



(43) International Publication Date  
24 September 2020 (24.09.2020)

(51) International Patent Classification:  
*H04N 19/597* (2014.01)

(21) International Application Number:  
PCT/US2020/016931

(22) International Filing Date:  
06 February 2020 (06.02.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
19305343.6 20 March 2019 (20.03.2019) EP  
19305344.4 20 March 2019 (20.03.2019) EP

(71) Applicant: **INTERDIGITAL VC HOLDINGS, INC.**  
[US/US]; 200 Bellevue Parkway, Suite 300, Wilmington,  
Delaware 19809 (US).

(72) Inventors: **LLACH, Joan**; TECHNICOLOR R&D  
FRANCE, 975 avenue des Champs Blancs, ZAC des  
Champs Blancs - CS 17616, 35576 CESSON-SEVIGNE  
(FR). **RICARD, Julien**; TECHNICOLOR R&D FRANCE,  
975 avenue des Champs Blancs, ZAC des Champs Blancs  
- CS 17616, 35576 CESSON-SEVIGNE (FR). **CHEVET,  
Jean-Claude**; TECHNICOLOR R&D FRANCE, 975 av-

enue des Champs Blancs, ZAC des Champs Blancs - CS  
17616, 35576 CESSON-SEVIGNE (FR).

(74) Agent: **SHEDD, Robert D.**; INTERDIGITAL VC  
HOLDINGS, INC., 200 Bellevue Parkway, Suite 300,  
Wilmington, Delaware 19809 (US).

(81) Designated States (*unless otherwise indicated, for every  
kind of national protection available*): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,  
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,  
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,  
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,  
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,  
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,  
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,  
SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR,  
TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every  
kind of regional protection available*): ARIPO (BW, GH,  
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,  
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,  
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,  
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,  
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: PROCESSING MISSING POINTS OF A POINT CLOUD

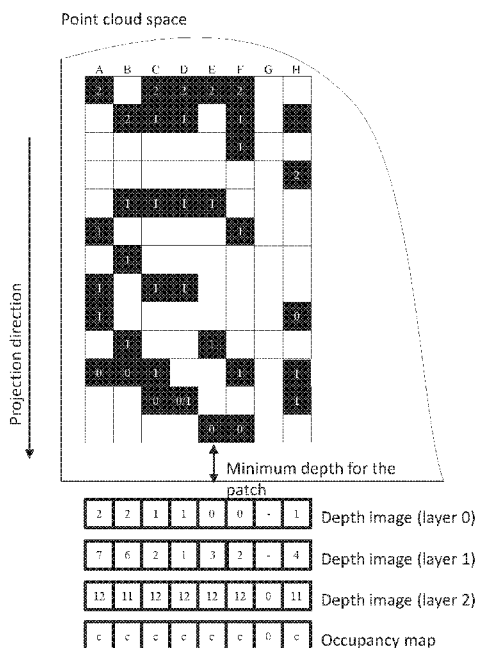


Fig. 10

(57) Abstract: At least one embodiment relates to a method comprising storing a depth value of at least two 3D samples of a point cloud frame in at least two depth images; and determining a minimum depth value and a maximum depth value at a 2D location of a projection plane from depth values stored at said 2D location in said at least two depth images and encoding a depth value of at least one additional 3D sample of the point cloud according to said minimal and maximal depth values, said at least one 3D sample and said at least one additional 3D sample being projected onto the projection plane along a same projection line.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
KM, ML, MR, NE, SN, TD, TG).

**Published:**

— *with international search report (Art. 21(3))*

## PROCESSING MISSING POINTS OF A POINT CLOUD

### Technical Field

5           At least one of the present embodiments relates generally to a processing of a point cloud.

### Background

10           The present section is intended to introduce the reader to various aspects of art, which may be related to various aspects of at least one of the present embodiments that is described and/or claimed below. This discussion is believed to be helpful in providing the reader with background information to facilitate a better understanding of the various aspects of at least one embodiment.

15           Point clouds may be used for various purposes such as culture heritage/buildings in which objects like statues or buildings are scanned in 3D in order to share the spatial configuration of the object without sending or visiting it. Also, it is a way to ensure preserving the knowledge of the object in case it may be destroyed; for instance, a temple by an earthquake. Such point  
20           clouds are typically static, colored and huge.

            Another use case is in topography and cartography in which using 3D representations allows for maps that are not limited to the plane and may include the relief. Google Maps is now a good example of 3D maps but uses meshes instead of point clouds. Nevertheless, point clouds may be a suitable  
25           data format for 3D maps and such point clouds are typically static, colored and huge.

            The automotive industry and the autonomous car are also domains in which point clouds may be used. Autonomous cars should be able to “probe”  
30           their environment to make good driving decisions based on the reality of their immediate neighbors. Typical sensors like LIDARs (Light Detection And Ranging) produce dynamic point clouds that are used by a decision engine. These point clouds are not intended to be viewed by a human being and they

are typically small, not necessarily colored, and dynamic with a high frequency of capture. These point clouds may have other attributes like the reflectance provided by the LIDAR as this attribute provides good information on the material of the sensed object and may help in making decisions.

5           Virtual Reality and immersive worlds have become hot topics recently and are foreseen by many as the future of 2D flat video. The basic idea is to immerse the viewer in an environment that surrounds the viewer, in contrast to a standard TV in which the viewer can only look at the virtual world in front of the viewer. There are several gradations in the immersivity depending on  
10 the freedom of the viewer in the environment. A point cloud is a good format candidate for distributing Virtual Reality (VR) worlds.

It is important in many applications to be able to distribute dynamic point clouds to an end-user (or store them in a server) by consuming only a reasonable amount of bit-rate (or storage space for storage applications) while  
15 maintaining an acceptable (or preferably very good) quality of experience. Efficient compression of these dynamic point clouds is a key point in order to make the distribution chain of many immersive worlds practical.

At least one embodiment has been devised with the foregoing in mind.

## 20 Summary

The following presents a simplified summary of at least one of the present embodiments in order to provide a basic understanding of some aspects of the present disclosure. This summary is not an extensive overview of an embodiment. It is not intended to identify key or critical elements of an  
25 embodiment. The following summary merely presents some aspects of at least one of the present embodiments in a simplified form as a prelude to the more detailed description provided elsewhere in the document.

According to a general aspect of at least one embodiment, there is provided a method comprising storing a depth value of at least two 3D samples  
30 of a point cloud frame in at least two depth images; determining a minimum depth value and a maximum depth value at a 2D location of a projection plane from depth values stored at said 2D location in said at least two depth images

and encoding a depth value of at least one additional 3D sample of the point cloud according to said minimal and maximal depth values, said at least one 3D sample and said at least one additional 3D sample being projected onto the projection plane along a same projection line.

5           According to an embodiment, encoding a depth value of at least one additional 3D sample of the point cloud according to said minimal and maximal depth values comprises encoding at least one additional 3D sample as a bit of a codeword, and skipping the encoding of depth value of a 3D sample projected along the projection line as a bit of the codeword when said depth  
10 value is already encoded as a value of a depth image.

          According to an embodiment, the length of the codeword is limited to a maximum value.

          According to an embodiment, the method further comprises signalling a first information indicating if the length of codeword is limited to a maximum  
15 length or not.

          According to an embodiment, the method further comprises signalling a second information representative of said maximum value.

          According to an embodiment, the method further comprises signalling a third information to indicate if the ordered bits of a codeword starting from  
20 the Least Significant Bit towards the Most Significant Bit represent a list of depth values starting either from the minimum depth value towards the maximal depth value or from the maximal depth value towards the minimal depth value.

          According to an embodiment, the length of the codeword equals at least  
25 a minimum value.

          According to an embodiment, the method further comprises signalling a fourth information indicating if the length of the codeword equals at least a minimum length or not.

          According to an embodiment, the method further comprises signalling  
30 a fifth information representative of said minimum value.

          According to a general aspect of at least one embodiment, there is provided a signal comprising a first information indicating if the length of a

codeword is limited to a maximum length or not, said codeword encoding a depth value of at least one additional 3D sample of a point cloud.

According to an embodiment, the signal further comprises a second information representative of said maximum value.

5           According to an embodiment, the signal further comprises a fourth information indicating if the length of the codeword equals at least a minimum length or not.

According to an embodiment, the signal further comprises a fifth information representative of said minimum value.

10           One or more of at least one of embodiment also provide a device, a computer program product and a non-transitory computer readable medium.

The specific nature of at least one of the present embodiments as well as other objects, advantages, features and uses of said at least one of the present embodiments will become evident from the following description of  
15 examples taken in conjunction with the accompanying drawings.

#### Brief Description of Drawings

In the drawings, examples of several embodiments are illustrated. The drawings show:

20           - **Fig. 1** illustrates a schematic block diagram of an example of a two-layer-based point cloud encoding structure in accordance with at least one of the present embodiments;

              - **Fig. 2** illustrates a schematic block diagram of an example of a two-layer-based point cloud decoding structure in accordance with at least one of  
25 the present embodiments;

              - **Fig. 3** illustrates a schematic block diagram of an example of an image-based point cloud encoder in accordance with at least one of the present embodiments;

              - **Fig. 3a** illustrates an example of a canvas comprising 2 patches and  
30 their 2D bounding boxes;

              - **Fig. 3b** illustrates an example of two in-between 3D samples located between two 3D samples along a projection line;

- **Fig. 4** illustrates a schematic block diagram of an example of an image-based point cloud decoder in accordance with at least one of the present embodiments;

5 - **Fig. 5** illustrates schematically an example of syntax of a bitstream representative of a base layer BL in accordance with at least one of the present embodiments;

- **Fig. 6** illustrates a schematic block diagram of an example of a system in which various aspects and embodiments are implemented;

10 - **Fig. 7** illustrates an example of such a 3D to 2D projection using V-PCC;

- **Fig. 8** illustrates a schematic block diagram of an example of a method for coding the depth values of 3D samples projected along a same projection line in accordance with at least one embodiment;

15 - **Fig. 9** illustrates a schematic block diagram of an example of a method for decoding the depth values of 3D samples projected along a same projection line in accordance with at least one embodiment;

- **Fig. 10** illustrates an example of determining pixel values of occupancy map in accordance with at least one embodiment;

20 - **Fig. 11** illustrates an example of determining pixel values of occupancy map in accordance with a variant;

- **Fig. 12** illustrates an example of a syntax element used for signalling the information INFO1 and INFO2 in accordance with at least one embodiment;

25 - **Fig. 13** illustrates an example of a syntax element used for signalling the information INFO3 in accordance with at least one embodiment; and

- **Fig. 14** illustrates an example of a syntax element used for signalling the information INFO4 and INFO5 in accordance with at least one embodiment.

30 Detailed Description

At least one of the present embodiments is described more fully hereinafter with reference to the accompanying figures, in which examples of

at least one of the present embodiments are shown. An embodiment may, however, be embodied in many alternate forms and should not be construed as limited to the examples set forth herein. Accordingly, it should be understood that there is no intent to limit embodiments to the particular forms disclosed. On the contrary, the disclosure is intended to cover all modifications, 5 equivalents, and alternatives falling within the spirit and scope of this application.

When a figure is presented as a flow diagram, it should be understood that it also provides a block diagram of a corresponding apparatus. Similarly, 10 when a figure is presented as a block diagram, it should be understood that it also provides a flow diagram of a corresponding method/process.

Similar or same elements of figures are referenced with the same reference numbers.

Some figures represent syntax tables widely used in V-PCC for defining 15 the structure of a bitstream that conforms with V-PCC. In those syntax tables, the term ‘...’ denotes unchanged portions of the syntax with respect to the original definition given in V-PCC and removed in the figures to facilitate reading. Bold terms in figures indicate that a value for this term is obtained by parsing a bitstream. The right column of the syntax tables indicates the number 20 of bits for encoding a data of a syntax element. For example,  $u(4)$  indicates that 4 bits are used for encoding a data,  $u(8)$  indicates 8 bits,  $ae(v)$  indicates a context adaptive arithmetic entropy coded syntax element.

The aspects described and contemplated below may be implemented in many different forms. **Figs. 1-14** below provide some embodiments, but 25 other embodiments are contemplated, and the discussion of **Figs. 1-14** does not limit the breadth of the implementations.

At least one of the aspects generally relates to point cloud encoding and decoding, and at least one other aspect generally relates to transmitting a bitstream generated or encoded.

30 More precisely, various methods and other aspects described herein may be used to modify modules, for example, the module PGM (step 3100), the encoder OMENC (step 3600), the geometry generation module GGM (step

4300), the occupancy map decoder OMDEC (step 4400) as shown in **Figs. 1-14**.

Moreover, the present aspects are not limited to MPEG standards such as MPEG-I part 5 that relates to the Point Cloud Compression, and may be applied, for example, to other standards and recommendations, whether pre-existing or future-developed, and extensions of any such standards and recommendations (including MPEG-I part 5). Unless indicated otherwise, or technically precluded, the aspects described in this application may be used individually or in combination.

10 In the following, image data refer to data, for example, one or several arrays of 2D samples in a specific image/video format. A specific image/video format may specify information pertaining to pixel values of an image (or a video). A specific image/video format may also specify information which may be used by a display and/or any other apparatus to visualize and/or decode an image (or video) for example. An image typically includes a first component, in the shape of a first 2D array of samples, usually representative of luminance (or luma) of the image. An image may also include a second component and a third component, in the shape of other 2D arrays of samples, usually representative of the chrominance (or chroma) of the image. Some  
15  
20 embodiments represent the same information using a set of 2D arrays of color samples, such as the traditional tri-chromatic RGB representation.

A pixel value is represented in one or more embodiments by a vector of **C** values, where **C** is the number of components. Each value of a vector is typically represented with a number of bits which may define a dynamic range  
25 of the pixel values.

An image block means a set of pixels which belong to an image. The pixel values of an image block (or image block data) refer to the values of the pixels which belong to this image block. An image block may have an arbitrary shape, although rectangles are common.

30 A point cloud may be represented by a dataset of 3D samples within a 3D volumetric space that have unique coordinates and that may also have one or more attributes.

A 3D sample of this data set may be defined by its spatial location ( $X$ ,  $Y$ , and  $Z$  coordinates in a 3D space) and possibly by one or more associated attributes such as a color, represented in the RGB or YUV color space for example, a transparency, a reflectance, a two-component normal vector or any  
5 feature representing a feature of this sample. For example, a 3D sample may be defined by 6 components ( $X$ ,  $Y$ ,  $Z$ ,  $R$ ,  $G$ ,  $B$ ) or equivalently ( $X$ ,  $Y$ ,  $Z$ ,  $y$ ,  $U$ ,  $V$ ) where ( $X$ ,  $Y$ ,  $Z$ ) defines the coordinates of a point in a 3D space and ( $R$ ,  $G$ ,  $B$ ) or ( $y$ ,  $U$ ,  $V$ ) defines a color of this 3D sample. The same type of attribute may be present multiple times. For example, multiple color attributes may provide color  
10 information from different points of view.

A point cloud may be static or dynamic depending on whether or not the cloud changes with respect to time. A static point cloud or an instance of a dynamic point cloud is usually denoted as a point cloud frame. It should be noticed that in the case of a dynamic point cloud, the number of points is  
15 generally not constant but, on the contrary, generally changes with time. More generally, a point cloud may be considered as dynamic if anything changes with time, such as, for example, the number of points, the position of one or more points, or any attribute of any point.

As an example, a 2D sample may be defined by 6 components ( $u$ ,  $v$ ,  $Z$ ,  $R$ ,  $G$ ,  $B$ ) or equivalently ( $u$ ,  $v$ ,  $Z$ ,  $y$ ,  $U$ ,  $V$ ). ( $u$ ,  $v$ ) defines the coordinates of a 2D  
20 sample in a 2D space of the projection plane.  $Z$  is the depth value of a projected 3D sample onto this projection plane. ( $R$ ,  $G$ ,  $B$ ) or ( $y$ ,  $U$ ,  $V$ ) defines a color of this 3D sample.

