All-around Depth from Small Motion with A Spherical Panoramic Camera

Sunghoon Im, Hyowon Ha, Francois Rameau, Hae-Gon Jeon, Gyeongmin Choe, In So Kweon
Korea Advanced Institute of Science and Technology (KAIST)

ECCV'16 EUROPEAN CONFERENCE ON COMPUTER VISION

Introduction

✓ **Motivation**

- For user-created VR contents, hand-held 360 VR cameras are released, but it cannot provide stereoscopic image (needs depth) which is essential for realistic VR contents.

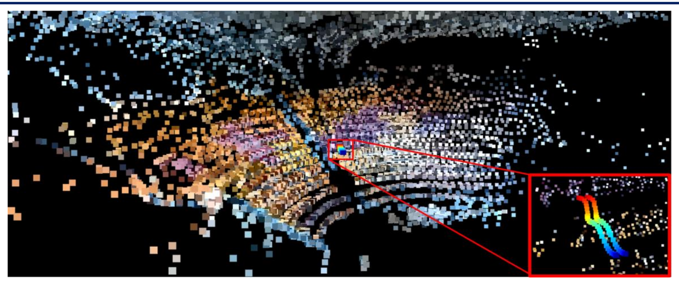
✓ **Objective**

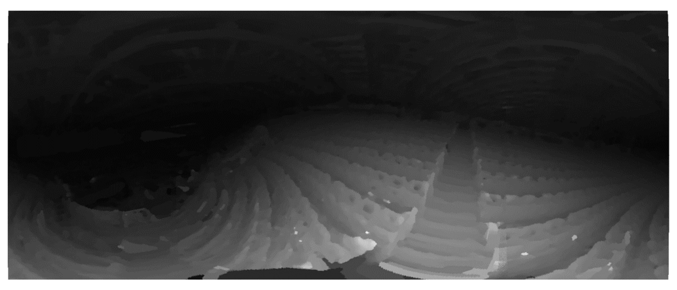
- Generate an all-around depth for 360 VR camera from 1-second small motion video.

✓ **Contribution**

- Unified bundle adjustment with frontal and rear camera whose residual computed on the unit sphere domain, instead of image domain.
- Sphere sweeping method on the basis of the unit sphere


Input image


3D points and camera pose


Depth map result


Stereoscopic panorama


360 VR Cameras

Small motion Bundle adjustment for SPC (360 VR Camera)

✓ **Unified bundle adjustment for both frontal and rear camera**

- Estimate approximately metric depth because of the pre-calibrated camera extrinsic.

✓ **Cost function is designed on the unit sphere domain, instead of image domain.**

- Omnidirectional cameras have two-projection model which increases the complexity of the cost function (hardly converges with a high-order model.
- The re-projection error is uniformly mapped on the sphere which is not the case in the image domain because of the no-linear resolution induced by fisheye lenses.

✓ **Bundle adjustment formulation**

- Find rotation (r), translation (t) and inverse depth (W^F, W^R),

Frontal camera — Rear camera

$$\underset{r,t,W}{\operatorname{argmin}} \sum_i \left(\sum_j \| \hat{\mathrm{X}}^F_{ij} - \left\langle \mathcal{P}^F_i \begin{bmatrix} \mathrm{X}^F_j \\ 1 \end{bmatrix} \right\rangle \|_{\mathrm{H}} + \sum_j \| \hat{\mathrm{X}}^R_{ij} - \left\langle \mathcal{P}^R_i \begin{bmatrix} \mathrm{X}^R_j \\ 1 \end{bmatrix} \right\rangle \|_{\mathrm{H}} \right)$$

point on the unit sphere — Projection on the unit sphere

$$\mathrm{X}_j = \hat{\mathrm{X}}_j / w_j, \ \mathrm{X} = \hbar^{-1}(\mathrm{x}, w)$$

$$\mathrm{x} = \hbar(\mathrm{X}) = \begin{bmatrix} \mathrm{X}/(\mathrm{Z} + \| \mathrm{X} \| \xi) \\ \mathrm{Y}/(\mathrm{Z} + \| \mathrm{X} \| \xi) \\ 1 \end{bmatrix}$$

