



(51) International Patent Classification:

H04N 13/128 (2018.01) *H04N 13/268* (2018.01)
H04N 13/161 (2018.01) *H04N 13/271* (2018.01)
H04N 13/178 (2018.01) *H04N 21/235* (2011.01)
H04N 19/156 (2014.01) *H04N 21/236* (2011.01)
H04N 19/176 (2014.01) *H04N 21/43* (2011.01)
H04N 19/597 (2014.01) *H04N 21/845* (2011.01)
H04N 19/98 (2014.01) *G06T 7/593* (2017.01)

(21) International Application Number:

PCT/FI2018/050008

(22) International Filing Date:

04 January 2018 (04.01.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1700302.1 09 January 2017 (09.01.2017) GB

(71) Applicant: **NOKIA TECHNOLOGIES OY** [FI/FI];
Karaportti 3, 02610 Espoo (FI).

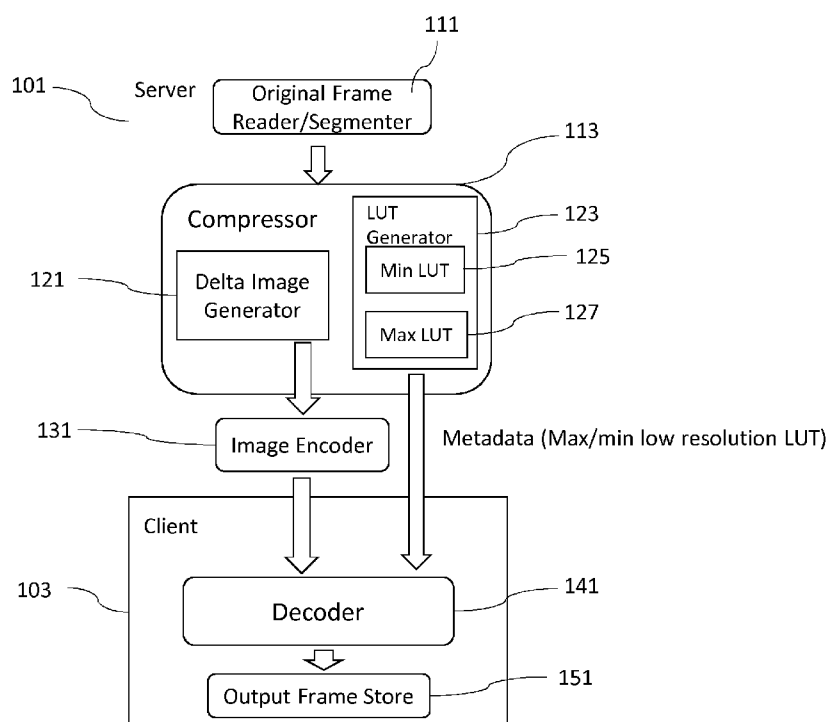
(72) Inventors: **PESONEN, Mika**; Lampuotinkatu 7 c 6, 33800
Tampere (FI). **RAJALA, Johannes**; Pohjantikantie 3,
39160 Julkujärvi (FI).

(74) Agent: **NOKIA TECHNOLOGIES OY** et al.; Ari Aarnio,
IPR Department, Karakaari 7, 02610 Espoo (FI).

(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, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(54) Title: METHOD AND APPARATUS FOR VIDEO DEPTH MAP CODING AND DECODING

Figure 1



(57) Abstract: A method for encoding at least one depth image frame within a sequence of depth image frames, the method comprising: segmenting each at least one depth image frame into an array of depth image blocks; determining for each depth image block a minimum depth image value and a maximum depth image value; generating a maximum depth image from the determined maximum depth image values; generating a minimum depth image from the determined minimum depth image values; and encoding values of the at least one depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

(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, 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))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

METHOD AND APPARATUS FOR VIDEO DEPTH MAP CODING AND DECODING

TECHNICAL FIELD

5 The present application relates generally to an apparatus, a method and a computer program for video coding and decoding.

BACKGROUND

10 This section is intended to provide a background or context to the invention that is recited in the claims. The description herein may include concepts that could be pursued, but are not necessarily ones that have been previously conceived or pursued. Therefore, unless otherwise indicated herein, what is described in this section is not prior art to the description and claims in this application and is not admitted to be prior art by inclusion in this section.

15 A video coding system may comprise an encoder that transforms an input video into a compressed representation suited for storage/transmission and a decoder that can uncompress the compressed video representation back into a viewable form. The encoder may discard some information in the original video sequence in order to represent the video in a more compact form, for example, to enable the storage/transmission of the video information at a lower bitrate than otherwise might
20 be needed.

Various technologies for providing three-dimensional (3D) video content are currently investigated and developed. Especially, studies have been focused on various multi-view applications wherein a viewer is provided with a limited number of input views, e.g. a mono or a stereo video plus supplementary data such as a depth
25 mapping of the images, is provided to a decoder side and all required views are then rendered (i.e. synthesized) locally by the decoder to be displayed on a display.

In the encoding of 3D video content, video compression systems, such as Advanced Video Coding standard H.264/AVC, the Multiview Video Coding MVC extension of H.264/AVC or scalable extensions of HEVC can be used.

30 However video decoders that support more than 8-bit values per pixel compression codecs are not commonly available for all platforms (Win, iOS/macOS, Android). Currently H.264 can be used for playing back video on all relevant platforms. However, H.264 supports only a maximum of 8-bit values for grey scale images and this is not sufficient to encode high quality depth maps. Typically depth maps require
35 16, 24 or even 32-bit values in order to avoid visual artefacts. Visual artefacts may

occur where objects intersect with each other. This is also known as z-fighting in computer graphics terms.

SUMMARY

5 According to a first aspect there is provided a method for encoding at least one depth image frame within a sequence of depth image frames, the method comprising: segmenting each at least one depth image frame into an array of depth image blocks; determining for each depth image block a minimum depth image value and a maximum depth image value; generating a maximum depth image from the determined
10 maximum depth image values; generating a minimum depth image from the determined minimum depth image values; and encoding values of the at least one depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

 The method may further comprise: combining the encoded values of the at least
15 one depth image frame with the determined blocks minimum depth image values and maximum depth image values; and transmitting or storing the combination of the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values.

 Generating a maximum depth image frame from the determined blocks
20 maximum depth image values may comprise applying one of a Gaussian blur filter or bilinear filter to the determined blocks maximum depth image values.

 Generating a minimum depth image frame from the determined blocks minimum depth image values may comprise applying one of a Gaussian blur filter or bilinear filter to the determined blocks minimum depth image values.

25 Determining for each depth image block a minimum depth image value and a maximum depth image value may comprise: determining for the sequence of depth image frames a minimum depth image value for each depth image block; and determining for the sequence of depth image frames a maximum depth image value for each depth image block.

30 Determining for each depth image block a minimum depth image value and a maximum depth image value may comprise: determining a block maximum depth image value as a maximum value from the block maximum depth image values from the block maximum depth image value and spatially neighbouring block maximum depth image values; and determining a block minimum depth image value as a

minimum value from the block minimum depth image values from the block minimum depth image value and spatially neighbouring block minimum depth image values.

The method may further comprise comparing, for each block, a sequence of determined blocks maximum depth image values to determine that the determined block maximum depth image value is equal to a maximum value from the sequence of determined blocks maximum depth image values.

The method may further comprise comparing, for each block, a sequence of determined blocks minimum depth image values to determine that the determined block minimum depth image value is equal to a minimum value from the sequence of determined blocks maximum depth image values.

Encoding values of the at least one depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions may comprise determining an encoded value of the at least one depth image frame by: determining a difference value by subtracting a corresponding location minimum depth image value from a location depth image value; and scaling the difference value by dividing the difference value by the difference between a corresponding location maximum image value and the corresponding location minimum image value.

According to a second aspect there is provided a method for decoding at least one depth image frame within a sequence of depth image frames, the method comprising: receiving an encoded depth image frame; receiving a depth image frame array of minimum values and array of maximum values; generating a maximum depth image from the depth image frame array of maximum values; generating a minimum depth image from the depth image frame array of maximum values; and decoding values of the encoded depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

Decoding values of the encoded depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions may comprise: rescaling a depth image value at a location to generate a rescaled depth image value by multiplying an encoded depth image frame at the location by the difference between a corresponding location maximum image value and a corresponding location minimum image value; and releveling the depth image value at the location to generate the decoded value by adding the corresponding location minimum image value to the rescaled depth image value.

According to a third aspect there is provided an apparatus for encoding at least one depth image frame within a sequence of depth image frames, the apparatus comprises: a segmenter configured to segment each at least one depth image frame into an array of depth image blocks; a minimum and maximum depth determiner
5 configured to determine for each depth image block a minimum depth image value and a maximum depth image value; an upscaler configured to generate a maximum depth image from the determined maximum depth image values and generate a minimum depth image from the determined minimum depth image values; and an encoder
10 configured to encode values of the at least one depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

The apparatus may further comprise a combiner configured to combine the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values.

15 The apparatus may comprise a transmitter configured to transmit the combination of the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values.

The apparatus may comprise a memory configured to store the combination of the encoded values of the at least one depth image frame with the determined blocks
20 minimum depth image values and maximum depth image values.

The upscaler configured to generate a maximum depth image frame from the determined blocks maximum depth image values may be configured to apply one of a Gaussian blur filter or bilinear filter to the determined blocks maximum depth image values.

25 The upscaler configured to generate a minimum depth image frame from the determined blocks minimum depth image values is configured to apply one of a Gaussian blur filter or bilinear filter to the determined blocks minimum depth image values.

The minimum and maximum depth determiner may be configured to: determine
30 for the sequence of depth image frames a minimum depth image value for each depth image block; and determine for the sequence of depth image frames a maximum depth image value for each depth image block.

The minimum and maximum depth determiner may be configured to: determine a block maximum depth image value as a maximum value from the block maximum
35 depth image values from the block maximum depth image value and spatially

neighbouring block maximum depth image values; and determine a block minimum depth image value as a minimum value from the block minimum depth image values from the block minimum depth image value and spatially neighbouring block minimum depth image values.

5 The apparatus may comprise a maximum value comparator configured to compare, for each block, a sequence of determined blocks maximum depth image values to determine that the determined block maximum depth image value is equal to a maximum value from the sequence of determined blocks maximum depth image values.

10 The apparatus may comprise a minimum value comparator configured to compare, for each block, a sequence of determined blocks minimum depth image values to determine that the determined block minimum depth image value is equal to a minimum value from the sequence of determined blocks maximum depth image values.

15 The encoder may comprise: a delta image generator configured to determine a difference value by subtracting a corresponding location minimum depth image value from a location depth image value; and a scaler configured to scale the difference value by dividing the difference value by the difference between a corresponding location maximum image value and the corresponding location minimum image value.

