment is obtained from the bitstream and comprises at least one overlap region with an adjacent first context segment
3502
Obtain a second context segment i according to the entropy segment and/or at least one entropy segment adjacent to the entropy segment i, wherein the second context segment represents an input prediction segment i corresponding to the entropy segment i and comprises at least one overlap region with an adjacent second context segment
3503
Input the first context segment i and the second context segment i to a context model to form the synthesis segment i
3504
Determine the reconstructed segment i according to the synthesis segment i by inputting the synthesis segment i into the decoding network, wherein the reconstructed segment i has no overlap region with its adjacent reconstructed segment

FIG.35

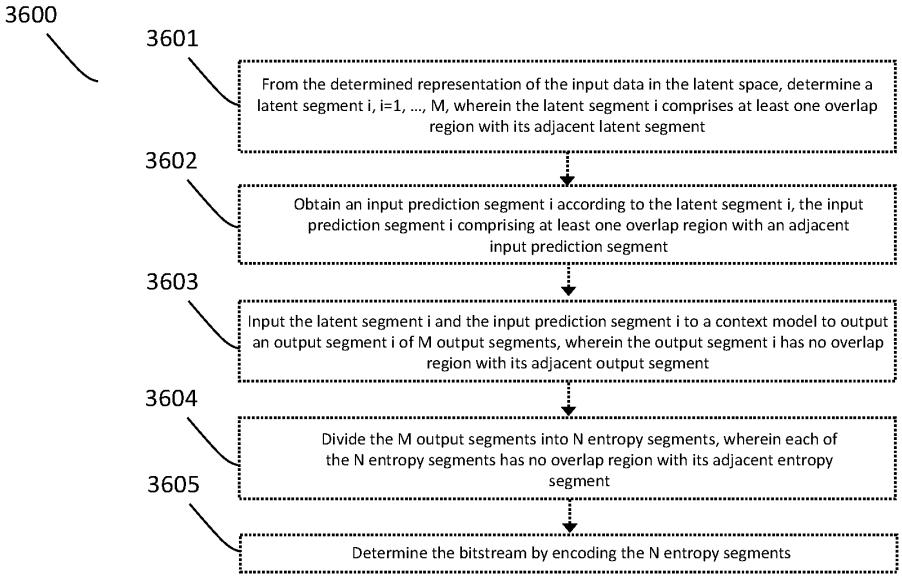

3600
3601
From the determined representation of the input data in the latent space, determine a latent segment i, i=1, ..., M, wherein the latent segment i comprises at least one overlap region with its adjacent latent segment
3602
Obtain an input prediction segment i according to the latent segment i, the input prediction segment i comprising at least one overlap region with an adjacent input prediction segment
3603
Input the latent segment i and the input prediction segment i to a context model to output an output segment i of M output segments, wherein the output segment i has no overlap region with its adjacent output segment
3604
Divide the M output segments into N entropy segments, wherein each of the N entropy segments has no overlap region with its adjacent entropy segment
3605
Determine the bitstream by encoding the N entropy segments

FIG.36

Start of Image
File Header
Entropy Encoded Data
End of Image

FIG. 37

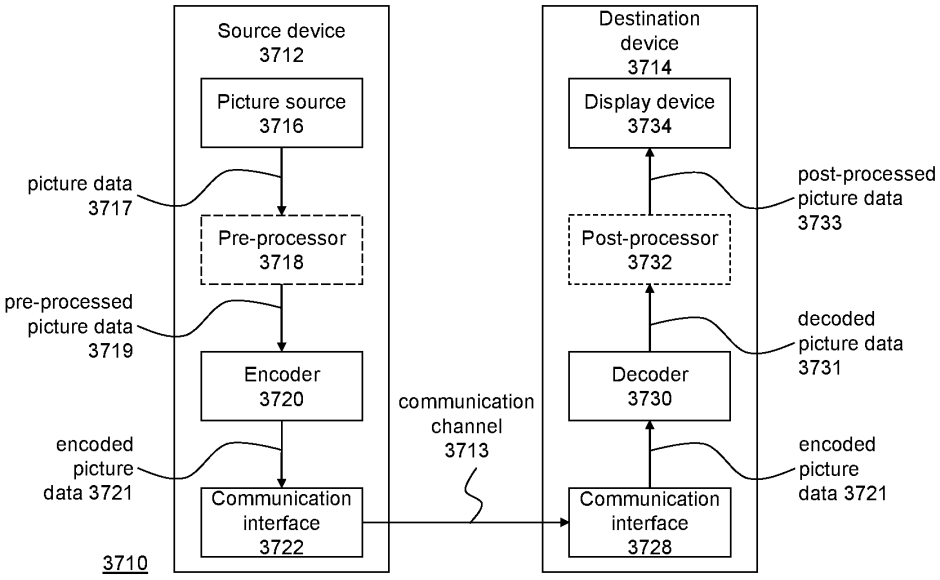

Source device
3712
Picture source
3716
picture data
3717
Pre-processor
3718
pre-processed
picture data
3719
Encoder
3720
encoded
picture
data 3721
Communication
interface
3722
3710
communication
channel
3713
Destination
device
3714
Display device
3734
post-processed
picture data
3733
Post-processor
3732
decoded
picture data
3731
Decoder
3730
encoded
picture
data 3721
Communication
interface
3728

