CREATION OF 360° LIGHT FIELDS USING CONCENTRIC MOSAICS WITH VARYING SLIT WIDTHS

Sebastian Knorr, Matthias Knoblauch, and Thomas Sikora

Communication Systems Group, Technical University of Berlin, Germany

ABSTRACT

In this paper, we present a detailed procedure to capture high spatial resolution 360° cylindrical light fields using the approach of concentric mosaics. We therefore use a DSLR camera which is mounted off-centered on a horizontal bar and rotated with a fixed radius. Due to the off-centered arrangement, a 360° image can be created by stitching together predefined columns of each image captured at a certain angle, the so called slit images. By changing the height of the camera after a full rotation, an array of 360° images with horizontal and vertical angular resolution can be created. We then evaluate the impact of varying slit image widths with respect to storage requirements, capture time and resulting image quality.

Index Terms **—** Omnidirectional content, light fields, concentric mosaics, virtual reality

1. INTRODUCTION

With the increasing commercialization of virtual reality (VR) as a medium, one of the main factors that drive the uptake of devices is content, in particular omnidirectional or 360° content. In contrast to traditional content, where the viewer perceives the world through a window i.e. a 2D screen, omnidirectional content allows a person to be present within the world by wearing a head-mounted display (HMD). However, omnidirectional content like e.g. 360° video only allows the user to turn the head when wearing an HMD, i.e. the user has only three degrees of freedom (3DOF) without the possibility to navigate within a captured 3-dimensional scene. Thus, capturing 3-dimensional omnidirectional content which allows the user to explore the scene from different viewpoints, i.e. with six degrees of freedom (6DOF), is currently a hot topic in research and in industry [1, 2, 3].

In addition to computer generated imagery (CGI) where 3D models are available and can be rendered in realtime from different perspectives using game engines like Unreal or Unity, light fields become more and more attractive for live action content creation with 6DOF. Light fields capture all light rays passing through a given volume of space [4]. Compared to traditional 2D imaging systems which capture the spatial intensity of the light rays, the 4D light fields also contain the angular direction of light rays. This additional information allows for multiple applications such as the reconstruction of the 3D geometry of a scene, creating new images from virtual view points, or changing the focus of an image after it is captured.

In this paper we present a detailed description of the hardware setup and procedure to capture a set of 360° cylindrical images, i.e. a 360° light field, with high spatial resolution using a single off-centered DSLR camera and the approach of concentric mosaics [5, 6], which is the first main contribution. Due to the off-centered arrangement, a 360° image can be created by stitching together predefined columns of each image captured at a certain angle, the so called slit images. By changing the height of the camera after a full rotation, an array of 360° images, the so-called sub-aperture images, with horizontal and vertical angular resolution can be created.

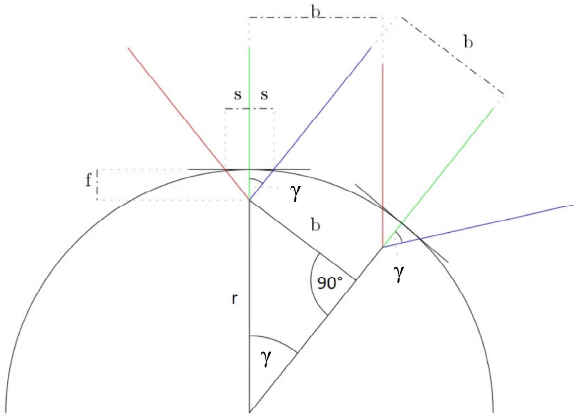


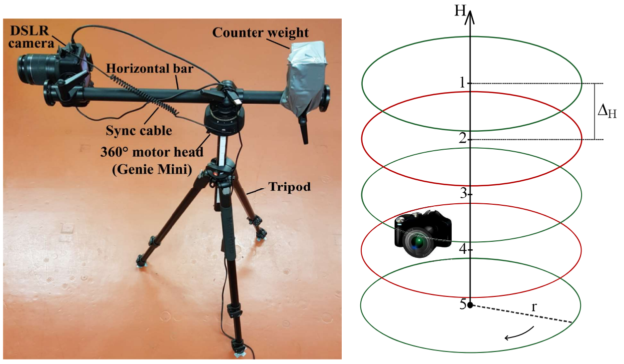
Figure 1: Schematic concept of concentric mosaics and geometric relationships for the calculation of column distances s.

