Citation for published version:
Bertel, T, Campbell, N & Richardt, C 2018, 'MegaParallax: 360° Panoramas with Motion Parallax' CVMP 2018, London, UK United Kingdom, 13/12/18 - 14/12/18, .

Publication date:
2018

Document Version
Peer reviewed version

[Link to publication](#)

University of Bath

MegaParallax: 360° Panoramas with Motion Parallax

Tobias Bertel
T.B.Bertel@bath.ac.uk
Neill D. F. Campbell
N.Campbell@bath.ac.uk
Christian Richardt
christian@richardt.name

Department of Computer Science,
University of Bath

Panoramic environments can be casually acquired by mobile phones. Unfortunately, representations suitable for *real-world* virtual reality (VR) must provide (at least) binocular disparity for stereo vision and supported head motion, e.g. rotation and translation. The common equirectangular representations of panoramic environments offer only monocular vision and only up to 2D eye rotations and are thus not suitable for VR experiences out of the box.

We propose a pipeline to capture, process and display cylindrical 360° panoramas with *motion parallax*, an important depth cue for monocular vision, suitable for *real-world* VR applications. We do not stitch a video into an equirectangular representation, but use the individual frames in an omnidirectional stereo panorama setup [3] and perform image-based rendering (IBR) without accurate scene proxy or per-image depth maps. The IBR formulation allows our method to synthesise novel viewpoints on-the-fly within the area of the captured camera circle, sampling input images efficiently per pixel in order to create a novel view, which in turn leads to motion parallax. Our approach makes it possible for casual consumers to capture and view high-quality 360° panoramas with binocular disparity and motion parallax. We assume static environments throughout the pipeline.

Our pipeline is depicted in Figure 1. a) we start from an input video captured with a consumer camera, and b) register each video frame on a circle. First, the camera poses of the individual video frames are estimated using structure-from-motion (SfM). We undistort the images before the SfM reconstruction to obtain more accurate geometry. The obtained pose $\mathbf{E}$ together with known intrinsic parameters $\mathbf{K}$ per input image I define a *viewpoint* $V = (\mathbf{P}, I)$, in which $\mathbf{P} = \mathbf{K} \cdot \mathbf{E}$ is the projection matrix. A frame i is identified with the viewpoint $V_i = (\mathbf{P}_i, I_i)$ after registration. Second, we fit a circle to the estimated camera poses. The centre is determined as the centroid of the poses, the normal by the average up-direction of the camera poses and the radius as the average distance of the camera centres to the centroid.

We compute dense correspondences between neighbouring pairs of viewpoints in form of optical flow in c). The baselines between neighbouring viewpoints is only a few millimetres and thus optical flow reliably estimates inter-frame pixel motion for static environments. Still, very close scene objects, which contain naturally more parallax in a pair of consecutive frames, can be hard to catch.

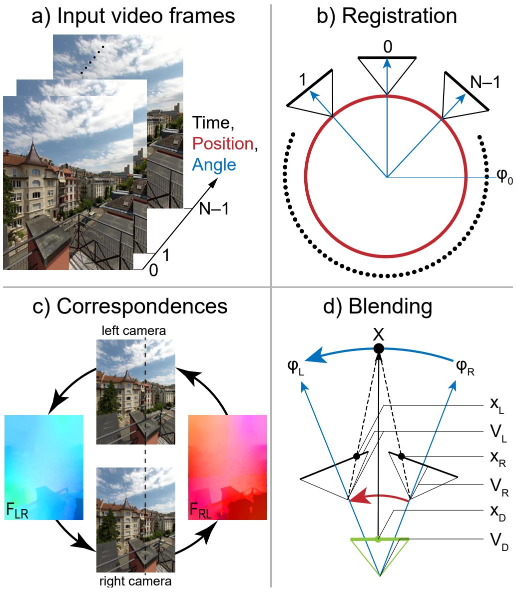

Figure 1: Overview of our approach to casually create and display VR content. See description in text.