20 According to a fourth aspect there is provided an apparatus for decoding at least one depth image frame within a sequence of depth image frames, the apparatus comprising: an input configured to receive an encoded depth image frame; an input configured to receive a depth image frame array of minimum values and array of maximum values; an upscaler configured to generate a maximum depth image from the depth image frame array of maximum values, and a minimum depth image from the depth image frame array of maximum values; and a decoder configured to decode values of the encoded depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

25 The decoder may comprise: a rescaler configured to rescale a depth image value at a location to generate a rescaled depth image value by multiplying an encoded depth image frame at the location by the difference between a corresponding location maximum image value and a corresponding location minimum image value; and a relever configured to relever the depth image value at the location to generate the decoded value by adding the corresponding location minimum image value to the rescaled depth image value.

30

35

Each depth image block may comprise at least one of: 16x16 depth image values; and 32x32 depth image values.

BRIEF DESCRIPTION OF THE DRAWINGS

5 For a more complete understanding of example embodiments of the present invention, reference is now made to the following descriptions taken in connection with the accompanying drawings in which:

Figure 1 shows schematically an example system employing some embodiments;

10 Figure 2 shows schematically a flow diagram of the operation of the device shown in Figure 1;

Figure 3 shows schematically a series of example read depth map frames according to some embodiments;

15 Figure 4 shows schematically an example block segmented frame generated from according to some embodiments;

Figure 5 shows schematically maximum and minimum depth values assigned for each block within the segmented frame as shown in Figure 4;

Figure 6 shows spatial neighbouring block maximum (and minimum) depth value comparison according to some embodiments;

20 Figure 7 shows an example depth value image;

Figure 8 shows an example upscaled version of a block segmented minimum depth value image generated from the example depth value image shown in Figure 7;

Figure 9 shows an example upscaled version of a block segmented maximum depth value image generated from the example depth value image shown in Figure 7;

25 Figure 10 shows an example maximum minus minimum depth value image based on the upscaled version of a block segmented minimum depth value image of Figure 8 and the example upscaled version of a block segmented maximum depth value image of Figure 9; and

30 Figure 11 shows an example compressed depth value image based on the image of claim 7 and processed according to some embodiments.

DETAILED DESCRIPTION OF SOME EXAMPLE EMBODIMENTS

In the following, several embodiments will be described in the context of one video coding arrangement and example. It is to be noted, however, that the invention
35 is not limited to this particular arrangement or example. In fact, the different

embodiments have applications widely in any environment where improvement of depth values compression is required. For example, the invention may be applicable to video coding systems like streaming systems, DVD players, digital television receivers, personal video recorders, systems and computer programs on personal
5 computers, handheld computers and communication devices, as well as network elements such as transcoders and cloud computing arrangements where video data is handled.

In the following, several embodiments are described using the convention of referring to (de)coding, which indicates that the embodiments may apply to decoding
10 and/or encoding.

Depth-enhanced video refers to texture video having one or more views associated with depth video having one or more depth views. A number of approaches may be used for representing of depth-enhanced video, including the use of video plus depth (V+D), multiview video plus depth (MVD), and layered depth video (LDV). In the
15 video plus depth (V+D) representation, a single view of texture and the respective view of depth are represented as sequences of texture picture and depth pictures, respectively. The MVD representation contains a number of texture views and respective depth views. In the LDV representation, the texture and depth of the central view are represented conventionally, while the texture and depth of the other views are
20 partially represented and cover only the dis-occluded areas required for correct view synthesis of intermediate views.

It is to be appreciated that the term 'depth view' may refer to a view that represents distance information of a texture sample from the camera sensor, disparity or parallax information between a texture sample and a respective texture sample in
25 another view, or similar information. A depth view may include depth images (which may be also referred to as depth image maps and depth view components) having one component, similar to the luma component of texture views. The other color components, similar to chroma components of texture views, may be absent in the depth views, and may be set to default values (e.g. by an encoder) and/or may be
30 omitted (e.g. by a decoder).

A depth map image may be considered to represent the values related to the distance of the surfaces of the scene objects from a reference location, for example a view point of an observer. A depth map image is an image that may include per-pixel depth information or any similar information. For example, each sample in a depth map
35 image represents the distance of the respective texture sample or samples from the

plane on which the camera lies. In other words, if the z axis is along the shooting axis of the cameras (and hence orthogonal to the plane on which the cameras lie), a sample in a depth map image represents the value on the z axis.

5 Since depth map images are generated containing a depth value for each pixel in the image, they can be depicted as gray-level images or images containing only the luma component. Alternatively chroma components of the depth map images may be set to a pre-defined value, such as a value indicating no chromaticity, e.g. 128 in typical 8-bit chroma sample arrays, where a zero chromaticity level is arranged into the middle of the value range. Alternatively, chroma components of depth map images may be used to contain other picture data, such as any type of monochrome auxiliary pictures, such as alpha planes.

A texture view component may be defined as a coded representation of the texture of a view in a single access unit. A texture view component in depth-enhanced video bitstream may be coded in a manner that is compatible with a single-view texture bitstream or a multi-view texture bitstream so that a single-view or multi-view decoder can decode the texture views even if it has no capability to decode depth views. For example, an H.264/AVC decoder may decode a single texture view from a depth-enhanced H.264/AVC bitstream. A texture view component may alternatively be coded in a manner that a decoder capable of single-view or multi-view texture decoding, such as H.264/AVC or MVC decoder, is not able to decode the texture view component for example because it uses depth-based coding tools. A depth view component may be defined as a coded representation of the depth of a view in a single access unit. A view component pair may be defined as a texture view component and a depth view component of the same view within the same access unit.

25 Depth-enhanced video may be coded in a manner where texture and depth are coded independently of each other. For example, texture views may be coded as one MVC bitstream and depth views may be coded as another MVC bitstream. Depth-enhanced video may also be coded in a manner where texture and depth are jointly coded. In a form of a joint coding of texture and depth views, some decoded samples of a texture picture or data elements for decoding of a texture picture are predicted or derived from some decoded samples of a depth picture or data elements obtained in the decoding process of a depth picture. Alternatively or in addition, some decoded samples of a depth picture or data elements for decoding of a depth picture are predicted or derived from some decoded samples of a texture picture or data elements obtained in the decoding process of a texture picture. In another option, coded video

data of texture and coded video data of depth are not predicted from each other or one is not coded/decoded on the basis of the other one, but coded texture and depth view may be multiplexed into the same bitstream in the encoding and demultiplexed from the bitstream in the decoding. In yet another option, while coded video data of texture is not predicted from coded video data of depth in e.g. below slice layer, some of the high-level coding structures of texture views and depth views may be shared or predicted from each other. For example, a slice header of coded depth slice may be predicted from a slice header of a coded texture slice. Moreover, some of the parameter sets may be used by both coded texture views and coded depth views.

Depth-enhanced video formats enable generation of virtual views or pictures at camera positions that are not represented by any of the coded views. Generally, any depth-image-based rendering (DIBR) algorithm may be used for synthesizing views.

Depth information can be obtained by various means. For example, depth of the 3D scene may be computed from the disparity registered by capturing cameras or colour image sensors. A depth estimation approach, which may also be referred to as stereo matching, takes a stereoscopic view as an input and computes local disparities between the two offset images of the view. Since the two input views represent different viewpoints or perspectives, the parallax creates a disparity between the relative positions of scene points on the imaging planes depending on the distance of the points. A target of stereo matching is to extract those disparities by finding or detecting the corresponding points between the images. Several approaches for stereo matching exist. For example, in a block or template matching approach each image is processed pixel by pixel in overlapping blocks, and for each block of pixels a horizontally localized search for a matching block in the offset image is performed. Once a pixel-wise disparity is computed, the corresponding depth value z is calculated by:

$$z = \frac{f \cdot b}{d + \Delta d},$$

where f is the focal length of the camera and b is the baseline distance between cameras. Further, d may be considered to refer to the disparity observed between the two cameras or the disparity estimated between corresponding pixels in the two cameras. The camera offset Δd may be considered to reflect a possible horizontal misplacement of the optical centres of the two cameras or a possible horizontal cropping in the camera frames due to pre-processing. However, since the algorithm is based on block matching, the quality of a depth-through-disparity estimation is content dependent and very often not accurate. For example, no straightforward solution for

depth estimation is possible for image fragments that are featuring very smooth areas with no textures or large level of noise. Another approach to represent the depth values of different views in the stereoscopic or multiview case is to report the disparity between pixels of each view to the adjacent view instead of the actual depth values.

- 5 The following equation shows how depth values are converted to disparity:

$$D = f \times l \times \left(\frac{d}{2^N - 1} \times \left(\frac{1}{Z_{near}} - \frac{1}{Z_{far}} \right) + \frac{1}{Z_{far}} \right)$$

where:

D = disparity value

- 10 f = focal length of capturing camera

l = translational difference between cameras

d = depth map value

N = number of bits representing the depth map values

- 15 Z_{near} and Z_{far} are the respective distances of the closest and farthest objects in the scene to the camera (mostly available from the content provider), respectively.

The semantics of depth map values may for example include the following:

1. Each luma sample value in a coded depth view component represents an inverse of real-world distance (Z) value, i.e. $1/Z$, normalized in the dynamic range of the luma samples, such as to the range of 0 to 255, inclusive, for 8-bit luma representation. The normalization may be done in a manner where the quantization $1/Z$ is uniform in terms of disparity.
2. Each luma sample value in a coded depth view component represents an inverse of real-world distance (Z) value, i.e. $1/Z$, which is mapped to the dynamic range of the luma samples, such as to the range of 0 to 255, inclusive, for 8-bit luma representation, using a mapping function $f(1/Z)$ or table, such as a piece-wise linear mapping. In other words, depth map values result in applying the function $f(1/Z)$.
3. Each luma sample value in a coded depth view component represents a real-world distance (Z) value normalized in the dynamic range of the luma samples, such as to the range of 0 to 255, inclusive, for 8-bit luma representation.
- 30 4. Each luma sample value in a coded depth view component represents a disparity or parallax value from the present depth view to another indicated or derived depth view or view position.

The semantics of depth map values may be indicated in the bit-stream, for example, within a video parameter set syntax structure, a sequence

parameter set syntax structure, a video usability information syntax structure, a picture parameter set syntax structure, a camera/depth/adaptation parameter set syntax structure, a supplemental enhancement information message, or anything alike.

5 An encoding system or any other entity creating or modifying a bitstream including coded depth image maps may create and include information on the semantics of depth samples and on the quantization scheme of depth samples into the bitstream. Such information on the semantics of depth samples and on the quantization scheme of depth samples may be for example included in a
10 video parameter set structure, in a sequence parameter set structure, or in a supplemental enhancement information (SEI) message.

