OPEN ACCESS Freely available online

ISSN 2737-5315
Volume 2, Issue 2
https://www.iikii.com.sg/journal/EIET
Educational Innovations and Emerging Technologies

Research papers

Combining 360° Video and Camera Mapping for Virtual Reality: An Innovative Solution

Hung Sun [1], Tsun-Hung Tsai [2,*] and Ke Jiang [3]

[1] Department of Digital Multimedia Design, China University of Technology, Taipei, Taiwan; arion3D@gm.cute.edu.tw
[2] Shih Chien University, Taipei, Taiwan;
[3] Sophon Ltd, Taichung, Taiwan.; jackykjiang@gmail.com
* Correspondence: otaru@g2.usc.edu.tw

Received: Apr 11, 2022; **Accepted:** May 11, 2022; **Published:** Jun 30, 2022

Abstract: Changing the framework of traditional video with limited viewing angles, 360° photo/video provides an immersive viewing experience. 360° video is one of the applications and the most important feature of virtual reality for immersion and the feeling of being in another space. However, when viewing 360° videos with a head-mounted display, the viewer feels like a fixed rotatable camera, and the viewer's movement does not change the viewing angle of the object, which greatly reduces the spatial immersion required for virtual reality. Therefore, we propose a solution that maintains high-quality graphics and low hardware demands and supports 6DoF head-mounted displays. Through the camera mapping function in the 3D animation software, 360° surround video is projected into a 3D sphere to create a simple 3D object that corresponds to the shape of the object in the image. With pristine video quality and a realistic 3D spatial perspective, it provides better virtual reality immersion than 360° video and is not as complex as a full 3D virtual reality environment. In the future, 3D scanning and photogrammetry can be integrated to reconstruct a more easily applied 3D virtual reality environment.

Keywords: 360° photos/videos, camera mapping, virtual reality, 3D modeling

1. Introduction

Virtual reality (VR) has been used for years as a medium for teaching and exhibiting (Freina and Ott, 2015). Compared to 3D VR, 360° photos or videos are easy to produce and convenient to watch. A 360° surround photo or video provides the immersive spherical images. Compared to traditional flat images with fixed and limited viewable angles, a global-shaped 360° x 180° video is a means to view 360° images with more information about the surrounding environment and a more immersive visual content experience. We refer to this medium as "360° video" for short (Kopf, 2016). In recent years, with the development of the Internet and social media such as Facebook and YouTube, 360° content support is available. The use of 360° photos or videos to present an immersive view is also becoming more common. These 360° images can be viewed via Web-Based or mobile devices with a mouse or finger movement and are commonly used in applications such as home sales, attraction tours, and map navigation. 360° videos are particularly suitable for sports broadcasts or heritage tours (Hebbel-Seeger and Horky, 2019: Škola et al.. 2020), where viewers can combine head-mounted displays with mobile devices or use VR headsets to view videos from different angles.

One of the most important features of virtual reality is immersion, the feeling of being in another space. In 360° video using a head-mounted display, the viewer feels like a fixed 360-degree rotating camera. The viewer's movement in space does not change the viewing angle of the object, which significantly reduces the spatial immersion required for virtual reality. Although 360° video with 3-DoF+ support is available, high bandwidth is required to transmit high-quality video with multiple viewing angles. Fully realistic VR environments in 3D require sophisticated 3D production techniques and the ability to integrate game engines. This is a technical threshold for practical applications.

Based on a 360° video shot in a room, the exhibited content requires 6-DoF 3D virtual scenes. Thus, we propose an innovative solution for 6-DoF 3D virtual scenes by combining the concept of Cube Map as Sky Box in the game engine and using the camera mapping function of 3D animation software. Then, the 360° video is projected on top of a 3D sphere to create simple objects with corresponding shapes based on the data on the image. Without the use of difficult 3D production techniques, the 3-Dof 360° video can be converted into a 3D virtual space supporting 6-DoF, while maintaining a 3D VR environment with high-quality but with low computing requirements.

2. Research Background

The study is based on the following: 360° video shot in the room as the original material, camera mapping function in 3D animation software to build 3D models to convert 2D images into 3D scenes with limited operation, and a full 3D virtual reality space. 3D room models are combined and imported into the Unity game engine to create a VR for exhibition.

