MegaParallax: 360° Panoramas with Motion Parallax

Tobias Bertel
University of Bath
T.B.Bertel@bath.ac.uk

Christian Richardt
University of Bath
christian@richardt.name

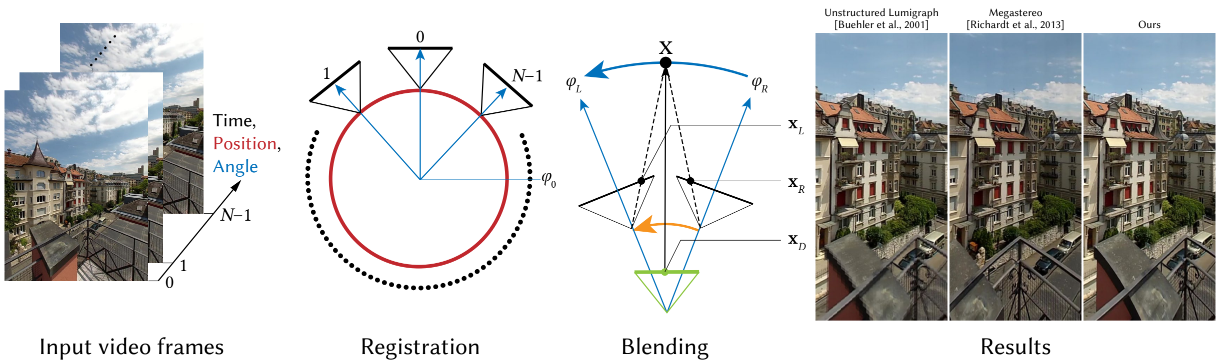

Figure 1: We propose a novel approach for creating high-quality 360° panoramas – with motion parallax – from a single input video sweep. We start from an input video captured with a consumer camera, and register each video frame on a circle using structure-from-motion. Using flow-based blending, we synthesise novel views (green camera) on the fly between each pair of captured images (black), which produces impressive motion parallax. Our results show correct perspective compared to Megastereo [Richardt et al. 2013] and avoid the ghosting artefacts of the Unstructured Lumigraph [Buehler et al. 2001].

ABSTRACT

Capturing 360° panoramas has become straightforward now that this functionality is implemented on every phone. However, it remains difficult to capture immersive 360° panoramas with motion parallax, which provide different views for different viewpoints. Alternatives such as omnidirectional stereo panoramas provide different views for each eye (binocular disparity), but do not support motion parallax, while Casual 3D Photography [Hedman et al. 2017] reconstructs textured 3D geometry that provides motion parallax but suffers from reconstruction artefacts. We propose a new image-based approach for capturing and rendering high-quality 360° panoramas with motion parallax. We use novel-view synthesis with flow-based blending to turn a standard monoscopic video into an enriched 360° panoramic experience that can be explored in real time. Our approach makes it possible for casual consumers to capture and view high-quality 360° panoramas with motion parallax.

CCS CONCEPTS

• **Computer graphics** → **Image-based rendering**; *Computational photography*;

KEYWORDS

image-based rendering, novel-view synthesis, plenoptic modeling

SIGGRAPH '18 Posters, August 12–16, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5817-0/18/08.
https://doi.org/10.1145/3230744.3230793

ACM Reference Format:
Tobias Bertel and Christian Richardt. 2018. MegaParallax: 360° Panoramas with Motion Parallax. In *Proceedings of SIGGRAPH '18 Posters.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3230744.3230793

1 INTRODUCTION

360° panoramas and videos are becoming increasingly widespread thanks to affordable consumer 360° cameras. However, the initial excitement about this medium is often met with disappointment as the imagery cannot provide any non-pictorial depth cues, such as stereo (binocular) disparity or motion parallax. Our goal is to lower the barriers for casual consumers to get their hands on high-quality 360° panoramas with motion parallax. We strongly believe that it is important to make this technology accessible more easily – if we provide the right tools, creative people will create incredible things.

The lack of depth cues is a principal limitation of any capture approach that only uses a single viewpoint, such as 360° cameras. Buehler et al.'s unstructured lumigraph rendering (ULR) [2001] creates new views by blending existing views dependent on their proximity, similarity of viewing direction and other considerations. The quality of results greatly depends on the number of input views and the quality of the geometry proxy, with poor proxies leading to blurry results due to ghosting artefacts. Davis et al. [2012] guide users during the capturing process of a light field to ensure that images are taken evenly over many angles while keeping the same distance to an object. This produces improved results for cameras looking at an object, but it is not applicable for panoramas looking *inside-out*. Omnidirectional stereo approaches [e.g. Richardt et al. 2013] produce multi-perspective stitched panoramas with stereo disparity but no motion parallax. Hedman et al. [2017] propose a pipeline for casual 3D photography, which reconstructs a textured

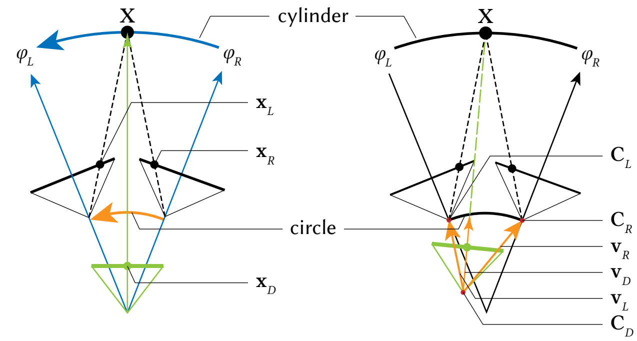

Figure 2: View synthesis. The intersection of the desired camera's forward direction with camera circle and cylinder finds the closest camera pair (L, R). The colour of pixel $\mathbf{x}_D$ is determined by flow-based blending between $\mathbf{x}_L$ and $\mathbf{x}_R$, the projections of the scene point X into cameras L and R.

