A 3DOF+ view rendering method for panoramic light field

Lyu, Xianqiang, Wang, Xue, Zhou, Guoqing, Wang, Qing

Xianqiang Lyu, Xue Wang, Guoqing Zhou, Qing Wang, "A 3DOF+ view rendering method for panoramic light field," Proc. SPIE 11550, Optoelectronic Imaging and Multimedia Technology VII, 115500B (10 October 2020); doi: 10.1117/12.2573864

SPIE.

Event: SPIE/COS Photonics Asia, 2020, Online Only

A 3DOF+ View Rendering Method For Panoramic Light Field

Xianqiang Lyu, Xue Wang, Guoqing Zhou, and Qing Wang

School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China

ABSTRACT

We propose a new method of rendering panoramic light field within a certain range, which provides an effective way for the acquisition of virtual reality data. Different from existing panoramic light field reconstruction algorithms, we propose the concept of the ray sphere to render a panoramic light field. Using the internal and external parameters of the cameras in the acquisition system, all the light field data are uniformly expressed in the ray sphere. The constructed ray sphere enables the rendering of a panoramic light field of any view point within a 3DOF+ space, which can not be achieved with correlation methods. In addition, we design and build an acquisition system to capture real scenes to verify the effectiveness of our method. Experimental results show that our method can get panoramic light field rendering of any view point on the horizontal plane within half the radius of the acquisition system, and can effectively process the light field video data.

Keywords: Panoramic light field, Ray sphere, View rendering, 3DOF+ space

1. INTRODUCTION

A panoramic image, as its name implies, is a complete 360° image made up of images taken by one or multiple cameras, giving people a three-dimensional feeling. Different from plane images, panoramic images can make the viewers feel as if they are in a real scene. Panoramic images, as an extension of the traditional image display form, have been paid more and more attention by scientific researchers in both academia and industry in recent years. For example, 3D panoramic watching experiences for online house hunting websites, panoramic aerial photographing by UAVs, etc.

Light field[1] is an important theoretical innovation comparing with traditional imaging methods, and it is also a hot topic in the computational imaging community. Traditional imaging process is the integration result of ray radiation in the three-dimensional space onto a two-dimensional sensor, so the scene depth information is missing, resulting in some negative effects such as color distortion and image blur. Light field images record both spacial and angular information of the scene simultaneously, and the scene depths can be obtained by decoupling the relationship between the spacial and angular information. Considering that a light field records the light information in a three-dimensional space, we propose that panoramic light field data can be used to establish a ray sphere (the light space of three-dimensional space), and can render panoramic light field from different viewpoints within a certain range. This ability of collecting data one time and multiple rendering embodies the unique advantages of light field for panoramic rendering, and is also what we focus on in this work.

Most of the conventional image-based panoramic rendering stitching algorithms follow similar procedures, including feature point acquisition, feature point matching, image registration and image fusion. These algorithms take pixel as processing object, and focus on the design and optimization of a fusion algorithm, which warp a target image onto a reference image domain. One of the most crucial and challenging steps is image fusing. Since homography is a simple and traditional image warping model which describes the parametric planar transformation based on planar scene assumption, these algorithms can not solve the parallax problem caused by different scene depths when the scene is not planar or the camera baseline is large. The images can not be stitched well for both foreground and background at different depths, also parallax artifacts arise in the vicinity of object boundaries.

Further author information: (Send correspondence to Xue Wang)
E-mail: xwang@nwpu.edu.cn, Telephone: +86 29 8843 1518
The work was supported by NSFC under Grants No.61531014 and No.61801396.

Optoelectronic Imaging and Multimedia Technology VII, edited by Qionghai Dai
Tsutomu Shimura, Zhenrong Zheng, Proc. of SPIE Vol. 11550, 115500B
 · doi: 10.1117/12.2573864

