# APPARATUSES, METHODS, AND COMPUTER-READABLE MEDIUM FOR ENCODING AND DECODING A VIDEO SIGNAL

## TECHNICAL FIELD

Generally, the present invention relates to the field of video coding. More specifically, the present invention relates to apparatuses and methods for encoding and decoding a video coding block of a video signal.

## BACKGROUND

A 360° panoramic video signal comprises a sequence that captures the scene in all directions around a single view point. This is usually achieved using multiple cameras arranged near each other and with different view directions. The captured content is then stitched together and can be thought of as a sphere around the view point with texture on it, which represents the 360° scene or panoramic video signal at a certain point in time. Such a sphere, however, cannot be processed directly using current video codecs, which generally require a flat (2D) rectangular image of a scene. Thus, in order to process a panoramic video signal using current video codecs the sphere has to be converted to such a 2D format.

For converting the sphere into a 2D format comprising one or more segments one of the following projections can be used.

An equirectangular projection defines a projection to a rectangle similar of the one that creates world maps from the globe of the earth. The sample density differs with different "heights". Extreme case are the poles of the sphere, they are mapped to lines. Geometrical distortions are spread all over the resulting image.

In a projection based on a cube the sphere is mapped to the faces or segments of a cube. Each segment looks like a usual 2D image with no visible geometrical distortions. However, there are strong geometrical distortions in the boundary regions of the segments. There are multiple ways to arrange the segments in a rectangular format.

In a projection based on an icosahedron the sphere is mapped to the faces or segments of an icosahedron. Each segment looks like a usual 2D image with no visible geometrical distortions. Also in this case there are strong geometrical distortions in the boundary regions of the segments of the icosahedron. These distortions, however, are not as strong

In a fourth possible implementation form of the decoding apparatus according to the second aspect as such or any one of its implementation forms, the decoding apparatus further comprises a reference frame buffer configured to store the first segment of the reference frame together with the extension region of the first segment of the reference frame.

In a fifth possible implementation form of the decoding apparatus according to the second aspect as such or any one of its implementation forms, the first segment and the second segment correspond to different faces of a cube or an icosahedron.

In a sixth possible implementation form of the decoding apparatus, the segment extension unit is configured to perform segment extension based on a flag/information at a block level for each video coding block, at a frame level, at a GOP (group of pictures) level, at a PPS (picture parameter set) level or at an SPS (sequence parameter set) level. Based on the status of the flag, segment extension operation can be enabled or disabled.

According to a third aspect the invention relates to a method for encoding a video signal, the video signal comprising video content which includes video content of a first image plane and video content of a second image plane, the video content being provided in the form of a plurality of successive frames, wherein each frame of the plurality of successive frames comprises a first segment which comprises image content of the first image plane and a second segment which comprises image content of the second image plane, the first segment and the second segment each comprising a plurality of pixels, each pixel being associated with a pixel value, wherein the method comprises:

generating an extension region of the first segment of a reference frame, the extension region comprising a plurality of extension pixels, wherein generating the extension region comprises, for each extension pixel, determining the pixel value of the extension pixel on the basis of a pixel value of a pixel of the second segment of the reference frame;

selecting a reference block among a set of candidate reference blocks, the set of candidate reference blocks comprising a candidate reference block located entirely or partly in the extension region of the first segment of the reference frame; and

encoding the pixel values of the pixels of a video coding block of the first segment of a current frame on the basis of the pixel values of the pixels of the video coding block and the pixel values of the pixels of the selected reference block.

Again, the first image plane and the second image plane may be angled relative to each other by an inter-plane angle (e.g., between 1° and 90°).

According to a fourth aspect the invention relates to a method for decoding an encoded video signal, the encoded video signal comprising video content which includes video

$$\mathbf{H_{B2A}} = \begin{pmatrix} \mathbf{K_A} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{R}(\theta) & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{K_B}^{-1} & \\ 0 \quad 0 & Z^{-1} \end{pmatrix} = \mathbf{K_A R}(\theta) \mathbf{K_B}^{-1}$$

wherein $\mathbf{K_A}$ and $\mathbf{K_B}$ denote camera calibration matrices for the first segment and the second segment of the reference frame, $\theta$ denotes the angle between the image plane of the first segment and the image plane of the second segment of the reference frame, $\mathbf{R}(\theta)$ denotes a rotation and $Z$ denotes a distance from a camera position $C$.

11. The decoding apparatus (200) of any one of claims 7 to 10, wherein the decoding apparatus (200) further comprises a reference frame buffer (209) configured to store the first segment of the reference frame together with the extension region of the first segment of the reference frame.

12. The decoding apparatus (200) of any one of claims 7 to 11, wherein the first segment and the second segment correspond to different faces of a cube or an icosahedron.

13. The decoding apparatus of any one of claims 7 to 12, wherein the segment extension unit (211) is configured to perform segment extension based on a flag/information at a block level for each video coding block, at a frame level, at a GOP (group of pictures) level, at a PPS (picture parameter set) level or at an SPS (sequence parameter set) level. Based on the status of the flag, segment extension operation can be enabled or disabled.

14. A method (800) for encoding a video signal, the video signal comprising video content which includes video content of a first image plane and video content of a second image plane, the video content being provided in the form of a plurality of successive frames, wherein each frame of the plurality of successive frames comprises a first segment which comprises image content of the first image plane and a second segment which comprises image content of the second image plane, the first segment and the second segment each comprising a plurality of pixels, each pixel being associated with a pixel value, wherein the method comprises:

generating an extension region of the first segment of a reference frame, the extension region comprising a plurality of extension pixels, wherein generating the extension region

comprises, for each extension pixel, determining the pixel value of the extension pixel on the basis of a pixel value of a pixel of the second segment of the reference frame;

selecting a reference block among a set of candidate reference blocks, the set of candidate reference blocks comprising a candidate reference block located entirely or partly in the extension region of the first segment of the reference frame; and

encoding the pixel values of the pixels of a video coding block of the first segment of a current frame on the basis of the pixel values of the pixels of the video coding block and the pixel values of the pixels of the selected reference block.

15. A method (900) for decoding an encoded video signal, the encoded video signal comprising video content which includes video content of a first image plane and video content of a second image plane, the video content being provided in an encoded form of a plurality of successive frames, wherein each frame of the plurality of successive frames comprises a first segment which comprises image content of the first image plane and a second segment which comprises image content of the second plane, the first segment and the second segment each comprising a plurality of pixels, each pixel being associated with a pixel value, wherein the method comprises:

generating an extension region of the first segment of a reference frame, the extension region comprising a plurality of extension pixels, wherein generating the extension region comprises, for each extension pixel, determining the pixel value of the extension pixel on the basis of a pixel value of a pixel of the second segment of the reference frame; and

reconstructing the pixel values of the pixels of a video coding block of the first segment of the current frame on the basis of an encoded video coding block of the first segment of a current frame and the pixel values of the pixels of a reference block, wherein the reference block is located entirely or partly in the first segment of the reference frame or entirely or partly in the extension region of the first segment of the reference frame.

16. A computer-readable medium comprising program code for causing a computer to perform the method (800) of claim 14 or the method (900) of claim 15 when executed on the computer.