enjoy Vienna

EUROGRAPHICS 2021

Rendering and Visualization in Mixed Reality

Markus Tatzgern, Shohei Mori, Christoph Ebner, David Mandl, Kasper Ladefoged, Peter Mohr and Denis Kalkofen

vrvis zentrum für virtual reality und visualisierung forschungs-gmbh

TU Graz

Keio University

Aalborg Universitet

FH Salzburg

Mixed Reality

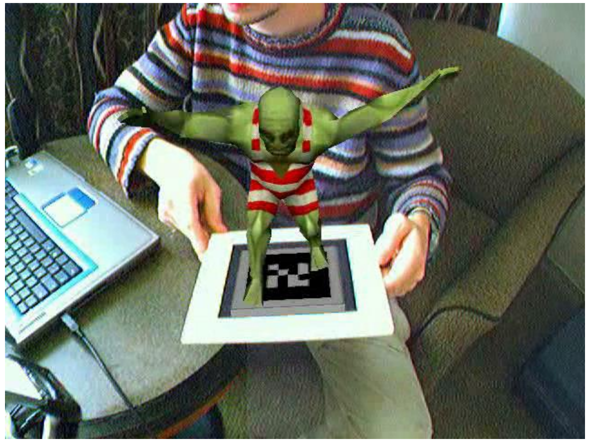
Real

+

Virtual

=

Real + Virtual



Syllabus

Part I – Visually Coherent Mixed Reality

- Light Estimation and Camera Simulation (David Mandl)
- Material Estimation (Kasper Ladefoged)
- Diminished Reality (Shohei Mori)

Part II – Dynamic Mixed Reality

- Perceptual issues (Markus Tatzgern)
- Displaying MR Environments (Christoph Ebner)
- Authoring dynamic MR Environments (Peter Mohr)

enjoy Vienna

EUROGRAPHICS 2021

Visual Coherence in Mixed Reality

David Mandl

Overview

Assume correct reconstruction & registration

What is needed for coherent rendering?

- Geometry of real scene
- Correct registration of virtual scene
- Light&Shadows
- Camera effects
- Material