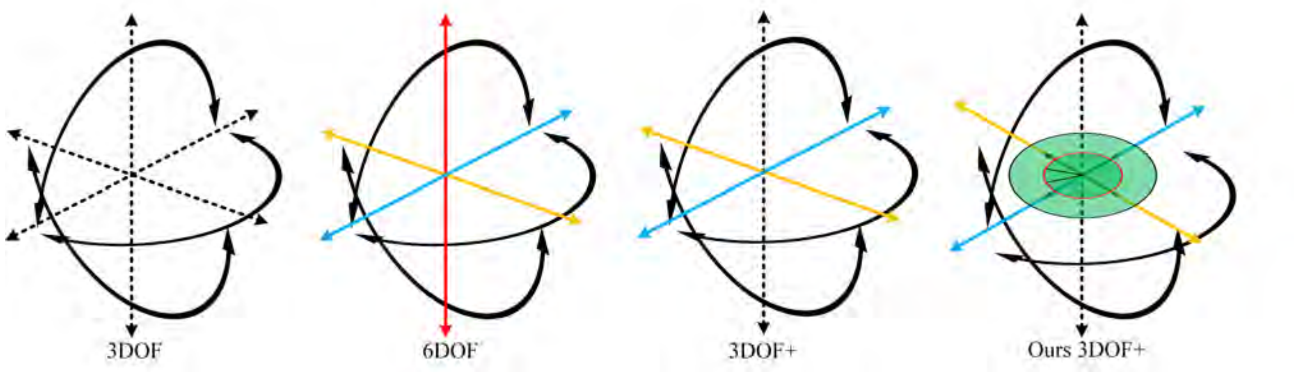

Figure 1. The illustration for the 3DOF+ concept used in this paper. Solid line/curve indicates the actions along this direction are taken into consideration. Dotted line/curve indicates the actions along this direction in not taken into consideration. The scope of synthesized new viewing points in this paper is 3DOF+ in a certain range.

By exploring rendering from the ray perspective, recently, researchers have proposed several successful algorithms for panoramic light field rendering[2,3] and light field stitching.[4] Focusing on rendering a panoramic light field and splicing adjacent light fields respectively, these methods solve the parallax problem of rendering a panoramic light field to a certain extent.

In this paper, we present a novel algorithm for rendering a 3DOF+ panoramic light field. The concept of 3DOF+ used in this paper is shown in figure 1. First of all, we register the light field camera system and build the ray sphere for ray re-sampling. Then the theory of 3DOF+ panoramic light field rendering is analyzed and deduced, which verifies the feasibility of rendering a panoramic light field in a certain range of space. Then the advantages and disadvantages of two rendering techniques are analyzed and compared, which are bicylinder to bicylinder and two-parallel-plane (TPP) to bicylinder. Finally, a rendering method from TPP to bicylinder is proposed to get promising synthesized results of a panoramic light field from new viewing positions.

We set up an experimental platform to collect panoramic light field video, and carry out experiments in real scenes. Experimental results show that the proposed algorithm can effectively render a 3DOF+ panoramic light field. The main contributions are summarized as follows,

(1) We propose the task of rendering a 3DOF+ panoramic light field, which lays the foundation for rendering a 6DOF panoramic light field. Rendering panoramic light field at any viewpoint position can provide high quality data for virtual reality devices.

(2) We propose the ray sphere representation which can be used for ray re-sampling.

(3) We propose an algorithm to render a 3DOF+ panoramic light field, and build a light field video capturing system for collecting real scene light field data for validation. Experimental results with real scene data verify the effectiveness of the proposed algorithm.

2. RELATED WORK

2.1 Panoramic Image

The image stitching and panoramic image rendering techniques are quite mature. Levin et al.[5] propose two representative methods for image stitching in the gradient domain. One is to evaluate the differences between the derivative of the fusion result and the derivative of the source image using a minimum cost function. Another is to use image gradient to optimize and infer fusion results. Zaragoza et al.[6] propose the classic As-Projective-As-Possible (APAP) algorithm, which not only emphasizes the global projection characteristics, but also allows the local non-projection deviation. APAP can effectively avoid the double shadow phenomenon in the overlapped areas of two images. This algorithm makes the stitched image look closer to reality in geometry and present less distortion. Chang et al.[7] propose a new parameterized transformation method, including the combination of

