INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2017/M40806
July 2017, Torino, Italy

Source **Poznań University of Technology**
Status **Input**
Title **Omnidirectional 6-DoF/3-DoF+ rendering**
Author **Krzysztof Wegner, Olgierd Stankiewicz, Adrian Dziembowski, Dawid Mieloch, Marek Domański**

1. Introduction

Omnidirectional video formats natively provide capability of changing direction of viewing but not the position of viewing. Such capability, of changing also the position of viewing, is key for the prospective Omnidirectional 6DoF/3DoF+ video technology.
In one of the previous documents [1] we have shown that it is possible to estimate depth from such content with the use of existing MPEG technology. Therefore, it is possible to attain omnidirectional video plus omnidirectional depth 3D representation. In the case of "flat" multiview video plus depth (MVD), depth allows rendering of virtual viewpoints and thus changing direction and position of viewing. However, during recent 118th MPEG meeting in Hobart it has been questioned whether the state-of-the-art rendering technology is capable of using omnidirectional video content with omnidirectional depth to render high quality Omnidirectional 6DoF/3DoF+ viewpoints with shifted position. In response to these concerns, in this document we show the technical background behind Omnidirectional 6DoF/3DoF+ rendering along with results attained with our proprietary rendering software.

2. Equi-rectangular projection format (ERP)

The Equi-rectangular projection format (ERP) projection format is the most widely used projection format for representing 360-degree video as a rectangular video. It is the default projection format described by JVET [2].

In order to obtain direction (φ, θ) in 3D space based on pixel 2D coordinate (m, n) in rectangular video of resolution $W \times H$ following equation are used (1):

$$\varphi = \left(\frac{m}{W} - 0.5\right) \cdot 2\pi$$
$$\theta = \left(0.5 - \frac{n}{H}\right) \cdot \pi \qquad (1)$$

Similarly conversion of direction in 3D space to pixel 2D position (m, n) is expressed as (2):

$$m = \left(\frac{\varphi}{2\pi} + 0.5\right) \cdot W$$
$$n = \left(0.5 - \frac{\theta}{\pi}\right) \cdot H \qquad (2)$$

3. Omnidirectional 6DoF/3DoF + rendering

The goal of the 360 degree rendering is to render omnidirectional virtual viewpoint based on single omnidirectional video accompanied with omnidirectional depth (OVD).

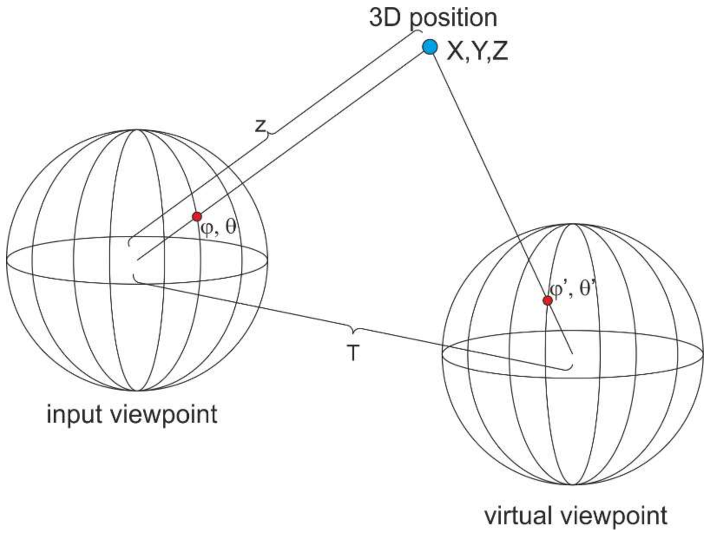

Figure 1. Omnidirectional virtual viewpoint rendering.

For the sake of brevity, let's assume for a moment, that both omnidirectional virtual viewpoint and input video are stored as rectangular video in equi-rectangular projection format. Of course, the presented considerations can be used for any projection format.