2.1. 360° Video

360° video/graphics, also known as panoramic, spherical, or omnidirectional video/graphics, is a new type of multimedia that provides an immersive experience (Kopf, 2016). A normal 360 camera is used to capture a wide-angle surround image through two 180° wide-angle lenses in front and back, and then, the image is processed by the built-in software to obtain a 360° surround image with the camera at the center of the sphere. The ultra-high resolution 360° video uses multiple cameras to capture images that are stitched together through the post-production software. 360° video uses Omnidirectional Media Format (OMAF) which only supports a simple VR with three degrees of freedom (3-DoF), and the viewer's movement is not reflected in the screen change. OMAF V.2, on the other hand, proposes a 3-DoF+ mode that allows lateral head movements (Hannuksela et al., 2019). To address the limitations of 3-DoF, previous studies have proposed solutions for the 6-DoF by transmitting multi-angle video (Jeong et al., 2020). However, providing 3DoF+ and 6DoF VR data requires compressing and streaming multiple videos to support users' body movements, which is challenging for High-Efficiency Video Coding (HEVC). Therefore, many studies have proposed various compression or prediction solutions (Xu et al., 2020). At present, the transmission rate of 5G already solves the problem of transmitting huge amounts of data, so multi-view 360° Video is a direction of VR development application. As the 360 camera can record the environment entirely, the story of the space at the moment is recorded and sent back to the viewer's eyes through images. The complete experience also needs to match the sound effect of the image at the moment, so the sound of the environment cannot be ignored. Ambisonic is a spherical omnidirectional surround sound technology (Snelson and Hsu, 2020) to provide the best 360 VR experience, and the 360 recording effect is used for the technology.

2.2. Camera Mapping

Camera mapping has been used for many scenes in movies. In addition to video effect synthesis, the background of 2D animation requires Camera mapping to produce a depth-of-field lens effect. Camera mapping is different from Mapping Camera, the latter is photogrammetry, which is commonly used for terrain and feature mapping of the unattended aerial vehicle (UAV) (Dhanda, 2019). Projection Mapping, on the other hand, is a form of artistic expression in which pre-designed motion images are projected onto the surface of real objects such as buildings with precise alignment (Song and Park, 2014).

Figure 1 illustrates the standard Camera mapping production concept. The camera on the left sees the image on the flat panel in front of it, and the actual 3D model is presented on the screen. The method is to first set the environment background as a 2D image, correct the camera's perspective and position according to the background image to match the original shooting point, and then build a 3D model to match the contours of the objects in the image. The 3D model created from the camera projection image only looks correct in the projection camera but is a distorted 3D model in other angles. Finally, the original image is set as a texture to be projected on the surface of the 3D model, and the result is shown in Fig. 2. The other camera can be moved within a limited range to run the mirror. Before rendering the output, the preview image of the viewfinder camera looks strange as shown in Fig. 3. Previous Camera mapping applications allow a 2D image to simulate a 3D space with the limited 3D Camera shot, which is further extended in this study to become a full 3D environment reconstruction.

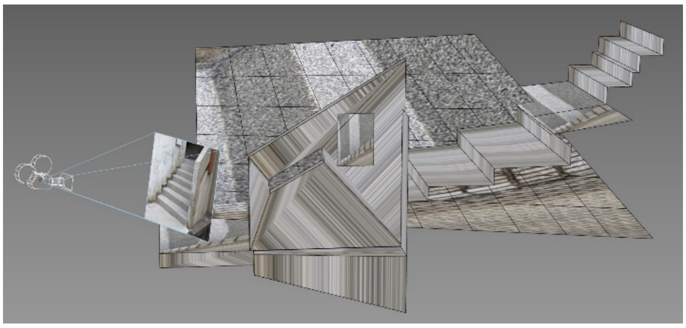

Fig. 1. Camera mapping screen and 3D model comparison.