 Alternatively or in addition to the above-described stereo view depth estimation, the depth value may be obtained using the time-of-flight (TOF) principle for example by using a camera which may be provided with a light source, for example an infrared
15 emitter, for illuminating the scene. Such an illuminator may be arranged to produce an intensity modulated electromagnetic emission for a frequency between e.g. 10–100 MHz, which may require LEDs or laser diodes to be used. Infrared light may be used to make the illumination unobtrusive. The light reflected from objects in the scene is detected by an image sensor, which may be modulated synchronously at the same
20 frequency as the illuminator. The image sensor may be provided with optics; a lens gathering the reflected light and an optical bandpass filter for passing only the light with the same wavelength as the illuminator, thus helping to suppress background light. The image sensor may measure for each pixel the time the light has taken to travel from the illuminator to the object and back. The distance to the object may be
25 represented as a phase shift in the illumination modulation, which can be determined from the sampled data simultaneously for each pixel in the scene.

 Alternatively or in addition to the above-described stereo view depth estimation and/or TOF-principle depth sensing, depth values may be obtained using a structured light approach which may operate for example approximately as follows. A light emitter,
30 such as an infrared laser emitter or an infrared LED emitter, may emit light that may have a certain direction in a 3D space (e.g. follow a raster-scan or a pseudo-random scanning order) and/or position within an array of light emitters as well as a certain pattern, e.g. a certain wavelength and/or amplitude pattern. The emitted light is reflected back from objects and may be captured using a sensor, such as an infrared

image sensor. The image/signals obtained by the sensor may be processed in relation to the direction of the emitted light as well as the pattern of the emitted light to detect a correspondence between the received signal and the direction/position of the emitted light as well as the pattern of the emitted light for example using a triangulation principle. From this correspondence a distance and a position of a pixel may be concluded.

It is to be understood that the above-described depth estimation and sensing methods are provided as non-limiting examples and embodiments may be realized with the described or any other depth estimation and sensing methods and apparatuses.

Figure 1 shows schematically an example system for compressing and decompressing the depth map values or depth images according to some embodiments.

Figure 1, for example shows an example server, or encoder/compressor, 101 and client, or decoder/decompressor, 103. The server 101 is configured to receive depth image frames (or the array of depth values) and output at least one compressed value range image.

In some embodiments of the server 101 comprises a frame reader/segmenter 111. The frame reader/segmenter 111 is configured to receive a sequence of depth image frames (such as those generated by the methods described above or generated by any suitable other method such as software stitching or Lidar). An example of a depth image frame is one which comprises 4096x2048 32-bit floating point depth values. It is understood that the frame size (4096x2048) is an example only and the frame may be any suitable size. Similarly the representation of the depth value within the frame, a 32-bit floating point representation is an example only and the depth value may be any suitable other representation. An example depth image frame sequence is shown in Figure 3 which shows a sequence of depth map frames from frame 1 301_1 , frame 2 301_2 , frame 3 301_3 to frame N 301_N . Each frame is a frame captured at a different time instant and in the example shown in Figure 3 shows the image frames progressing in time t. A further example of a depth image is shown in Figure 7 which shows an example image where there are foreground objects located against a varying background.

The frame reader/segmenter 111 having received the sequence of frames may segment or split the depth image frames into blocks. For example in some embodiments the reader/segmenter 111 can be configured to split the 4096x2048

depth values per frame into non-overlapping blocks of 16x16 depth values. The 16x16 block depth value blocks are example only and it is understood that any suitable block size may be used. For example the block may be a 32x32 block. In such a manner the 4096x2048 depth image frame may be represented as 256x128 blocks. These blocks
 5 can be passed to a compressor 113 and in some embodiments a look up table (LUT) generator 123 which is part of the compressor 113.

An example of the blocks generated by the reader/segmenter 111 is shown in Figure 4 which shows frame 1 301₁, segmented into 16x16 value blocks. For example Figure 4 shows for frame 1 301₁ an upper left block 401_{1,1,1}, an upper right block
 10 401_{1,1,256}, a lower left block 401_{1,128,1} and a lower right block 401_{1,128,256}, where the first subscript reference is the frame number, the second subscript reference the block row number and the third subscript the block column number.

In some embodiments the server 101 comprises a compressor 113. The compressor 113 in some embodiments comprises Look up table (LUT) generator 123.

15 The LUT generator 123 may be configured to receive each of the depth value blocks. The LUT generator 123 may comprise a minimum value LUT generator 125 and a maximum value LUT generator 127. In some embodiments the functionality of the maximum and minimum value LUT generators is performed by the same functional element.

20 The LUT generator 123 may be configured to receive each block and determine a minimum and a maximum depth value for each block. Having determined a minimum and a maximum depth value for each block these values may be passed to the minimum value LUT generator 125 and the maximum value LUT generator 127 and used to create a low resolution minimum lookup table (256x128) of minimum values
 25 and maximum lookup table (256x128) of maximum values. An example of such tables are shown in Figure 5 which shows the block 401_{1,1,1} from frame 1 301 and the generation of the low resolution minimum value lookup table (256x128) 503 and a minimum value 503_{1,1,1} associated with the block 401_{1,1,1} and maximum value lookup table (256x128) 501 and a maximum value 501_{1,1,1} associated with the block 401_{1,1,1}.

30 In some embodiments the LUT generator 123 is configured to determine the minimum/maximum value for a sequence of frames for each block. Thus for example the LUT generator 123 may be configured to compare the minimum value block 401_{1,1,1} from frame 1 301, against minimum value blocks 401_{1-T,1,1} to 401_{1+T,1,1} to determine a minimum value for the range of blocks over times -T to +T and set the minimum value
 35 block to the minimum value for all of the compared blocks. Similarly the LUT generator

123 may be configured to compare the maximum value block from frame 1, against maximum value blocks over a range of values from $-T$ to $+T$ to determine a maximum value for the range of blocks over times $-T$ to $+T$ and set the maximum value block to the maximum value for all of the compared blocks. The range of frame values may be any suitable range of values and is not limited to $-T$ to $+T$.

In some embodiments the look up table generator 123 (and the minimum value LUT generator 125 and the maximum LUT generator 127) is then configured to compare each block minimum value to spatial neighbouring block minimum values and where a neighbouring block value is less than the current block value to set the current block value to the neighbouring block value. In other words checking whether the minimum value for a block is equal to the minimum value for all the neighbouring blocks.

This operation is required in order to limit over/under flow of the values once Gaussian or bilinear filtering has been applied. Furthermore in some circumstances the depth values will fit inside the temporal sequence min value range once the upscaled min LUTs are generated (as described later).

A similar check of each block maximum value against spatial neighbouring block maximum values for each block is also performed in some embodiments. In this check each block maximum value is compared to spatial neighbouring block maximum values and the current block maximum value is set to the neighbouring block maximum value where the neighbouring block value is greater than the current block maximum value.

An example of the spatial neighbouring block check is shown in Figure 6 which shows an example block maximum value $501_{t,x,y}$ and column neighbouring blocks maximum values $501_{t,x,y-1}$ and $501_{t,x,y+1}$, row neighbouring blocks maximum values $501_{t,x-1,y}$ and $501_{t,x+1,y}$, first diagonal neighbouring blocks maximum values $501_{t,x-1,y-1}$ and $501_{t,x+1,y+1}$ and second diagonal neighbouring blocks maximum values $501_{t,x+1,y-1}$ and $501_{t,x-1,y+1}$. In other words each spatial neighbouring check may be performed by comparing the block maximum value $501_{t,x,y}$ against block maximum values $501_{t,x,y-1}$, $501_{t,x,y+1}$, $501_{t,x-1,y}$, $501_{t,x+1,y}$, $501_{t,x-1,y-1}$, $501_{t,x+1,y+1}$, $501_{t,x+1,y-1}$ and $501_{t,x-1,y+1}$ and setting the block maximum value to the maximum of all of the values.

The LUT generator 123 (and the minimum value LUT generator 125 and the maximum value LUT generator 127) can further be configured to upscale the low resolution 256×128 value images to form a full resolution 4096×2048 image for both of the maximum and minimum value images. The upscaling may be achieved using any suitable upscaling method. For example in some embodiments the upscaling is

performed using Gaussian filters or bilinear filters to interpolate between the block low resolution values to the full resolution values.

An example of an upscaled minimum distance image based on the image shown in Figure 7 is shown in Figure 8. An example of an upscaled maximum distance image based on the image shown in Figure 7 is shown in Figure 9.

Having determined the upscaled minimum distance image and the upscaled maximum distance image these can be passed to a delta image generator 121.

The compressor 113 may in some embodiments comprise a delta image generator 121. The delta image generator receives the upscaled minimum distance image and the upscaled maximum distance image and generates a delta or difference image which comprises for each of the pixels the difference value between the upscaled minimum distance image and the upscaled maximum distance image. The delta image may be passed to an image encoder 131.

An example of the delta image based on the image shown in Figure 7 is shown in Figure 10. In this image it is possible to see where there are areas of large differences and small difference. Thus it may be possible to use the available small dynamic range (such as allowed by the use of 8 bits) efficiently for encoding or compressing the depth values in such a way that the small difference areas may have smaller quantization steps when compared to the areas where there are large differences (and require larger quantization steps).

The LUT generator 123 (and the minimum value LUT generator 125 and the maximum LUT generator 127) can further be configured to verify that the determined maximum and minimum values are the maximum and minimum values not only for the frame being analysed but all of the frames in the current sequence. In other words the LUT generator verifies that the maximum and minimum are the maximum and minimum for a range of temporal separated frames.

Where the minimum or maximum values are exceeded across the sequence range of frames then the range is extended to cover the minimum or maximum value. This may in some embodiments require the regeneration of the upscaled maximum depth image, upscaled minimum depth image and the delta image. In some corner cases bilinear/Gaussian upscaling may provide out of range values and therefore validation of all the blocks for a sequence of frames to determine that they fit inside the range is performed. If there are overflows the range may be extended a bit until all the values are inside.

The LUT generator 123 (and the minimum value LUT generator 125 and the maximum LUT generator 127) may then output the block (low resolution) maximum and minimum lookup table values for this sequence of frames to the image encoder 131 to be encoded as meta data. Furthermore the upscaled minimum distance image and the upscaled maximum distance image may also be output to the image encoder 131.

The server 101 further comprises an image encoder 131. The image encoder 131 may be configured to receive the original image frames from the reader/segmenter 111, the delta image from the delta image generator 121 and the block (low resolution) maximum and minimum lookup table values for this sequence of frames, the upscaled minimum distance images and the upscaled maximum distance images from the LUT generator 123.

The image encoder may then encode the actual depth sequence image frames to video frames. For example in some embodiments each depth video frame pixel may be represented as

$$\text{compressedOutput} = (\text{depthValue} - \text{lutMIN}) / (\text{lutMAX} - \text{lutMIN}) * 255.0f$$

where compressed output is the 8-bit output value which will be used in the video encoding, depthValue represents the original (32-bit floating point) value that is to be compressed and lutMIN/lutMAX values are the upscaled (Gaussian or bilinear filtered) image values.

An example of the compressed output value image based on the image shown in Figure 7 is shown in Figure 11.

With respect to Figure 2 the flow diagram of the operation of the server is shown in further detail.

The first operation is to read the (4096x2048 32-bit floating point) depth frames and segment the depth frames into (16x16) blocks.

The operation of reading the depth frames and segmenting these frames is shown in Figure 2 by step 201.