For every pixel (m, n) in the rectangular video we have obtain appropriate direction (φ, θ) in 3D space through (e1).

From the depth map associated with video, we get depth value for the processed point (m, n). Given the depth map format used [4] we can easily calculate distance of given point from the center of the omnidirectional viewpoint.

For direction (φ, θ) and the distance z, exact 3D position the point can be calculated (3):

$$X = z \cdot cos(\theta) \cdot cos(\varphi)$$
$$Y = z \cdot sin(\theta) \qquad (3)$$
$$Z = -z \cdot cos(\theta) \cdot sin(\varphi)$$

In order to allow changing position of virtual viewpoint (and thus allow 6DoF/3DoF+) the entire scene can be shifted in any direction by $T = [T_x \quad T_y \quad T_z]$ as expressed in following equation (4):

$$X' = X + T_x$$
$$Y' = Y + T_y \quad (4)$$
$$Z' = Z + T_z$$

For the obtained position (X', Y', Z') in virtual viewpoint coordinate system, we can calculate direction (φ', θ') relative to the virtual viewpoint center (5):

$$\varphi' = tan^{-1}\left(-\frac{Z}{X}\right)$$
$$\theta' = sin^{-1}\left(\frac{Y}{\sqrt{X^2 + Y^2 + Z^2}}\right) \quad (5)$$

In the end, appropriate pixel position (m', n) in the output rectangular video (representing the virtual viewpoint) can be obtained from direction (φ, θ) in 3D space through equation (e1).

Finally, using projection equation appropriate pixel coordinate in rectangle video can be obtained. In our case we use (e2).

4. Results

In order to assess the proposed approach, we have used omnidirectional 360-degree "Dancer360" video sequence [3] (see Fig. 2) with omnidirectional depth estimated with method proposed in [1]. First, we have used view synthesis technique implemented in MPEG View Synthesis Reference Software (VSRS), but we have identified problems with inpainting. Those problems are presented in Figure 3 (top) and Figure 4 (top), for small and big change of position of viewing, respectively. In VSRS inpainting is performed with morphological filtering of generated virtual depth image which is not enough to fill large holes. Therefore, for further tests, we have used our proprietary view rendering software from Poznan University of Technology, which uses more sophisticated inpainting schemes. As it can be seen in Figure 3 (bottom) and Figure 4 (bottom), it provides much higher quality of synthesized video.

Figure 2. Original 360-degree omnidirectional frame from "Dancer360" sequence [3].