**Fig. 1** illustrates a schematic block diagram of an example of a two-  
25 layer-based point cloud encoding structure 1000 in accordance with at least one of the present embodiments.

The two-layer-based point cloud encoding structure 1000 may provide a bitstream  $B$  representative of an input point cloud frame IPCF. Possibly, said input point cloud frame IPCF represents a frame of a dynamic point cloud.  
30 Then, a frame of said dynamic point cloud may be encoded by the two-layer-based point cloud encoding structure 1000 independently from another frame.

Basically, the two-layer-based point cloud encoding structure 1000 may provide ability to structure the bitstream B as a Base Layer BL and an Enhancement Layer EL. The base layer BL may provide a lossy representation of an input point cloud frame IPCF and the enhancement layer EL may provide  
5 a higher quality (possibly lossless) representation by encoding isolated points not represented by the base layer BL.

The base layer BL may be provided by an image-based encoder 3000 as illustrated in **Fig. 3**. Said image-based encoder 3000 may provide geometry/texture images representing the geometry/attributes of 3D samples  
10 of the input point cloud frame IPCF. It may allow isolated 3D samples to be discarded. The base layer BL may be decoded by an image-based decoder 4000 as illustrated in **Fig. 4** that may provide an intermediate reconstructed point cloud frame IRPCF.

Then, back to the two-layer-based point cloud encoding 1000 in **Fig. 1**,  
15 a comparator COMP may compare the 3D samples of the input point cloud frame IPCF to the 3D samples of the intermediate reconstructed point cloud frame IRPCF in order to detect/locate missed/isolated 3D samples. Next, an encoder ENC may encode the missed 3D samples and may provide the enhancement layer EL. Finally, the base layer BL and the enhancement layer  
20 EL may be multiplexed together by a multiplexer MUX so as to generate the bitstream B.

According to an embodiment, the encoder ENC may comprise a detector that may detect and associate a 3D reference sample R of the intermediate reconstructed point cloud frame IRPCF to a missed 3D samples  
25 M.

For example, a 3D reference sample R associated with a missed 3D sample M may be its nearest neighbor of M according to a given metric.

According to an embodiment, the encoder ENC may then encode the spatial locations of the missed 3D samples M and their attributes as differences  
30 determined according to spatial locations and attributes of said 3D reference samples R.

In a variant, those differences may be encoded separately.

For example, for a missed 3D sample M, with spatial coordinates  $x(M)$ ,  $y(M)$  and  $z(M)$ , a x-coordinate position difference  $Dx(M)$ , a y-coordinate position difference  $Dy(M)$ , a z-coordinate position difference  $Dz(M)$ , a R-attribute component difference  $Dr(M)$ , a G-attribute component difference  $Dg(M)$  and the B-attribute component difference  $Db(M)$  may be calculated as follows:

$$Dx(M)=x(M)-x(R),$$

where  $x(M)$  is the x-coordinate of the 3D sample M, respectively R in a geometry image provided by **Fig. 3**,

$$Dy(M)=y(M)-y(R)$$

where  $y(M)$  is the y-coordinate of the 3D sample M, respectively R in a geometry image provided by **Fig. 3**,

$$Dz(M)=z(M)-z(R)$$

where  $z(M)$  is the z-coordinate of the 3D sample M, respectively R in a geometry image provided by **Fig. 3**,

$$Dr(M)=R(M)-R(R).$$

where  $R(M)$ , respectively  $R(R)$  is the r-color component of a color attribute of the 3D sample M, respectively R,

$$Dg(M)=G(M)-G(R).$$

where  $G(M)$ , respectively  $G(R)$  is the g-color component of a color attribute of the 3D sample M, respectively R,

$$Db(M)=B(M)-B(R).$$

where  $B(M)$ , respectively  $B(R)$  is the b-color component of a color attribute of the 3D sample M, respectively R.

**Fig. 2** illustrates a schematic block diagram of an example of a two-layer-based point cloud decoding structure 2000 in accordance with at least one of the present embodiments.

The behavior of the two-layer-based point cloud decoding structure 2000 depends on its capabilities.

A two-layer-based point cloud decoding structure 2000 with limited capabilities may access only the base layer BL from the bitstream B by using a de-multiplexer DMUX, and then may provide a faithful (but lossy) version

IRPCF of the input point cloud frame IPCF by decoding the base layer BL by a point cloud decoder 4000 as illustrated in **Fig. 4**.

A two-layer-based point cloud decoding structure 2000 with full capabilities may access both the base layer BL and the enhancement layer EL from the bitstream B by using the de-multiplexer DMUX. The point cloud decoder 4000, as illustrated in **Fig. 4**, may determine the intermediate reconstructed point cloud frame IRPCF from the base layer BL. The decoder DEC may determine a complementary point cloud frame CPCF from the enhancement layer EL. A combiner COM then may combine together the intermediate reconstructed point cloud frame IRPCF and the complementary point cloud frame CPCF to therefore provide a higher quality (possibly lossless) representation (reconstruction) CRPCF of the input point cloud frame IPCF.

**Fig. 3** illustrates a schematic block diagram of an example of an image-based point cloud encoder 3000 in accordance with at least one of the present embodiments.

The image-based point cloud encoder 3000 leverages existing video codecs to compress the geometry and texture (attribute) information of a dynamic point cloud. This is accomplished by essentially converting the point cloud data into a set of different video sequences.

In particular embodiments, two videos, one for capturing the geometry information of the point cloud data and another for capturing the texture information, may be generated and compressed using existing video codecs. An example of an existing video codec is the HEVC Main profile encoder/decoder (ITU-T H.265 Telecommunication standardization sector of ITU (02/2018), series H: audiovisual and multimedia systems, infrastructure of audiovisual services - coding of moving video, High efficiency video coding, Recommendation ITU-T H.265).

Additional metadata that are used to interpret the two videos are typically also generated and compressed separately. Such additional metadata includes, for example, an occupancy map OM and/or auxiliary patch information PI.

The generated video bitstreams and the metadata may be then multiplexed together so as to generate a combined bitstream.

It should be noted that the metadata typically represents a small amount of the overall information. The bulk of the information is in the video bitstreams.

5 An example of such a point cloud coding/decoding process is given by the Test model Category 2 algorithm (also denoted V-PCC) that implements the MPEG draft standard as defined in ISO/IEC JTC1/SC29/WG11 MPEG2019/w18180 (January 2019, Marrakesh).

10 In step 3100, a module PGM may generate at least one patch by decomposing 3D samples of a data set representative of the input point cloud frame IPCF to 2D samples on a projection plane using a strategy that provides best compression.

A patch may be defined as a set of 2D samples.

15 For example, in V-PCC, a normal at every 3D sample is first estimated as described, for example, in Hoppe et al. (Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, Werner Stuetzle. Surface reconstruction from unorganized points. *ACM SIGGRAPH 1992 Proceedings*, 71-78). Next, an initial clustering of the input point cloud frame IPCF is obtained by associating each 3D sample with one of the six oriented planes of a 3D bounding box  
20 encompassing the 3D samples of the input point cloud frame IPCF. More precisely, each 3D sample is clustered and associated with an oriented plane that has the closest normal (that is maximizes the dot product of the point normal and the plane normal). Then the 3D samples are projected to their associated planes. A set of 3D samples that forms a connected area in their  
25 plane is referred as a connected component. A connected component is a set of at least one 3D sample having similar normal and a same associated oriented plane. The initial clustering is then refined by iteratively updating the cluster associated with each 3D sample based on its normal and the clusters of its nearest neighboring samples. The final step consists of generating one  
30 patch from each connected component, that is done by projecting the 3D samples of each connected component onto the oriented plane associated with said connected component. A patch is associated with auxiliary patch

information PI that represents auxiliary patch information defined for each patch to interpret the projected 2D samples that correspond to the geometry and/or attribute information.

In V-PCC, for example, the auxiliary patch information PI includes 1) information indicating one of the six oriented planes of a 3D bounding box encompassing the 3D samples of a connected component; 2) information relative to the plane normal; 3) information determining the 3D location of a connected component relative to a patch represented in terms of depth, tangential shift and bi-tangential shift; and 4) information such as coordinates (u0, v0, u1, v1) in a projection plane defining a 2D bounding box encompassing a patch.

In step 3200, a patch packing module PPM may map (place) at least one generated patch onto a 2D grid (also called canvas) without any overlapping in a manner that typically minimizes the unused space, and may guarantee that every TxT (for example, 16x16) block of the 2D grid is associated with a unique patch. A given minimum block size TxT of the 2D grid may specify the minimum distance between distinct patches as placed on this 2D grid. The 2D grid resolution may depend on the input point cloud size and its width W and height H and the block size T may be transmitted as metadata to the decoder.

The auxiliary patch information PI may further include information relative to an association between a block of the 2D grid and a patch.

In V-PCC, the auxiliary information PI may include a block to patch index information (*BlockToPatch*) that determines an association between a block of the 2D grid and a patch index.

**Fig. 3a** illustrates an example of a canvas C comprising 2 patches P1 and P2 and their associated 2D bounding boxes B1 and B2. Note that two bounding boxes may overlap in the canvas C as illustrated on **Fig. 3a**. The 2D grid (the splitting of the canvas) is only represented inside the bounding box but the splitting of the canvas also occurs outside those bounding boxes. A bounding box associated with a patch can be split into TxT blocks, typically T=16.

TxT blocks containing 2D samples belonging to a patch may be considered as *occupied* blocks. Each *occupied* block of the canvas is represented by a particular pixel value (for example 1) in the occupancy map OM and each *unoccupied* block of the canvas is represented by another particular value, for example 0. Then, a pixel value of the occupancy map OM may indicate whether a TxT block of the canvas is *occupied*, that is contains 2D samples that belong to a patch.

In **Fig. 3a**, an *occupied* block is represented by a white block and light grey blocks represent *unoccupied* blocks. The image generation processes (steps 3300 and 3400) exploit the mapping of the at least one generated patch onto the 2D grid computed during step 3200, to store the geometry and texture of the input point cloud frame IPCF as images.

In step 3300, a geometry image generator GIG may generate at least one geometry image GI from the input point cloud frame IPCF, the occupancy map OM and the auxiliary patch information PI. The geometry image generator GIG may exploit the occupancy map information in order to detect (locate) the occupied blocks and thus the non-empty pixels in the geometry image GI.

A geometry image GI may represent the geometry of the input point cloud frame IPCF and may be a monochromatic image of WxH pixels represented, for example, in YUV420-8bit format.

In order to better handle the case of multiple 3D samples being projected (mapped) to a same 2D sample of the projection plane (along a same projection direction (line)), multiple images, referred to as layers, may be generated. Thus, different depth values  $D_1, \dots, D_n$  may be associated with a 2D sample of a patch and multiple geometry images may then be generated.

In V-PCC, 2D samples of a patch are projected onto two layers. A first layer, also called the near layer, stores the depth values  $D_0$  associated with the 2D samples with smaller depths. A second layer, referred to as the far layer, stores the depth values  $D_1$  associated with the 2D samples with larger depths. For example, the information stored by the second depth image may be within an interval  $[0, \Delta]$  corresponding to depth values in the range  $[D_0,$

$D_0 + \Delta]$ , where  $\Delta$  is a user-defined parameter that describes the surface thickness.

By this way, the second layer may contain significant contour-like high frequency features. Thus, it clearly appears that the second depth image may be difficult to code by using a legacy video coder and, therefore, the depth values may be poorly reconstructed from said decoded second depth image, which results on a poor quality of the geometry of the reconstructed point cloud frame.

According to an embodiment, the geometry image generating module GIG may code (derive) depth values associated with 2D samples of the first and second layers by using the auxiliary patch information PI.

In V-PCC, the location of a 3D sample in a patch with a corresponding connected component may be expressed in terms of depth  $\delta(u, v)$ , tangential shift  $s(u, v)$  and bi-tangential shift  $r(u, v)$  as follows:

$$\begin{aligned} \delta(u, v) &= \delta_0 + g(u, v) \\ s(u, v) &= s_0 - u_0 + u \\ r(u, v) &= r_0 - v_0 + v \end{aligned}$$

where  $g(u, v)$  is the luma component of the geometry image,  $(u, v)$  is a pixel associated with the 3D sample on a projection plane,  $(\delta_0, s_0, r_0)$  is the 3D location of the corresponding patch of a connected component to which the 3D sample belongs and  $(u_0, v_0, u_1, v_1)$  are the coordinates in said projection plane defining a 2D bounding box encompassing the projection of the patch associated with said connected component.

Thus, a geometry image generating module GIG may code (derive) depth values associated with 2D samples of a layer (first or second or both) as a luma component  $g(u, v)$  given by:  $g(u, v) = \delta(u, v) - \delta_0$ . It is noted that this relationship may be employed to reconstruct 3D sample locations  $(\delta_0, s_0, r_0)$  from a reconstructed geometry image  $g(u, v)$  with the accompanying auxiliary patch information PI.

According to an embodiment, a projection mode may be used to indicate if a first geometry image GI0 may store the depth values of the 2D samples of either the first or second layer and a second geometry image GI1

may store the depth values associated with the 2D samples of either the second or first layer.

For example, when a projection mode equals 0, then the first geometry image GI0 may store the depth values of 2D samples of the first layer and the second geometry image GI1 may store the depth values associated with 2D samples of the second layer. Reciprocally, when a projection mode equals 1, then the first geometry image GI0 may store the depth values of 2D samples of the second layer and the second geometry image GI1 may store the depth values associated with 2D samples of the first layer.

10 According to an embodiment, a frame projection mode may be used to indicate if a fixed projection mode is used for all the patches or if a variable projection mode is used in which each patch may use a different projection mode.

The projection mode and/or the frame projection mode may be transmitted as metadata.

A frame projection mode decision algorithm may be provided, for example, in section 2.2.1.3.1 of V-PCC.

20 According to an embodiment, when the frame projection indicates that a variable projection mode may be used, a patch projection mode may be used to indicate the appropriate mode to use to (de-)project a patch.

A patch projection mode may be transmitted as metadata and may be, possibly, an information included in the auxiliary patch information PI.

A patch projection mode decision algorithm is provided, for example in section 2.2.1.3.2 of V-PCC.

25 According to an embodiment of step 3300, the pixel value in a first geometry image, for example GI0, corresponding to a 2D sample (u,v) of a patch, may represent the depth value of least one in-between 3D sample defined along a projection line corresponding to said 2D sample (u,v). More precisely, said in-between 3D samples reside along a projection line and share the same coordinates of the 2D sample (u,v) whose depth value D1 is coded in a second geometry image, for example GI1. Further, the said in-between 3D samples may have depth values between the depth value D0 and a depth

30

value  $D_1$ . A designated bit may be associated with each said in-between 3D samples which is set to 1 if the in-between 3D sample exists and 0 otherwise.

**Fig. 3b** illustrates an example of two in-between 3D samples  $P_{i1}$  and  $P_{i2}$  located between two 3D samples  $P_0$  and  $P_1$  along a projection line  $PL$ . The 3D samples  $P_0$  and  $P_1$  have respectively depth values equal to  $D_0$  and  $D_1$ . The depth values  $D_{i1}$  and  $D_{i2}$  of respectively the two in-between 3D samples  $P_{i1}$  and  $P_{i2}$  are greater than  $D_0$  and lower than  $D_1$ .

Then, all said designated bits along said projection line may be concatenated to form a codeword, denoted Enhanced-Occupancy map (EOM) codeword hereafter. As illustrated in **Fig. 3b**, assuming an EOM codeword of 8 bits of length, 2 bits equal 1 to indicate the location of the two 3D samples  $P_{i1}$  and  $P_{i2}$ . Finally, all the EOM codewords may be packed in an image, for example, the occupancy map  $OM$ . In that case, at least one patch of the canvas may contain at least one EOM codeword. Such a patch is denoted reference patch and a block of a reference patch is denoted a EOM reference block. Thus, a pixel value of the occupancy map  $OM$  may equal to a first value, for example 0, to indicate an *unoccupied* block of the canvas, or another value, for example greater than 0, to indicate either an *occupied* block or a EOM reference block of the canvas.