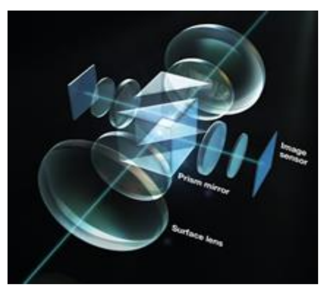
$\|\cdot\|_{\mathrm{H}}$: Huber loss function

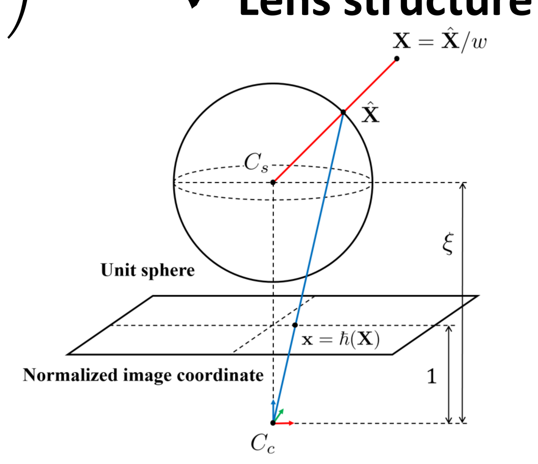
$$\mathcal{P}^F_i = [R(r_i)|t_i],$$

$$\mathcal{P}^R_i = \boldsymbol{P} \begin{bmatrix} \mathcal{P}^F_i \\ m \end{bmatrix} \begin{bmatrix} \boldsymbol{P} \\ m \end{bmatrix}^{-1},$$

$$m = [0\ 0\ 0\ 1]$$

$\boldsymbol{P}$: **pre-calibrated camera extrinsic**


✓ **Lens structure**



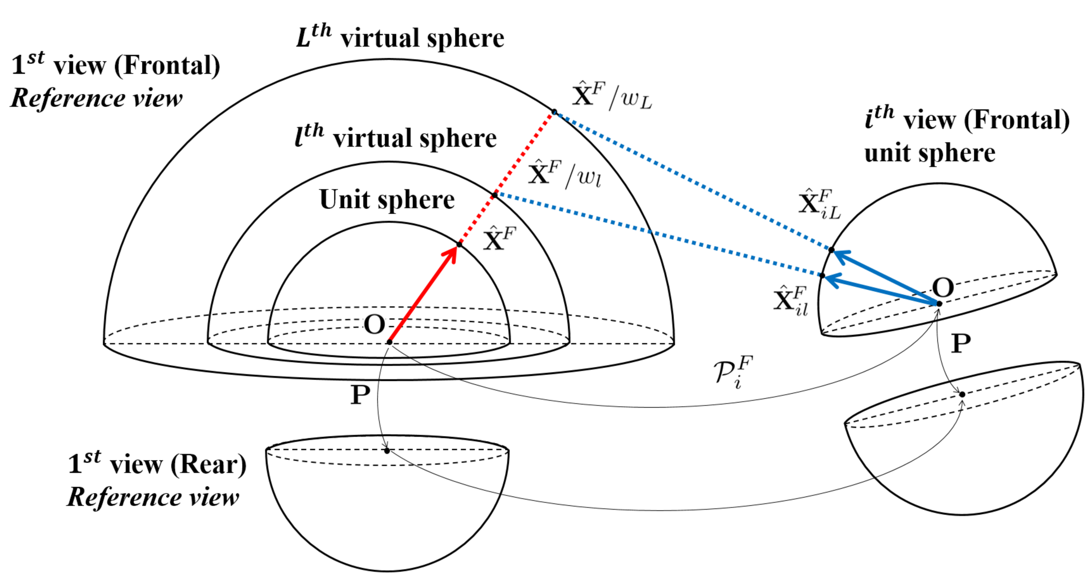
✓ **Omnidirectional camera model**

Sphere sweeping for dense matching

✓ **Warp images via the successive virtual spheres**

- Covers all-around 3D points only with positive depth representation
- Find a label that has the highest color consistency
- Use both frontal and rear images for matching