In general, the capture of light field data needs a large amount of storage, which even increases for 360° light fields. Thus, our second contribution is to reduce the storage requirements by increasing the slit width which reduces the amount of captured images per 360° rotation. However, increasing the slit width also increases geometrical distortions when stitching the slit images together. Thus, in this paper, we also evaluate the quality of the 360° images, the storage requirements and the capture time in relation to the slit width. The entire data set, i.e. the captured source data as well as the stitched 360° cylindrical image data set, is publicly available with this paper[1], which is the third main contribution, and can be used for research on 360° light field representation and rendering, e.g. [7]. To our knowledge, it is the first time that a high spatial resolution 360° light field data set was captured with a single DSLR camera.

2. RELATED WORK

A spherical light field was captured in [1] using a DSLR camera with a fish-eye lens that is placed off-centered on a rotating beam. The method is closely related to the approach made in this paper, but differs in the used lens and the low spatial resolution of the sub-aperture images. A spherical light field was also captured by Google VR in [2, 3], which uses 16 Go Pro cameras that are attached to a rotating arc. Each individual Go Pro samples the light field from different viewpoints due to its rotation, similar to [1]. However, the setup becomes more complex and ex-

[1] http://tinyurl.com/y2gvqonv



(a) (b)

Figure 2: (a) Hardware setup and (b) Creation of vertical sub-aperture panorama images ($r = 36.8cm$ and $\Delta_H = 2.3cm$).

pensive, due to its requirement of many components. Birklbauer and Bimber [7, 8] present a novel approach to recording and computing panorama light fields using a rotating plenoptic camera. Their approach is the first that processes ray entries directly and does not require depth reconstruction or matching of image features. Finally, *SpinVr* [9] presents a method that aims to enable live streaming of 360° 3D video content. It makes use of two fish eye lenses that allow the capturing of two perspectives, enabling the 3D display of the scene. However, their system only allows the capturing of stereoscopic 3D content but not light fields.

3. PROPOSED METHOD

3.1. Concentric mosaics

The approach of using panoramic or concentric mosaics in order to create panoramic images was introduced by Peleg and Herman in [5] and later by Shum and He in [6]. The idea is based on sampling the surface of a cylinder using slit images, i.e. single columns extracted from traditional 2D images which were captured off-centered with a given radius. Fig. 1 illustrates schematically the principle of concentric mosaics. The green lines represent light rays passing through the camera centers and the center column of the image planes of camera models with focal length f and off-center camera radius r. By rotating a single camera with radius r around around the optical axis with small steps, the cylindrical plane can be sampled by stitching together the center columns of each image I_k captured at angle γ_k.

Shum and He also showed in [6] that stereoscopic panoramas, i.e. panoramas with horizontal parallax, can be created with the concentric mosaic approach by selecting different columns for each of the panoramas. This is also illustrated in Fig. 1. For instance, the red lines represent light rays passing through the camera centers and columns of the image planes with distance s from the center column. If the camera rotates by an angle γ around the center of the cylinder, the green ray, which is used to create panorama P_C is parallel to the red ray, which is used to create the stereoscopic right panorama P_R. Panoramas P_C and P_R would form a stereoscopic panorama with baseline b. It is obvious that multiple panoramas with different baselines can be extracted from images captured with an off-center rotating camera.

3.2. Hardware setup and technical specifications

In order to capture an array of 360° cylindrical light field images with a single DSLR camera, we setup the hardware as illustrated in Fig. 2a. The setup consists of a tripod with 360° motor head *Genie Mini* [10], a horizontal bar *Vanguard multi-mount 6*, a DSLR camera *Canon 700D* [11] with zoom lens EF-S 18-55mm, a sync cable between camera and Genie Mini, and a counter weight for the camera in order to rotate the camera parallel to the ground plane.