According to an example, when  $D_1 - D_0 \leq 1$  then the pixel value of the occupancy map  $OM$  may indicate an *occupied* block and when  $D_1 - D_0 > 1$  then the pixel value map  $OM$  may indicate an EOM reference block of the canvas.

The locations of pixels in the occupancy map  $OM$  that indicates EOM reference blocks and the values of the bits of a EOM codeword that are obtained from the values of those pixels, indicate the 3D coordinates of the in-between 3D samples.

In step 3400, a texture image generator TIG may generate at least one texture image  $TI$  from the input point cloud frame IPCF, the occupancy map  $OM$ , the auxiliary patch information  $PI$  and a geometry of a reconstructed point cloud frame derived from at least one decoded geometry image DGI, output of a video decoder VDEC (step 4200 in **Fig. 4**).

A texture image TI may represent the texture of the input point cloud frame IPCF and may be an image of WxH pixels represented, for example, in YUV420-8bit format.

5 The texture image generator TG may exploit the occupancy map information in order to detect (locate) the occupied blocks and thus the non-empty pixels in the texture image.

The texture image generator TIG may be adapted to generate and associate a texture image TI with each geometry image/layer DGI.

10 According to an embodiment, the texture image generator TIG may code (store) the texture (attribute) values T0 associated with 2D samples of the first layer as pixel values of a first texture image TI0 and the texture values T1 associated with the 2D samples of the second layer as pixel values of a second texture image TI1.

15 Alternatively, the texture image generating module TIG may code (store) the texture values T1 associated with 2D samples of the second layer as pixel values of the first texture image TI0 and the texture values D0 associated with the 2D samples of the first layer as pixel values of the second geometry image GI1.

20 For example, colors of 3D samples may be obtained as described in section 2.2.3, 2.2.4, 2.2.5, 2.2.8 or 2.5 of V-PCC.

The texture values of two 3D samples are stored in either the first or second texture images. But, the texture values of in-between 3D samples cannot be stored neither in this first texture image TI0 nor the second texture image TI1 because the locations of the projected in-between 3D samples correspond to *occupied* blocs that are already used for storing a texture value of another 3D sample (P0 or P1) as illustrated in **Fig. 3b**. The texture values of in-between 3D samples are thus stored in EOM texture blocks located elsewhere in either the first or second texture image in locations procedurally defined (section 9.4.5 of V-PCC). In brief, this process determines locations of *unoccupied* blocks in a texture image and stored the texture values associated with in-between 3D samples as pixel values of said *unoccupied* blocks of the texture image, denoted EOM texture blocks.

25  
30

According to an embodiment, a padding process may be applied on the geometry and/or texture image. The padding process may be used to fill empty space between patches to generate a piecewise smooth image suited for video compression.

5 An image padding example is provided in sections 2.2.6 and 2.2.7 of V-PCC.

In step 3500, a video encoder VENC may encode the generated images /layers TI and GI.

10 In step 3600, an encoder OMENC may encode the occupancy map as an image as detailed, for example, in section 2.2.2 of V-PCC. Lossy or lossless encoding may be used.

According to an embodiment, the video encoder ENC and/or OMENC may be a HEVC-based encoder.

15 In step 3700, an encoder PIENC may encode the auxiliary patch information PI and possibly additional metadata such as the block size T, the width W and height H of the geometry/texture images.

According to an embodiment, the auxiliary patch information may be differentially encoded (as defined, for example in section 2.4.1 of V-PCC).

20 In step 3800, a multiplexer may be applied to the generated outputs of steps 3500, 3600 and 3700, and as a result these outputs may be multiplexed together so as to generate a bitstream representative of the base layer BL. It should be noted that the metadata information represents a small fraction of the overall bitstream. The bulk of the information is compressed using the video codecs.

25 **Fig. 4** illustrates a schematic block diagram of an example of an image-based point cloud decoder 4000 in accordance with at least one of the present embodiments.

In step 4100, a de-multiplexer DMUX may applied to demultiplex the encoded information of the bitstream representative of the base layer BL.

30 In step 4200, a video decoder VDEC may decode encoded information to derive at least one decoded geometry image DGI and at least one decoded texture image DTI.

In step 4300, a decoder OMDEC may decode encoded information to derive a decoded occupancy map DOM.

According to an embodiment, the video decoder VDEC and/or OMDEC may be a HEVC-based decoder.

5 In step 4400, a decoder PIDEDEC may decode encoded information to derive auxiliary patch information DPI.

Possibly, metadata may also be derived from the bitstream BL.

In step 4500, a geometry generating module GGM may derive the geometry RG of a reconstructed point cloud frame IRPCF from the at least one  
10 decoded geometry image DGI, the decoded occupancy map DOM, the decoded auxiliary patch information DPI and possible additional metadata.

The geometry generating module GGM may exploit the decoded occupancy map information DOM in order to locate the non-empty pixels in the at least one decoded geometry image DGI.

15 Said non-empty pixels belong to either *occupied* blocks or EOM reference blocks depending on pixels values of the decoded occupancy information DOM and value of D1-D0 as explained above.

According to an embodiment of step 4500, the geometry generating module GGM may derive two of the 3D coordinates of in-between 3D samples  
20 from coordinates of non-empty pixels.

According to an embodiment of step 4500, when said non-empty pixels belong to said EOM reference block, the geometry generating module GGM may derive the third of the 3D coordinates of in-between 3D samples from the bit values of the EOM codewords.

25 For example, according to the example of **Fig. 3b**, the EOM codeword EOMC is used for determining the 3D coordinates of in-between 3D samples  $P_{i1}$  and  $P_{i2}$ . The third coordinate of the in-between 3D sample  $P_{i1}$  may be derived, for example, from D0 by  $D_{i1}=D0+3$  and the third coordinate of the reconstructed 3D sample  $P_{i2}$  may be derived, for example, from D0 by  
30  $D_{i2}=D0+5$ . The offset value (3 or 5) is the number of intervals between D0 and D1 along the projection line.

According to an embodiment, when said non-empty pixels belong to an *occupied* block, the geometry generating module GGM may derive the 3D coordinates of reconstructed 3D samples from coordinates of non-empty pixels, values of said non-empty pixels of one of the at least one decoded geometry image DGI, the decoded auxiliary patch information, and possibly,  
 5 from additional metadata.

The use of non-empty pixels is based on 2D pixel relationship with 3D samples. For example, with the said projection in V-PCC, the 3D coordinates of reconstructed 3D samples may be expressed in terms of depth  $\delta(u, v)$ ,  
 10 tangential shift  $s(u, v)$  and bi-tangential shift  $r(u, v)$  as follows:

$$\delta(u, v) = \delta_0 + g(u, v)$$

$$s(u, v) = s_0 - u_0 + u$$

$$r(u, v) = r_0 - v_0 + v$$

where  $g(u, v)$  is the luma component of a decoded geometry image DGI,  $(u, v)$   
 15 is a pixel associated with a reconstructed 3D sample,  $(\delta_0, s_0, r_0)$  is the 3D location of a connected component to which the reconstructed 3D sample belongs and  $(u_0, v_0, u_1, v_1)$  are the coordinates in a projection plane defining a 2D bounding box encompassing the projection of a patch associate with said connected component.

20 In step 4600, a texture generating module TGM may derive the texture of the reconstructed point cloud frame IRPCF from the geometry RG and the at least one decoded texture image DTI.

According to an embodiment of step 4600, the texture generating module TGM may derive the texture of non-empty pixels that belong to a EOM  
 25 reference block from a corresponding EOM texture block. The locations of a EOM texture blocks in a texture image are procedurally defined (section 9.4.5 of V-PCC)

According to an embodiment of step 4600, the texture generating module TGM may derive the texture of non-empty pixels that belong to an  
 30 *occupied* block directly as pixel values of either the first or second texture image.

**Fig. 5** illustrates schematically an example syntax of a bitstream representative of a base layer BL in accordance with at least one of the present embodiments.

The bitstream comprises a Bitstream Header SH and at least one Group  
5 Of Frame Stream GOFS.

A group of frame stream GOFS comprises a header HS, at least one  
syntax element OMS representative of an occupancy map OM, at least one  
syntax element GVS representative of at least one geometry image (or video),  
at least one syntax element TVS representative of at least one texture image  
10 (or video) and at least one syntax element PIS representative of auxiliary patch  
information and other additional metadata.

In a variant, a group of frame stream GOFS comprises at least one  
frame stream.

**Fig. 6** shows a schematic block diagram illustrating an example of a  
15 system in which various aspects and embodiments are implemented.

System 6000 may be embodied as one or more devices including the  
various components described below and is configured to perform one or more  
of the aspects described in this document. Examples of equipment that may  
form all or part of the system 6000 include personal computers, laptops,  
20 smartphones, tablet computers, digital multimedia set top boxes, digital  
television receivers, personal video recording systems, connected home  
appliances, connected vehicles and their associated processing systems,  
head mounted display devices (HMD, see-through glasses), projectors  
(beamers), "caves" (system including multiple displays), servers, video  
25 encoders, video decoders, post-processors processing output from a video  
decoder, pre-processors providing input to a video encoder, web servers, set-  
top boxes, and any other device for processing a point cloud, a video or an  
image or other communication devices. Elements of system 6000, singly or in  
combination, may be embodied in a single integrated circuit, multiple ICs,  
30 and/or discrete components. For example, in at least one embodiment, the  
processing and encoder/decoder elements of system 6000 may be distributed  
across multiple ICs and/or discrete components. In various embodiments, the

system 6000 may be communicatively coupled to other similar systems, or to other electronic devices, via, for example, a communications bus or through dedicated input and/or output ports. In various embodiments, the system 6000 may be configured to implement one or more of the aspects described in this document.

The system 6000 may include at least one processor 6010 configured to execute instructions loaded therein for implementing, for example, the various aspects described in this document. Processor 6010 may include embedded memory, input output interface, and various other circuitries as known in the art. The system 6000 may include at least one memory 6020 (for example a volatile memory device and/or a non-volatile memory device). System 6000 may include a storage device 6040, which may include non-volatile memory and/or volatile memory, including, but not limited to, Electrically Erasable Programmable Read-Only Memory (EEPROM), Read-Only Memory (ROM), Programmable Read-Only Memory (PROM), Random Access Memory (RAM), Dynamic Random Access Memory (DRAM), Static Random Access Memory (SRAM), flash, magnetic disk drive, and/or optical disk drive. The storage device 6040 may include an internal storage device, an attached storage device, and/or a network accessible storage device, as non-limiting examples.

The system 6000 may include an encoder/decoder module 6030 configured, for example, to process data to provide encoded data or decoded data, and the encoder/decoder module 6030 may include its own processor and memory. The encoder/decoder module 6030 may represent module(s) that may be included in a device to perform the encoding and/or decoding functions. As is known, a device may include one or both of the encoding and decoding modules. Additionally, encoder/decoder module 6030 may be implemented as a separate element of system 6000 or may be incorporated within processor 6010 as a combination of hardware and software as known to those skilled in the art.

Program code to be loaded onto processor 6010 or encoder/decoder 6030 to perform the various aspects described in this document may be stored

in storage device 6040 and subsequently loaded onto memory 6020 for execution by processor 6010. In accordance with various embodiments, one or more of processor 6010, memory 6020, storage device 6040, and encoder/decoder module 6030 may store one or more of various items during the performance of the processes described in this document. Such stored items may include, but are not limited to, a point cloud frame, encoded/decoded geometry/texture videos/images or portions of the encoded/decoded geometry/texture video/images, a bitstream, matrices, variables, and intermediate or final results from the processing of equations, formulas, operations, and operational logic.

In several embodiments, memory inside of the processor 6010 and/or the encoder/decoder module 6030 may be used to store instructions and to provide working memory for processing that may be performed during encoding or decoding.

In other embodiments, however, a memory external to the processing device (for example, the processing device may be either the processor 6010 or the encoder/decoder module 6030) may be used for one or more of these functions. The external memory may be the memory 6020 and/or the storage device 6040, for example, a dynamic volatile memory and/or a non-volatile flash memory. In several embodiments, an external non-volatile flash memory may be used to store the operating system of a television. In at least one embodiment, a fast external dynamic volatile memory such as a RAM may be used as working memory for video coding and decoding operations, such as for MPEG-2 part 2 (also known as ITU-T Recommendation H.262 and ISO/IEC 13818-2, also known as MPEG-2 Video), HEVC (High Efficiency Video coding), or VVC (Versatile Video Coding).

The input to the elements of system 6000 may be provided through various input devices as indicated in block 6130. Such input devices include, but are not limited to, (i) an RF portion that may receive an RF signal transmitted, for example, over the air by a broadcaster, (ii) a Composite input terminal, (iii) a USB input terminal, and/or (iv) an HDMI input terminal.

In various embodiments, the input devices of block 6130 may have

associated respective input processing elements as known in the art. For example, the RF portion may be associated with elements necessary for (i) selecting a desired frequency (also referred to as selecting a signal, or band-limiting a signal to a band of frequencies), (ii) down-converting the selected  
5 signal, (iii) band-limiting again to a narrower band of frequencies to select (for example) a signal frequency band which may be referred to as a channel in certain embodiments, (iv) demodulating the down-converted and band-limited  
10 include one or more elements to perform these functions, for example, frequency selectors, signal selectors, band-limiters, channel selectors, filters, downconverters, demodulators, error correctors, and de-multiplexers. The RF portion may include a tuner that performs various of these functions, including, for example, down-converting the received signal to a lower frequency (for  
15 example, an intermediate frequency or a near-baseband frequency) or to baseband.

In one set-top box embodiment, the RF portion and its associated input processing element may receive an RF signal transmitted over a wired (for example, cable) medium. Then, the RF portion may perform frequency  
20 selection by filtering, down-converting, and filtering again to a desired frequency band.

Various embodiments rearrange the order of the above-described (and other) elements, remove some of these elements, and/or add other elements performing similar or different functions.

25 Adding elements may include inserting elements in between existing elements, such as, for example, inserting amplifiers and an analog-to-digital converter. In various embodiments, the RF portion may include an antenna.

Additionally, the USB and/or HDMI terminals may include respective interface processors for connecting system 6000 to other electronic devices  
30 across USB and/or HDMI connections. It is to be understood that various aspects of input processing, for example, Reed-Solomon error correction, may be implemented, for example, within a separate input processing IC or within

processor 6010 as necessary. Similarly, aspects of USB or HDMI interface processing may be implemented within separate interface ICs or within processor 6010 as necessary. The demodulated, error corrected, and demultiplexed stream may be provided to various processing elements, including, for example, processor 6010, and encoder/decoder 6030 operating in combination with the memory and storage elements to process the data stream as necessary for presentation on an output device.

Various elements of system 6000 may be provided within an integrated housing. Within the integrated housing, the various elements may be interconnected and transmit data therebetween using suitable connection arrangement 6140, for example, an internal bus as known in the art, including the I2C bus, wiring, and printed circuit boards.

The system 6000 may include communication interface 6050 that enables communication with other devices via communication channel 6060. The communication interface 6050 may include, but is not limited to, a transceiver configured to transmit and to receive data over communication channel 6060. The communication interface 6050 may include, but is not limited to, a modem or network card and the communication channel 6060 may be implemented, for example, within a wired and/or a wireless medium.

Data may be streamed to the system 6000, in various embodiments, using a Wi-Fi network such as IEEE 802.11. The Wi-Fi signal of these embodiments may be received over the communications channel 6060 and the communications interface 6050 which are adapted for Wi-Fi communications. The communications channel 6060 of these embodiments may be typically connected to an access point or router that provides access to outside networks including the Internet for allowing streaming applications and other over-the-top communications.