The system may then further determine the minimum and maximum depth values for each (16x16) block and for all of the depth frame images for a sequence of image and creates a low resolution minimum/maximum lookup table.

The operation of determining the minimum/maximum values for the (16x16) blocks and creating a low resolution minimum value image and low resolution maximum value image is shown in Figure 2 by step 203.

The system further comprises on a block by block basis comparing the minimum values to spatially neighbouring block minimum values. Furthermore on a block by block basis comparing the maximum values to spatially neighbouring block maximum values.

5 The operation of comparing the values to specially neighbouring block values is shown in Figure 2 by step 205.

The system further comprises upscaling the minimum and maximum low resolution images (using a Gaussian blur or bilinear filter) to generate an original resolution minimum and maximum depth image.

10 The operation of upscaling the minimum and maximum low resolution images (using a Gaussian blur or bilinear filter) to generate original resolution depth images is shown in Figure 2 by step 207.

The system further comprises verifying that depth values for a block for the time sequence of image frames are within the minimum and maximum range.

15 The operation of verifying that depth values for a block for the time sequence of image frames are within the minimum and maximum range is shown in Figure 2 by step 209.

Where the determined value for a block is outside the maximum/minimum range then the system further extends the block value range so that all depth values for the sequence fit inside the range.

20 The operation of extending the minimum/maximum range is shown in Figure 2 by step 210.

Where the determined value for a block is within the range then the next operation is to store/output the minimum/maximum values as video metadata.

25 The operation of storing the minimum/maximum values as video meta data is shown in Figure 2 by step 211.

Furthermore the system comprises encoding the depth image frames using the upscaled minimum/maximum images.

30 The operation of encoding or compressing the depth image frames using the upscaled minimum/maximum image values is shown in Figure 2 by step 213.

The system then encodes the depth images with the maximum/minimum meta data and outputs this combination.

35 The operation of combining the encoded depth image frames with the minimum/maximum lookup table information as meta data is shown in Figure 2 by step 215.

Decompression of the compressed/encoded sequence may be performed by the client 103 shown in Figure 1.

The client 103 comprises a decoder 141 configured to receive the encoded depth images and the maximum and minimum LUT block values.

5 The decoder 141 may then read or extract the low-resolution minimum/maximum value look-up tables (LUT). The minimum/maximum value LUTs may then be upsampled using the same algorithm as performed in the LUT generator 123 to generate upsampled maximum and minimum depth image frames.

10 Then using the compressed images (compressedValue) and the upsampled look-up tables a decompressed or decoded (floating point) depth value can be determined. For example the decompressed or decoded (floating point) depth value can be generated by implementing the following formula:

$$\text{depthValue} = (\text{compressedValue} * (\text{lutMax} - \text{lutMin})) / 255.0 + \text{lutMin}.$$

15 The decoder 141 may then pass the decoded depth frame data to a output frame store 151.

In some embodiments the client may comprise a output frame store 151 configured to receive the decoded depth frame data from the decoder 141 and store/output the depth frame.

20 Embodiments of the present invention may be implemented in software, hardware, application logic or a combination of software, hardware and application logic. In an example embodiment, the application logic, software or an instruction set is maintained on any one of various conventional computer-readable media. In the context of this document, a "computer-readable medium" may be any media or means that can contain, store, communicate, propagate or transport the instructions for use
25 by or in connection with an instruction execution system, apparatus, or device, such as a computer, with one example of a computer described and depicted in Figure 1. A computer-readable medium may comprise a computer-readable storage medium that may be any media or means that can contain or store the instructions for use by or in connection with an instruction execution system, apparatus, or device, such as a
30 computer.

If desired, the different functions discussed herein may be performed in a different order and/or concurrently with each other. Furthermore, if desired, one or more of the above-described functions may be optional or may be combined.

35 Although the above examples describe embodiments operating within a codec within an electronic device, it would be appreciated that the invention as described

below may be implemented as part of any video codec. Thus, for example, embodiments of the invention may be implemented in a video codec which may implement video coding over fixed or wired communication paths.

In general, the various embodiments of the invention may be implemented in hardware or special purpose circuits, software, logic or any combination thereof. For example, some aspects may be implemented in hardware, while other aspects may be implemented in firmware or software which may be executed by a controller, microprocessor or other computing device, although the invention is not limited thereto. While various aspects of the invention may be illustrated and described as block diagrams, flow charts, or using some other pictorial representation, it is well understood that these blocks, apparatuses, systems, techniques or methods described herein may be implemented in, as non-limiting examples, hardware, software, firmware, special purpose circuits or logic, general purpose hardware or controller or other computing devices, or some combination thereof.

The embodiments of this invention may be implemented by computer software executable by a data processor of the mobile device, such as in the processor entity, or by hardware, or by a combination of software and hardware. Further in this regard it should be noted that any blocks of the logic flow as in the Figures may represent program steps, or interconnected logic circuits, blocks and functions, or a combination of program steps and logic circuits, blocks and functions. The software may be stored on such physical media as memory chips, or memory blocks implemented within the processor, magnetic media such as hard disk or floppy disks, and optical media such as for example DVD and the data variants thereof, CD.

The various embodiments of the invention can be implemented with the help of computer program code that resides in a memory and causes the relevant apparatuses to carry out the invention. For example, a terminal device may comprise circuitry and electronics for handling, receiving and transmitting data, computer program code in a memory, and a processor that, when running the computer program code, causes the terminal device to carry out the features of an embodiment. Yet further, a network device may comprise circuitry and electronics for handling, receiving and transmitting data, computer program code in a memory, and a processor that, when running the computer program code, causes the network device to carry out the features of an embodiment.

The memory may be of any type suitable to the local technical environment and may be implemented using any suitable data storage technology, such as

semiconductor-based memory devices, magnetic memory devices and systems, optical memory devices and systems, fixed memory and removable memory. The data processors may be of any type suitable to the local technical environment, and may include one or more of general purpose computers, special purpose computers, microprocessors, digital signal processors (DSPs) and processors based on multi-core processor architecture, as non-limiting examples.

Embodiments of the inventions may be practiced in various components such as integrated circuit modules. The design of integrated circuits is by and large a highly automated process. Complex and powerful software tools are available for converting a logic level design into a semiconductor circuit design ready to be etched and formed on a semiconductor substrate.

Programs, such as those provided by Synopsys Inc., of Mountain View, California and Cadence Design, of San Jose, California automatically route conductors and locate components on a semiconductor chip using well established rules of design as well as libraries of pre-stored design modules. Once the design for a semiconductor circuit has been completed, the resultant design, in a standardized electronic format (e.g., Opus, GDSII, or the like) may be transmitted to a semiconductor fabrication facility or "fab" for fabrication.

The foregoing description has provided by way of exemplary and non-limiting examples a full and informative description of the exemplary embodiment of this invention. However, various modifications and adaptations may become apparent to those skilled in the relevant arts in view of the foregoing description, when read in conjunction with the accompanying drawings and the appended claims. However, all such and similar modifications of the teachings of this invention will still fall within the scope of this invention.

Claims:

1. A method for encoding at least one depth image frame within a sequence of depth image frames, the method comprising:
 - 5 segmenting each at least one depth image frame into an array of depth image blocks;
determining for each depth image block a minimum depth image value and a maximum depth image value;
generating a maximum depth image from the determined maximum depth
10 image values;
generating a minimum depth image from the determined minimum depth image values; and
encoding values of the at least one depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding
15 positions.
2. The method as claimed in claim 1, comprising:
 - combining the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values;
20 and
transmitting or storing the combination of the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values.
- 25 3. The method as claimed in any of claims 1 or 2, wherein generating a maximum depth image frame from the determined blocks maximum depth image values comprises applying one of a Gaussian blur filter or bilinear filter to the determined blocks maximum depth image values.
- 30 4. The method as claimed in any of claims 1 to 3, wherein generating a minimum depth image frame from the determined blocks minimum depth image values comprises applying one of a Gaussian blur filter or bilinear filter to the determined blocks minimum depth image values.

5. The method as claimed in any of claims 1 to 4, wherein determining for each depth image block a minimum depth image value and a maximum depth image value comprises:

determining for the sequence of depth image frames a minimum depth image value for each depth image block; and

determining for the sequence of depth image frames a maximum depth image value for each depth image block.

6. The method as claimed in any of claims 1 to 5, wherein determining for each depth image block a minimum depth image value and a maximum depth image value comprises:

determining a block maximum depth image value as a maximum value from the block maximum depth image values from the block maximum depth image value and spatially neighbouring block maximum depth image values; and

determining a block minimum depth image value as a minimum value from the block minimum depth image values from the block minimum depth image value and spatially neighbouring block minimum depth image values.

7. The method as claimed in any of claims 1 to 6, further comprising comparing, for each block, a sequence of determined blocks maximum depth image values to determine that the determined block maximum depth image value is equal to a maximum value from the sequence of determined blocks maximum depth image values.

8. The method as claimed in any of claims 1 to 7, further comprising comparing, for each block, a sequence of determined blocks minimum depth image values to determine that the determined block minimum depth image value is equal to a minimum value from the sequence of determined blocks maximum depth image values.

9. The method as claimed in any of claims 1 to 8, wherein encoding values of the at least one depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions comprises determining an encoded value of the at least one depth image frame by:

determining a difference value by subtracting a corresponding location minimum depth image value from a location depth image value; and

scaling the difference value by dividing the difference value by the difference between a corresponding location maximum image value and the corresponding location minimum image value.

10. A method for decoding at least one depth image frame within a sequence of depth image frames, the method comprising:

receiving an encoded depth image frame;

receiving a depth image frame array of minimum values and array of maximum values;

generating a maximum depth image from the depth image frame array of maximum values;

generating a minimum depth image from the depth image frame array of maximum values; and

decoding values of the encoded depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

11. The method for decoding as claimed in claim 10, wherein decoding values of the encoded depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions comprises:

rescaling a depth image value at a location to generate a rescaled depth image value by multiplying an encoded depth image frame at the location by the difference between a corresponding location maximum image value and a corresponding location minimum image value; and

releveling the depth image value at the location to generate the decoded value by adding the corresponding location minimum image value to the rescaled depth image value.

12. An apparatus for encoding at least one depth image frame within a sequence of depth image frames, the apparatus comprises:

a segmenter configured to segment each at least one depth image frame into an array of depth image blocks;

a minimum and maximum depth determiner configured to determine for each depth image block a minimum depth image value and a maximum depth image value;

an upscaler configured to generate a maximum depth image from the determined maximum depth image values and generate a minimum depth image from the determined minimum depth image values; and

an encoder configured to encode values of the at least one depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

13. The apparatus as claimed in claim 12, further comprising a combiner configured to combine the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values.

14. The apparatus as claimed in any of claims 12 and 13, further comprising a transmitter configured to transmit the combination of the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values.

15. The apparatus as claimed in any of claims 12 to 14, further comprising a memory configured to store the combination of the encoded values of the at least one depth image frame with the determined blocks minimum depth image values and maximum depth image values.