To capture the cylindrical panoramas with a maximum vertical angular field of view (AFOV), we set the focal length $f = 18mm$ and rotated the camera by $90°$. The vertical AFOV α_v of the 360° image depends on the sensor size and the focal length f of the lens, and can be determined with

$$\alpha_v = 2 \cdot tan^{-1}(\frac{l_v}{2 \cdot f}), \quad (1)$$

where l_v is the vertical sensor size, which eventually resulted in a vertical AFOV of $\alpha_v = 63.5°$.

Table 1 shows the select-able image quality types, their file size and their vertical and horizontal resolution, as specified in the user manual of the EOS 700D [11].

3.3. Determination of relevant parameters

The most important parameter to be determined is the amount of required images for a full 360° rotation. It needs to be considered, that the vertical resolution of the cylindrical panorama is fixed with the selected orientation and capture mode. Aiming for the creation of 360° images, the horizontal AFOV equals 360°. The horizontal resolution of the panorama, i.e. the width w_P of a 360° image, equals the amount of slit images N used for its creation, i.e. equal to the captured images if a slit image consists of only one column, and can be calculated with

$$w_P = \frac{h \cdot \alpha_h}{\alpha_v}, \quad (2)$$

where h is the vertical resolution, i.e. height, of the slit image and the horizontal AFOV is $\alpha_h = 360°$.

The Genie Mini motor head can be synchronized with the EOS 700D camera, but the capture speed is limited to a minimum of one picture per second. Table 1 also shows the required number of slit images with a slit width of one column, the resulting capture time and the required total storage size for a full 360° rotation for different capture modes.

Due to a limited storage size and the very time-consuming capturing process, we selected the *fine S1* capture mode as a good trade-off between high resolution and picture quality on one side and the duration and storage size on the other side. According to this selection, the recording of one full 360° rotation takes about 4 hours with a total of 14695 images with 2592×1728 pixel resolution, resulting in cylindrical 360° images with 14695×2592 pixel resolution. The remaining two important parameters for capturing a cylindrical 360° light field are Δ_H, i.e. the difference of the height between rotation planes in order to create vertical sub-aperture panoramas as illustrated in Fig. 2b, and the distance s between slits, i.e. the columns which are used to create horizontal sub-aperture panoramas as illustrated in Fig. 1. We set $\Delta_H = 2.3cm$ in order to allow to capture at least 5 vertical sub-aperture panoramas with the used tripod including the telescopic rod to vary the height.

Based on these settings, s can be determined as follows:

$$s = f \cdot tan(\gamma), \quad (3)$$

with

$$\gamma = sin^{-1}(\frac{b}{r}) \quad (4)$$

using the geometrical relationship as illustrated in Fig. 1, where b is the baseline between parallel rays through the camera centers

Image quality	File size (in MB)	Width w (in pixels)	Height h (in pixels)	Number of slit images N	Duration t (in hours)	Total storage size per rotation (in GB)
fine L	6,4	5184	3456	29,390	8.164	183.89
normal L	3,2	5184	3456	29,390	8.164	91.84
fine M	3,4	3456	2304	19,593	5.443	65.06
normal M	1,7	3456	2304	19,593	5.443	32.53
fine S1	**2,2**	**2592**	**1728**	**14,695**	**4.082**	**31.57**
normal S1	1,1	2592	1728	14,695	4.082	14.86
S2	1,3	1920	1280	10,885	3.024	13.82
S3	0,3	720	480	4,082	1,134	1.2
RAW + fine L	29,9	5184	3456	29,390	8.164	858.15
RAW	23,5	5184	3456	29,390	8.164	674.47

Table 1: Technical specifications and determined parameters for Canon EOS 700D with Canon EF-S Zoom Lenses 18-55mm. The selected image quality and the corresponding parameters are marked in **bold**.

at different angle γ, and $r = 36.8cm$ is the distance between the center of rotation and the camera center of the rotating camera. In order to ensure symmetrical sub-aperture panoramas in horizontal and vertical direction, b needs to be identical to Δ_H, i.e. $b = \Delta_H = 2.3cm$. As the sensor size and image resolution are given with $22.3 \times 14.9mm$ and 2592×1728 pixel, respectively, s can be calculated in pixel coordinates and results in $s = 131$ pixel.