Mixed Reality

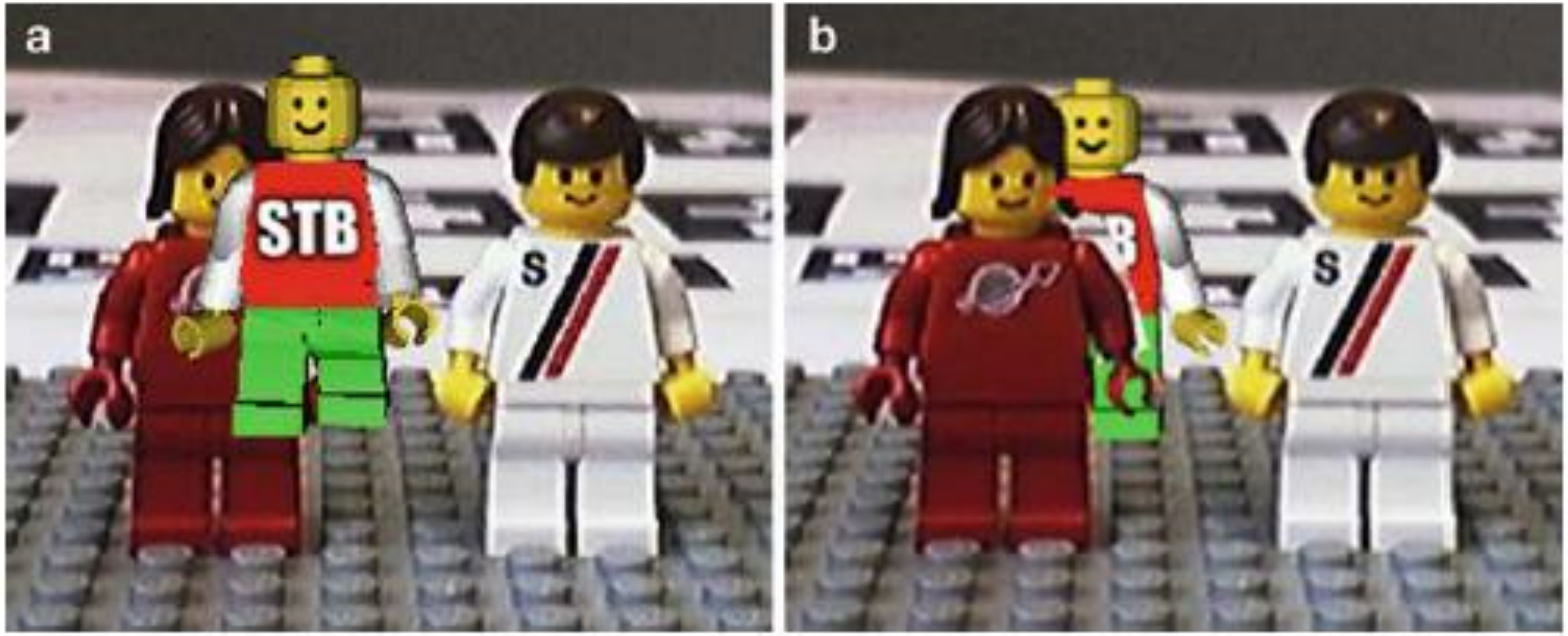
Camera Registration

- Extrinsic parameter
- Perspective camera

Registered Cameras

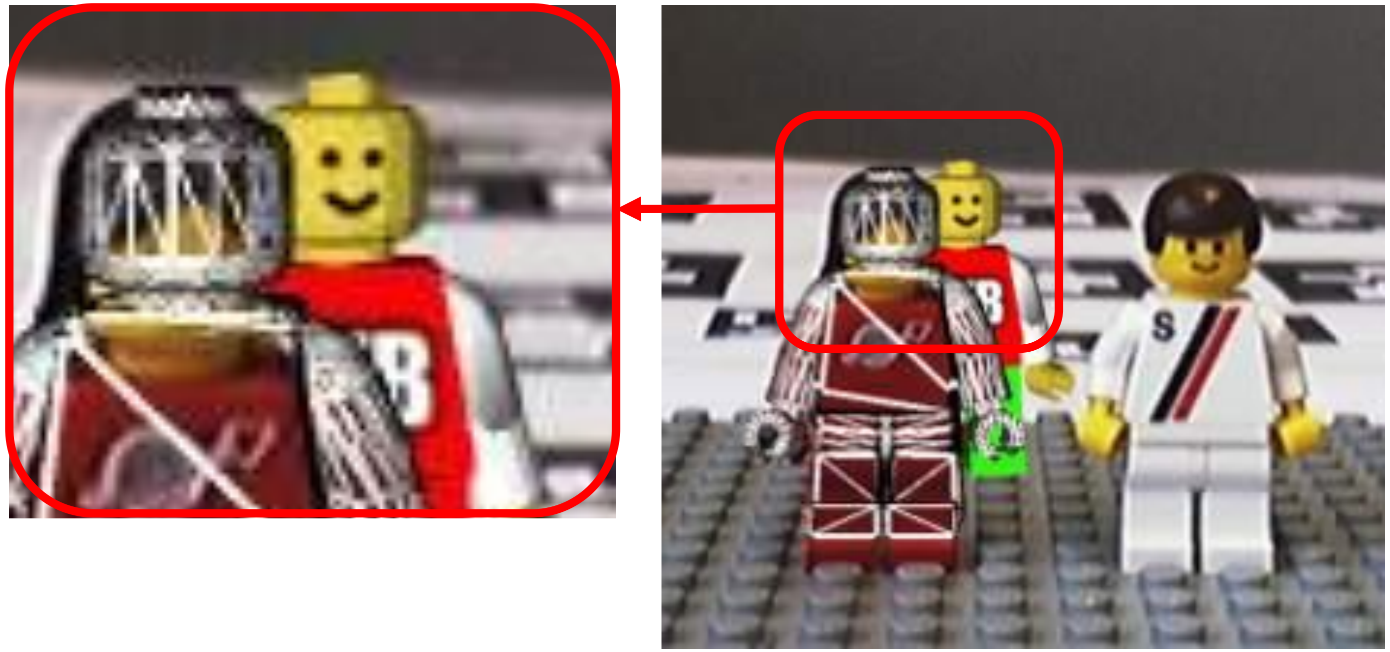
Occlusions

Occlusion



Occlusion handling

- Need model of the real object (Phantom Object)