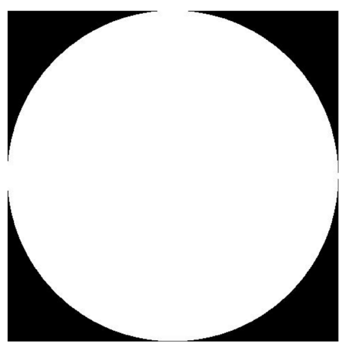
✓ **Illustration on the sphere sweeping**

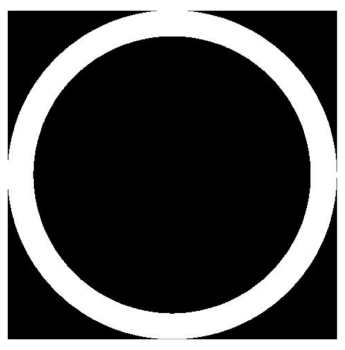


✓ **Warped image and visibility**

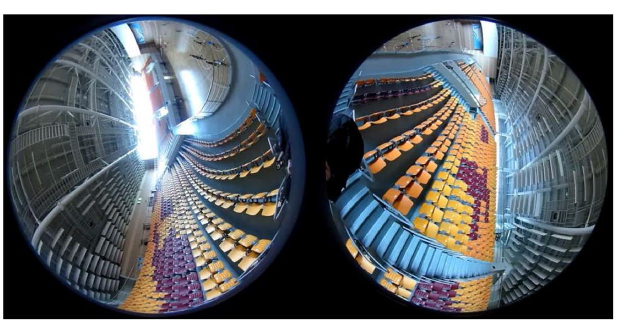

(a) F to F warping

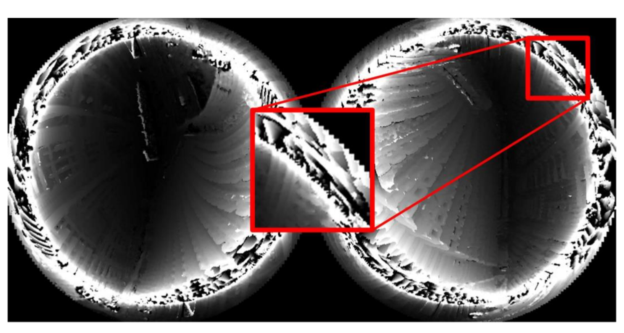

(b) R to F warping

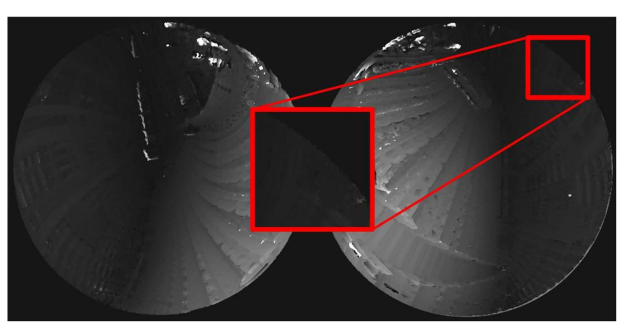

(c) Mask of (a)


(d) Mask of (b)

✓ **Comparison on dense matching method**


Input reference image

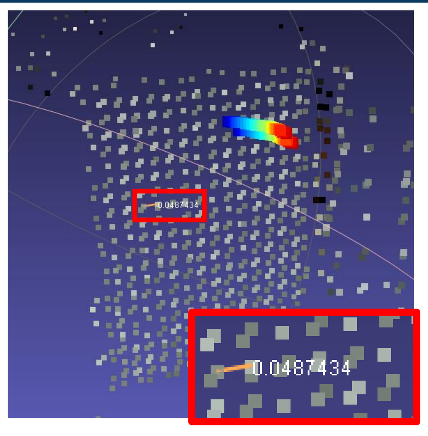

Depth from plane sweeping


Depth from dual sphere sweep

Quantitative evaluation & Qualitative result

✓ **The average reconstructed scale value**

Checkerboard size	10cm	5cm	2cm
1st trial	13.9cm	5.1cm	3.3cm
2nd trial	10.9cm	7.7cm	1.9cm
3rd trial	9.5cm	6.1cm	2.5cm