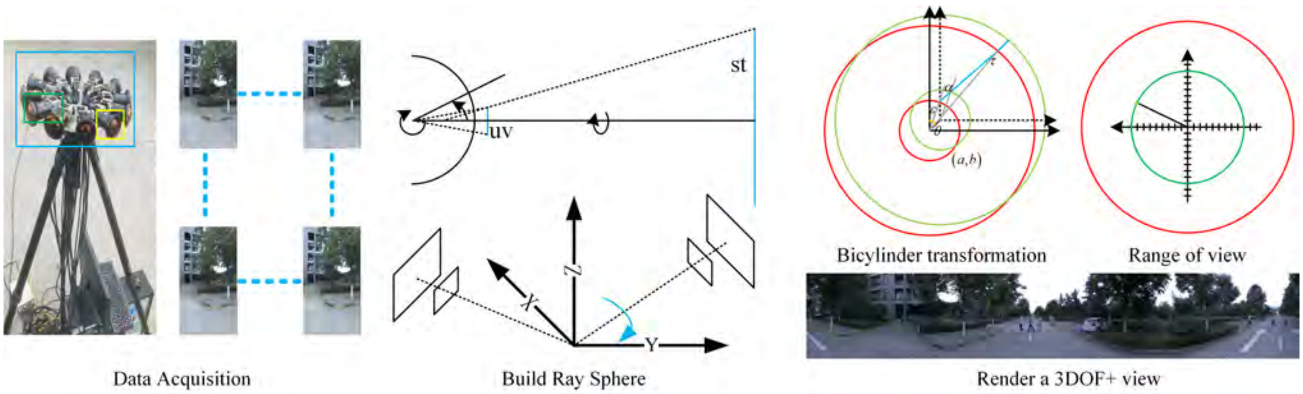

Figure 2. The pipeline of our method for rendering a 3DOF+ panoramic light field. The whole pipeline consists of three parts: data acquisition, building ray sphere and rendering a 3DOF+ view. In the process of data acquisition, multiple self-developed field cameras are used to collect the light field of 360° scenes. Then, all light field cameras are registered in the same world coordinate system to build the ray sphere. Finally a 3DOF+ view of panoramic light field can be obtained by re-sampling the rays in the ray sphere.

global projective transformation, transition region transformation and similar transformation. This transformation method takes the advantages of both projection and similarity transformation, and provides good alignment accuracy of projection transformation, while retaining a single image perspective as the similarity transformation.

The above mentioned methods are essentially the optimization of the stitching algorithm, which can not fundamentally solve the parallax problem in the stitching process. Light field records both the spatial and angular information of light rays, which provides the possibility to fundamentally solve the shortcomings of traditional image stitching process. To avoid the parallax problem caused by changing depths, Birklbauer et al.[2] propose a panoramic light field rendering method based on the mutual transformation between the focusing stack and the light field, which is equivalent to performing image stitching at the same depth layer. Based on,[2] Birklbauer et al.[3] explore a panoramic light field rendering method which doesn't rely on accurate depth estimation and feature point matching. It is the first time to render a panoramic light field from the perspective of ray processing. Guo et al.[4] consider light field stitching from the perspective of light field transformation. They put forward the transformation matrix of light space for the first time, and deduce the transformation relationship between light rays and their corresponding transformation matrix in a Planck coordinate system.

Compared with existing panoramic light field rendering algorithms, our algorithm can not only render a panoramic light field, but also can render panoramic light fields at novel viewpoints within a certain range. In addition, our method can deal with light field video data, which is not supported by existing algorithms.

2.2 View Synthesis

In the fields of computer vision and computer graphics, viewpoint synthesis and image-based rendering (IBR) are classic and difficult tasks. Chen et al.[8] carry out a detailed description and in-depth research on these problems. A large number of conventional view synthesis methods depend on the structure information of the scene. For example, several works[9,10] reconstruct the 3D structure of the scene, and use image warping and blending to obtain synthesized viewpoints. With the development of deep neural networks, the learning-based view synthesis methods achieve better visual effects. Flynn et al.[11] firstly apply deep learning techniques to view synthesis from sets of real-world, natural imagery. Zhou et al.[12] propose a novel light field representation, multiple plane images (MPIs), which can be learned from two input images. They can use predicted MPIs to synthesize the image of the desired view. Mildenhall et al.[13] further improve the prediction network of MPIs, which enables the model to process multiple viewpoint images captured by mobile devices. Using four adjacent groups of MPIs, this method first synthesize the MPIs of novel views, and then obtain the synthesis result of target views. Flynn et al.[14] use a learned gradient descent algorithm to generate MPIs from a set of images taken with sparsely located cameras.