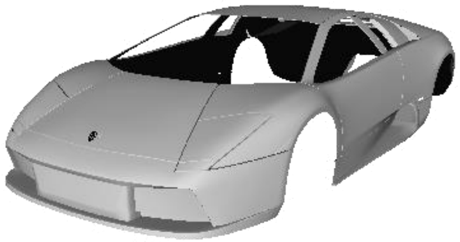
Occlusion

- Requires model of the environment

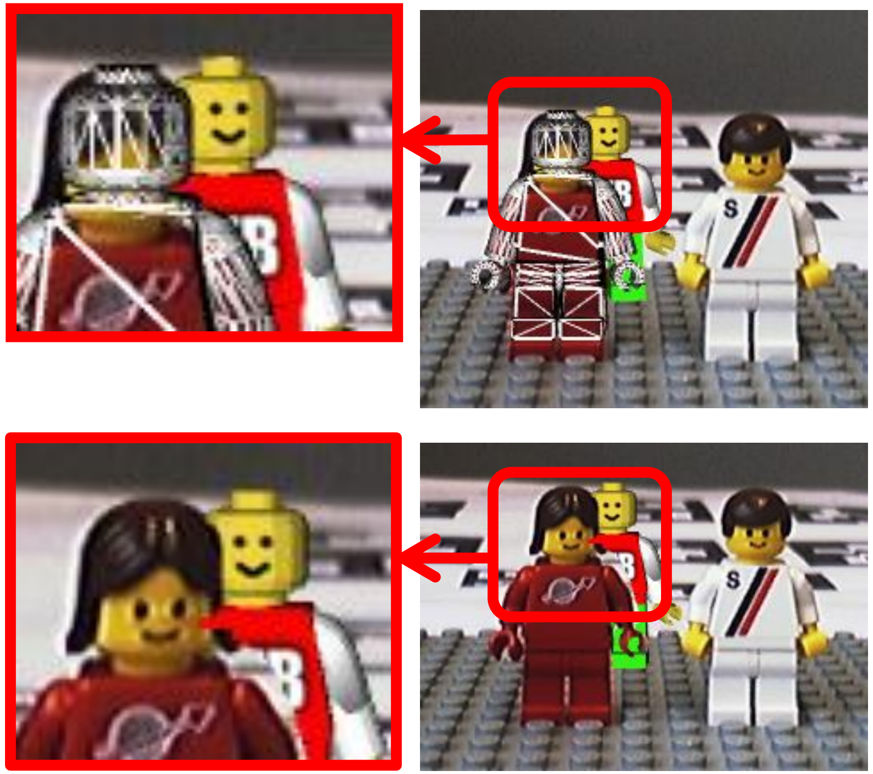
Phantom Rendering

- Render registered virtual representations (Phantoms) of real objects
- Occlusions handled by graphics hardware

1. Draw Video
2. Disable writing to color buffer
3. Render phantoms
   ➔ fills the depth buffer
4. Enable writing to color buffer
5. Draw virtual objects



Problems of Phantom Rendering



- Requires accurate
  - Model
  - Registration

Lighting

Most important aspect

- Full light simulation in AR is hard!
- Need all information
  - Geometry
  - Material
  - Light sources
- Many unknowns!
- Online vs Offline
- Local vs Global

How to get light information?

There are two main Categories

- Measurements
  - Light is measured using additional physical sensors in the scene
  - Measured light is applied using a physical accurate model
  - For example: Sperical light probes, 360° cameras, light sensors….
- Estimation
  - Light parameters or Lightsources are directly estimated on the input image
  - A parametric lighting model is used to render the synthetic scene
  - For example: Spherical harmonics, Parametric sun model
  - HDR Lightprobe estimation, Light position estimation, …

Measured Lighting

- Physical Lightprobes
- 360° Cameras
- Lux meters
- …

Physical Lightprobes

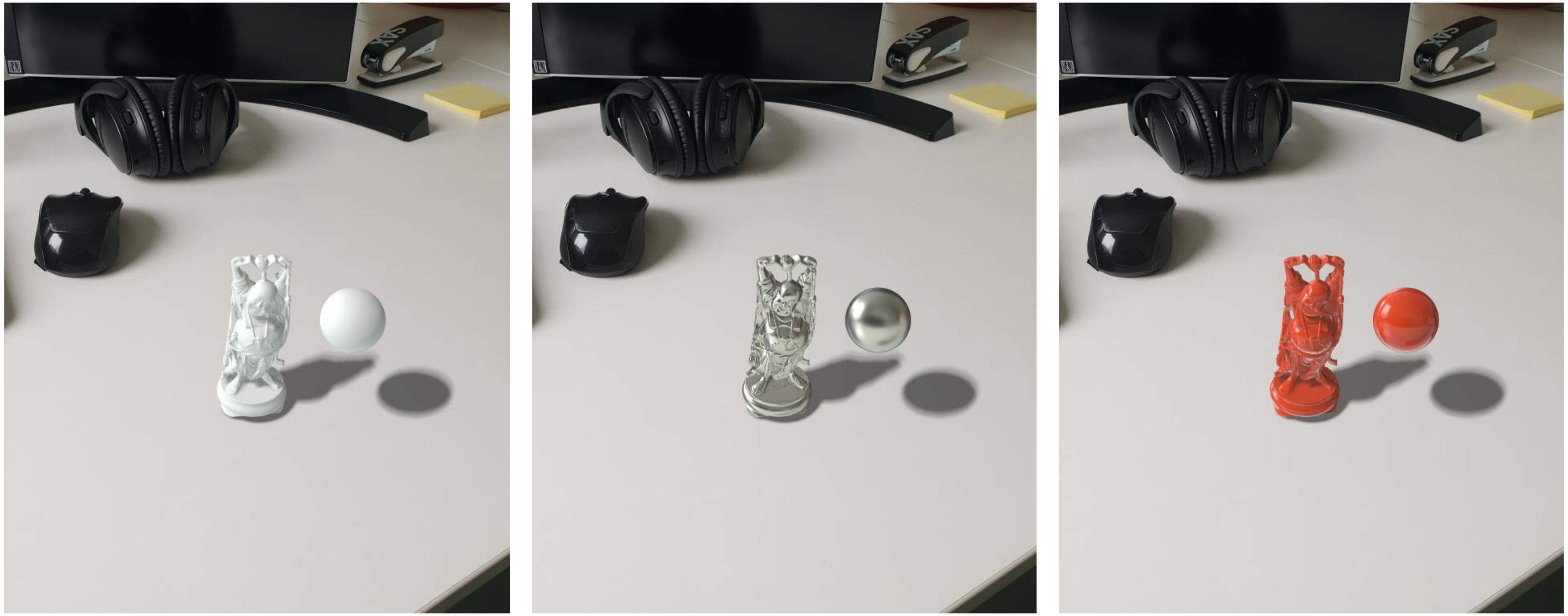
- Mirror balls [1,2]
- Capture surrounding radiance
- Use to illuminate virtual scene