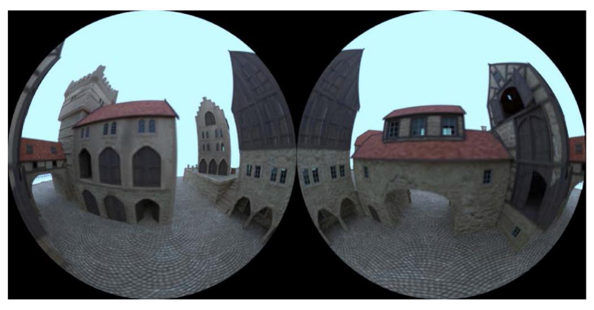


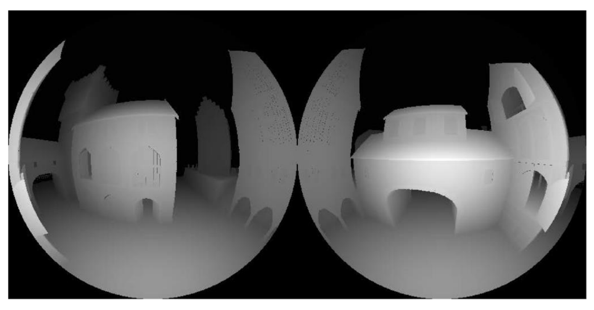
✓ **Re-projection error percentage w.r.t the number of iteration**

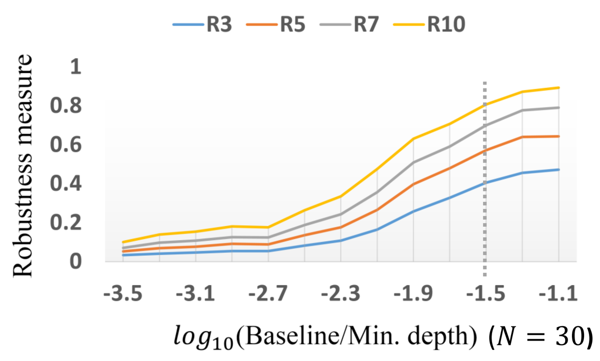
# of iteration	Initial	1	2	3	4
Proposed	**100%**	**48.7 %**	**7.9 %**	**4.4 %**	**3.8 %**
Standard	100%	74.3 %	67.8 %	64.4 %	61.6 %

Proposed: Residuals are computed **on the unit sphere**
Standard: Residuals are computed **on the image domain**

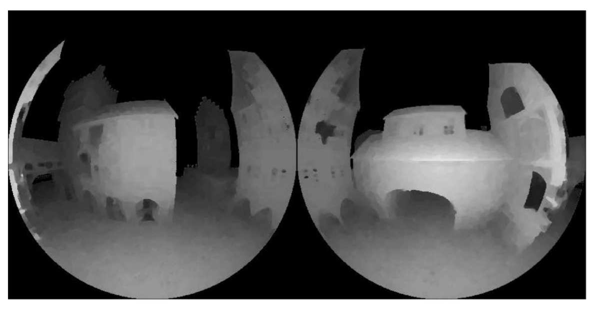
✓ **Analysis on the magnitude of baseline and the number of images**


Averaged image of input


Ground truth depth map




Error map


Our depth map

✓ **Qualitative result**

Application

✓ **Stereoscopic images for VR head-mounted display & anaglyph**

Discussion

✓ **The reconstruction is up to scale when:**

- Only pure translation or only z-axis rotation
- Zero baseline between frontal and rear camera

✓ **General rotation matrix can be adapted to bundle adjustment for spherical sensor**

- Only small angle approximated rotation matrix is working for pin-hole camera [1-3].
- The proposed bundle adjustment may have the potential to be generalized to any type of motion.

References

[1] F. Yu and D. Gallup. 3d reconstruction from accidental motion. *CVPR* 2014

[2] S. Im, H. Ha, G. Choe, H.-G. Jeon, K. Joo, and I. S. Kweon. High quality structure from small motion for rolling shutter cameras. *ICCV* 2015

[3] H. Ha, S. Im, J. Park, H.-G. Jeon, I. S. Kweon. High-quality depth from uncalibrated small motion clip. *CVPR 2016*

[4] N. Joshi and C. L. Zitnick. Micro-baseline stereo. Technical Report MSR-TR-2014-73, Microsoft Research, 2014