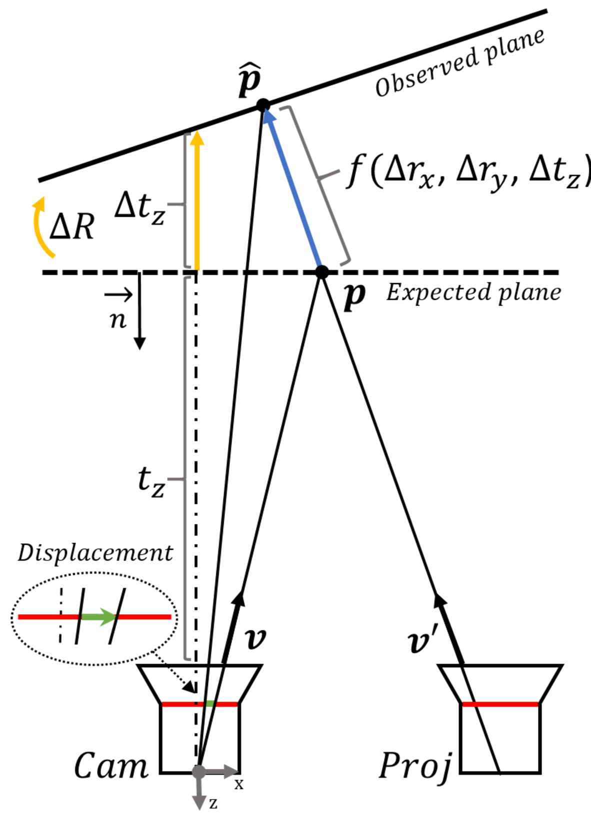
Fraunhofer Heinrich-Hertz-Institut
Vision and Imaging Technologies Department
*Computer Vision & Graphics Group*

Visual Computing Group
*Department of Computer Science*

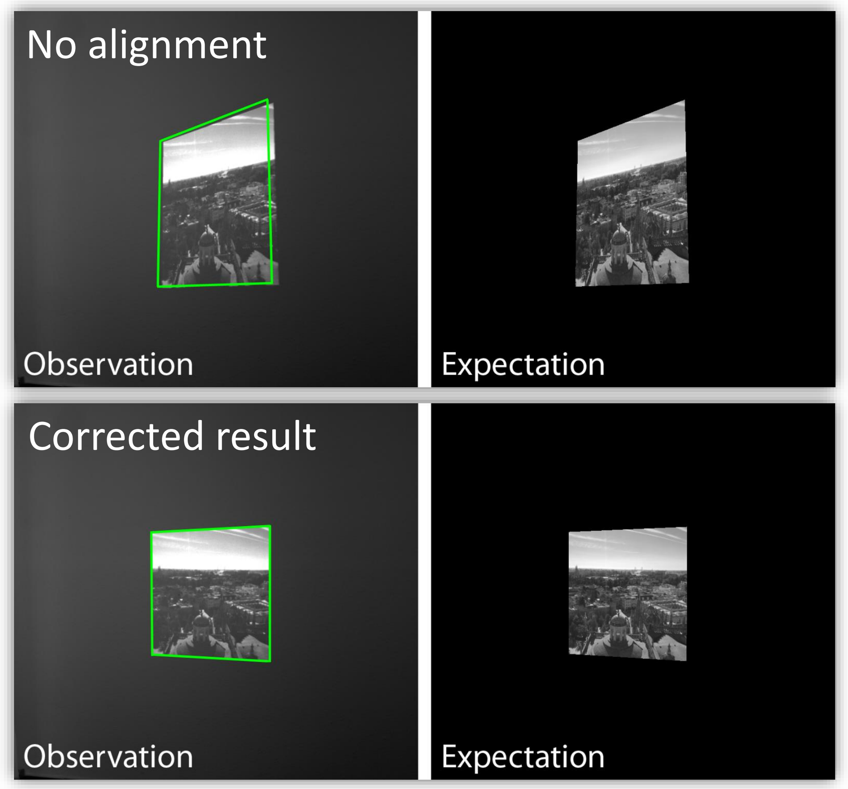
# Markerless Closed-Loop Projection Plane Tracking for Mobile Projector-Camera Systems