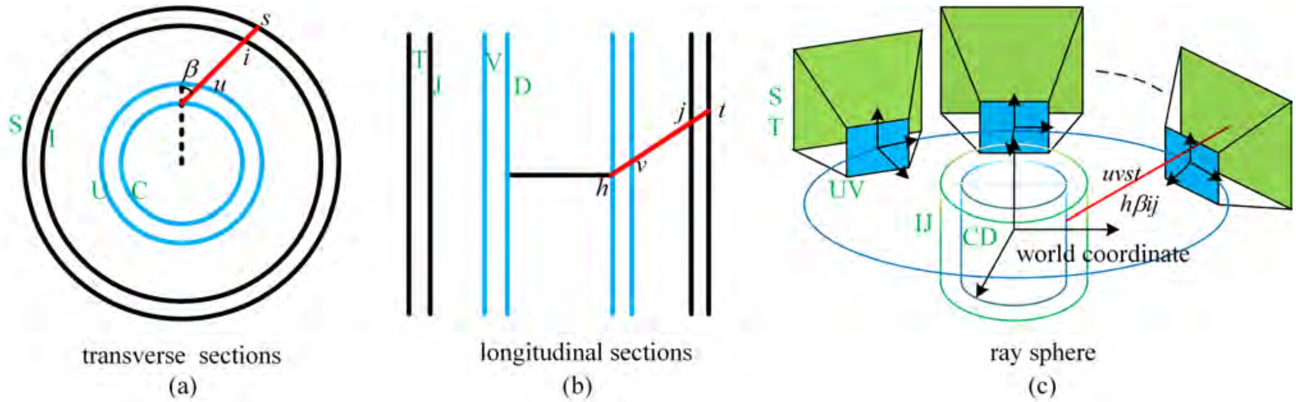

Figure 3. Bicylinder parameterization and ray sphere construction. The red line represents a light ray. The UV/ST and IJ/CD notations correspond to the biplane light field model and the bicylinder light field model respectively.

Some works focus on view synthesis within a light field. Kalantari et al.[15] warp the input views to novel views with estimated depth maps. Levin et al.[16] adopt a Gaussian priori for rendering a 4D light field from a 3D focusing stack. Yoon et al.[17] train multiple CNNs to improve both spatial and angular resolutions. However, this method only synthesize one new view between two or four input views. Wu et al.[18] train a residual based network to improve angular resolution of EPIs. Guo et al.[19] propose a learning based method to reconstruct multiple novel viewpoint images between two mutually independent light fields. The above view synthesis methods mainly synthesize a limited number of views according to existing views, and have certain requirements for the view positions. Moreover, these methods only deal with single or two light fields and not feasible for rendering a 3DOF+ panoramic light field.

3. 3DOF+ VIEW RENDERING FOR PANORAMIC LIGHT FIELD

The pipeline of the proposed algorithm is shown in Figure 2. The whole process is divided into three modules: data acquisition, ray sphere building and 3DOF+ view rendering. This section describes related concepts and derivations for all the three modules.

3.1 Data Acquisition

In the process of data acquisition, multiple cameras are evenly distributed on the acquisition disk to synchronously acquire 360°scene data. The data is preprocessed to prepare for subsequent steps. We estimate the internal and external parameters of the multiple camera system, and use these parameters to represent all the light field cameras in a unified world coordinate system. The collection of all light field data constitutes the ray sphere, which can be used for subsequent ray re-sampling. When rendering a panoramic light field at a new viewpoint location, we need to estimate the correspondences of required rays in the ray sphere, and fuse the traced rays to represent the required rays.

3.2 Ray Sphere Building

In this paper, the bicylinder parameterization method is used to represent the panoramic light field. Figure 3 shows the idea of bicylinder parameterization, in which two sub-figures (a) and (b) are the transverse and longitudinal sections of a bicylinder model respectively. The ST and IJ notations represent the image planes of a plane light field and a panoramic light field respectively. Similarly, the UV and CD notations represent the view plane of a plane light field and a panoramic light field respectively. The key to rendering a panoramic light field is to obtain the ray correspondence between a plane light field and a bicylinder light field. For example, $uvst$ and $h\beta ij$ represent the same ray described in different light field models.