16. The apparatus as claimed in any of claims 12 to 15, wherein the upscaler is configured to apply one of a Gaussian blur filter or bilinear filter to the determined blocks maximum depth image values.

17. The apparatus as claimed in any of claims 12 to 16, wherein the upscaler is configured to apply one of a Gaussian blur filter or bilinear filter to the determined blocks minimum depth image values.

18. The apparatus as claimed in any of claims 12 to 17 wherein the minimum and maximum depth determiner is configured to:

determine for the sequence of depth image frames a minimum depth image value for each depth image block; and

determine for the sequence of depth image frames a maximum depth image value for each depth image block.

19. The apparatus as claimed in any of claims 12 to 18, wherein the minimum and maximum depth determiner is configured to:

determine a block maximum depth image value as a maximum value from the block maximum depth image values from the block maximum depth image value and spatially neighbouring block maximum depth image values; and

determine a block minimum depth image value as a minimum value from the block minimum depth image values from the block minimum depth image value and spatially neighbouring block minimum depth image values.

20. The apparatus as claimed in any of claims 12 to 19, further comprising a maximum value comparator configured to compare, for each block, a sequence of determined blocks maximum depth image values to determine that the determined block maximum depth image value is equal to a maximum value from the sequence of determined blocks maximum depth image values.

21. The apparatus as claimed in any of claims 12 to 20, further comprising a minimum value comparator configured to compare, for each block, a sequence of determined blocks minimum depth image values to determine that the determined block minimum depth image value is equal to a minimum value from the sequence of determined blocks maximum depth image values.

22. The apparatus as claimed in any of claims 12 to 21, wherein the encoder comprises:

a delta image generator configured to determine a difference value by subtracting a corresponding location minimum depth image value from a location depth image value; and

a scaler configured to scale the difference value by dividing the difference value by the difference between a corresponding location maximum image value and the corresponding location minimum image value.

23. An apparatus for decoding at least one depth image frame within a sequence of depth image frames, the apparatus comprising:

an input configured to receive an encoded depth image frame;

an input configured to receive a depth image frame array of minimum values and array of maximum values;

5 an upscaler configured to generate a maximum depth image from the depth image frame array of maximum values, and a minimum depth image from the depth image frame array of maximum values; and

a decoder configured to decode values of the encoded depth image frame based on values of the maximum depth image and values of the minimum depth image at corresponding positions.

10

24. The apparatus as claimed in claim 23, wherein the decoder comprises:

a rescaler configured to rescale a depth image value at a location to generate a rescaled depth image value by multiplying an encoded depth image frame at the location by the difference between a corresponding location maximum image value and a corresponding location minimum image value; and a

15

releveler configured to relevel the depth image value at the location to generate the decoded value by adding the corresponding location minimum image value to the rescaled depth image value.

20

25. The method as claimed in claims 1 to 11 or the apparatus as claimed in claims 12 to 24, wherein each depth image block comprises at least one of:

16x16 depth image values; and

32x32 depth image values.

25

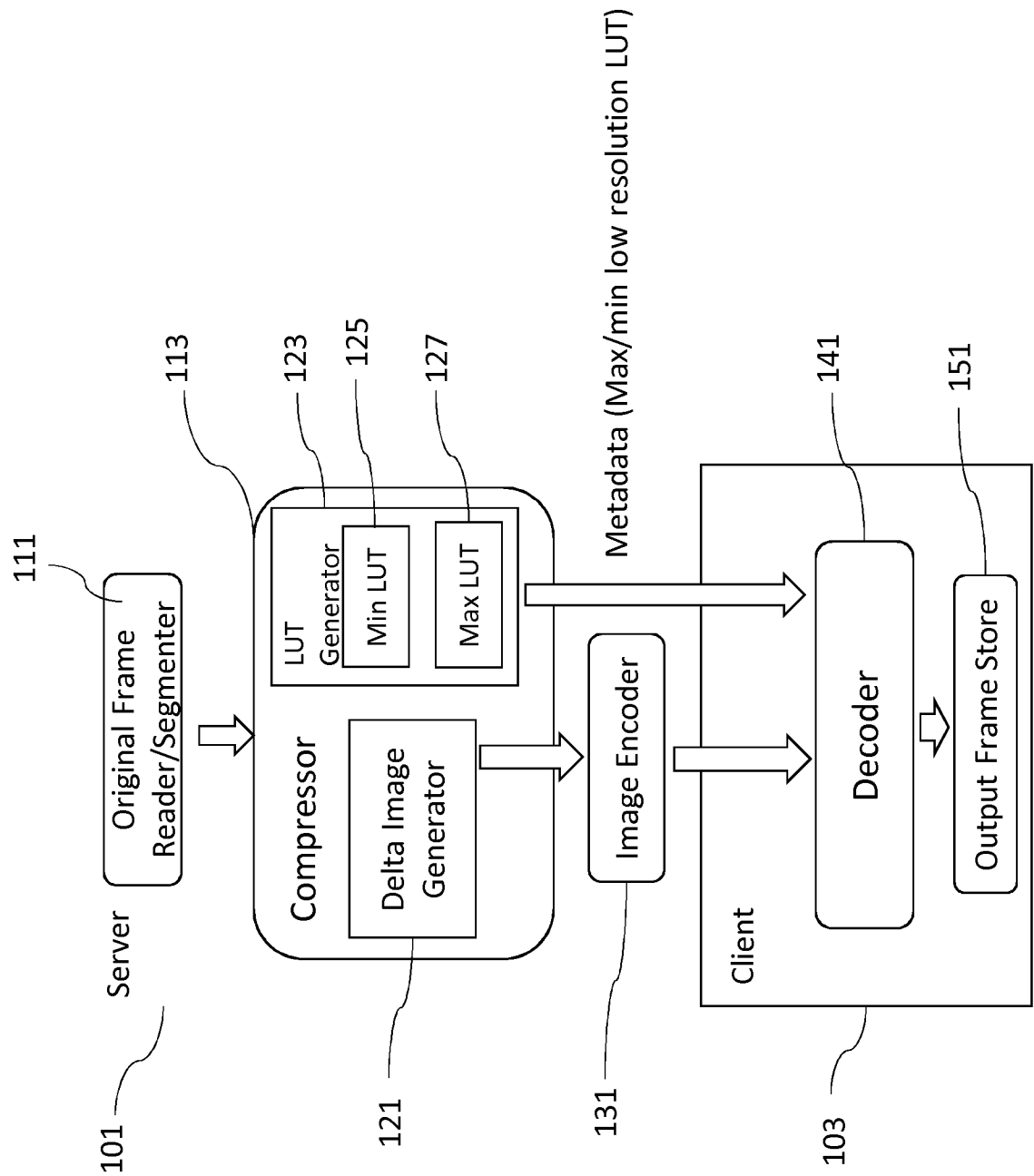


Figure 1

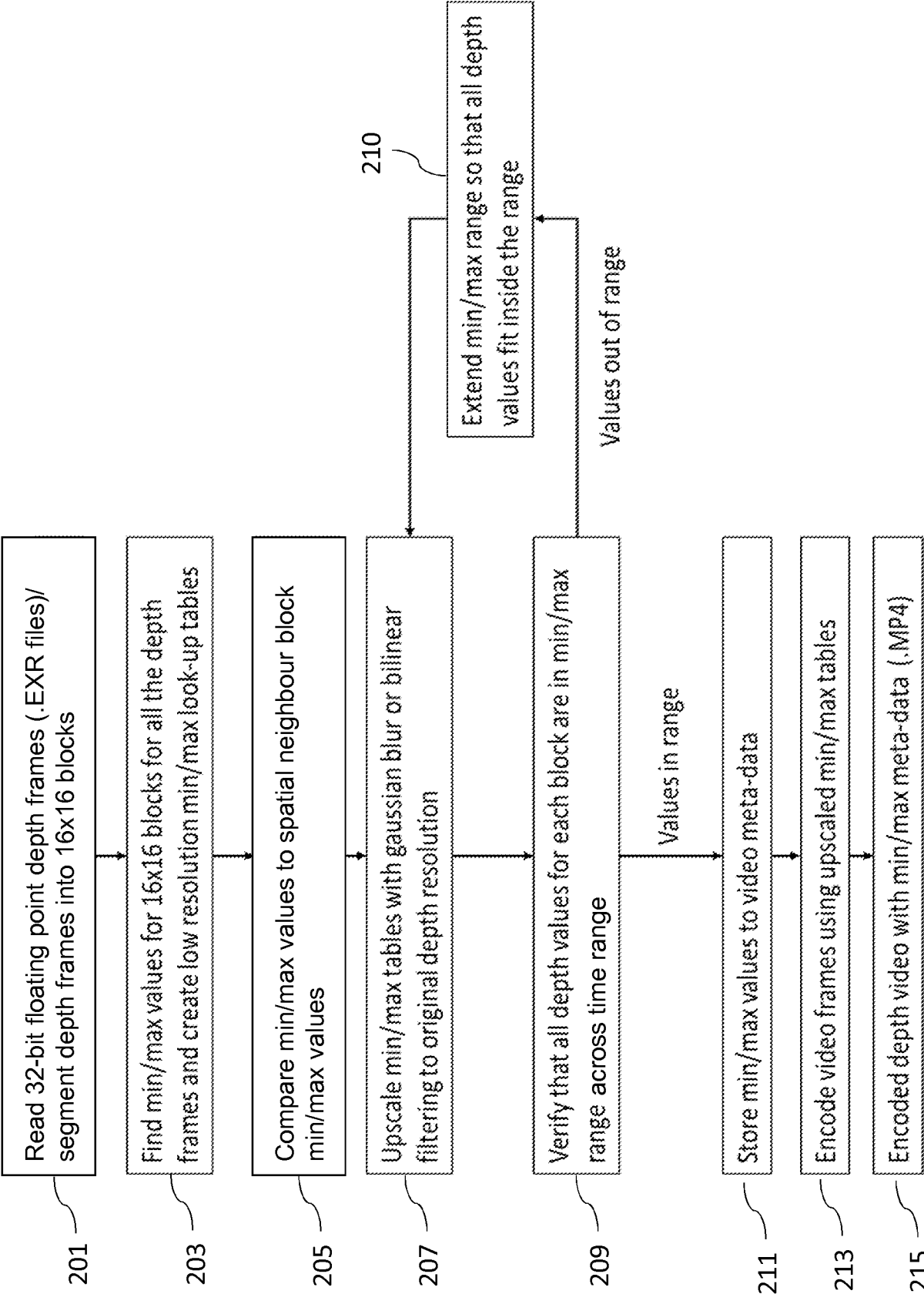
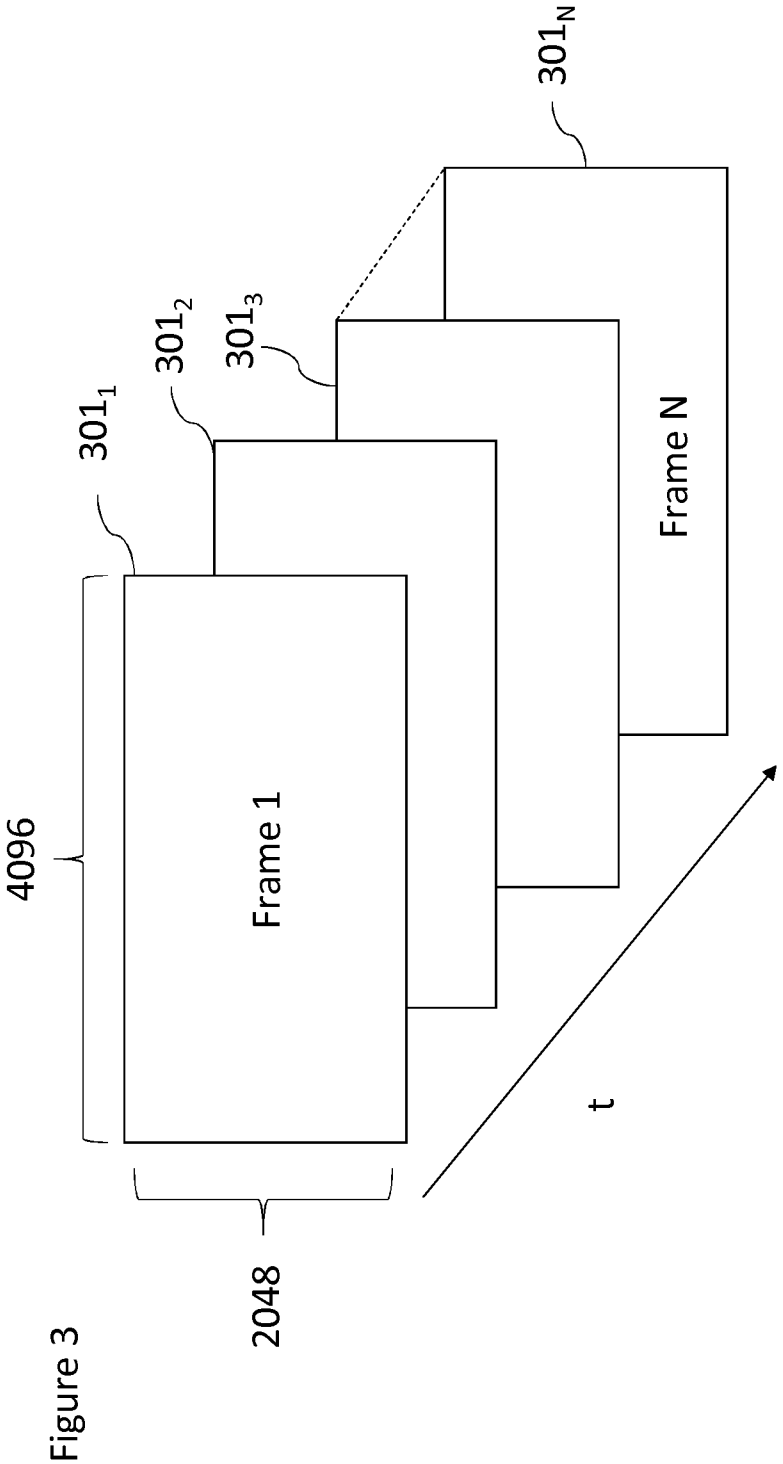