Other embodiments may provide streamed data to the system 6000 using a set-top box that delivers the data over the HDMI connection of the input block 6130.

Still other embodiments may provide streamed data to the system 6000 using the RF connection of the input block 6130.

It is to be appreciated that signaling may be accomplished in a variety of ways. For example, one or more syntax elements, flags, and so forth may be used to signal information to a corresponding decoder in various embodiments.

5           The system 6000 may provide an output signal to various output devices, including a display 6100, speakers 6110, and other peripheral devices 6120. The other peripheral devices 6120 may include, in various examples of embodiments, one or more of a stand-alone DVR, a disk player, a stereo system, a lighting system, and other devices that provide a function based on  
10           the output of the system 3000.

          In various embodiments, control signals may be communicated between the system 6000 and the display 6100, speakers 6110, or other peripheral devices 6120 using signaling such as AV.Link (Audio/Video Link), CEC (Consumer Electronics Control), or other communications protocols that  
15           enable device-to-device control with or without user intervention.

          The output devices may be communicatively coupled to system 6000 via dedicated connections through respective interfaces 6070, 6080, and 6090.

          Alternatively, the output devices may be connected to system 6000  
20           using the communications channel 6060 via the communications interface 6050. The display 6100 and speakers 6110 may be integrated in a single unit with the other components of system 6000 in an electronic device such as, for example, a television.

          In various embodiments, the display interface 6070 may include a  
25           display driver, such as, for example, a timing controller (T Con) chip.

          The display 6100 and speaker 6110 may alternatively be separate from one or more of the other components, for example, if the RF portion of input 6130 is part of a separate set-top box. In various embodiments in which the display 6100 and speakers 6110 may be external components, the output  
30           signal may be provided via dedicated output connections, including, for example, HDMI ports, USB ports, or COMP outputs.

          As explained above, in V-PCC, 3D samples of a point cloud frame are

projected onto two layers (3D to 2D projection): A first layer, also called the near layer, may store, for example, the depth values D0 associated with the 3D samples with smaller depths. A second layer, referred to as the far layer, may store, for example, the depth values D1 associated with the 3D samples with larger depths.

**Fig. 7** illustrates an example of such a 3D to 2D projection using V-PCC.

For simplicity of representation, the drawing is done using a 2D sample cloud frame and a 1D projection plane. The 3D samples (solid squares) are projected downwards, therefore all the black 3D samples in a same column are projected to a same location of the projection plane. Two layers are used for storing the depth values of projected 3D samples: the near layer D0 and the far layer D1.

Since two layers are used, only two 3D samples can be projected per column: 3D samples projected to the near, respectively the far, layer are marked as '0', respectively '1'. Each layer value represents distance values from the projection plane to 3D samples. Moreover, in V-PCC, the depth values of in-between 3D samples (located between the two layers) are coded using EOM codeword as explained above. These in-between 3D samples are marked 'l' in **Fig. 7**. Black squares with no label refer to projected 3D samples that are not considered in the method of **Fig. 8** and **9**.

The near layer D0 (first depth image) and the far layer D1 (second depth image) stores the locations of projected 3D samples from a projection plane. The occupancy map (image) stores either the value '0' (*unoccupied* block of the canvas), '1' (*occupied* block of the canvas) or 'c' (value of an EOM codeword).

**Fig. 8** illustrates a schematic block diagram of an example of a method for coding the depth values of 3D samples projected along a same projection line in accordance with at least one embodiment.

In step 8100, a module determines if at least one 3D sample is projected along a projection line PL corresponding to a location in the occupancy map OM defined by the coordinates (u,v).

If no 3D sample is projected along the projection line PL, then in step 8200, the value of the pixel located at the coordinates (u,v) in the occupancy map equals a first value that indicates an *unoccupied* block in the canvas. For example, said first value equals 0.

5 Step 8200 is followed by step 8300 that considers another location in the occupancy map. The process ends when all locations have been considered.

If at least one 3D sample is projected along the projection line PL, in step 8400, the value of the pixel located at the coordinates (u,v) in the  
10 occupancy map equals a second value that indicates an *occupied* block in the canvas. For example, said second value equals 1.

In step 8500, a module checks if at least one additional 3D sample is projected along the projection line PL.

Note that, in V-PCC, a first additional 3D sample is a third 3D sample  
15 projected on the projection line PL because the depth values of the two first projected 3D samples are stored in the near and far layers.

In V-PCC, in step 8500, the module checks if the location of a 3D sample is greater than  $D0(u,v)$  and lower than  $D1(u,v)$ , and second, if  $D1(u,v) - D0(u,v) > 1$ .

20 If no additional 3D sample is projected along the projection line PL, step 8500 is followed by step 8300.

If at least one additional 3D sample is projected along the projection line PL, then in step 8600, a module initializes a length  $N(u,v)$  of a (N-1)-bits length PCM code  $PCM(u,v)$  used to encode the depth value of said at least one  
25 additional 3D sample.

In V-PCC,  $N(u,v) = D1(u,v) - D0(u,v)$ .

In step 8700, a module generates the (N-1)-bits length PCM code  $PCM(u,v)$  to encode depth value of at least one 3D sample projected along the projection line PL.

30 In V-PCC, a i-th bit of  $PCM(u,v)$  equals 1 to indicate if a location  $D0(u,v) + i$  is *occupied* by a projected 3D sample and 0 to indicate an *unoccupied* location.

In step 8800, a module computes a EOM codeword  $C(u,v)$  from  $PCM(u,v)$  as follows:

$$C(u, v) = 1 + (2^{(N-1)} - 1 - PCM \text{ code})$$

In step 8900, a module assigns the EOM codeword to the value of the  
5 pixel in the occupancy map OM that is located at the coordinates  $(u,v)$ :  
 $OM(u,v)=C(u,v)$ .

Step 8900 is followed by step 8300.

**Fig. 9** illustrates a schematic block diagram of an example of a method  
for decoding the depth values of 3D samples projected along a same projection  
10 line in accordance with at least one embodiment.

In step 9100, a module considers a location of an occupancy map (OM  
or DOM) at the coordinates  $(u,v)$ . When all the location of the occupancy map  
have been considered, the process ends.

In step 9200, a module determines if the value  $OM(u,v)$  of the pixel  
15 located at the coordinates  $(u,v)$  in the occupancy map indicates no 3D sample  
is projected at this location along a projection line PL.  $OM(u,v)$  may also equal  
another value indicating an *occupied* block in the canvas. In other terms, said  
another value indicates that a depth value of at least one projected 3D sample  
along said projection line PL is stored as a pixel value in a layer (near and/or  
20 far layer), and possibly, that a EOM codeword  $C(u,v)$  representing depth  
values of additional 3D samples is stored in the occupancy map at the location  
 $(u,v)$ .

In V-PCC, in step 9200,  $OM(u,v)=0$  indicates that no 3D sample is  
projected at a location  $(u,v)$  and another location is considered in step 9100.  
25  $OM(u,v)=1$  indicates an *occupied* block in the canvas and  $D1(u,v) - D0(u,v) >$   
1 indicates that a depth value of at least one additional 3D sample may be  
stored.

If the value  $OM(u,v)$  of the pixel located at the coordinates  $(u,v)$  in the  
occupancy map indicates no 3D sample is projected along said projection line  
30 PL, step 9200 is followed by step 9100 that considers another location in the  
occupancy map.

Otherwise, in step 9300, a module initializes a length  $N(u,v)$  of a  $(N-1)$ -bits length PCM code  $PCM(u,v)$  used to decode the depth value of said at least one additional 3D sample.

In V-PCC,  $N(u,v)=D1(u,v)-D0(u,v)$ .

5 In step 9400, a module generates the  $(N-1)$ -bits length PCM code  $PCM(u,v)$  to decode depth value of at least one 3D sample projected along the projection line PL from the pixel value of the occupancy map  $OM(u,v)$  at the coordinates  $(u,v)$ .

In V-PCC, the PCM code  $PCM(u,v)$  is obtained from the EOM codeword  
10  $C(u,v)$  by:

$$PCM \text{ code } (u,v) = 1 + (2^{(N-1)} - 1 - C(u, v))$$

In step 9500, a module derives the depth value of at least one additional 3D sample from the PCM code  $PCM(u,v)$ .

In V-PCC, the depth value of a  $i$ -th 3D sample projected from  $D0$  along  
15 the projection line PL equals  $D0(u,v) + i$ .

In V-PCC, computing EOM codewords requires that the difference between the depth value  $D1(u,v)$  stored in the far layer and the depth value  $D0(u,v)$  stored in the near layer be greater than 1 (for example for initializing  $N(u,v)$  in step 8500 of **Fig. 8** and step 9200 of **Fig. 9**). This is a constraint  
20 especially when multiple layers are used for storing depth values of 3D samples.

As shown in column H of **Fig. 7**, the depth value  $D1(u,v)$ , black square marked '1', is lower than a depth value  $D0(u,v)$ , black square marked '0'. Then, in V-PCC, an in-between 3D sample, black square with no label, is not  
25 considered by the method of **Fig. 8** and **Fig. 9** because the difference between  $D1(u,v)$  et  $D0(u,v)$  is negative.

According to a general aspect of at least one embodiment, a minimum depth value  $Dmin(u,v)$  and a maximum depth value  $Dmax(u,v)$  at a location  $(u,v)$  are determined from the depth values stored at said location  $(u,v)$  in all  
30 the layers (depth images), and then encoding/decoding the depth values of at least one additional 3D sample projected on a projection line PL according to said minimal and maximal depth values  $Dmin(u,v)$  and  $Dmax(u,v)$ .



Then, according to an embodiment of step 9300, the module may initialize the length  $N(u,v)$  of a  $(N-1)$ -bits length PCM code  $PCM(u,v)$  as follows:

$$N(u,v) = D_{\max}(u,v) - D_{\min}(u,v).$$

According to an embodiment of step 8700, the module may skip the  
 5 encoding of depth value of a 3D sample projected along the projection line PL as a bit of the PCM code  $PCM(u,v)$  when said depth value is already encoded as a layer value (value of a depth image).

For example, the module may calculate a current location  $D_{\min}(u,v)+i$  for each  $i$ -th bit of  $PCM(u,v)$  and may check if a pixel in the layers (depth image)  
 10 store such depth value as projected 3D sample. In that case, the  $i$ -th bit of  $PCM(u,v)$  is set to 0. Otherwise, 1.

This embodiment of step 8700 is also advantageous because it may extend the covered area as further explained in relation with **Fig. 11**.

**Fig. 11** illustrates an example of determining pixel values of occupancy  
 15 map  $OM(u,v)$  in accordance with said variant.

For example, column A, the bits  $b_0..b_8$  refer respectively to the locations 3..11 of the PCM code relative to this column A. In this example, the depth value of the location 7 of column A is already encoded in the layer 1 and not encoded as a EOM codeword. Thus, the bits  $b_5, \dots, b_8$  of the PCM code may  
 20 now refer to the locations 8, ..., 12, extending the area to 11 locations from the location 2 (the layer D0). The area for capturing the in-between is thus extended by skipping.

Note, that white squares filled out with multiple points refer to locations of 3D sample whose depth values is already stored in a layer. For example,  
 25 the location 4 of column H refer to a depth value of a 3D sample stored in the layer 1 (marked '1' in **Fig. 11**).

According to an embodiment of the method, the length of the PCM code may be limited to a maximum value  $MV$ .

For example, the maximum value  $MV$  is the bit depth of the occupancy  
 30 map  $OM$ . Typically, the maximum value  $MV$  equals 8 bits when the occupancy map  $OM$  is coded as image data.

This embodiment avoids quantizing (clipping) of the PCM codes stored as pixel value of the occupancy map and thus reduces possible loss of depth value precision. This embodiment also limits the length of the PCM code to reduce the bitrate required to code it – at the expense of missing some  
5 intermediate points.

According to an embodiment of the method, illustrated in **Fig. 8 and 9**, an information INFO1 indicating if the length of PCM code is limited to a maximum length or not may be signalled.

According to this embodiment, said information INFO1 may be added  
10 to a bitstream, possibly transmitted over a network, and said information may be extracted from a bitstream, possibly received from communication means.

According to this embodiment, a module may obtain said information INFO1 from a bitstream indicating if the length of PCM code is limited to a maximum length or not. The length of PCM code may be derived from the two  
15 layers as in V-PCC both in step 8600 of **Fig. 8** and step 9300 of **Fig. 9**, and if said information INFO1 indicates that the length of PCM code is limited to a maximum length, then the length of PCM code is limited by a maximum value MV that may be known at the encoding and decoding side.

According to an embodiment, an information INFO2 representative of  
20 said maximum value MV may be signalled.

According to this embodiment, said information INFO2 may be added to a bitstream, possibly transmitted over a network, said information may then be extracted from a bitstream, possibly received from communication means.

According to embodiment, the information INFO1 and/or INFO2 may be  
25 signalled at the sequence level. Thus, a single value of the information INFO1 and a single value of the information INFO2 may be signalled for encoding a sequence of point cloud frame.

**Fig. 12** illustrates an example of a syntax element used for signalling the information INFO1 and INFO2 in accordance with at least one  
30 embodiment.

In this example, the information INFO1 is a syntax element '**sps\_enhanced\_occupancy\_map\_code\_max\_bit\_count\_flag**' and the

information INFO2 is a syntax element  
'**sps\_enhanced\_occupancy\_map\_code\_max\_bit\_count\_minus1**'  
signalled in a syntax element "**sequence\_parameter\_set**".

According to embodiment, the information INFO1 and/or INFO2 may be  
5 signalled in the occupancy parameter set.

According to a variant of said embodiment, the information INFO1  
and/or INFO2 may be signalled as a syntax element of **geometry parameter  
set**.

According to embodiment of the method, the information INFO1 and/or  
10 INFO2 may be signalled on per frame basis.

This embodiment requires higher bit rate because the information  
INFO1 and/or INFO2 are sent more often but allows better adjusting the  
encoding/decoding methods for each frame.

In a variant of said embodiment, the information INFO1 and/or INFO2  
15 may be signaled as a syntax element of the **geometry\_frame\_params unit**  
(section 7.3.18 of V-PCC).

In a variant of said embodiment, the information INFO1 and/or INFO2  
may be signaled as a syntax element of the **occupancy\_frame\_params unit**  
analogous to the geometry frame params unit, with the corresponding  
20 **occupancy\_frame\_parameter\_set unit** and the corresponding patch  
sequence unit type (e.g. PSD\_OFPS, occupancy frame parameter set; see  
table 7-4 of V-PCC).

According to embodiment of the method, the information INFO1 and/or  
INFO2 may be signalled on per patch basis.

25 This embodiment requires higher bit rate because the information  
INFO1 and/or INFO2 are sent more often but allows better adjusting the  
encoding/decoding methods for each patch.

According to a variant of said embodiment, the information INFO1  
and/or INFO2 may be signalled as a syntax element of the **patch\_data unit**  
30 and **delta\_patch\_data unit**.

According to a variant of said embodiment, the information INFO1 and/or INFO2 may be signalled as a syntax element of **geometry\_patch\_params unit** (section 7.3.22 of V-PCC).

5 According to a variant of said embodiment, the information INFO1 and/or INFO2 may be signalled as a syntax element of the **occupancy\_patch\_params unit**, with the corresponding **occupancy\_patch\_parameter\_set unit**.

According to an embodiment, a third information INFO3 may be signaled to indicate if the ordered bits of a PCM code starting from the Least Significant Bit (LSB)  $b_0$  towards the Most Significant Bit (MSB)  $b_{N-1}$  represent  
10 a list of depth values starting either from  $D_{min}(u,v)$  towards  $D_{max}(u,v)$  or from  $D_{max}$  towards  $D_{min}$ .