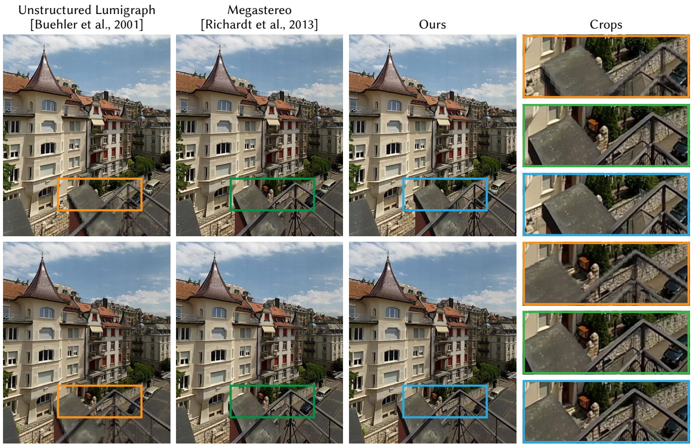

Figure 2: Comparison of synthetic views for two viewpoints, one per row. Our approach shows artefact-free motion parallax. Description in text.

We next synthesise a desired novel view V_D (shown in green in Figure 1d) within the camera ring. We use a plane attached to the OpenGL camera to create rasterised world points, so only directions with respect to the desired camera are considered during view synthesis. The rendering itself can be broken down into the following steps:

1. obtain a world point X from the OpenGL rasterisation stage,
2. determine the camera pair which encloses X, i.e. V_L and V_R,
3. project X into the viewpoints, i.e. $x_i = \mathbf{P}_i X$ for $i \in \{L, R\}$ (note that X is available per pixel of the desired viewpoint x_D), and
4. determine the colour of x_D by linearly interpolating x_L and x_R using depth-compensating flow-based blending.

We apply our method on the ‘ROOFTOP’ dataset [3], which was captured with a GoPro HD HERO2, at 960×1280 pixels and 48 fps, on a rotary stage with 1.2 m radius (see Figure 2). Note that unstructured lumigraph rendering [2] (left) shows seams and blurry artefacts, and Megastereo (centre) does not support motion parallax. Our approach (right) produces high-quality views with motion parallax by combining the best of two worlds, general image-based view synthesis performed on a dense set of calibrated viewpoints, commonly used in omnidirectional approaches.

Our results show the same non-linear perspective for close scene objects caused by vertical distortion which is commonly observed in state-of-the-art commercial VR content generation systems, e.g. Google Jump [1]. We need to capture and store 200–400 frames on a circle to perform artefact-free viewpoint synthesis without scene geometry. Thus, our approach supports 1D camera manifolds Our approach would benefit from accurate proxy geometry, firstly to increase the space in which novel views can be synthesised and secondly to address vertical distortion for close scene objects not well represented with the used planar proxy surface. The computational bottleneck of our approach is the structure-from-motion computation. Structure-from-motion is also known to not be robust for our desired narrow-baseline inside-out capturing scenario.

This work was funded by EU H2020 MSCA FIRE (665992), EPSRC CDE (EP/L016540/1), RCUK grant CAMERA (EP/M023281/1), a UKRI Innovation Fellowship (EP/S001050/1) and a Rabin Ezra Scholarship.

[1] R. Anderson, D. Gallup, J. T. Barron, J. Kontkanen, N. Snavely, C. Hernandez, S. Agarwal, and S. M. Seitz. Jump: Virtual reality video. *ACM Trans. Graph.*, 35(6):198:1–13, 2016.

[2] C. Buehler, M. Bosse, L. McMillan, S. Gortler, and M. Cohen. Unstructured lumigraph rendering. In *SIGGRAPH*, 2001.

[3] C. Richardt, Y. Pritch, H. Zimmer, and A. Sorkine-Hornung. Megastereo: Constructing high-resolution stereo panoramas. In *CVPR*, 2013.