Figure 2



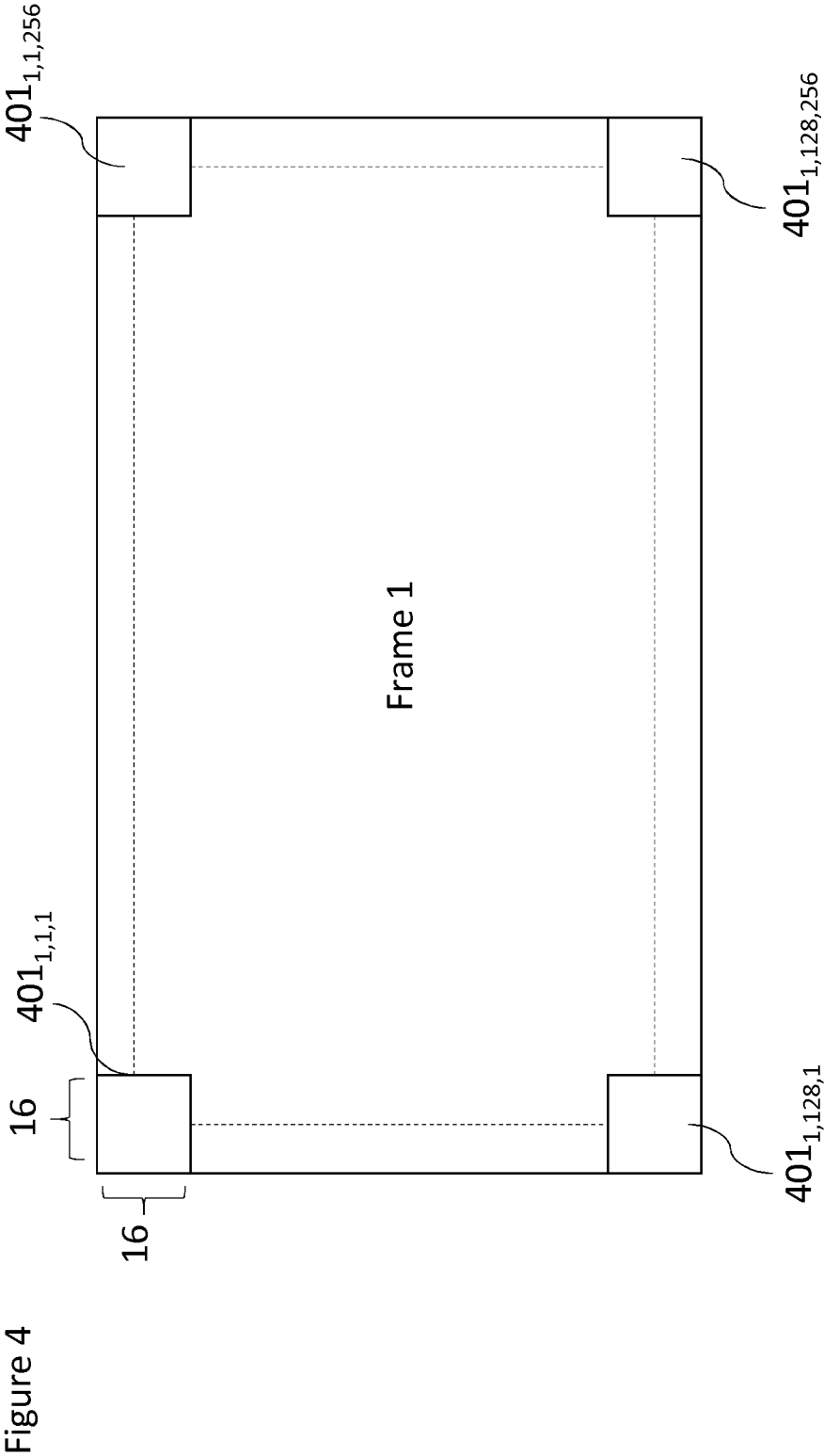


Figure 5

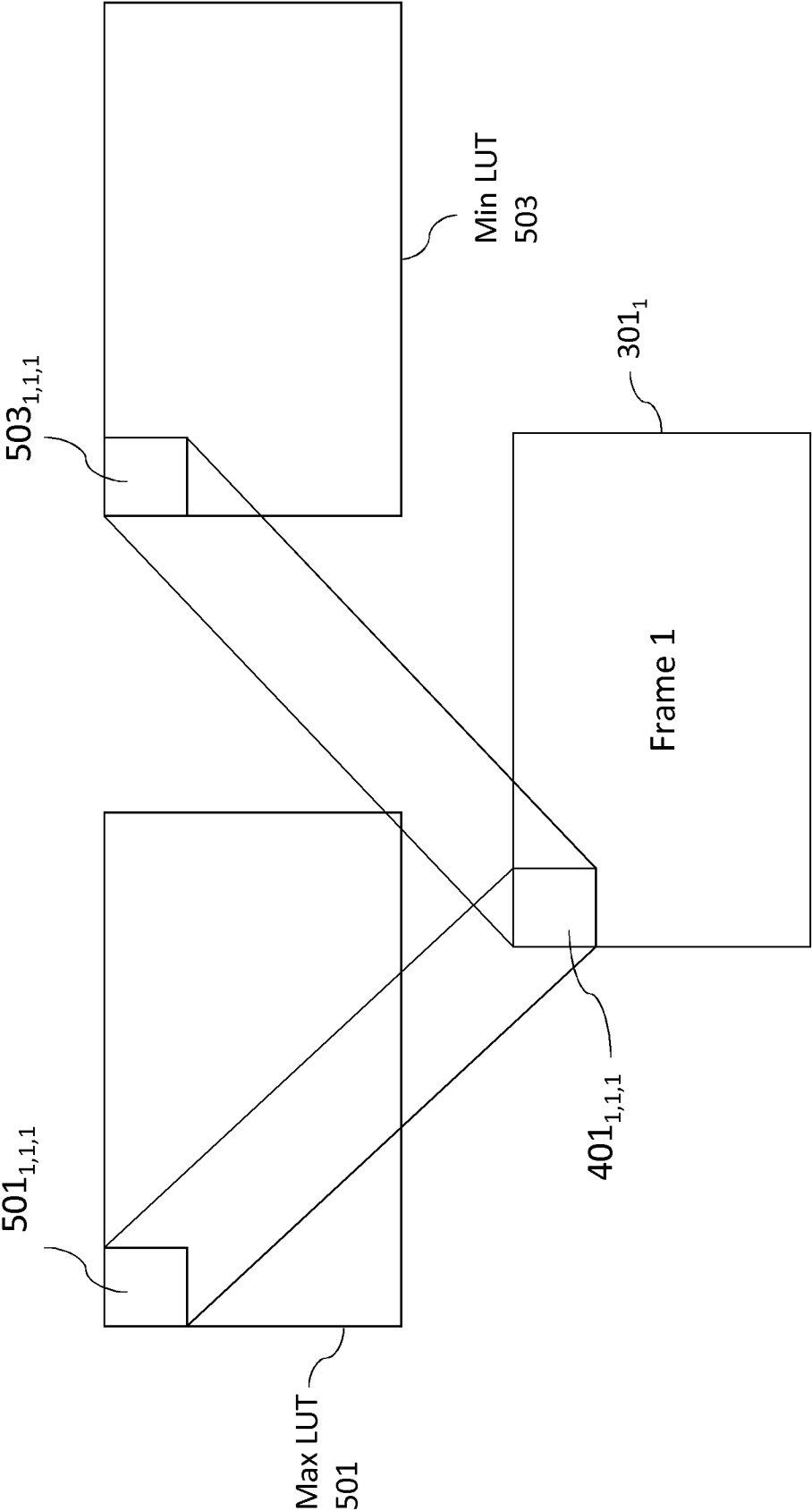
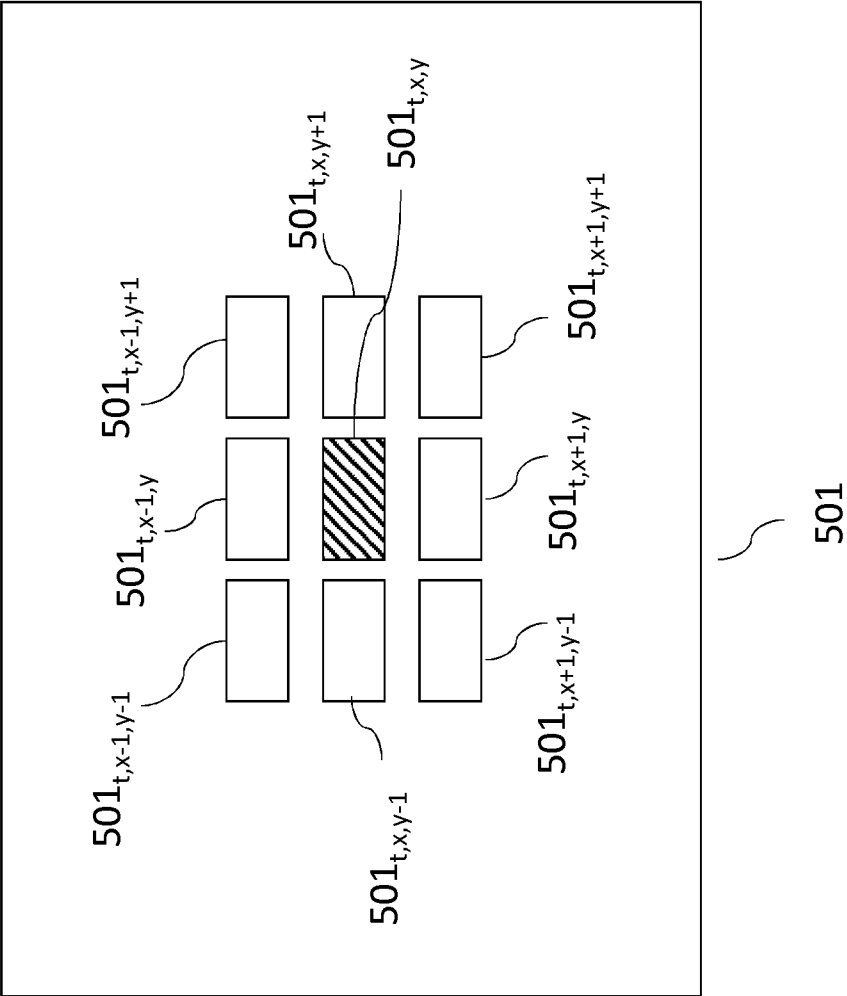