Said embodiment switches the sense of the ordered list of bits of the PCM code  $PCM(u,v)$ : the LSB of the PCM code corresponds to the location  
15 closest to  $D_{min}(u,v)$ , while the MSB corresponds to the location closest to  $D_{max}(u,v)$ , or reciprocally, the LSB of the PCM code corresponds to the location closest to  $D_{max}(u,v)$ , while the MSB corresponds to the location closest to  $D_{min}(u,v)$ ,

This embodiment is especially advantageous when multiple layers are used for storing the depth values of 3D samples because the size of patches  
20 are particularly large.

For example, column F of **Fig. 10**, in-between 3D samples at locations 10 and 11 would be coded in PCM bits  $b_9$  and  $b_{10}$ , which means none of them could be coded in the 10-bit video typically used for the occupancy map. By  
25 switching the ordered list of bits of the PCM code, the depth values of those in-between 3D samples are now encoded by bits  $b_0$  and  $b_1$ , respectively, and therefore all of them can now be encoded in the 10-bits long EOM codeword.

According to an embodiment, the information INFO3 may be signaled at the sequence level.

30 **Fig. 13** illustrates an example of a syntax element used for signalling the information INFO3 in accordance with at least one embodiment.

In this example, the information INFO3 is a syntax element **sps\_enhanced\_occupancy\_map\_code\_direction\_flag** that may equal to 0 to indicate a first order of the bits of PCM code, for example from 'Dmin-to-Dmax' and to 1 indicate a second order of the bits of PCM code, for example  
5 from 'Dmax-to-Dmin'.

According to alternative, the information INFO3 may be signaled on per frame basis or on per patch basis.

According to an embodiment, the length of the PCM code may equal at least a minimum value  $mV$ .

10 For example, the minimum value  $mV$  equals 2.

This embodiment allows using the EOM codewords even a single layer is used (Otherwise, the length of th PCM code is null.). It may also allow encoding with additional 3D samples that are not in-between 3D samples, that is 3D samples having depth values lower than  $D_{min}(u,v)$  or greater than  
15  $D_{max}(u,v)$ .

According to an embodiment, an information INFO4 indicating if the length of PCM code equals at least a minimum length  $mL$  or not may be signalled.

20 According to this embodiment, said information INFO4 may be added to a bitstream, possibly transmitted over a network, and said information may be extracted from a bitstream, possibly received from communication means.

According to this embodiment, a module may obtain said information INFO4 from a bitstream indicating that the length of PCM code is not equal to a minimum length  $mL$  or not. The length of PCM code may be derived from the  
25 two layers as in V-PCC both in step 8600 of **Fig. 8** and step 9300 of **Fig. 9**, and if the information INFO4 indicates that the length of PCM code is set to the minimum length  $mL$  if lower than  $ML$ .

According to an embodiment, an information INFO5 representative of said minimum length  $mL$  may be signalled.

30 According to this embodiment, said information INFO5 may be added to a bitstream, possibly transmitted over a network, said information may then be extracted from a bitstream, possibly received from communication means.

According to embodiment, the information INFO4 and/or INFO5 may be signalled at the sequence level. Thus, a single value of the information INFO4 and a single value of the information INFO5 may be signalled for encoding a sequence of point cloud frame.

5           **Fig. 14** illustrates an example of a syntax element used for signalling the information INFO4 and INFO5 in accordance with at least one embodiment.

10           In this example, the information INFO4 is a syntax element ‘**sps\_enhanced\_occupancy\_map\_code\_min\_bit\_count\_flag**’ and the information INFO5 is a syntax element ‘**sps\_enhanced\_occupancy\_map\_code\_min\_bit\_count\_minus2**’ signalled in a syntax element “**sequence\_parameter\_set**”.

According to embodiment, the information INFO4 and/or INFO5 may be signalled in the occupancy parameter set.

15           According to embodiment of the method, the information INFO4 and/or INFO5 may be signalled on per frame basis.

This embodiment requires higher bit rate because the information INFO4 and/or INFO5 are sent more often but allows better adjusting the encoding/decoding methods for each frame.

20           According to embodiment of the method, the information INFO4 and/or INFO5 may be signalled on per patch basis.

This embodiment requires higher bit rate because the information INFO4 and/or INFO5 are sent more often but allows better adjusting the encoding/decoding methods for each patch.

25           In **Fig. 1-14**, various methods are described herein, and each of the methods includes one or more steps or actions for achieving the described method. Unless a specific order of steps or actions is required for proper operation of the method, the order and/or use of specific steps and/or actions may be modified or combined.

30           Some examples are described with regard to block diagrams and operational flowcharts. Each block represents a circuit element, module, or portion of code which includes one or more executable instructions for

implementing the specified logical function(s). It should also be noted that in other implementations, the function(s) noted in the blocks may occur out of the indicated order. For example, two blocks shown in succession may, in fact, be executed substantially concurrently or the blocks may sometimes be executed  
5 in the reverse order, depending on the functionality involved.

The implementations and aspects described herein may be implemented in, for example, a method or a process, an apparatus, a computer program, a data stream, a bitstream, or a signal. Even if only discussed in the context of a single form of implementation (for example, discussed only as a  
10 method), the implementation of features discussed may also be implemented in other forms (for example, an apparatus or computer program).

The methods may be implemented in, for example, a processor, which refers to processing devices in general, including, for example, a computer, a microprocessor, an integrated circuit, or a programmable logic device.  
15 Processors also include communication devices.

Additionally, the methods may be implemented by instructions being performed by a processor, and such instructions (and/or data values produced by an implementation) may be stored on a computer readable storage medium. A computer readable storage medium may take the form of a  
20 computer readable program product embodied in one or more computer readable medium(s) and having computer readable program code embodied thereon that is executable by a computer. A computer readable storage medium as used herein may be considered a non-transitory storage medium given the inherent capability to store the information therein as well as the  
25 inherent capability to provide retrieval of the information therefrom. A computer readable storage medium may be, for example, but is not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, or device, or any suitable combination of the foregoing. It is to be appreciated that the following, while providing more specific examples  
30 of computer readable storage mediums to which the present embodiments may be applied, is merely an illustrative and not an exhaustive listing as is readily appreciated by one of ordinary skill in the art: a portable computer

diskette; a hard disk; a read-only memory (ROM); an erasable programmable read-only memory (EPROM or Flash memory); a portable compact disc read-only memory (CD-ROM); an optical storage device; a magnetic storage device; or any suitable combination of the foregoing.

5           The instructions may form an application program tangibly embodied on a processor-readable medium.

Instructions may be, for example, in hardware, firmware, software, or a combination. Instructions may be found in, for example, an operating system, a separate application, or a combination of the two. A processor may be  
10 characterized, therefore, as, for example, both a device configured to carry out a process and a device that includes a processor-readable medium (such as a storage device) having instructions for carrying out a process. Further, a processor-readable medium may store, in addition to or in lieu of instructions, data values produced by an implementation.

15           An apparatus may be implemented in, for example, appropriate hardware, software, and firmware. Examples of such apparatus include personal computers, laptops, smartphones, tablet computers, digital multimedia set top boxes, digital television receivers, personal video recording systems, connected home appliances, head mounted display devices (HMD,  
20 see-through glasses), projectors (beamers), "caves" (system including multiple displays), servers, video encoders, video decoders, post-processors processing output from a video decoder, pre-processors providing input to a video encoder, web servers, set-top boxes, and any other device for processing a point cloud, a video or an image or other communication devices.  
25 As should be clear, the equipment may be mobile and even installed in a mobile vehicle.

Computer software may be implemented by the processor 6010 or by hardware, or by a combination of hardware and software. As a non-limiting example, the embodiments may be also implemented by one or more  
30 integrated circuits. The memory 6020 may be of any type appropriate to the technical environment and may be implemented using any appropriate data storage technology, such as optical memory devices, magnetic memory

devices, semiconductor-based memory devices, fixed memory, and removable memory, as non-limiting examples. The processor 6010 may be of any type appropriate to the technical environment, and may encompass one or more of microprocessors, general purpose computers, special purpose  
5 computers, and processors based on a multi-core architecture, as non-limiting examples.

As will be evident to one of ordinary skill in the art, implementations may produce a variety of signals formatted to carry information that may be, for example, stored or transmitted. The information may include, for example,  
10 instructions for performing a method, or data produced by one of the described implementations. For example, a signal may be formatted to carry the bitstream of a described embodiment. Such a signal may be formatted, for example, as an electromagnetic wave (for example, using a radio frequency portion of spectrum) or as a baseband signal. The formatting may include, for  
15 example, encoding a data stream and modulating a carrier with the encoded data stream. The information that the signal carries may be, for example, analog or digital information. The signal may be transmitted over a variety of different wired or wireless links, as is known. The signal may be stored on a processor-readable medium.

20 The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. As used herein, the singular forms "a", "an", and "the" may be intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms "includes/comprises" and/or "including/comprising"  
25 when used in this specification, may specify the presence of stated, for example, features, integers, steps, operations, elements, and/or components but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof. Moreover, when an element is referred to as being "responsive" or "connected"  
30 to another element, it may be directly responsive or connected to the other element, or intervening elements may be present. In contrast, when an

element is referred to as being "directly responsive" or "directly connected" to other element, there are no intervening elements present.

It is to be appreciated that the use of any of the symbol/term "/", "and/or", and "at least one of", for example, in the cases of "A/B", "A and/or B" and "at least one of A and B", may be intended to encompass the selection of the first listed option (A) only, or the selection of the second listed option (B) only, or the selection of both options (A and B). As a further example, in the cases of "A, B, and/or C" and "at least one of A, B, and C", such phrasing is intended to encompass the selection of the first listed option (A) only, or the selection of the second listed option (B) only, or the selection of the third listed option (C) only, or the selection of the first and the second listed options (A and B) only, or the selection of the first and third listed options (A and C) only, or the selection of the second and third listed options (B and C) only, or the selection of all three options (A and B and C). This may be extended, as is clear to one of ordinary skill in this and related arts, for as many items as are listed.

Various numeric values may be used in the present application, for example. The specific values may be for example purposes and the aspects described are not limited to these specific values.

It will be understood that, although the terms first, second, etc. may be used herein to describe various elements, these elements are not limited by these terms. These terms are only used to distinguish one element from another. For example, a first element could be termed a second element, and, similarly, a second element could be termed a first element without departing from the teachings of this application. No ordering is implied between a first element and a second element.

Reference to "one embodiment" or "an embodiment" or "one implementation" or "an implementation", as well as other variations thereof, is frequently used to convey that a particular feature, structure, characteristic, and so forth (described in connection with the embodiment/implementation) is included in at least one embodiment/implementation. Thus, the appearances of the phrase "in one embodiment" or "in an embodiment" or "in one implementation" or "in an implementation", as well any other variations,

appearing in various places throughout this application are not necessarily all referring to the same embodiment.

Similarly, reference herein to “in accordance with an embodiment/example/implementation” or “in an embodiment/example/implementation”, as well as other variations thereof, is frequently used to convey that a particular feature, structure, or characteristic (described in connection with the embodiment/example/implementation) may be included in at least one embodiment/example/implementation. Thus, the appearances of the expression “in accordance with an embodiment/example/implementation” or “in an embodiment/example/implementation” in various places in the specification are not necessarily all referring to the same embodiment/example/implementation, nor are separate or alternative embodiment/examples/implementation necessarily mutually exclusive of other embodiments/examples/implementation.

Reference numerals appearing in the claims are by way of illustration only and shall have no limiting effect on the scope of the claims. Although not explicitly described, the present embodiments/examples and variants may be employed in any combination or sub-combination.

When a figure is presented as a flow diagram, it should be understood that it also provides a block diagram of a corresponding apparatus. Similarly, when a figure is presented as a block diagram, it should be understood that it also provides a flow diagram of a corresponding method/process.

Although some of the diagrams include arrows on communication paths to show a primary direction of communication, it is to be understood that communication may occur in the opposite direction to the depicted arrows.

Various implementations involve decoding. “Decoding”, as used in this application, may encompass all or part of the processes performed, for example, on a received point cloud frame (including possibly a received bitstream which encodes one or more-point cloud frames) in order to produce a final output suitable for display or for further processing in the reconstructed point cloud domain. In various embodiments, such processes include one or

more of the processes typically performed by an image-based decoder. In various embodiments, such processes also, or alternatively, include processes performed by a decoder of various implementations described in this application, for example,

5           As further examples, in one embodiment “decoding” may refer only to entropy decoding, in another embodiment “decoding” may refer only to differential decoding, and in another embodiment “decoding” may refer to a combination of entropy decoding and differential decoding. Whether the phrase “decoding process” may be intended to refer specifically to a subset of  
10 operations or generally to the broader decoding process will be clear based on the context of the specific descriptions and is believed to be well understood by those skilled in the art.

          Various implementations involve encoding. In an analogous way to the above discussion about “decoding”, “encoding” as used in this application may  
15 encompass all or part of the processes performed, for example, on an input point cloud frame in order to produce an encoded bitstream. In various embodiments, such processes include one or more of the processes typically performed by an image-based decoder.

          As further examples, in one embodiment “encoding” may refer only to  
20 entropy encoding, in another embodiment “encoding” may refer only to differential encoding, and in another embodiment “encoding” may refer to a combination of differential encoding and entropy encoding. Whether the phrase “encoding process” may be intended to refer specifically to a subset of operations or generally to the broader encoding process will be clear based on  
25 the context of the specific descriptions and is believed to be well understood by those skilled in the art.

          Note that the syntax elements as used herein, for example  
**sps\_enhanced\_occupancy\_map\_code\_max\_bit\_count\_flag,**  
**sps\_enhanced\_occupancy\_map\_code\_max\_bit\_count\_minus1,**  
30 **sps\_enhanced\_occupancy\_map\_code\_direction\_flag,**  
**sps\_enhanced\_occupancy\_map\_code\_min\_bit\_count\_flag,**  
**sps\_enhanced\_occupancy\_map\_code\_min\_bit\_count\_minus2,**       are

descriptive terms. As such, they do not preclude the use of other syntax element names.

Various embodiments refer to rate distortion optimization. In particular, during the encoding process, the balance or trade-off between the rate and distortion is usually considered, often given the constraints of computational complexity. The rate distortion optimization may be usually formulated as minimizing a rate distortion function, which is a weighted sum of the rate and of the distortion. There are different approaches to solve the rate distortion optimization problem. For example, the approaches may be based on an extensive testing of all encoding options, including all considered modes or coding parameters values, with a complete evaluation of their coding cost and related distortion of the reconstructed signal after coding and decoding. Faster approaches may also be used, to save encoding complexity, in particular with computation of an approximated distortion based on the prediction or the prediction residual signal, not the reconstructed one. A mix of these two approaches may also be used, such as by using an approximated distortion for only some of the possible encoding options, and a complete distortion for other encoding options. Other approaches only evaluate a subset of the possible encoding options. More generally, many approaches employ any of a variety of techniques to perform the optimization, but the optimization is not necessarily a complete evaluation of both the coding cost and related distortion.

Additionally, this application may refer to “determining” various pieces of information. Determining the information may include one or more of, for example, estimating the information, calculating the information, predicting the information, or retrieving the information from memory.

Further, this application may refer to “accessing” various pieces of information. Accessing the information may include one or more of, for example, receiving the information, retrieving the information (for example, from memory), storing the information, moving the information, copying the information, calculating the information, determining the information, predicting the information, or estimating the information.

Additionally, this application may refer to “receiving” various pieces of information. Receiving is, as with “accessing”, intended to be a broad term. Receiving the information may include one or more of, for example, accessing the information, or retrieving the information (for example, from memory).

5 Further, “receiving” is typically involved, in one way or another, during operations such as, for example, storing the information, processing the information, transmitting the information, moving the information, copying the information, erasing the information, calculating the information, determining the information, predicting the information, or estimating the information.