By varying the height of the tripod with $\Delta_H = 2.3cm$ and selecting image columns with distance $s = 131$ pixel, and by keeping the environment unchanged, we captured a 5×5 array of cylindrical 360° images within 4 days.

4. EXPERIMENTAL RESULTS

In this section, we show some results of the proposed method of capturing a cylindrical 360° light field data set using the selected image capture mode. In order to allow the reduction of the recording time, we also performed two additional experiments. First, we captured a 8×8 light field data set by using the video capure mode of the DSLR camera and visually compared the outcome between these 2 capture modes. Secondly, we varied the slit width between 1 and 50 columns and evaluated the quality degradation caused by geometrical distortions.

4.1. Image vs. video capture mode

Fig. 3 visually compares two panoramas created with the proposed method introduced in Section 3, whereas the top panorama was captured in video mode of the DSLR camera and the bottom one was captured in the image mode. As the frame resolution in video capture mode is full HD, the resulting panoramas have a resolution of 10885×1920 pixels. Furthermore, the frame rate in the video capture mode is 29.59 fps which reduces the capture time to 367.86 seconds for a full 360° rotation instead of more than 4 hours in the image capture mode.

Fig. 3b shows close-ups of these panoramas. Both capture modes provide high-resolution panoramas with no geometrical distortions. However, as expected, panoramas generated with the image capture mode provide better illumination properties, resulting in slightly higher contrast and less compression artifacts, which are partly visible in the close-up of the video capture mode.

4.2. Slit width variation

An alternative approach to reduce the capture time and the storage requirements is the increase of the slit width in the image capture mode, i.e. increasing the number of columns in each captured image. For instance, a slit width of 10 columns would reduce the storage size and the capture time by a factor of 10. However, the cylindrical surface would not be sampled accurately, and geometrical distortions might be introduced. Thus, we evaluate the quality of the generated panoramas by varying the slit width between 1 and 50 columns and using the slit width of 1 column as the reference, i.e. ground truth.

Fig. 4 shows the PSNR and the SSIM scores dependent on the number of used columns. Obviously, both PSNR and SSIM decrease with an increase of the slit width. However, even for a slit width of 50 columns, the PSNR and the SSIM with 39.6dB and 0.955, respectively, still have very large values. Fig. **??** also illustrates the SSIM map of a small window of the panorama with relative high textured areas, where black areas indicate a very low SSIM score. Fig. **??** depicts the SSIM scores of this window. Here, the SSIM score drops below a critical value of 0.9 at a slit

(a) Panorama with video (top) and image capture mode (bottom).

(b) Close-up for video (left) and image capture mode (right).

Figure 3: Visual comparison of panoramas captured in video and image capture modes.

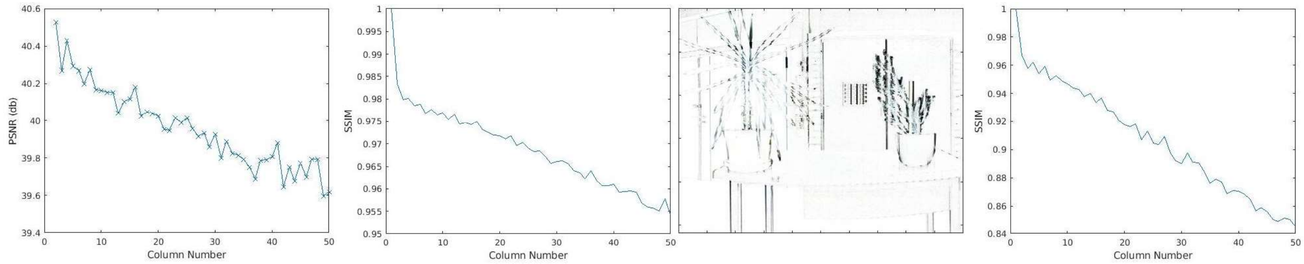


Figure 4: Objective evaluation of the slit width variation. From left to right: PSNR and SSIM scores of a full 360° panorama, SSIM map of a small window of the panorama and SSIM scores of the window.