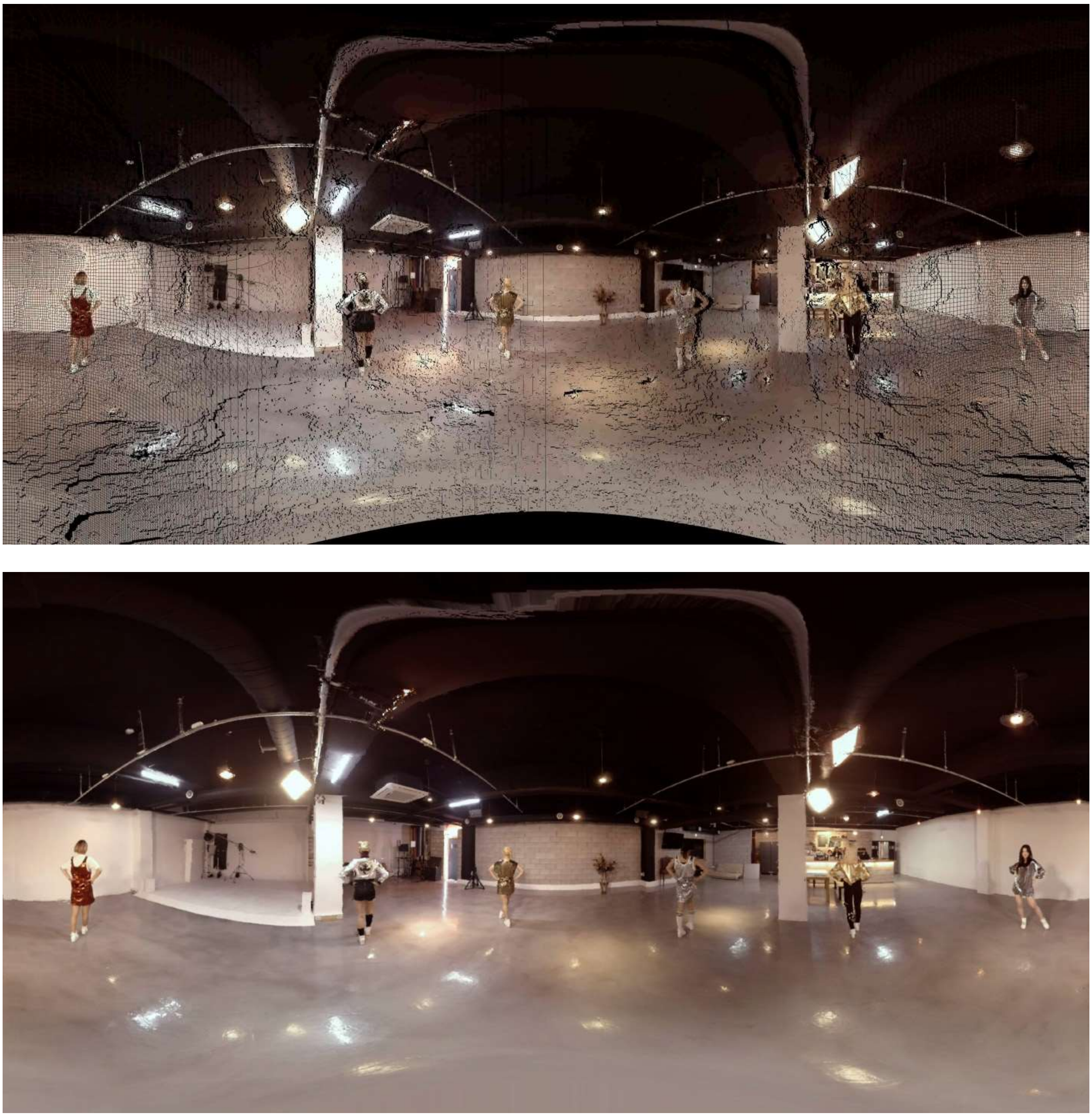

Figure 3. Example of virtual viewpoint rendering for 360-degree “Dancer360” sequence [3] - small change of position of viewing: VSRS (top), and with Poznan University of Technology proprietary rendering software (bottom).