Figure 3(c) shows the brief construction process of a ray sphere. We express all the light field data in the same world coordinate system. All rays constitute the data space, which we call the ray sphere. When building the ray sphere, we need to use the internal and external parameters of the light field camera system. The

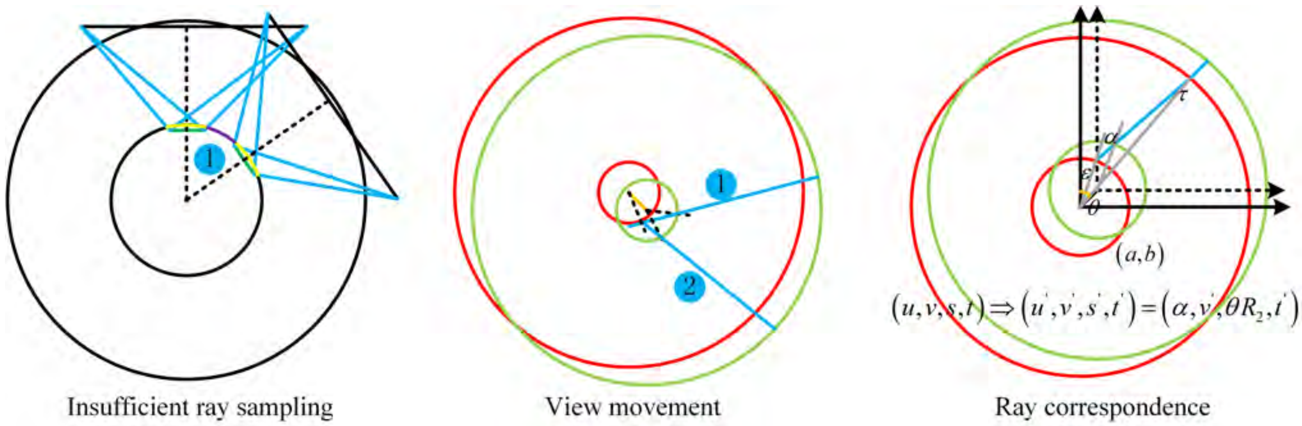

Figure 4. Three issues in 3DOF+ panoramic field rendering, the lack of rays caused by insufficient ray sampling, the lack of ray caused by view movement, and correspondence estimation of light rays. The labels 1 and 2 indicate two cases of ray missing: corresponding rays not been sampled and beyond the field of view of the original light field respectively.

internal parameters of the light field camera are calculated using the algorithm proposed by Zhang et al.[20] For external parameter estimation, we use three angles combined with a distance to determine the camera position and attitude information. With estimated camera system parameters, we can build a complete ray sphere, which is further used in the following ray tracing process.

3.3 Ray Re-sampling

In the process of using the ray sphere to render a 3DOF+ panoramic light field, three issues need to be solved, which are lacking rays caused by insufficient sampling, lacking rays caused by view movement, and correspondence estimation of light rays (4). For internal parameter calibration, the view plane of a plane light field is much smaller than that of a bicylinder light field, which would lead to insufficient ray sampling. We use the purple arc to represent the field of view of the bicylinder light field which can not be sampled by adjacent light field cameras. The rays in this range are missing and can not be collected by adjacent light field. This phenomenon is caused by the contradiction between the design requirements of the acquisition system and internal parameters of the light field camera. To overcome this, we adopt a compromise solution, that is, when the acquisition system is designed, the camera distribution should be as compact as possible, and the view plane of the light field should be magnified equivalently. With the above two-step processing, insufficient sampling can be improved to a certain extent.

When rendering a 3DOF+ panoramic light field, we allow the new view to move within a certain range. If the viewpoint moves beyond of the range, there are some rays that cannot be traced. As shown in the middle in Figure 4, two red circles represent the horizontal section of the original bicylinder light field. After moving to a new view position, the green bicylinder light field is obtained. If the moving range is too large, there will be some light rays that cannot be traced and the rays are beyond the view range of the original light field, which are labelled as 1 and 2 respectively. The reasonable moving range is affected by the radius of the view point surface of the bicylinder light field $radius1$, which is $\frac{1}{2}radius1$. Subsequent experiments show that the range can guarantee better rendering quality, that is, less ray loss and a larger FOV.