Fig. 2. Camera 01 is used for projecting alignment images

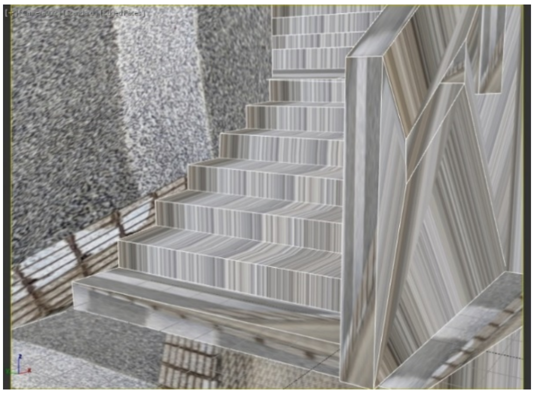

Fig. 3. Camera 02 is used to take pictures with the camera

2.3. Virtual Reality

Virtual reality (VR) has been developed for many years for the technical growth of hardware and software until head-mounted display technology advances. In 2012, after the release of the Oculus DK1, VR began to be widely known. However, due to the price of the device and the use of limits, VR still cannot become popular as smartphones. One of the most important features of virtual reality is immersion which is the feeling of being in another space (Freina and Ott, 2015; Škola et al.. 2020). In 360-degree surrounding VR with a head-mounted display, the viewer views a fixed, immobile, and rotating view of the surroundings, although, with a 360-directional view, the user's body in space does not move to change the distance or angle of the viewing object. Companies currently are developing VR devices considering Room-scale as the standard for VR. Room-scale VR defines a field that allows VR users to move freely. Physically moving in space helps the user perceive real-world movement and makes the virtual environment look more realistic. However, the technology requires to create true VR involves VR scripts, 3D production capabilities, interactive programming, and ambient sound design, and the use of VR also requires dedicated hardware, software, and space. They are still expensive, not easy to use, and constantly evolving. Fortunately, the use of inside-out technology supporting room-scale 6-DoF monolithic HMD is affordable for users, gradually being used by the VR community. At present, these monolithic HMDs are limited by hardware specifications, budget, and memory capacity, and are still unable to handle the high number of 3D geometries, high-resolution mapping, and complex picture effects. The method proposed in this study can construct a 360° 3D video into a 3D scene as a 3D virtual environment supporting 6-DoF while maintaining the visual information of the 360° video.

3. Materials and Methods

The traditional camera mapping application is to simulate a single plane of the image into 3D space with limited camera movement. It is commonly used in film, television, and 2D animation. The technological requirement for this study is to convert only 3-DoF 360° video images to restore the 3D filming room. The 3D software used is Maya which is one of the mainstream software in the 3D animation industry. Camera mapping is called Camera Projection in Maya, and with a different name, the function and application are the same. The original media is a 15-minute long with 360° video, and the appearance of the room is shown in Fig. 4.

Fig. 4. 360° video of the actual room.

3.1. Camera mapping in 360° video

First, a sphere is created in the 3D animation software, the 360° video is set as the surface texture, and a 2D video is converted into a 3D sphere shape. The subsequent concept originates from the sky box function in the game engine, which is a cube map composed of six images that looks like the sky background of a sphere. Therefore, in the opposite way, six cameras are placed at the center of the 360° video sphere, corresponding to the six view angles of Top, Button, Front, Back, Left, and Right. The result is shown in Fig. 5.

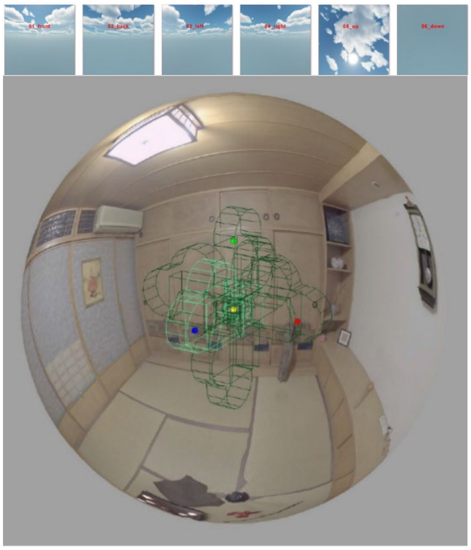

Fig. 5. Six cameras in the sphere corresponding to the six faces of the cube.