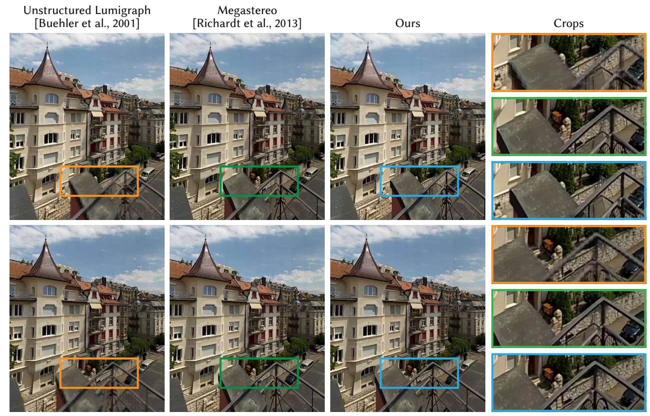

Figure 3: Comparison of synthetic views for two viewpoints, one per row. Note that unstructured lumigraph rendering (left) shows seams and blurry artefacts, and Megastereo (centre) does not support motion parallax. Our approach (right) produces high-quality views with motion parallax.

3D geometry that can be viewed from any viewpoint, providing both stereo and motion parallax. However, 3D reconstruction remains fragile and prone to artefacts, e.g. for thin or distant objects in a scene. Most recently, Luo et al. [2018] introduced a 360° scene representation for motion parallax that synthesises new viewpoints using a combination of disparity and optical flow fields. Their capture process requires a robot arm capturing more than 4,000 images over two hours, which is prohibitive for casual consumers.

In contrast, we propose a simple acquisition process to casually capture 360° panoramas and an approach for image-based view synthesis that produces appealing motion parallax in real time.

2 OUR APPROACH

We start by calibrating the frames of the input video intrinsically and extrinsically using structure-from-motion, and undistorting the images into pinhole images. All input frames are registered to a circle to reflect the camera path. We use the circle's polar angle to impose an order on all frames, and precompute bidirectional optical flow between all pairs of adjacent images. Novel views are then rendered in real time as follows (see diagrams in Figure 2).

For a particular desired view D, the principal ray (in green) intersects both the camera circle (in red), yielding a polar angle φ that determines the *closest* camera pair (L, R), and cylindrical imaging surface (in blue). A point $\mathbf{X}$ on the cylinder is projected into the cameras L and R, onto image coordinates $\mathbf{x}_L$ and $\mathbf{x}_R$, respectively. The colour of the corresponding projection in the desired view, $\mathbf{x}_D$, is determined by a linear interpolation of left and right colours guided by optical flow [Richardt et al. 2013], and weighted with respect to the position $\mathbf{C}_D$ of the desired camera.

All results shown here are created with OpenGL on an Intel i7 with 4 GHz, 10 GB RAM with an NVIDIA GTX 960. We apply our method on the 'Rooftop' dataset [Richardt et al. 2013], which was captured with a GoPro HD HERO2, at 960×1280 pixels and 48 fps, on a ring with a radius of 0.8 m. Processing takes 2–3 hours in total, of which 2 hours are required by the sparse reconstruction performed by COLMAP. Optical flow (using OpenCV's Brox GPU flow) takes 1.5 s per flow on a resolution of 480×640 pixels. The viewpoint synthesis, which demonstrates impressing parallax, runs at 400 fps to render a single viewpoint in a resolution of 1920×1080 pixels.

3 CONCLUSION

We presented a new solution for generating and displaying high-quality 360° panoramas with motion parallax from just a single input video. Our method is easy to adopt and produces convincing results despite not using any explicit proxy geometry. However, our approach would benefit from the availability of accurate proxy geometry, which would naturally lead to more translational freedom of novel viewpoints. A computational bottleneck of our approach is the reconstruction of extrinsic camera geometry, which also suffers from instabilities to the nature of the camera set-up on a ring looking outward.

ACKNOWLEDGMENTS

This work was supported by EU Horizon 2020 MSCA grant FIRE (665992), the EPSRC Centre for Doctoral Training in Digital Entertainment (EP/L016540/1), RCUK grant CAMERA (EP/M023281/1), a Rabin Ezra Scholarship and an NVIDIA Corporation GPU Grant.

REFERENCES

Chris Buehler, Michael Bosse, Leonard McMillan, Steven Gortler, and Michael Cohen. 2001. Unstructured Lumigraph Rendering. In *Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*. 425–432. https://doi.org/10.1145/383259.383309

Abe Davis, Marc Levoy, and Fredo Durand. 2012. Unstructured Light Fields. *Computer Graphics Forum (Proceedings of Eurographics)* 31, 2 (May 2012), 305–314. https://doi.org/10.1111/j.1467-8659.2012.03009.x

Peter Hedman, Suhib Alsisan, Richard Szeliski, and Johannes Kopf. 2017. Casual 3D Photography. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 36, 6 (November 2017), 234:1–15. https://doi.org/10.1145/3130800.3130828

Bicheng Luo, Feng Xu, Christian Richardt, and Jun-Hai Yong. 2018. Parallax360: Stereoscopic 360° Scene Representation for Head-Motion Parallax. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1545–1553. https://doi.org/10.1109/TVCG.2018.2794071

Christian Richardt, Yael Pritch, Henning Zimmer, and Alexander Sorkine-Hornung. 2013. Megastereo: Constructing High-Resolution Stereo Panoramas. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)*. 1256–1263. https://doi.org/10.1109/CVPR.2013.166