Figure 6



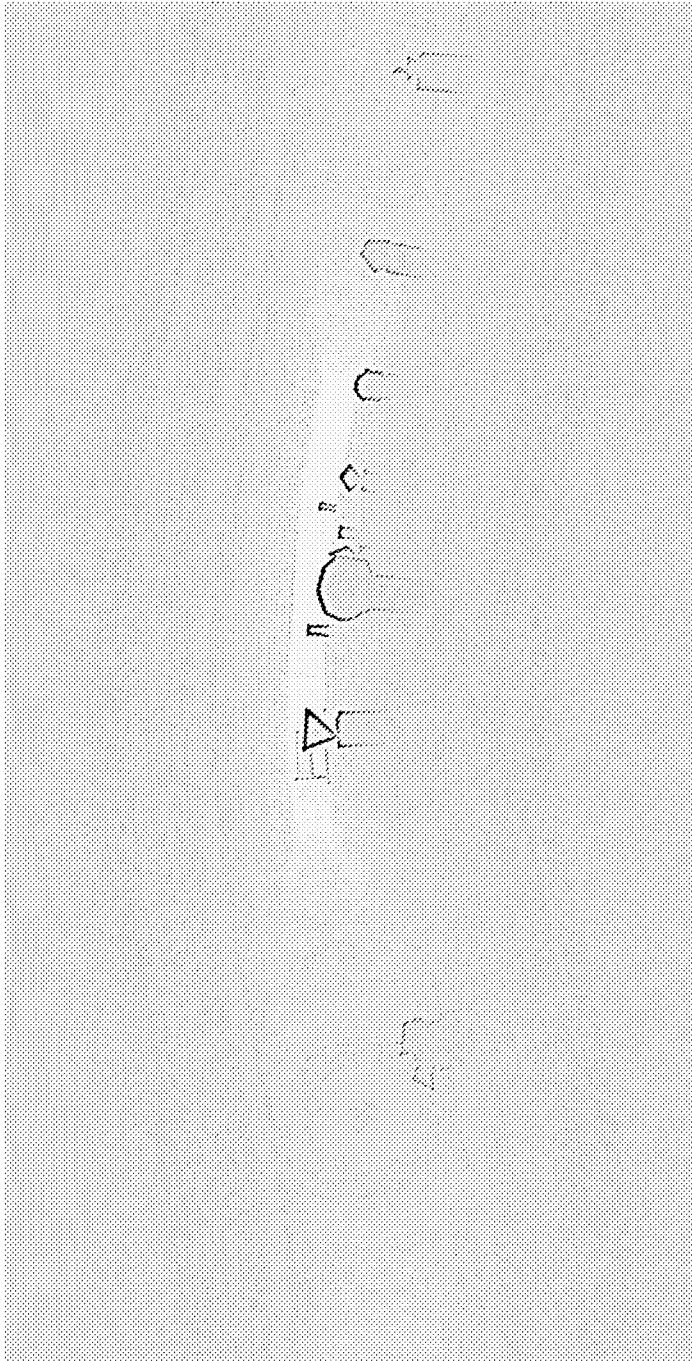


Figure 7

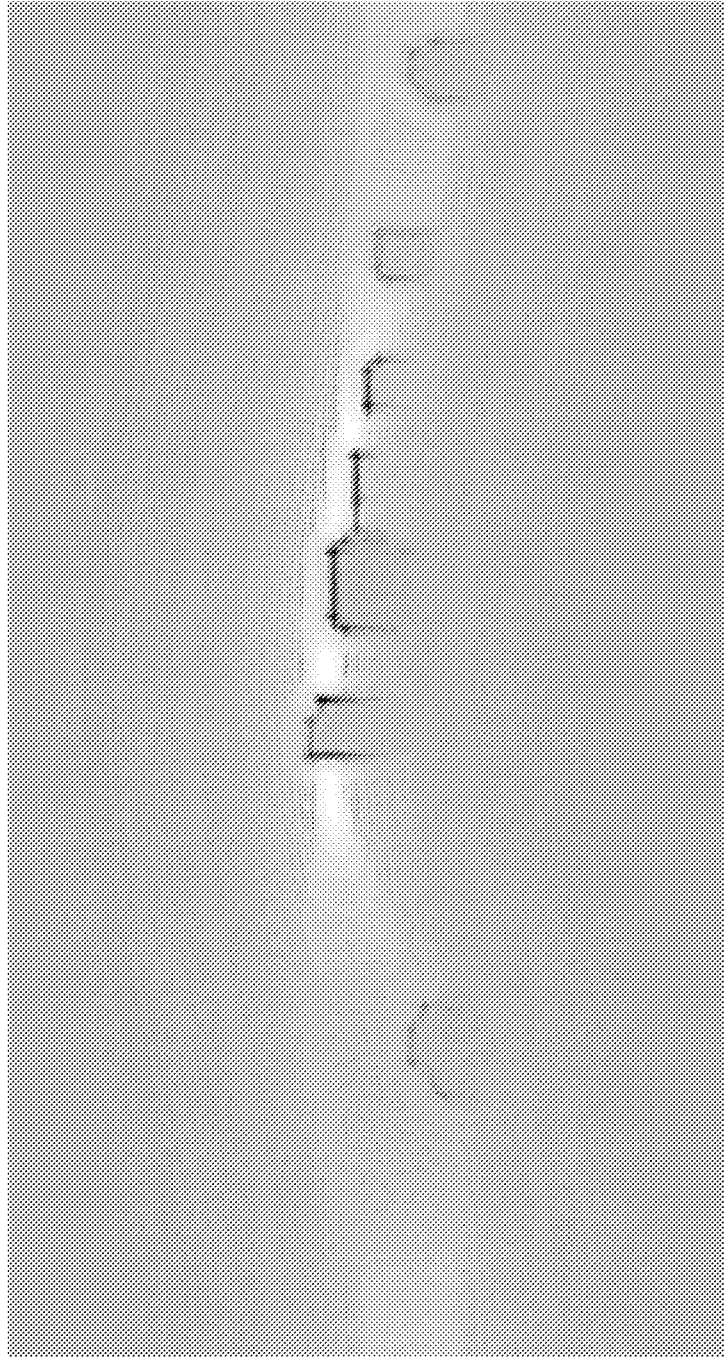


Figure 8



Figure 9

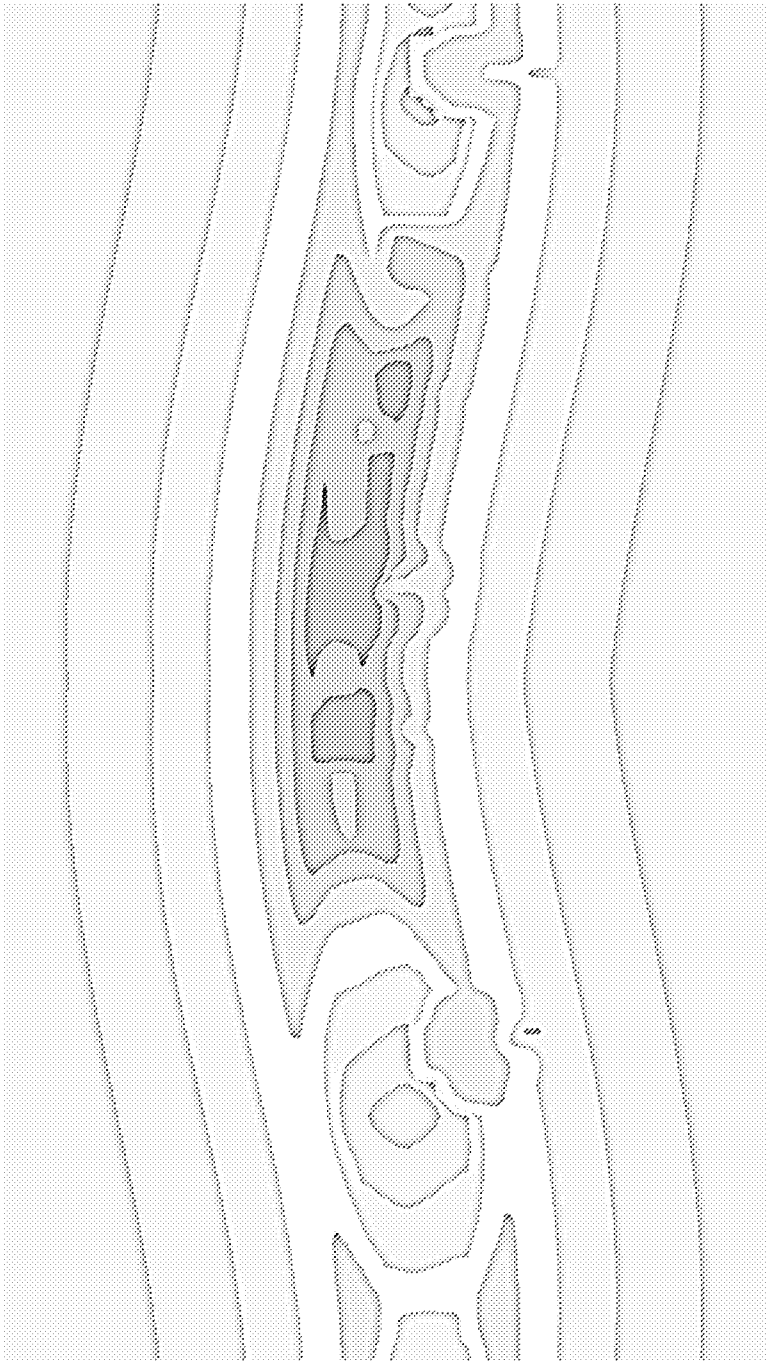


Figure 10

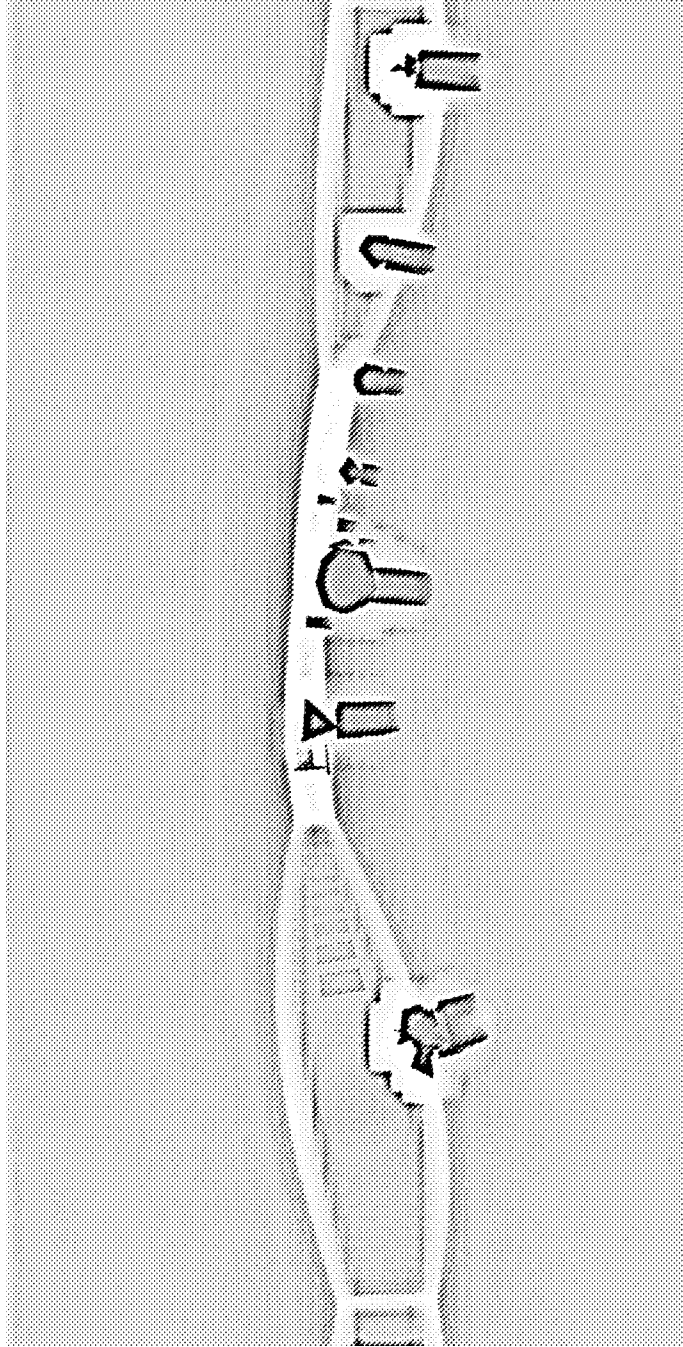


Figure 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2018/050008

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

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)

IPC: H04N, G06T, H03M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
FI, SE, NO, DK

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

EPODOC, EPO-Internal full-text databases, Full-text translation databases from Asian languages, WPIAP, COMPDX, INSPEC, TDB, NPL, Internet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KARADIMITRIOU, K et al. MIN-MAX COMPRESSION METHODS FOR MEDICAL IMAGE DATABASES. In: ACM SIGMOD Record. ACM Digital Library, New York, NY, USA: ACM [online], March 1997, Vol. 26, No. 1, Pages 47-52, [retrieved on 2018-04-16]. Retrieved from < https://dl.acm.org/citation.cfm?doid=248603.248613 >, <DOI:10.1145/248603.248613> sections I, III.A, III.B; figures 3-4	1-24
X	EP 0926898 A1 (SONY CORP [JP]) 30 June 1999 (30.06.1999) abstract; paragraphs [0040], [0053]-[0057], [0061]-[0063], [0076], [0093], [0110], [0119], [0152], [0188]; figures 1-4, 6, 9-10	1-5, 10, 12-18, 23, 25
X	US 6411295 B1 (HUNG DANIEL D [US] et al.) 25 June 2002 (25.06.2002) abstract; column 2, lines 35-63; column 6, lines 57-67	1, 10, 12, 23

☒ Further documents are listed in the continuation of Box C.
 ☒ See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search

23 April 2018 (23.04.2018)

Date of mailing of the international search report

07 May 2018 (07.05.2018)

 Name and mailing address of the ISA/FI
 Finnish Patent and Registration Office
 FI-00091 PRH, FINLAND

Facsimile No. +358 29 509 5328

Authorized officer

Janne Viljas

Telephone No. +358 29 509 5000

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2018/050008

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP H06165144 A (SONY CORP) 10 June 1994 (10.06.1994) figures 1, 4, 7 & abstract [online] EPOQUENET EPODOC & WPI & machine translation into English by EPO & Google [online] [retrieved 16.4.2018], paragraphs [0007]-[0008], [0011]-[0012], [0014], [0017]-[0018], [0020]	1, 10, 12, 23
X	JP 2012119901 A (UNIV CHUO) 21 June 2012 (21.06.2012) figures 3-5, 7 & abstract [online] EPOQUENET EPODOC & WPI & machine translation into English by ThomsonReuters [online] EPOQUENET TXPJPEA, paragraphs [0051]-[0059]	1, 10, 12, 23
X	US 2010295922 A1 (CHEUNG GENE [JP] et al.) 25 November 2010 (25.11.2010) abstract; paragraphs [0020]-[0022], [0033]-[0034], [0054]-[0055]; figure 2	1, 10, 12, 23