The purpose of ray tracing is to obtain ray correspondences between the original light field and the new view light field, which is the key process of rendering a 3DOF+ panoramic light field. As shown in the left of Figure 4, the original light field moves (a, b) to get a new view panoramic light field. A ray (u, v, s, t) in the original light field is represented as $(u^{'}, v^{'}, s^{'}, t^{'})$ in the new viewpoint position, or as $(\alpha, v^{'}, \theta R_2, t^{'})$, where α and θ represent the angles, and R_2 represents the image plane radius of the bicylinder light field. A ray can be determined by two points, so we only need to calculate the four angles $\alpha, \theta, \varepsilon, \tau$ to calculate the corresponding relations of rays. By using the Law of Sines and formula simplification, four angles can be obtained, as shown in Eq. (1,2,3.4). Then the corresponding relations of rays can be calculated.

$$\tau \approx arctan\frac{\Delta u}{dis_{uvst}} \tag{1}$$

$$sin\alpha = \frac{R_2}{R_1}sin\tau \tag{2}$$

$$\theta = \frac{s\Delta s}{R_2} \tag{3}$$

$$\varepsilon = \theta - \alpha + \tau \tag{4}$$

where Δu represents the physical size of a pixel, dis_{uvst} denotes the distance between two planes which can be computed during camera calibration, R_1 and R_2 represent the radius of two bicylinders respectively.

The algorithm of the proposed method is shown in Alg 1. We use n original light fields as the input. We need to estimate the internal and external parameters of the camera system, and then calculate the key parameters of the bicylinder light field. For each original light field, its ray coverage needs to be calculated. After constructing the ray sphere, we re-sample each ray in the 3DOF+ panoramic light field. For the case where there are multiple corresponding rays, we adopt a weighted fusion strategy for multiple rays. Finally, we can render a 3DOF+ panoramic light field with different view positions.

Algorithm 1 3DOF+ Panoramic Light Field Rendering Algorithm.

Input: Original light fields $LF_{1,2,...,n}$, where n represents the number of cameras in the acquisition system
Output: A panoramic Light Field $PF_{(\angle,D)}$
1: Obtain intrinsic parameters and calculate bicylinder parameters
2: Estimate external parameters to determine the camera's attitude
3: **for** $i = 1$ to n **do**
4: Initialize horizontal rotation angle
5: Calculate the ray coverage and enlarge it by 5 degrees leftward and rightward
6: **end for**
7: Calculate the position of the new view panoramic light field, using angle and distance $\angle$ and D
8: **for** Every ray $(h\beta ij)$ in the panoramic light field **do**
9: Ray tracing to obtain corresponding original ray (u, v, s, t)
10: Using a weighted fusion strategy for multiple rays
11: **end for**
12: The panoramic light field with different view positions $\angle, D$ is rendered.

4. EXPERIMENTS AND RESULTS

4.1 Device Configuration of Real Scene

After the above theoretical analysis, we design and build a panoramic light field system for collecting real scene data, as shown in Fig. 2. We conduct several experiments in real scenes to validation the performance of the proposed algorithm. The whole acquisition system uses 12 self-developed light field cameras, which are fixed on the acquisition platform through the base. The diameter of the whole platform is $26cm$, and each light field camera is fixed every 30°. The axial distance between a light field camera and the center of the platform is about 12-15cm. Due to the influence of sensor size and camera lens, the FOV of a light field camera is about 90°longitudinally and 60°horizontally. Using the 12 field cameras for data collection can ensure nearly 50% overlapping between adjacent light fields. The highest acquisition frame rate of each camera is 30 frames. The data captured with the 12 light field cameras are collected synchronously using a synchronous acquisition equipment. We use our own toolbox to decode the raw data and get the light field data for subsequent rendering.

After each frame data of each camera is decoded, the number of effective viewpoints is 7×7. The resolution of each viewpoint is 456×241, which is processed by $2\times$ super-resolution to 912×482. Then the size of a single