Figure 5: From left to right: close-up stitched with slit image width of 1, 9 and 28 columns.

width of 28 columns. However, when comparing the image quality visually (see Fig. 5), we found out that a slit width of 9 columns provides subjectively good results while a slit width of 28 columns produces visible distortions at object edges.

5. CONCLUSION AND OUTLOOK

In this paper we presented a detailed description of the hardware setup and procedure to capture a set of 360° cylindrical images, i.e. a 360° light field, with high spatial resolution using a single off-centered DSLR camera and the approach of concentric mosaics. As the capturing of the panorama light field data set is time-consuming and requires a large storage size, we evaluated two alternatives: the capturing of the data set using the video capture mode and the variation of the slit width in the image capture mode of the DSLR camera.

The video capture mode delivers a slightly smaller resolution and a small degradation of the picture quality in terms of contrast and compression artifacts. For the slit width variation, the objective measures do not give sufficient information about the actual picture quality as a visual (subjective) inspection shows that a slit width of more than 9 columns already introduces visible geometrical artifacts for this particular input scene.

Finally, the entire data set with more than 73,475 frames captured in image mode of the DSLR camera and the resulting 5×5 cylindrical 360° light field data set is publicly available with this paper.

Future work should consider the evaluation of different objective quality metrics which are more appropriate for omnidirectional content such as e.g. video multimethod assessment fusion (VMAF) [12] or the Voronoi-based objective quality metric VI-VMAF [13] in order to find an adequate threshold for the slit width objectively, the development of new quality metrics and the recording of additional cylindrical as well as spherical 360° light field data sets using the aforementioned approach.

6. REFERENCES

[1] P. Debevec, G. Downing, M. Bolas, H.-Y. Peng, and J. Urbach, "Spherical light field environment capture for virtual reality using a motorized pan/tilt head and offset camera," in *ACM SIGGRAPH 2015 Posters*, New York, NY, USA, 2015, SIGGRAPH '15, pp. 30:1–30:1, ACM.

[2] "Experimenting with Light Fields," https://www.blog.google/products/google-ar-vr/experimenting-light-fields/, Mar. 2018.

[3] R. S. Overbeck, D. Erickson, D. Evangelakos, M. Pharr, and P. Debevec, "A System for Acquiring , Processing , and Rendering Panoramic Light Field Stills for Virtual Reality," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, pp. 197:1—-197:15, 2018.

[4] M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '96*, New Orleans, USA, 1996, pp. 31–42, ACM Press.

[5] S. Peleg and J. Herman, "Panoramic mosaics by manifold projection," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1997, pp. 338–343.

[6] H.-Y. Shum and L.-W. He, "Rendering with concentric mosaics," in *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '99*, New York, NY, USA, 1999, pp. 299–306.

[7] C. Birklbauer, S. Opelt, and O. Bimber, "Rendering gigaray light fields," *Computer Graphics Forum*, vol. 32, no. 2 PART4, pp. 469–478, 2013.

[8] C. Birklbauer and O. Bimber, "Panorama light-field imaging," *Computer Graphics Forum*, vol. 33, no. 2, pp. 43–52, May 2014.

[9] R. Konrad, D. G. Dansereau, A. Masood, and G. Wetzstein, "SpinVR: Towards live-streaming 3D virtual reality video," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–12, Nov. 2017.

[10] Syrp, "Genie mini," https://syrp.co.

[11] Canon, "EOS Canon 700D," https://www.canon.co.uk.

[12] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara, "Toward a practical perceptual video quality metric," https://medium.com/netflix-techblog/toward-a-practical-perceptual- video-quality-metric-653f208b9652, Jan. 2019.

[13] S. Croci, C. Ozcinar, E. Zerman, J. Cabrera, and A. Smolic, "Voronoi-based objective quality metrics for omnidirectional video," in *11th International Conference on Quality of Multimedia Experience (QoMEX 2019)*, 2019.