Niklas Gard, Peter Eisert

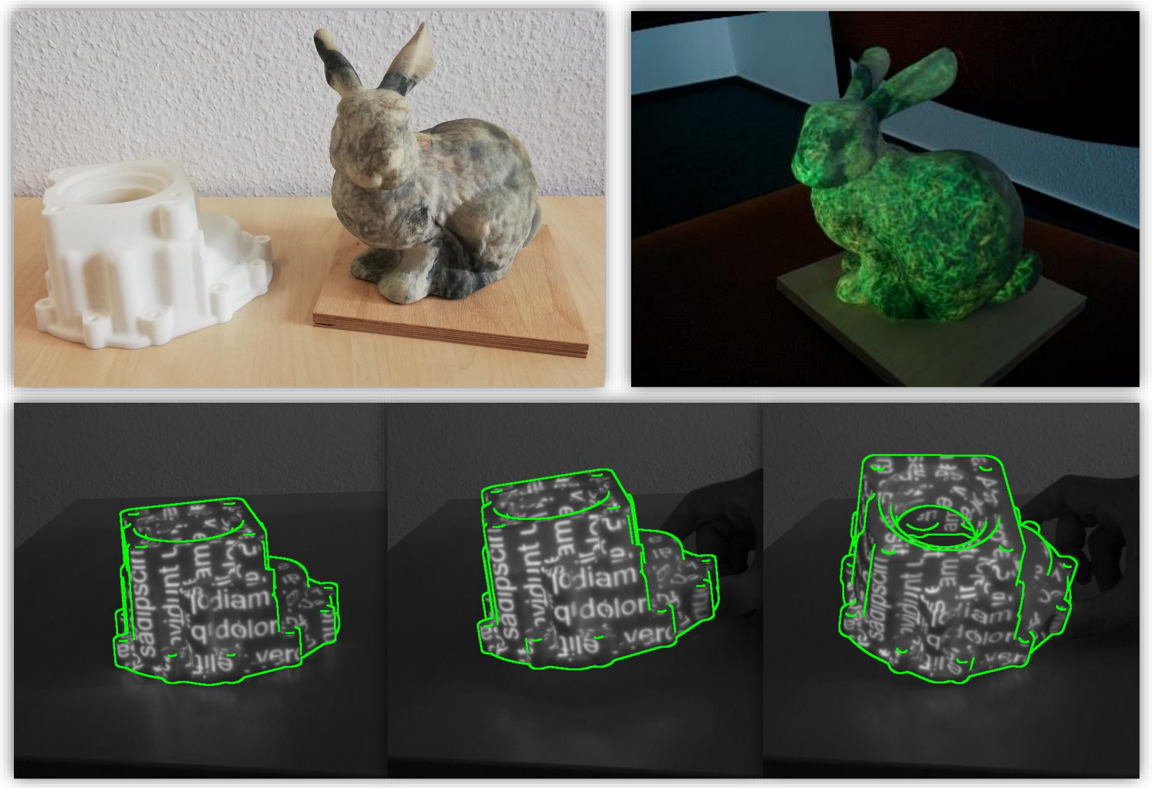
Intrinsically and extrinsically calibrated projector-camera system.




Geometric model: the movement of the projection plane is estimated from image displacements.




A distortion of the projection is visible, if the camera image (observation) and the rendered expectation mismatch. It is removed by aligning both images.



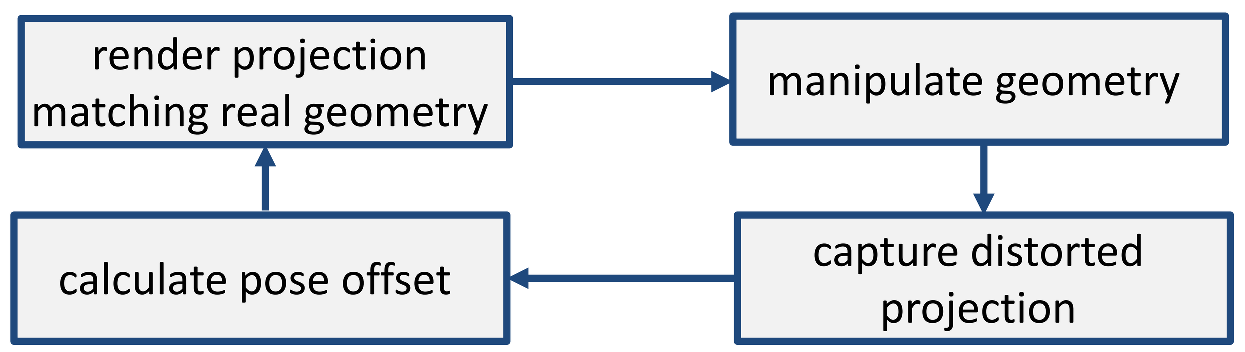
Extension of approach to 6d tracking of complex objects.

## Motivation

- Augmented reality assistance without glasses is enabled by a combined system of a projector and a camera.
- Geometrically undistorted projections on arbitrary geometries can be created by compensating captured distortions of the projected content.
- Scene information is acquired by determining the geometric relation between projected image and captured image.
- Modern small and mobile projectors allow sharp projection without focus adjustment.

## Closed-Loop Tracking

- Self correction mechanism:




- Fixes **metric size**, **aspect ratio** and **angles of projection** while projection plane or projector is moving.
- Three degrees of freedom for planar projection plane: normal vector $\boldsymbol{n}$ and distance $\boldsymbol{t} = [0, 0, t_z]^T$.
- Linearized rotation matrix $\Delta\boldsymbol{R}$ compensates small offsets between frames.