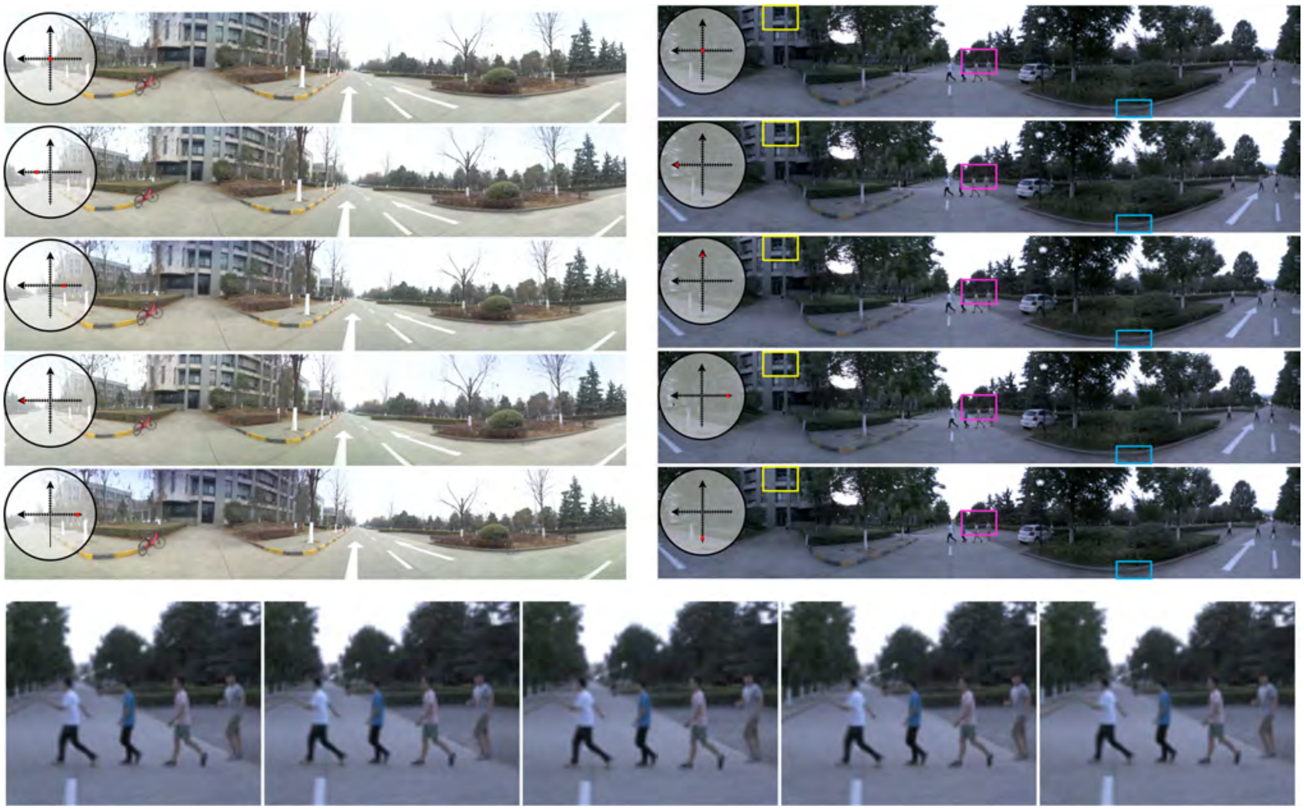

Figure 5. Experimental results of panoramic light field rendering at different views.

light filed data is about $60MB$. One frame data from the acquisition system occupies $720MB$ storage space approximately. If the acquisition frame rate is set to 30 fps, the amount of data per second will exceed $20GB$, which is a huge amount of data to be processed. After rendering, our algorithm get a new viewpoint image with a resolution of 4400×780, which is cropped to get a better visual effect.

4.2 3DOF+ Panoramic Light Field Rendering

We conduct several real experiments to certify the performance of the proposed 3DOF+ panoramic light field rendering method. In the experimental verification using real data, the polar coordinates are used to represent the position of new viewpoint, that is, radius and angle are used to represent any position within the rendering range. Fig. 5 shows the rendering results of a panoramic light field at different viewpoint positions and the differences between panoramic light fields caused by the changing the viewpoint position. The upper left corner of each sub-figure is the schematic diagram of the viewpoint position, and the red dot indicates the current viewpoint position. In order to emphasize the difference between rendering results caused by moving viewpoint, we mark several local regions with different color frame lines in the same position of the five groups of panoramic field images on the right. The bottom shows the local enlarged images for comparing between the rendering results of the panoramic light field with different viewpoint positions.

We also verify the algorithm on different scenarios and light field video data. An open natural scene and a multiple people scene are collected to test the performance of the algorithm. The frame rate for capturing light field data is set to 25 fps. In the dynamic light field, there are multiple pedestrians, and more complex occlusions are included. The rendered panoramic light field has no obvious inter frame jumping and stitching seam artifacts. Also the rendering results of both foreground and background are consistent. As shown in Fig. 6, our algorithm has achieved good visual effects on different scenes and also for light field video, which proves the effectiveness of our algorithm.

Figure 6. Experimental results of different scenes and frames.