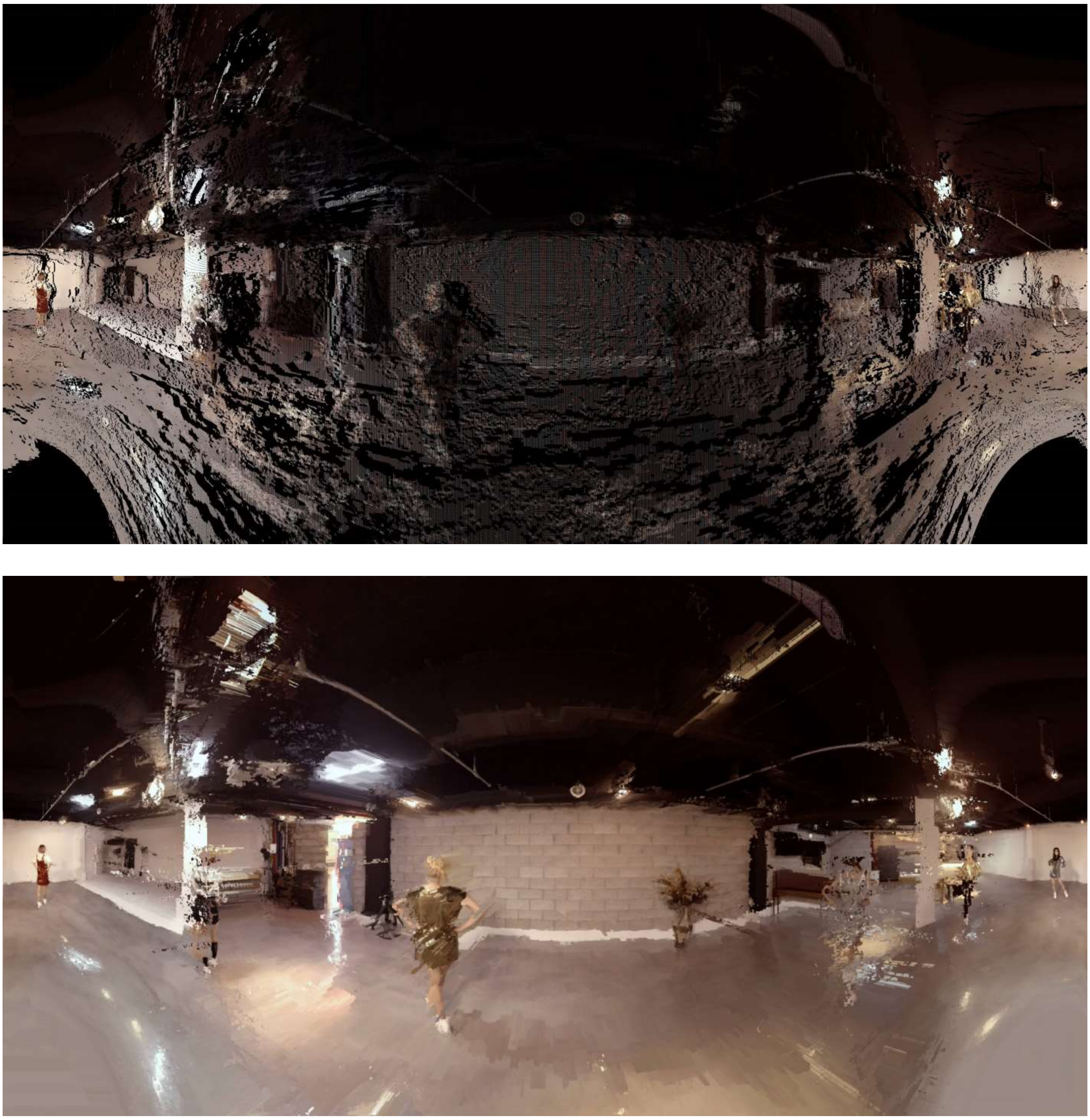

Figure 4. Example of virtual viewpoint rendering for 360-degree “Dancer360” sequence [3] - big change of position of viewing: VSRS (top), and with Poznan University of Technology proprietary rendering software (bottom).

The results will be presented during the meeting on a Head-Mounted Display (HMD) device.

5. Conclusions

We have shown that it is possible to render high quality omnidirectional video with varying position of viewing from omnidirectional video + depth content.

6. References

[1] Krzysztof Wegner, Olgierd Stankiewicz, Tomasz Grajek, Marek Domański, "Depth estimation from circular projection of 360 degree 3D video", ISO/IEC JTC1/SC29/WG11, MPEG2017/m40596, April 2017, Hobart, Australia

[2] Y. Ye, E. Alshina, J. Boyce, Algorithm descriptions of projection format conversion and video quality metrics in 360Lib, Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 6th Meeting: doc. JVET F-1003, Hobart, AU, 31 March – 7 April 2017.

[3] G. Bang, G. S. Lee, N. Ho H., "Test materials for 360 3D video application discussion", ISO/IEC JTC1/SC29/WG11 MPEG2016/M37810 February 2016, San Diego, USA

[4] T. Senoh, K. Yamamoto, R. Oi, T. Mishina, M. Okui: Consideration of depth format. ISO/IEC JTC1/SC29/WG11 MPEG m15047, Antalya, Turkey, Jan. 2007.