3.2. 3D space reconstruction

The camera mapping function is set for each camera to reconstruct the 3D object shape on the screen one by one. Because of the perspective, the field of view between cameras partially overlaps, so it is similar to the 3-point positioning method to slowly correct the model's wireframe from different camera images and establish the correct object shape ratio. Figure 6 shows the camera image before 3D editing, and Fig. 7 shows the process of manually changing the 3D model structure. The figures show that the distorted image becomes gradually corrected and restored. The concept is similar to the way to draw 360° images by hand using Photoshop, except that the tool used to create them is 3D polygon editing. It is possible to turn a 360° image drawn by Photoshop into a full 3D scene with a hand-painted style. The model's wireframe is edited through the six camera angles to obtain a low-poly 3D scene after all the adjustments are completed. The ceiling, floor, and walls in the 360° video are rendered back to a flat surface, and the furniture and furnishings in the room are also corrected to be three-dimensional objects. The final cubic room structure corresponds to the 360° video as shown in Fig. 8.

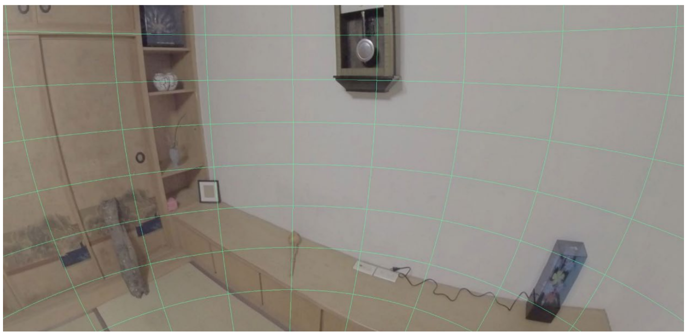

Fig. 6. Mesh before 3D editing is the original shape of the sphere.

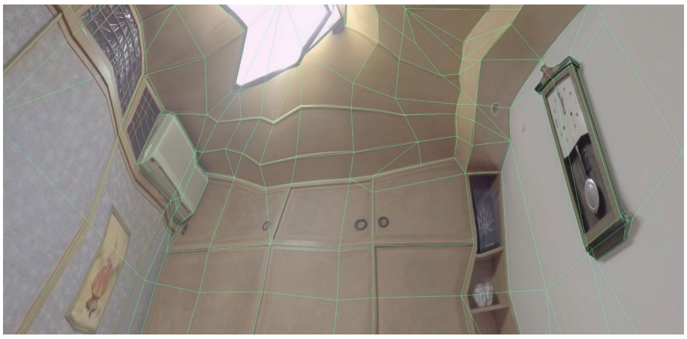

Fig. 7. Reconstructing 3D space from the camera by Camera mapping alignment.

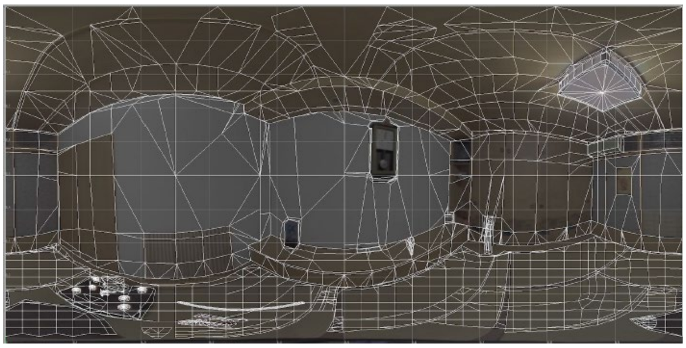

Fig. 8. Completed 3D model wireframe.

3.3. 3D room reconstruction from a sphere

The 3D room, including furniture, cabinets, wall clocks, candles, and others are reconstructed with six cameras and corrected for perspective, and the 360°video body is restored to a normal room model through this process. The results are shown in Fig. 9 which reveals that the 3D wireframe of the floor is part of a sphere-flattened 3D wireframe. The completed 3D rooms are reproduced as four rooms, being closely connected as shown in Fig. 10. The 3D models of clocks and furnishings in different rooms are adjusted in size and position to produce minor differences according to the exhibition needs.

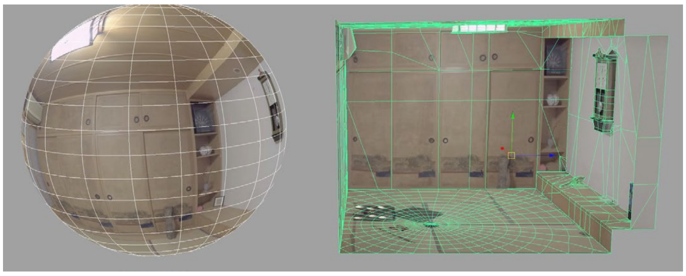

Fig. 9. Left: 360° video of the sphere space, right: reduced to a cubic room model

Fig. 10. 3D models of the interior space are connected without seams.