5. CONCLUSION

In this paper, we first propose the task of rendering a 3DOF+ panoramic light field. In order to solve this problem, we propose the concept of ray sphere, and use the ray sphere for ray tracing. Afterwards, the key problems in the rendering process and the model derivation are analyzed. Finally, several experiments for real scenes are carried out. Experimental results show the effectiveness of our algorithm and its ability to process video data. In the future, we will explore the method of 6DOF panoramic light field rendering.

REFERENCES

[1] Levoy, M. and Hanrahan, P., "Light field rendering," in [*Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*], 31–42, ACM (1996).

[2] Birklbauer, C., Opelt, S., and Bimber, O., "Rendering gigaray light fields," in [*Computer Graphics Forum*], **32**, 469–478 (2013).

[3] Birklbauer, C. and Bimber, O., "Panorama light-field imaging," in [*Computer Graphics Forum*], **33**, 43–52 (2014).

[4] Guo, X., Yu, Z., Kang, S. B., Lin, H., and Yu, J., "Enhancing light fields through ray-space stitching," *IEEE Transactions on Visualization and Computer Graphics* **22**(7), 1852–1861 (2016).

[5] Levin, A., Zomet, A., Peleg, S., and Weiss, Y., "Seamless image stitching in the gradient domain," in [*8th European Conference on Computer Vision*], **3024**, 377–389 (2004).

[6] Zaragoza, J., Chin, T.-J., Tran, Q.-H., Brown, M., and Suter, D., "As-projective-as-possible image stitching with Moving DLT," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* **36**(7), 1285–1298 (2014).

[7] Chang, C., Sato, Y., and Chuang, Y., "Shape-preserving half-projective warps for image stitching," in [*2014 IEEE Conference on Computer Vision and Pattern Recognition*], **1**, 3254–3261 (2014).

[8] Chen, S. E. and Williams, L., "View interpolation for image synthesis," 381390 (1998).

[9] Vedula, S., Baker, S., Rander, P., Collins, R., and Kanade, T., "Three-dimensional scene flow," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **27**(3), 475–480 (2005).

[10] Zitnick, C. L., Kang, S. B., Uyttendaele, M., Winder, S., and Szeliski, R., "High-quality video view interpolation using a layered representation," *Acm Transactions on Graphics* **23**(3), 600 (2004).

[11] Flynn, J., Neulander, I., Philbin, J., and Snavely, N., "Deep stereo: Learning to predict new views from the world's imagery," in [*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], 5515–5524 (2016).

[12] Zhou, T., Tucker, R., Flynn, J., Fyffe, G., and Snavely, N., "Stereo magnification: Learning view synthesis using multiplane images," *ACM Trans. Graph.* **37**(4) (2018).

[13] Mildenhall, B., Srinivasan, P. P., Ortiz-Cayon, R., Kalantari, N. K., Ramamoorthi, R., Ng, R., and Kar, A., "Local light field fusion: Practical view synthesis with prescriptive sampling guidelines," *ACM Transactions on Graphics (TOG)* (2019).

[14] Flynn, J., Broxton, M., Debevec, P., DuVall, M., Fyffe, G., Overbeck, R., Snavely, N., and Tucker, R., "Deepview: View synthesis with learned gradient descent," in [*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], (2019).

[15] Kalantari, N. K., Wang, T. C., and Ramamoorthi, R., "Learning-based view synthesis for light field cameras," *Acm Transactions on Graphics* **35**(6cd), 193.1–193.10 (2016).

[16] Levin, A. and Durand, F., "Linear view synthesis using a dimensionality gap light field prior," *The IEEE Computer Society Conference on Computer Vision and Pattern Recognition* , 1831 – 1838 (07 2010).

[17] Yoon, Y., Jeon, H. G., Yoo, D., Lee, J. Y., and Kweon, I. S., "Learning a deep convolutional network for light-field image super-resolution," in [*IEEE International Conference on Computer Vision Workshop (ICCVW)*], 24–32 (2015).

[18] Wu, G., Zhao, M., Wang, L., Dai, Q., Chai, T., and Liu, Y., "Light field reconstruction using deep convolutional network on epi," in [*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], 1638–1646 (2017).

[19] Guo, M., Zhu, H., Zhou, G., and Wang, Q., "Dense light field reconstruction from sparse sampling using residual network," 50–65 (2018).

[20] Zhang, Q., Zhang, C., Ling, J., Wang, Q., and Yu, J., "A generic multi-projection-center model and calibration method for light field cameras," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **41**(11), 2539–2552 (2019).