## Projector-based Optical-Flow

*How to manipulate the pose in the synthetic image to match with the captured image?*

- Estimate optical-flow between camera image $I$ and synthetically rendered image $\hat{I}$ → **analysis-by-synthesis** approach.
- $\boldsymbol{p}$ moves along viewing ray $\boldsymbol{v}'$ of projector if plane moves:

$$\hat{\boldsymbol{p}} = \boldsymbol{p} + \boldsymbol{v}' \cdot f(\Delta t_z, \Delta r_x, \Delta r_y)$$

- $f$ follows from line-plane intersection:

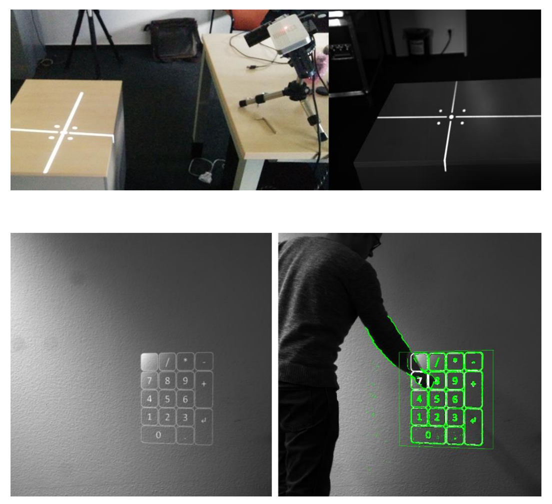
$$f(\Delta t_z, \Delta r_x, \Delta r_y) = \frac{(\Delta \boldsymbol{t} - (\boldsymbol{p} - \boldsymbol{t}))^T \Delta \boldsymbol{R} \boldsymbol{n}}{\boldsymbol{v}'^T \Delta \boldsymbol{R} \boldsymbol{n}} \qquad \Delta\boldsymbol{R} = \begin{pmatrix} 1 & 0 & \Delta r_y \\ 0 & 1 & -\Delta r_x \\ -\Delta r_y & \Delta r_x & 1 \end{pmatrix}$$

- Approximate movement of $\hat{\boldsymbol{p}}$ in camera image space with first order Taylor expansion.
- Reshape optical-flow equation to solve for motion offset:

$$\frac{\partial \hat{I}}{\partial X} u_m + \frac{\partial \hat{I}}{\partial Y} v_m \approx a_0\, \Delta t_z + a_1 \Delta r_x + a_2 \Delta r_y \approx \hat{I} - I$$

### Implementation Details

- Initialization by projection and detection of known pattern
- Flow estimation on binarized Sobel images with locally adaptive thresholds
- Iteratively reweighted least squares scheme with Charbonnier penalty



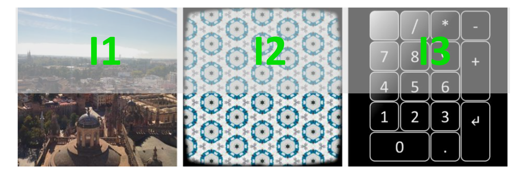
## Results

### Synthetic data

- Simulation of projection with side length of 50 cm on moving plane with distance between 0.9m and 1.8 m.

| | I 1 | | | I 2 | | | I 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Test | $r_x$ | $r_y$ | $t_z$ | $r_x$ | $r_y$ | $t_z$ | $r_x$ | $r_y$ | $t_z$ |
| **T 1** | 0.27 | 0.50 | 0.20 | 0.08 | 0.15 | 0.08 | 0.30 | 0.23 | 0.09 |
| **T 2** | 0.21 | 0.13 | 0.13 | 0.06 | 0.14 | 0.08 | 0.21 | 0.18 | 0.08 |
| **T 3** | 0.35 | 0.15 | 0.20 | 0.06 | 0.11 | 0.10 | 0.22 | 0.20 | 0.09 |
| **T 4** | 0.80 | 0.13 | 0.39 | 0.40 | 0.32 | 0.07 | 0.19 | 0.21 | 0.08 |
| **T 5** | 2.67 | 4.04 | 0.80 | 0.43 | 0.80 | 0.30 | 1.57 | 6.95 | 0.85 |
| **T 6** | 0.57 | 0.40 | 0.20 | 0.34 | 0.19 | 0.26 | 0.34 | 0.21 | 0.12 |

Mean error over sequence given in cm ($t_z$) or deg. ($r_x, r_y$)




**T1**: Small movement (± 1°, 1cm)
**T2**: Med. movement (± 5°, 5cm)
**T3**: Large movement (± 10°, 10cm)
**T4**: T2 + ambient lighting
**T5**: T2 + Gaussian blur (7x7)
**T6**: T2 + occlusion

### Real data

- Live tracking runs with about 10 fps.
- Locally adaptive binarization compensates illumination variations.
- Dynamic content tracking is possible (e.g. the content of a webbrowser).

## Outlook and Conclusion

- General model can be extended to solve similar problems.
  - Estimation of 6d pose of a more complex object
  - Optimization of the extrinsic parameters of the system assuming a known object pose in camera space
  - Estimation of lighting properties
- Parallizable on GPU due to pixel based calculation.
- Higher frame rates need advanced synchronization.

Funded by
Bundesministerium für Bildung und Forschung