10 Also, as used herein, the word “signal” refers to, among other things, indicating something to a corresponding decoder. For example, in certain embodiments the encoder signals a particular information INFO1, INFO2, INFO3, INFO4 and/or INFO5. In this way, in an embodiment the same parameter may be used at both the encoder side and the decoder side. Thus,  
15 for example, an encoder may transmit (explicit signaling) a particular parameter to the decoder so that the decoder may use the same particular parameter. Conversely, if the decoder already has the particular parameter as well as others, then signaling may be used without transmitting (implicit signaling) to simply allow the decoder to know and select the particular  
20 parameter. By avoiding transmission of any actual functions, a bit savings is realized in various embodiments. It is to be appreciated that signaling may be accomplished in a variety of ways. For example, one or more syntax elements, flags, and so forth are used to signal information to a corresponding decoder in various embodiments. While the preceding relates to the verb form of the  
25 word “signal”, the word “signal” may also be used herein as a noun.

A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made. For example, elements of different implementations may be combined, supplemented, modified, or removed to produce other implementations. Additionally, one of  
30 ordinary skill will understand that other structures and processes may be substituted for those disclosed and the resulting implementations will perform at least substantially the same function(s), in at least substantially the same

way(s), to achieve at least substantially the same result(s) as the implementations disclosed. Accordingly, these and other implementations are contemplated by this application.

**CLAIMS**

1. A method comprising encoding a depth value of at least one 3D sample of  
5 a point cloud frame based on a depth value of at least two other 3D samples,  
said at least one 3D sample and said at least two other 3D samples being  
projected onto the projection plane along a same projection line.
2. The method of claim 1, wherein depth value of said at least two other 3D  
10 sample are obtained from at least two depth images.
3. The method of claim 1, wherein encoding a depth value of at least one 3D  
sample based on the depth value of said at least two other 3D samples  
comprises encoding the depth value of said at least one 3D sample based on  
15 minimal and maximal depth values determined from the depth values of said  
at least two 3D samples.
4. The method of claim 3, wherein encoding the depth value of said at least  
one 3D sample based on said minimal and maximal depth values comprises  
20 encoding a depth value of at least one 3D sample as a bit of a codeword, and  
skipping the encoding of depth value of a 3D sample projected along the  
projection line as a bit of the codeword when said depth value is already  
encoded as a pixel value of a depth image.
- 25 5. The method of claim 4, wherein the codeword length is limited to a maximum  
length.
6. The method of claim 4 or 5 further comprising transmitting a first information  
indicating if the codeword length is limited to a maximum length or not.  
30
7. The method of claim 6 further comprising transmitting a second information  
representative of said maximum length.

8. The method of claim 4 further comprising transmitting a third information to indicate if the ordered bits of a codeword starting from a Least Significant Bit towards a Most Significant Bit represent a list of depth values starting either  
5 from the minimum depth value towards the maximal depth value or from the maximal depth value towards the minimal depth value.
9. The method of claim 4, wherein the codeword length equals at least a  
10 minimum length.
10. The method of claim 4 further comprising transmitting a fourth information indicating if the codeword length equals at least a minimum length or not.
11. The method of claim 9 or 10 further comprising transmitting a fifth  
15 information representative of said minimum length.
12. A method comprising decoding a depth value of at least one 3D sample of a point cloud based on depth value of at least two other 3D samples, said at least one 3D sample and said at least two other 3D sample being projected  
20 onto the projection plane along a same projection line.
13. The method of claim 12, wherein depth value of said at least two other 3D sample are obtained from at least two depth images.
- 25 14. The method of claim 12, wherein decoding a depth value of at least one 3D sample based on the depth value of said at least two other 3D samples comprises decoding the depth value of said at least one 3D sample based on minimal and maximal depth values determined from the depth values of said at least two 3D samples.
- 30 15. The method of claim 14, wherein decoding the depth value of said at least one 3D sample based on said minimal and maximal depth values comprises

decoding a depth value of at least one 3D sample from a bit of a codeword, and skipping the decoding of depth value of a 3D sample projected along the projection line from a bit of the codeword when said depth value is already encoded as a pixel value of a depth image.

5

16. The method of claim 15, wherein the codeword length is limited to a maximum length.

17. The method of claim 15 or 16 further comprising receiving a first  
10 information indicating if the codeword length is limited to a maximum length or not.

18. The method of claim 17 further comprising receiving a second information representative of said maximum length.

15

19. The method of claim 15 further comprising receiving a third information to indicate if the ordered bits of a codeword starting from a Least Significant Bit towards a Most Significant Bit represent a list of depth values starting either from the minimum depth value towards the maximal depth value or from the  
20 maximal depth value towards the minimal depth value.

20. The method of claim 15, wherein the codeword length equals at least a minimum length.

25 21. The method of claim 15 further comprising receiving a fourth information indicating if the codeword length equals at least a minimum length or not.

22. The method of claim 19 or 20 further comprising receiving a fifth information representative of said minimum length.

30

23. An apparatus comprising one or more processors configured to encoding a depth value of at least one 3D sample of a point cloud frame based on a

depth value of at least two other 3D samples, said at least one 3D sample and said at least two other 3D samples being projected onto the projection plane along a same projection line.

5 24. The apparatus of claim 23, wherein depth value of said at least two other 3D sample are obtained from at least two depth images.

25. The apparatus of claim 23, wherein encoding a depth value of at least one 3D sample based on the depth value of said at least two other 3D samples  
10 comprises encoding the depth value of said at least one 3D sample based on minimal and maximal depth values determined from the depth values of said at least two 3D samples.

26. The apparatus of claim 25, wherein encoding the depth value of said at  
15 least one 3D sample based on said minimal and maximal depth values comprises encoding a depth value of at least one 3D sample as a bit of a codeword, and skipping the encoding of depth value of a 3D sample projected along the projection line as a bit of the codeword when said depth value is already encoded as a pixel value of a depth image.

20 27. The apparatus of claim 25, wherein the codeword length is limited to a maximum length.

28. The apparatus of claim 26 or 27, wherein said one or more processors are  
25 further configured to transmitting a first information indicating if the codeword length is limited to a maximum length or not.

29. The apparatus of claim 28, wherein said one or more processors are further  
30 configured to transmitting a second information representative of said maximum length.

30. The apparatus of claim 27, wherein said one or more processors are further configured to transmitting a third information to indicate if the ordered bits of a codeword starting from a Least Significant Bit towards a Most Significant Bit represent a list of depth values starting either from the minimum depth value  
5 towards the maximal depth value or from the maximal depth value towards the minimal depth value.

31. The apparatus of claim 27, wherein the codeword length equals at least a minimum length.  
10

32. The apparatus of claim 27, wherein said one or more processors are further configured to transmitting a fourth information indicating if the codeword length equals at least a minimum length or not.

15 33. The apparatus of claim 31 or 32, wherein said one or more processors are further configured to transmitting a fifth information representative of said minimum length.

34. An apparatus comprising one or more processors configured to decoding  
20 a depth value of at least one 3D sample of a point cloud based on depth value of at least two other 3D samples, said at least one 3D sample and said at least two other 3D sample being projected onto the projection plane along a same projection line.

25 35. The apparatus of claim 34, wherein depth value of said at least two other 3D sample are obtained from at least two depth images.

36. The apparatus of claim 34, wherein decoding a depth value of at least one 3D sample based on the depth value of said at least two other 3D samples  
30 comprises decoding the depth value of said at least one 3D sample based on minimal and maximal depth values determined from the depth values of said at least two 3D samples.

37. The apparatus of claim 36, wherein decoding the depth value of said at least one 3D sample based on said minimal and maximal depth values comprises decoding a depth value of at least one 3D sample from a bit of a  
5 codeword, and skipping the decoding of depth value of a 3D sample projected along the projection line from a bit of the codeword when said depth value is already encoded as a pixel value of a depth image.

38. The apparatus of claim 37, wherein the codeword length is limited to a  
10 maximum length.

39. The apparatus of claim 37 or 38, wherein said one or more processors are further configured to receiving a first information indicating if the codeword length is limited to a maximum length or not.  
15

40. The apparatus of claim 39, wherein said one or more processors are further configured to receiving a second information representative of said maximum length.

20 41. The apparatus of claim 37, wherein said one or more processors are further configured to receiving a third information to indicate if the ordered bits of a codeword starting from a Least Significant Bit towards a Most Significant Bit represent a list of depth values starting either from the minimum depth value towards the maximal depth value or from the maximal depth value towards the  
25 minimal depth value.

42. The apparatus of claim 37, wherein the codeword length equals at least a minimum length.

43. The apparatus of claim 37, wherein said one or more processors are further  
30 configured to receiving a fourth information indicating if the codeword length equals at least a minimum length or not.

44. The apparatus of claim 41 or 42, wherein said one or more processors are further configured to receiving a fifth information representative of said minimum length.
- 5 45. A bitstream comprising image data representing a codeword whose bits encode depth value of at least one 3D sample of a point cloud projected along a projection line.
46. The bitstream of claim 45 further comprising a first information indicating if  
10 the length of a codeword is limited to a maximum length.
47. The bitstream of claim 45 or 46 further comprising a second information representative of said maximum length.
- 15 48. The bitstream of claim 45 further comprising a third information to indicate if the ordered bits of a codeword starting from a Least Significant Bit towards a Most Significant Bit represent a list of depth values starting either from the minimum depth value towards the maximal depth value or from the maximal depth value towards the minimal depth value.
- 20 49. The bitstream of claim 45 further comprising a fourth information indicating if the length of the codeword equals at least a minimum length or not.
50. The bitstream of claim 45 further comprising a fifth information  
25 representative of said minimum length.
51. A computer program product including instructions which, when the program is executed by one or more processors, causes the one or more processors to carry out a method comprising:
- 30       accessing depth value of at least two 3D samples of a point cloud frame located at a same 2D location of a projection plane; and  
          encoding a depth value of at least one other 3D sample of the point cloud based on the depth value of said at least two 3D samples, said at least

two 3D samples and said at least one other 3D sample being projected onto the projection plane along a same projection line.

52. A non-transitory computer-readable medium including instructions  
5 for causing one or more processors to perform encoding a depth value of at least one 3D sample of a point cloud frame based on a depth value of at least two other 3D samples, said at least one 3D sample and said at least two other 3D samples being projected onto the projection plane along a same projection line.

10

53. A computer program product including instructions which, when the program is executed by one or more processors, causes the one or more processors to carry out a method comprising encoding a depth value of at least one 3D sample of a point cloud frame based on a depth value of at least two  
15 other 3D samples, said at least one 3D sample and said at least two other 3D samples being projected onto the projection plane along a same projection line.

54. A non-transitory computer-readable medium including instructions for  
20 causing one or more processors to perform decoding a depth value of at least one 3D sample of a point cloud based on depth value of at least two other 3D samples, said at least one 3D sample and said at least two other 3D sample being projected onto the projection plane along a same projection line.

25 55. A computer program product including instructions which, when the program is executed by one or more processors, causes the one or more processors to carry out a method comprising decoding a depth value of at least one 3D sample of a point cloud based on depth value of at least two other 3D samples, said at least one 3D sample and said at least two other 3D sample  
30 being projected onto the projection plane along a same projection line.