3.4. Equipment and Space

The VR device used in the exhibition was the HTC Vive Pro, a 6-DoF room-scale VR HMD that supports a maximum space of 10 x 10 m. It is installed in a simple exhibition space, and the cable connecting to the PC is suspended in the center of the room ceiling not to interfere with the audience's activities as shown in Fig. 11. The exhibited content does not need to interact with the virtual environment, so the audience does not need to hold a controller. The duration of the exhibition is set at 18 min, and the length of the VR content is 15 min at the end of which the audience does not enter the black sphere space.

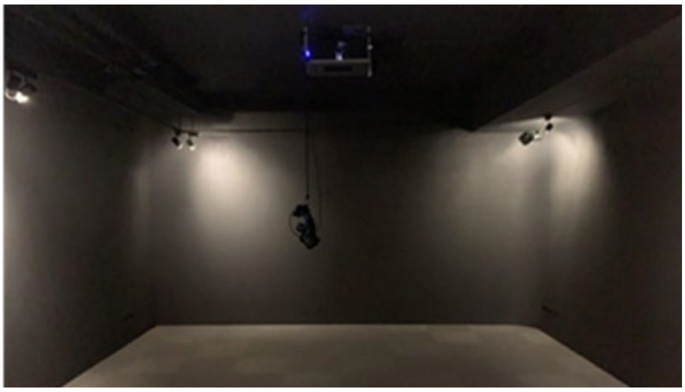

Fig. 11. Virtual reality art exhibition space.

4. Results and Discussion

The 360° video is mapped by Camera mapping as a reference, and the 3D indoor scene is created by the reference. Then, the surface texture is restored by Camera mapping, and the 3-DoF 360° video is converted into a 3D virtual space with 6-DoF room scale. Figure 12 presents the concept of a seamless connection between the rooms and the front, back, left, and right of the rooms, and the audience can walk and move freely through these four virtual rooms with the HTC Vive Pro HMD.

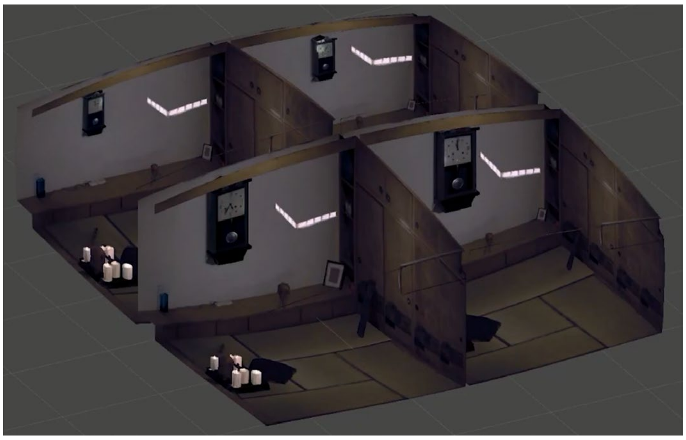

Fig. 12. Audience walks freely through the four rooms.

As the real space of the original shooting is conducted in a dimension of 4 x 4 m, the virtual space is reduced to 2 x 2 m, and the audience has the feeling of expanding body in this virtual space. When the audience moves to the border, they travel to another room that looks the same. After the 3D production modification, the real objects in the scene look familiar and subtly different, thus creating a change in immersion. Finally, when the set viewing time is reached, the environment is transformed into a black sphere. The audience in the black sphere feels the movement of disconnection from the surrounding environment unlike the sense of movement in a virtual room, where the perception of movement is invalid. Through the medium of VR, the audience experiences failure and disconnection between mobile perception and visualization, which is the moment of Tao that the artist intends to present.

Since the surface texture of the 3D scenes is derived from the 360° video, the movement of the hands of the clock, the pendulum of the clock, and the waving of the candle flame are expressed through the video.

The main rules of the Camera mapping process are as follows:

(1) The object to be projected are stationary and not animated.

(2) 3D scenes do not need to add UVW mapping.

(3) It is possible to use dynamic textures in the scene.