FIG. 38

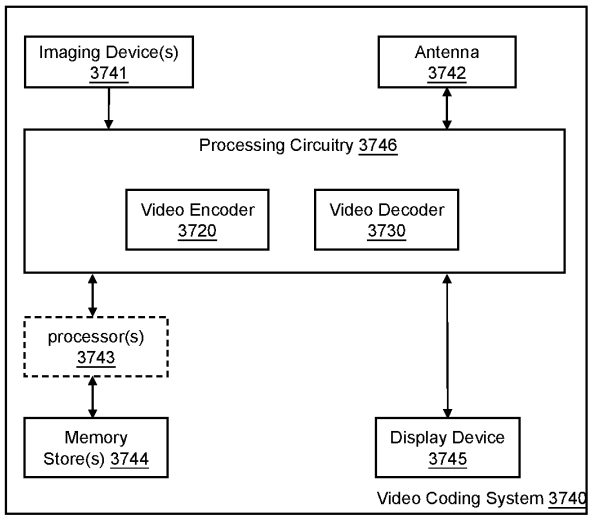

Imaging Device(s)
3741
Antenna
3742
Processing Circuitry 3746
Video Encoder
3720
Video Decoder
3730
processor(s)
3743
Memory
Store(s) 3744
Display Device
3745
Video Coding System 3740

FIG. 39

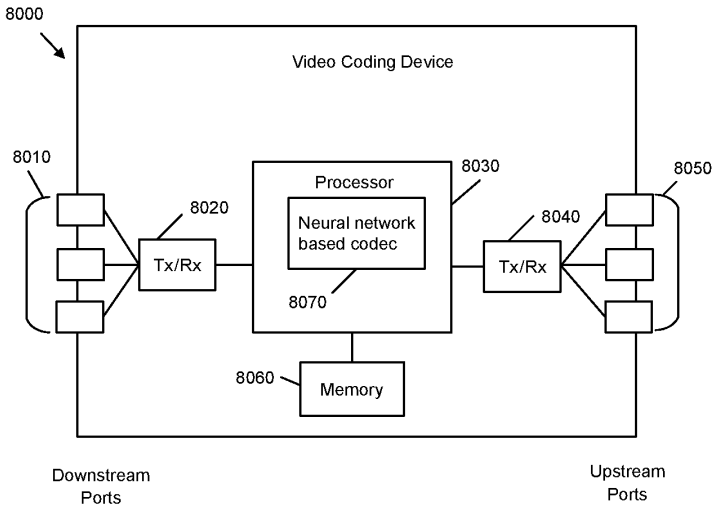

8000
Video Coding Device
8010
8020
Tx/Rx
Processor
Neural network based codec
8070
8030
8040
Tx/Rx
8050
8060
Memory
Downstream Ports
Upstream Ports

FIG. 40

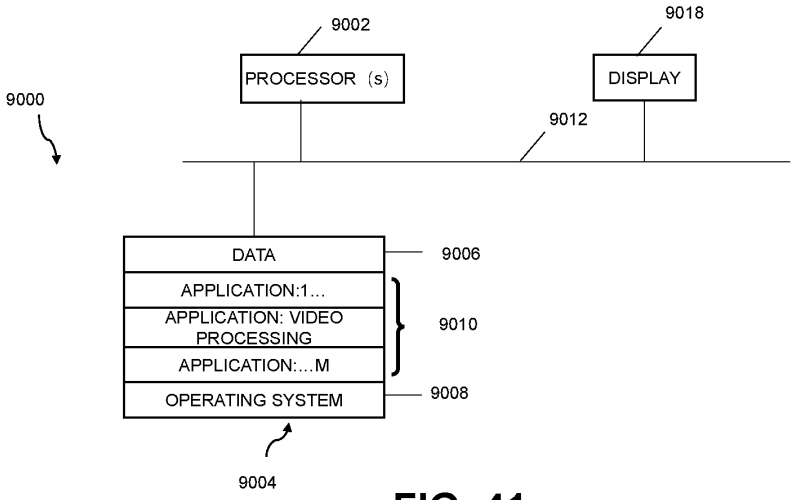

9000
9002
PROCESSOR（s）
9018
DISPLAY
9012
DATA
9006
APPLICATION:1...
APPLICATION: VIDEO PROCESSING
APPLICATION:...M
9010
OPERATING SYSTEM
9008
9004

FIG. 41

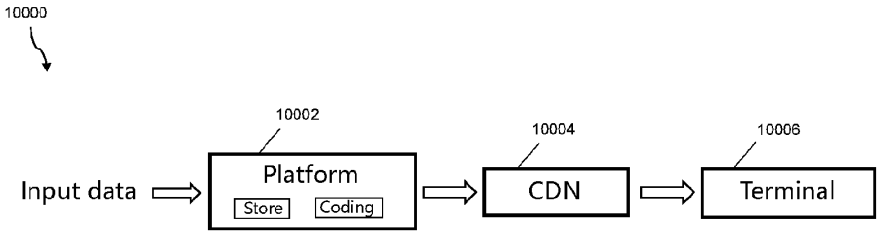

10000
10002
10004
10006
Input data
Platform
Store
Coding
CDN
Terminal

FIG. 42