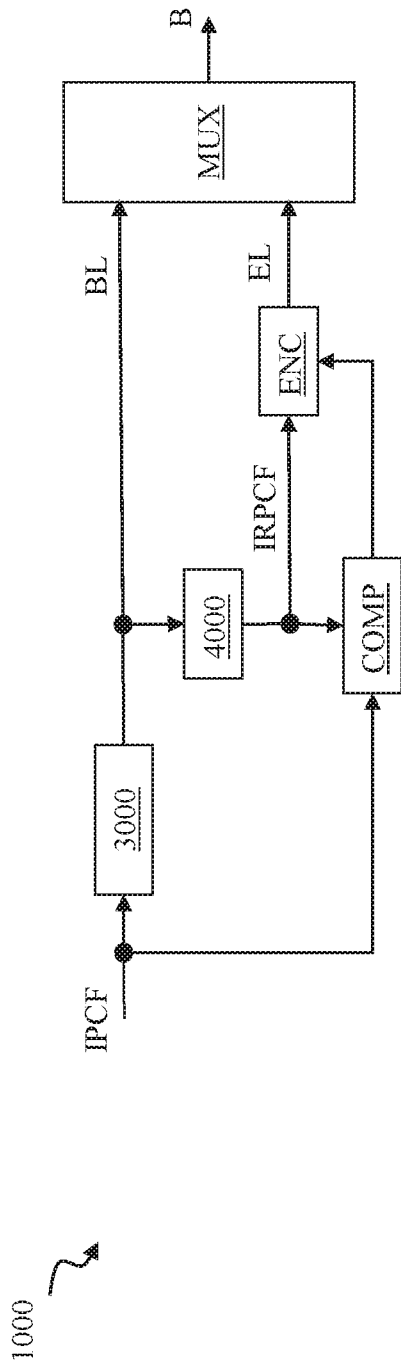


Fig. 1

2000

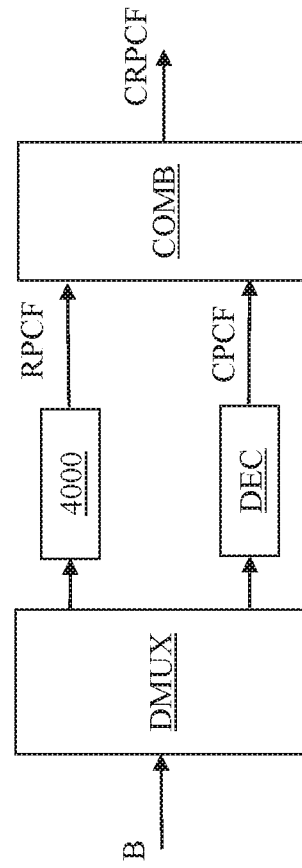


Fig. 2

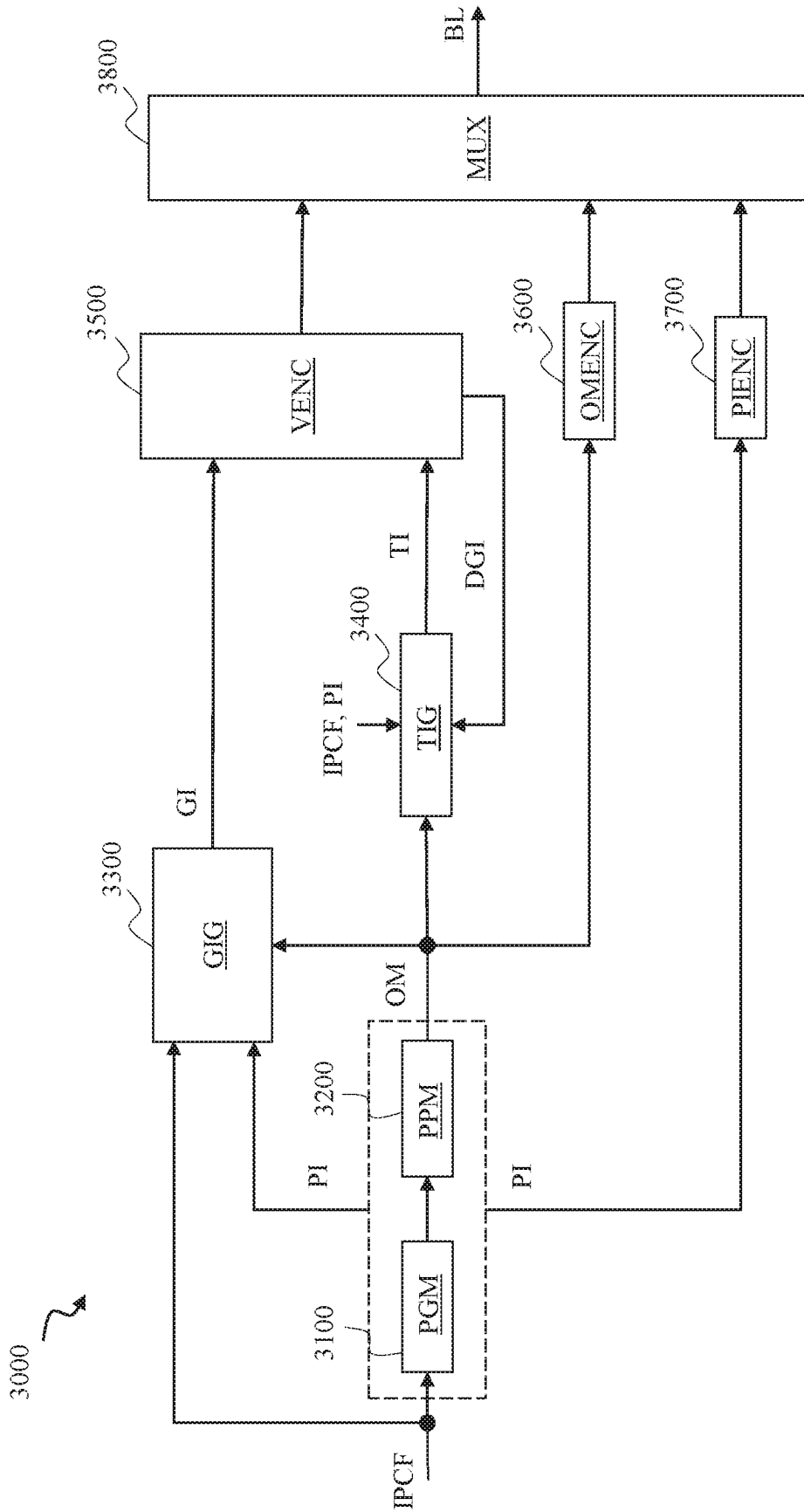


Fig. 3

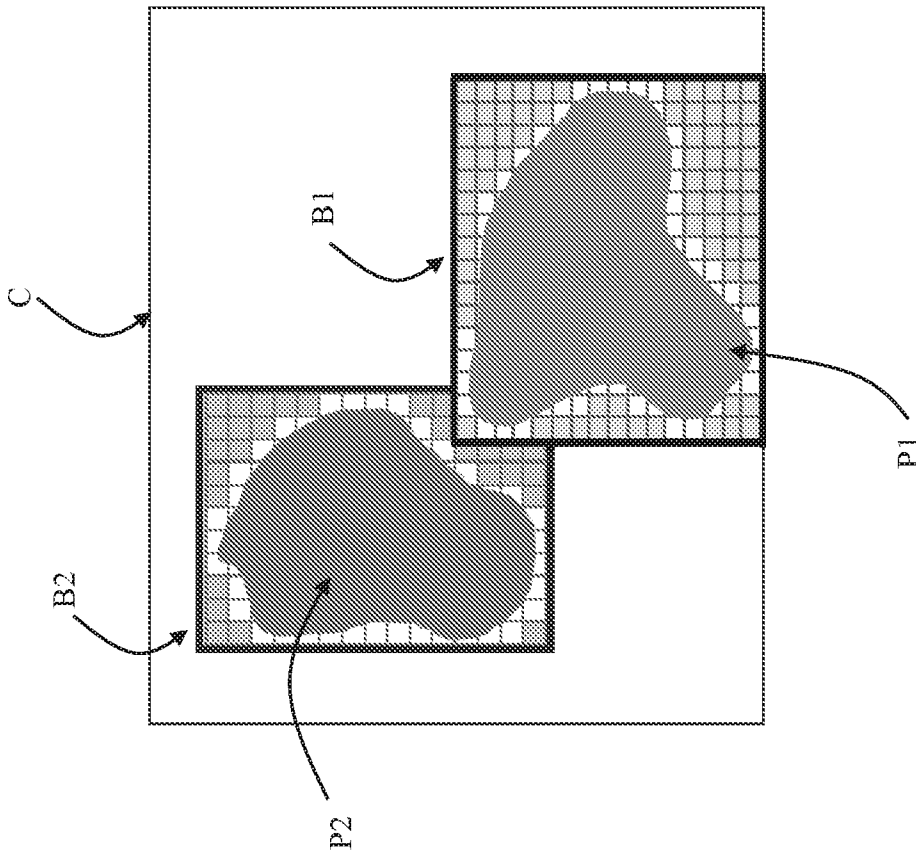


Fig. 3a

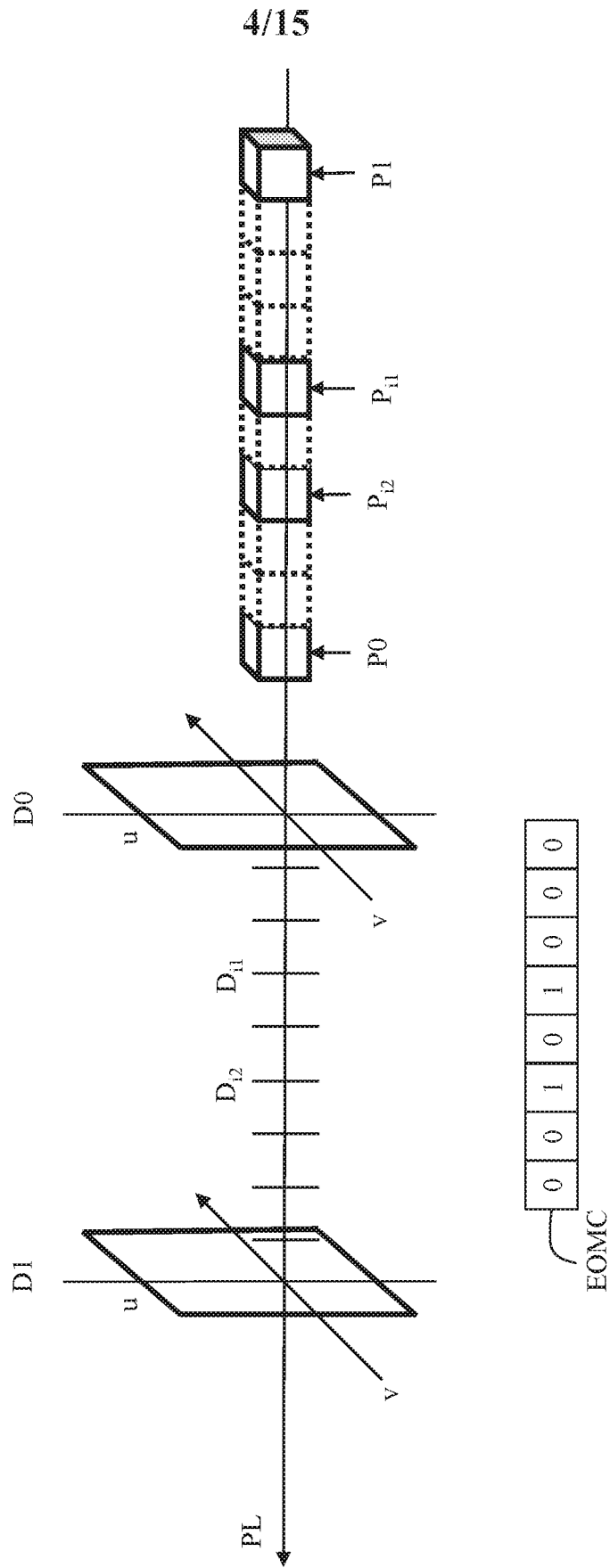


Fig. 3b

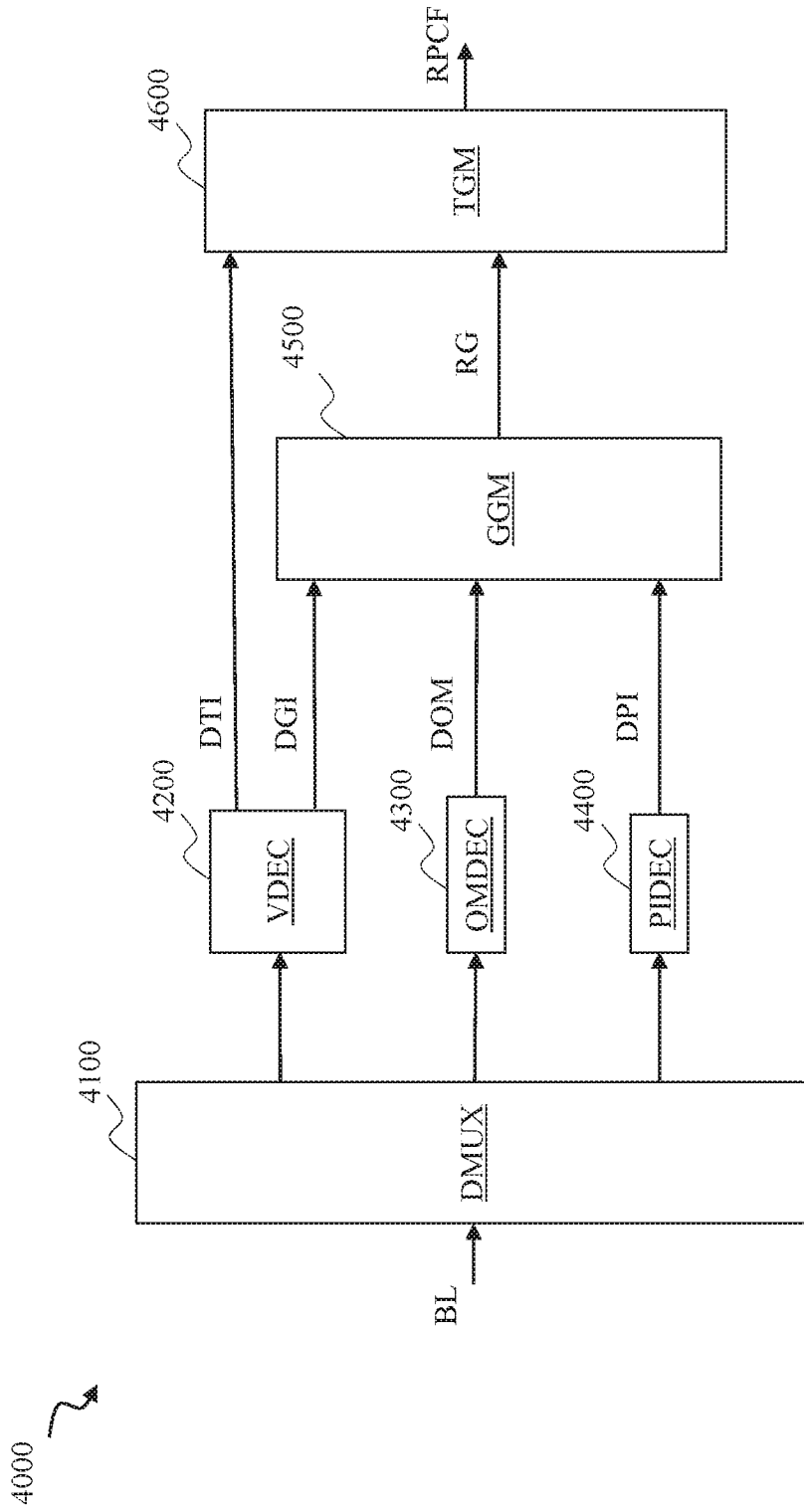


Fig. 4

5000 ↗

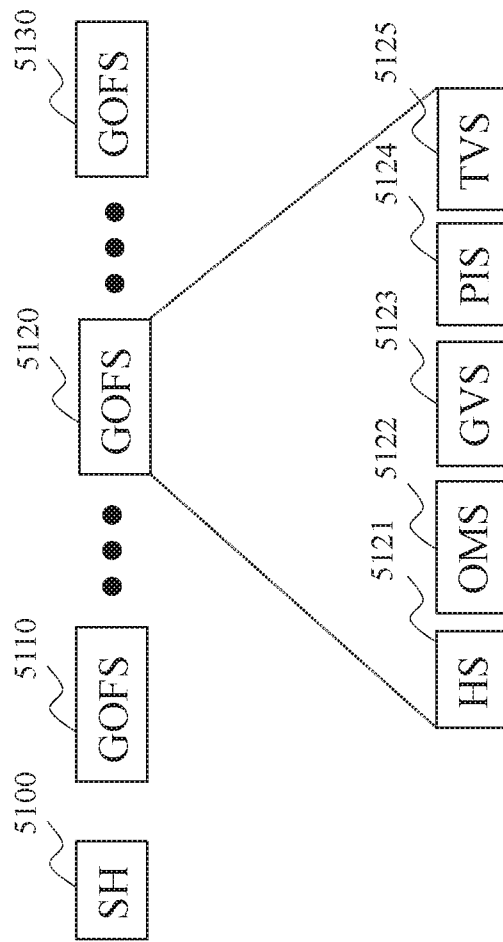


Fig. 5

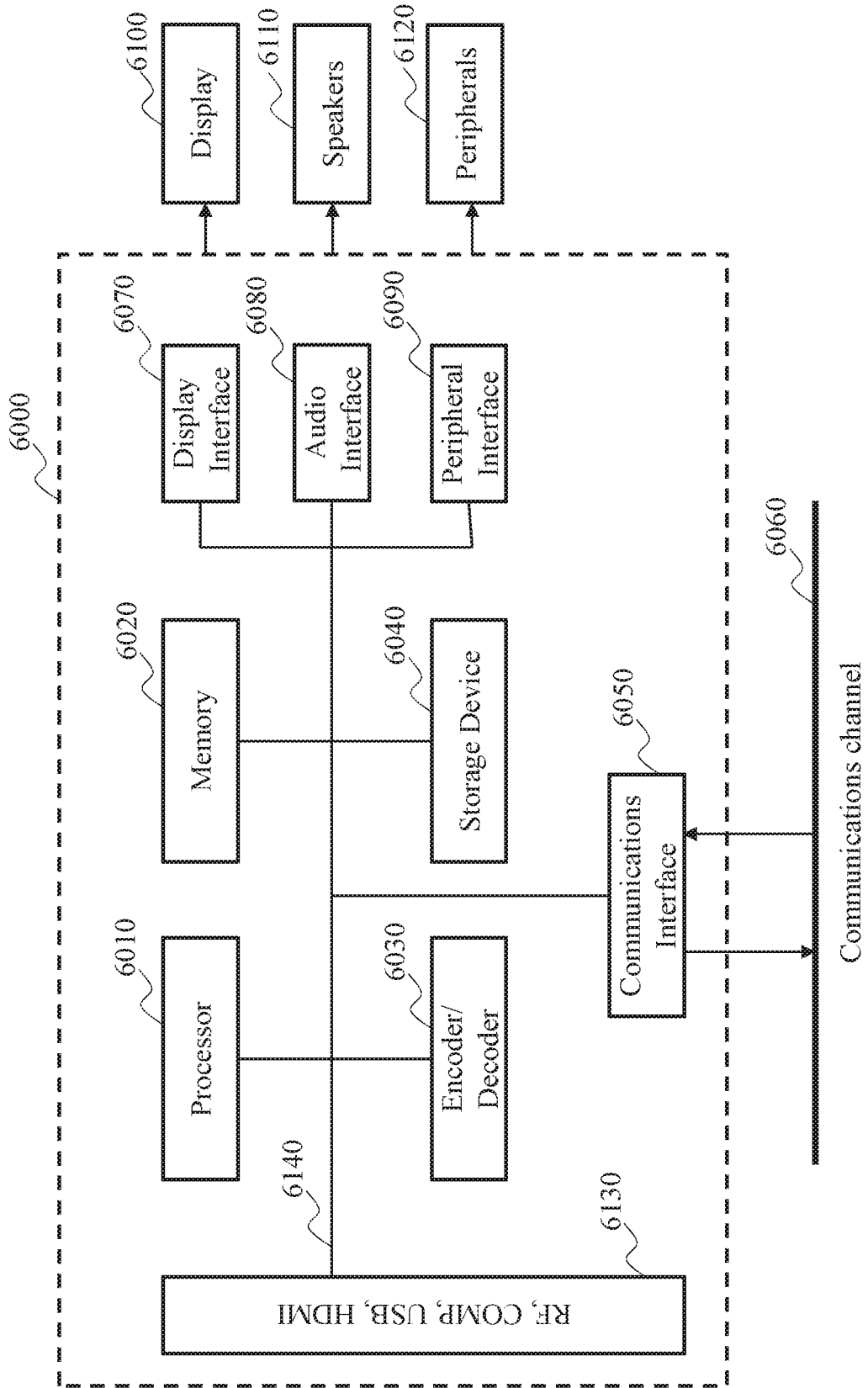


Fig. 6

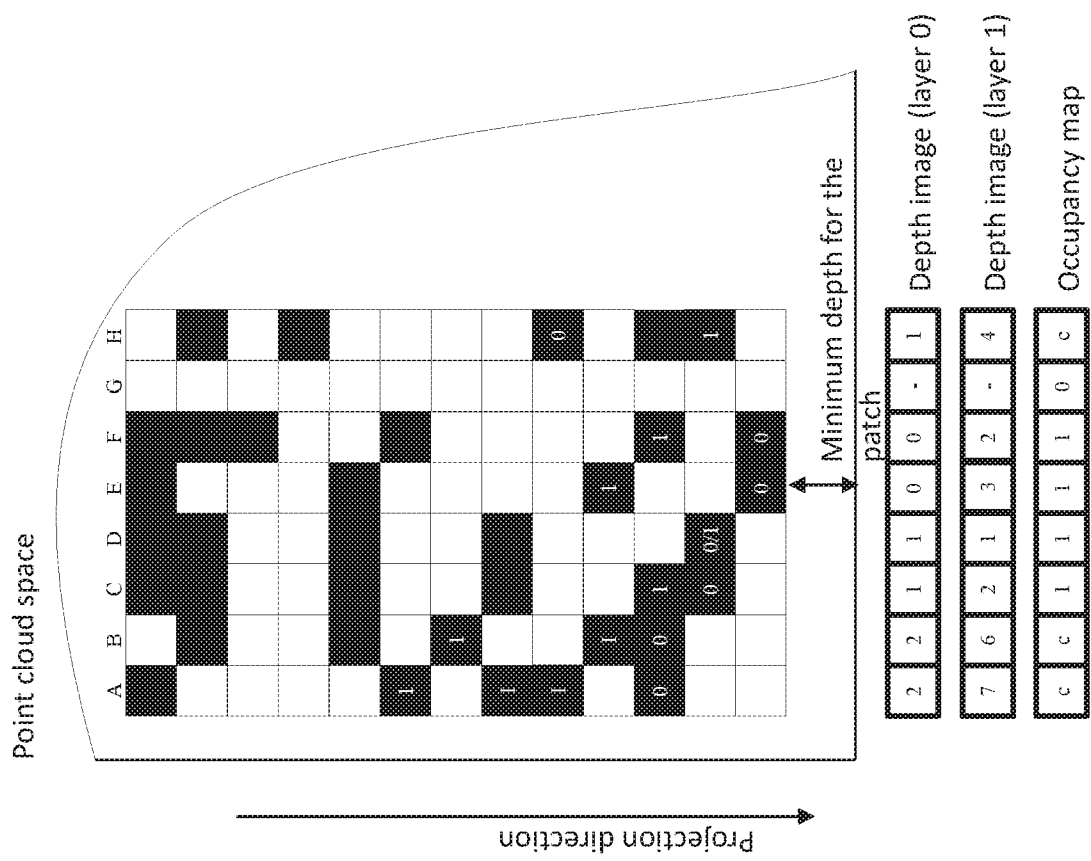


Fig. 7

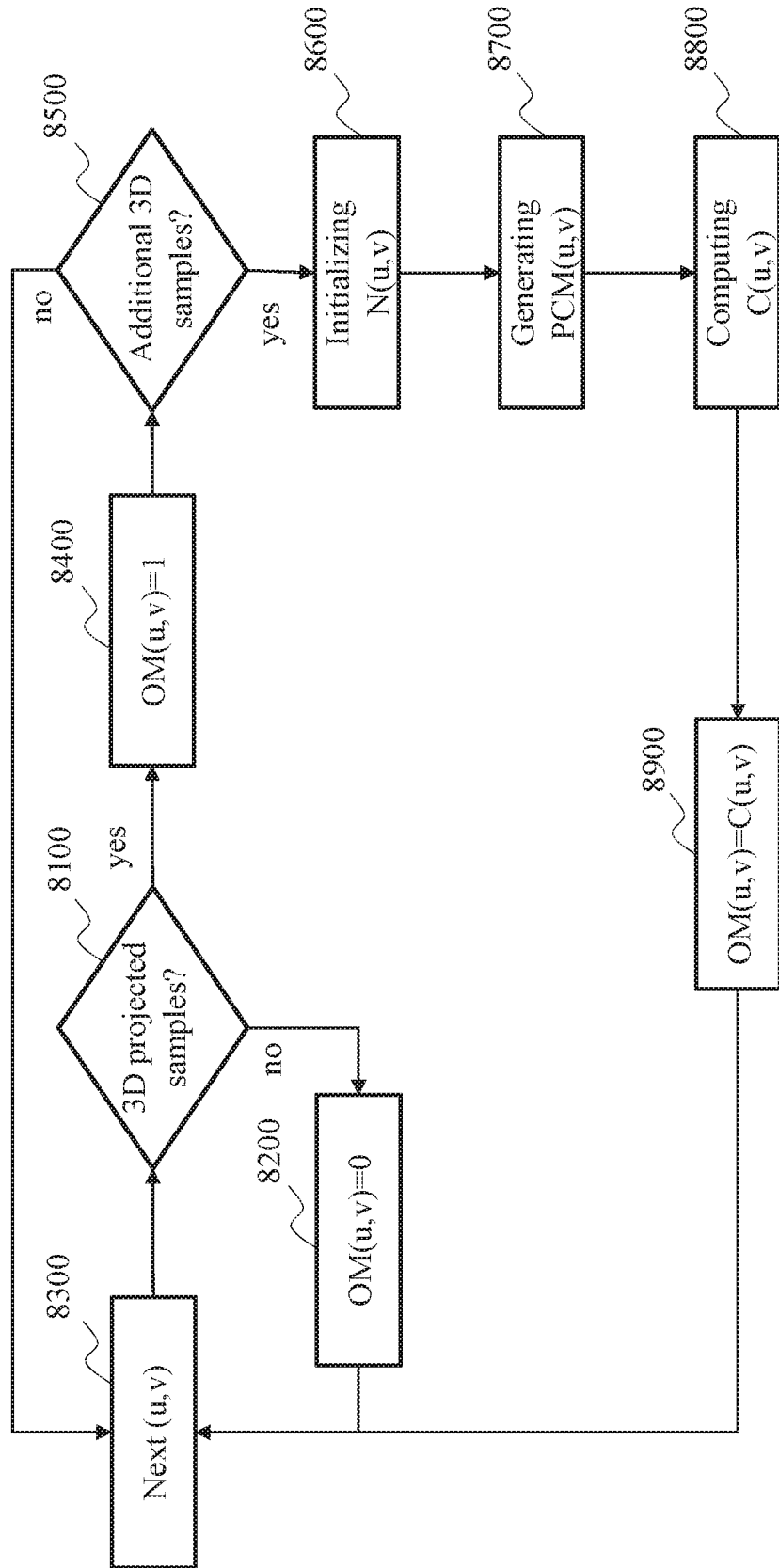


Fig. 8

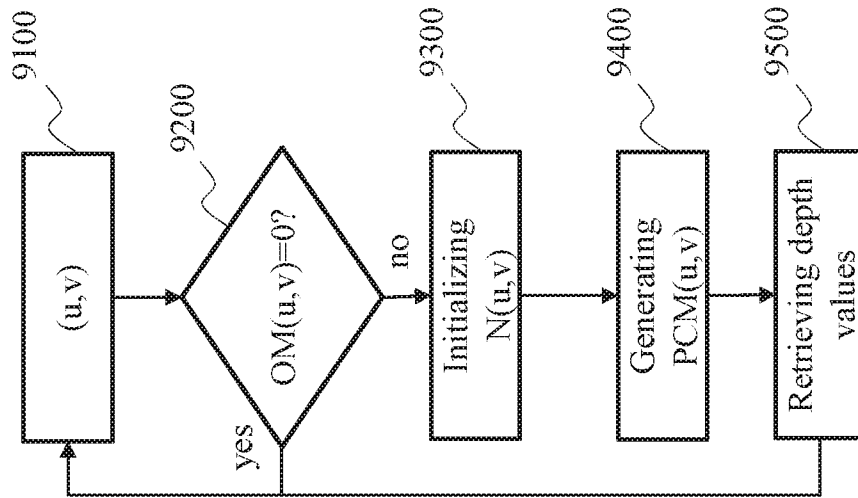


Fig. 9



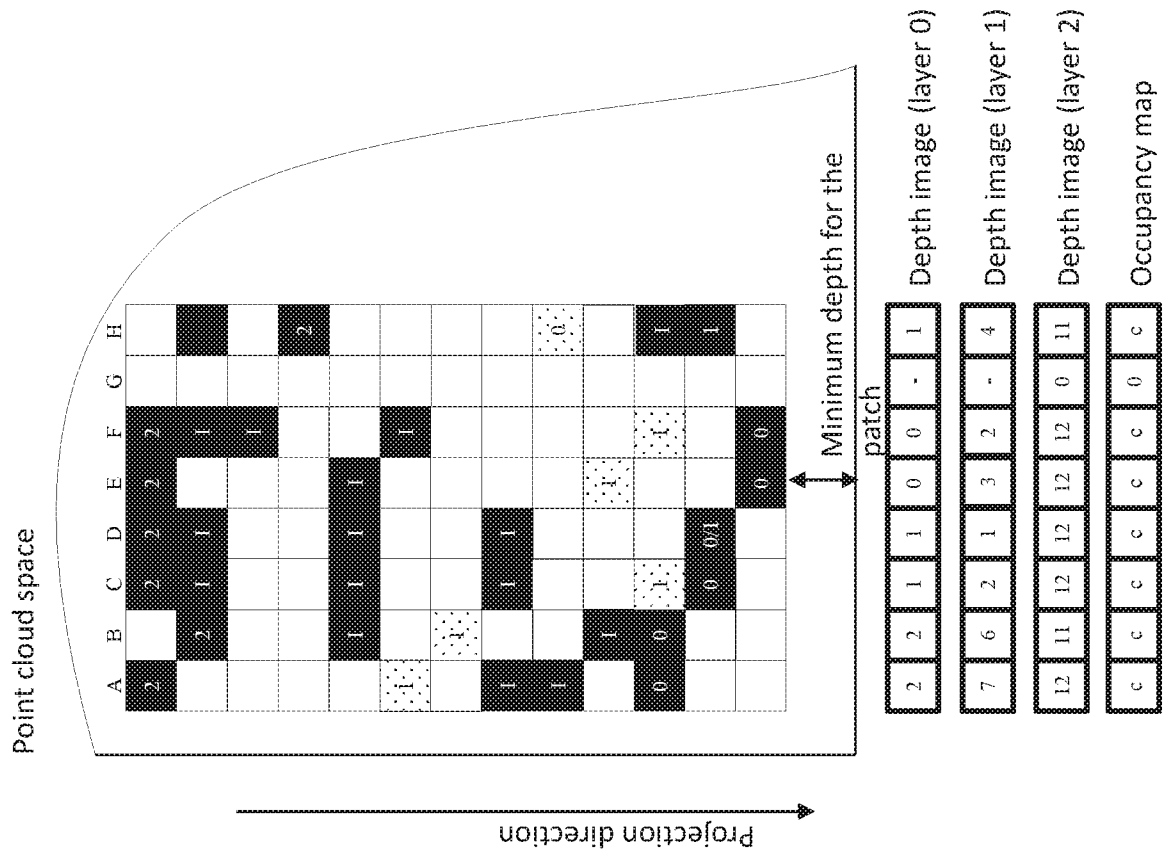


Fig. 11

	Descriptor
sequence_parameter_set() {	
profile_tier_level()	
sps_sequence_parameter_set_id	u(4)
sps_frame_width	u(16)
sps_frame_height	u(16)
sps_avg_frame_rate_present_flag	u(1)
if( sps_avg_frame_rate_present_flag )	
sps_avg_frame_rate	u(16)
sps_enhanced_occupancy_map_for_depth_flag	u(1)
if( sps_enhanced_occupancy_map_for_depth_flag )	
sps_enhanced_occupancy_map_code_max_bit_count_flag	u(1)
if( sps_enhanced_occupancy_map_code_max_bit_count_flag )	
sps_enhanced_occupancy_map_code_max_bit_count_minus1	u(4)
sps_layer_count_minus1	u(4)
...	
occupancy_parameter_set()	
geometry_parameter_set()	
...	
}	

Fig. 12

	Descriptor
sequence_parameter_set() {	
profile_tier_level()	
sps_sequence_parameter_set_id	u(4)
sps_frame_width	u(16)
sps_frame_height	u(16)
sps_avg_frame_rate_present_flag	u(1)
if( sps_avg_frame_rate_present_flag )	
sps_avg_frame_rate	u(16)
sps_enhanced_occupancy_map_for_depth_flag	u(1)
if( sps_enhanced_occupancy_map_for_depth_flag )	
sps_enhanced_occupancy_map_direction_flag	u(1)
sps_layer_count_minus1	u(4)
...	
occupancy_parameter_set()	
geometry_parameter_set()	
...	
}	

Fig. 13

	Descriptor
sequence_parameter_set() {	
profile_tier_level()	
sps_sequence_parameter_set_id	u(4)
sps_frame_width	u(16)
sps_frame_height	u(16)
sps_avg_frame_rate_present_flag	u(1)
if( sps_avg_frame_rate_present_flag )	
sps_avg_frame_rate	u(16)
sps_enhanced_occupancy_map_for_depth_flag	u(1)