However, this method is a complete reconstruction of the space at the time of framing and shooting, so the UVW of the 3D room model is aligned with the 360° video of the sphere, and the surface texture still needs to be processed by UVW mapping.

5. Conclusions

We presented the result of the VR technology research of "TAO in a Moment" at the Wu Yi-Lin Solo Exhibition (Wu, 2020). We used the traditional Camera mapping technology to convert the image of a sphere to a 3D model of a cube using 360° video as a reference and overcome the limitation of 360° video to 3-DoF to realize a virtual reality space with 6-DoF in a room-scale. The finished model has a low polygon count and a highly realistic appearance that can be imported into various game engines to significantly reduce the computational load on the computer. This approach is appropriate for constructing realistic virtual environments for inside-out monolithic VR head-mounted displays. In the future, the LiDAR function of mobile devices can be used to scan the indoor environment to create point cloud data to convert 3D models (Luetzenburg et al., 2021) and then can reduce the surface and retopology of the model. Photo modeling can be applied to restore 3D objects and integrate the production of surroundings and objects to reproduce a more complex and realistic 3D virtual reality space.

Author Contributions: Hung Sun and Tsun-Hung Tsai: Conceptualization, Hung Sun, Tsun-Hung Tsai and Ke Jiang: Methodology, Ke Jiang: Production, Tsun-Hung Tsai and Ke Jiang: Validation, Hung Sun: Writing-original draft preparation, Hung Sun and Tsun-Hung Tsai: Writing-review and editing.

Funding: This research was funded by Digital Art Foundation.

Acknowledgments: Thanks to artist Wu Yi-Lin for his creative ideas.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Dhanda, A., Ortiz, M. R., Weigert, A., Paladini, A., Min, A., Gyi, M., ... and Quintero, M. S. (2019). Recreating cultural heritage environments for VR using photogrammetry. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 305-310.
2. Freina, L., and Ott, M. (2015, April). A literature review on immersive virtual reality in education: state of the art and perspectives. In *The international scientific conference elearning and software for education* (Vol. 1, No. 133, pp. 10-1007).
3. Hannuksela, M. M., Wang, Y. K., and Hourunranta, A. (2019, March). An overview of the OMAF standard for 360 video. In *2019 Data compression conference (DCC)* (pp. 418-427). IEEE.
4. Hebbel-Seeger, A., and Horky, T. (2019). 360-Grad-Foto/Video in der Social Media-Kommunikation im Sport. In *Sportkommunikation in digitalen Medien* (pp. 179-195). Springer VS, Wiesbaden.
5. Jeong, J. B., Lee, S., Ryu, I. W., Le, T. T., and Ryu, E. S. (2020, October). Towards viewport-dependent 6dof 360 video tiled streaming for virtual reality systems. In *Proceedings of the 28th ACM International Conference on Multimedia* (pp. 3687-3695).
6. Kopf, J. (2016). 360 video stabilization. *ACM Transactions on Graphics (TOG)*, *35*(6), 1-9.
7. Luetzenburg, G., Kroon, A., and Bjørk, A. A. (2021). Evaluation of the Apple iPhone 12 Pro LiDAR for an application in geosciences. *Scientific reports*, *11*(1), 1-9.
8. Škola, F., Rizvić, S., Cozza, M., Barbieri, L., Bruno, F., Skarlatos, D., and Liarokapis, F. (2020). Virtual reality with 360-video storytelling in cultural heritage: Study of presence, engagement, and immersion. *Sensors*, *20*(20), 5851.
9. Snelson, C., and Hsu, Y. C. (2020). Educational 360-degree videos in virtual reality: A scoping review of the emerging research. *TechTrends*, *64*(3), 404-412.
10. Song, M. J., and Park, J. W. (2014). A Study on the Types of Camera Working of Projection Mapping Contents. *The Journal of the Korea Contents Association*, *14*(8), 1-12.
11. 【Taipei】TAO in a Moment - Wu Yi-Lin Solo Exhibition. (2020, June 12). Retrieved July 7, 2022, from https://www.digiarts.org.tw/DigiArts/NewsPage/171887508199000/En
12. Xu, M., Li, C., Zhang, S., and Le Callet, P. (2020). State-of-the-art in 360 video/image processing: Perception, assessment and compression. *IEEE Journal of Selected Topics in Signal Processing*, *14*(1), 5-26.