A	US 2016360178 A1 (TSUKAGOSHI IKUO [JP]) 08 December 2016 (08.12.2016) abstract; paragraphs [0038], [0167]-[0169], [0217]; figures 3-6	
A	US 2015110170 A1 (LIN JIAN-LIANG [TW] et al.) 23 April 2015 (23.04.2015) abstract; paragraphs [0011], [0035]-[0038], [0080]	
A	US 2014003711 A1 (NGAN KING NGI [HK] et al.) 02 January 2014 (02.01.2014) abstract; paragraphs [0021]-[0024], [0035]-[0036]	
A	US 2013202194 A1 (GRAZIOSI DANILLO BRACCO [US] et al.) 08 August 2013 (08.08.2013) abstract	
A	US 2015195573 A1 (AFLAKI BENI PAYMAN [FI] et al.) 09 July 2015 (09.07.2015) abstract; paragraphs [0290]-[0291], [0310]	

INTERNATIONAL SEARCH REPORT
Information on Patent Family Members

International application No.
PCT/FI2018/050008

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
EP 0926898 A1	30/06/1999	CN 1231100 A	06/10/1999
		CN 1178509 C	01/12/2004
		EP 1313326 A2	21/05/2003
		JP H1132330 A	02/02/1999
		JP 4240554 B2	18/03/2009
		KR 20000068542 A	25/11/2000
		KR 100574184 B1	27/04/2006
		US 6480630 B1	12/11/2002
		US 2002181793 A1	05/12/2002
		US 6636640 B2	21/10/2003
		US 2004086190 A1	06/05/2004
		US 6990244 B2	24/01/2006
		US 2004086189 A1	06/05/2004
		US 7072491 B2	04/07/2006
		WO 9903283 A1	21/01/1999
.....			
US 6411295 B1	25/06/2002	None	
.....			
JP H06165144 A	10/06/1994	JP 3532222 B2	31/05/2004
.....			
JP 2012119901 A	21/06/2012	None	
.....			
US 2010295922 A1	25/11/2010	CN 101978697 A	16/02/2011
		CN 101978697 B	13/02/2013
		EP 2238764 A1	13/10/2010
		WO 2009094036 A1	30/07/2009
.....			
US 2016360178 A1	08/12/2016	US 9924151 B2	20/03/2018
		CN 103503446 A	08/01/2014
		CN 103503446 B	26/04/2017
		EP 2688304 A1	22/01/2014
		WO 2013129158	30/07/2015
		JP 6003901 B2	05/10/2016
		JP 2017022741 A	26/01/2017
		KR 20140138527 A	04/12/2014
		US 2014071236 A1	13/03/2014
		US 9451234 B2	20/09/2016
		WO 2013129158 A1	06/09/2013
.....			

INTERNATIONAL SEARCH REPORT
Information on Patent Family Members

International application No.
PCT/FI2018/050008

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
US 2015110170 A1	23/04/2015	US 9503751 B2 CN 104618724 A CN 104618724 B JP 2015080207 A US 2017041623 A1	22/11/2016 13/05/2015 10/11/2017 23/04/2015 09/02/2017
US 2014003711 A1	02/01/2014	None	
US 2013202194 A1	08/08/2013	None	
US 2015195573 A1	09/07/2015	WO 2015104456 A1	16/07/2015

CLASSIFICATION OF SUBJECT MATTER

IPC

H04N 13/128 (2018.01)**H04N 13/161** (2018.01)**H04N 13/178** (2018.01)**H04N 19/159** (2014.01)**H04N 19/176** (2014.01)**H04N 19/597** (2014.01)**H04N 19/98** (2014.01)**H04N 13/268** (2018.01)**H04N 13/271** (2018.01)**H04N 21/235** (2011.01)**H04N 21/236** (2011.01)**H04N 21/43** (2011.01)**H04N 21/845** (2011.01)**G06T 7/593** (2017.01)