if( sps_enhanced_occupancy_map_for_depth_flag )	
sps_enhanced_occupancy_map_code_min_bit_count_flag	u(1)
if( sps_enhanced_occupancy_map_code_min_bit_count_flag )	
sps_enhanced_occupancy_map_code_max_bit_count_flag	u(4)
sps_enhanced_occupancy_map_code_min_bit_count_minus2	u(4)
sps_layer_count_minus1	
...	
occupancy_parameter_set()	
geometry_parameter_set()	
...	
}	

Fig. 14

INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2020/016931

A. CLASSIFICATION OF SUBJECT MATTER  
INV. H04N19/597  
ADD.  
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED  
Minimum documentation searched (classification system followed by classification symbols)  
H04N  
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	J Llach ET AL: "Generalized Enhanced Occupancy Map for Depth", , no. m47895 26 March 2019 (2019-03-26), XP055622266, Retrieved from the Internet: URL:URL:http://phenix.int-evry.fr/mpeg [retrieved on 2019-09-13] the whole document ----- -/--	1-54

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search

14 May 2020

Date of mailing of the international search report

20/05/2020

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Raeymaekers, Peter

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2020/016931

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	<p>RAJAN JOSHI (SAMSUNG) ET AL: "[V-PCC] Clarification of tool behavior for more than two layers", 126. MPEG MEETING; 20190325 - 20190329; GENEVA; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), , no. m47541 20 March 2019 (2019-03-20), XP030211604, Retrieved from the Internet: URL:http://phenix.int-evry.fr/mpeg/doc_end_user/documents/126_Geneva/wg11/m47541-v1-m47541.zip m47541.DOCX [retrieved on 2019-03-20] the whole document</p> <p style="text-align: center;">-----</p>	1-54
X	<p>Zakharchenko: "V-PCC Codec description", , no. N18190 13 March 2019 (2019-03-13), XP055622270, Retrieved from the Internet: URL:http://phenix.int-evry.fr/mpeg/ [retrieved on 2019-09-13] Section 2.2.1.6</p> <p style="text-align: center;">-----</p>	1-54