360° Cameras

- Used to capture panoramic images of the scene
- Multiple cameras, image is stitched
- Usually used for image-based lighting (IBL)

Image-based lighting

- Lightprobes are directly used to shade objects [1]
- Can be used for diffuse and specular materials



High dynamic range (HDR)

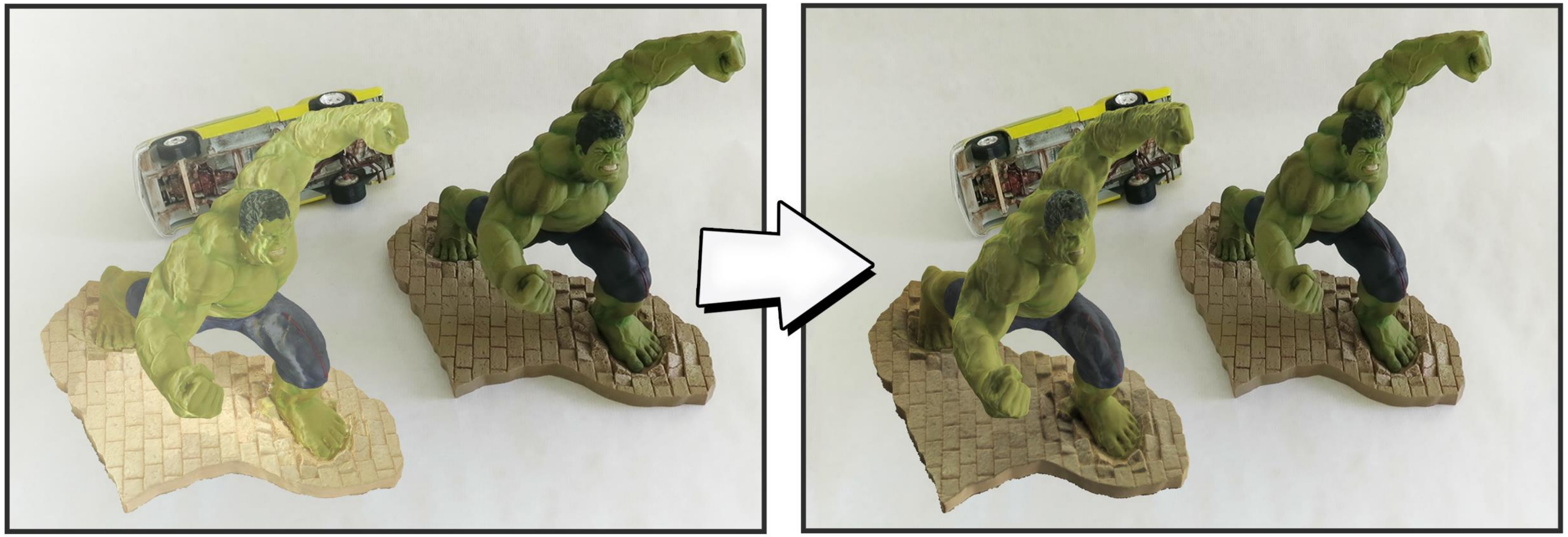
- Physical plausible lighting [3]
- HDR environment map
- Lookup incoming radiance
- LDR vs HDR
- Exposure bracketing

Estimated Lighting

- Indoor vs Outdoor
- Parametric models
- Implicit lightprobes
- Learned lightprobes
- Global vs local

Learned Lightprobes

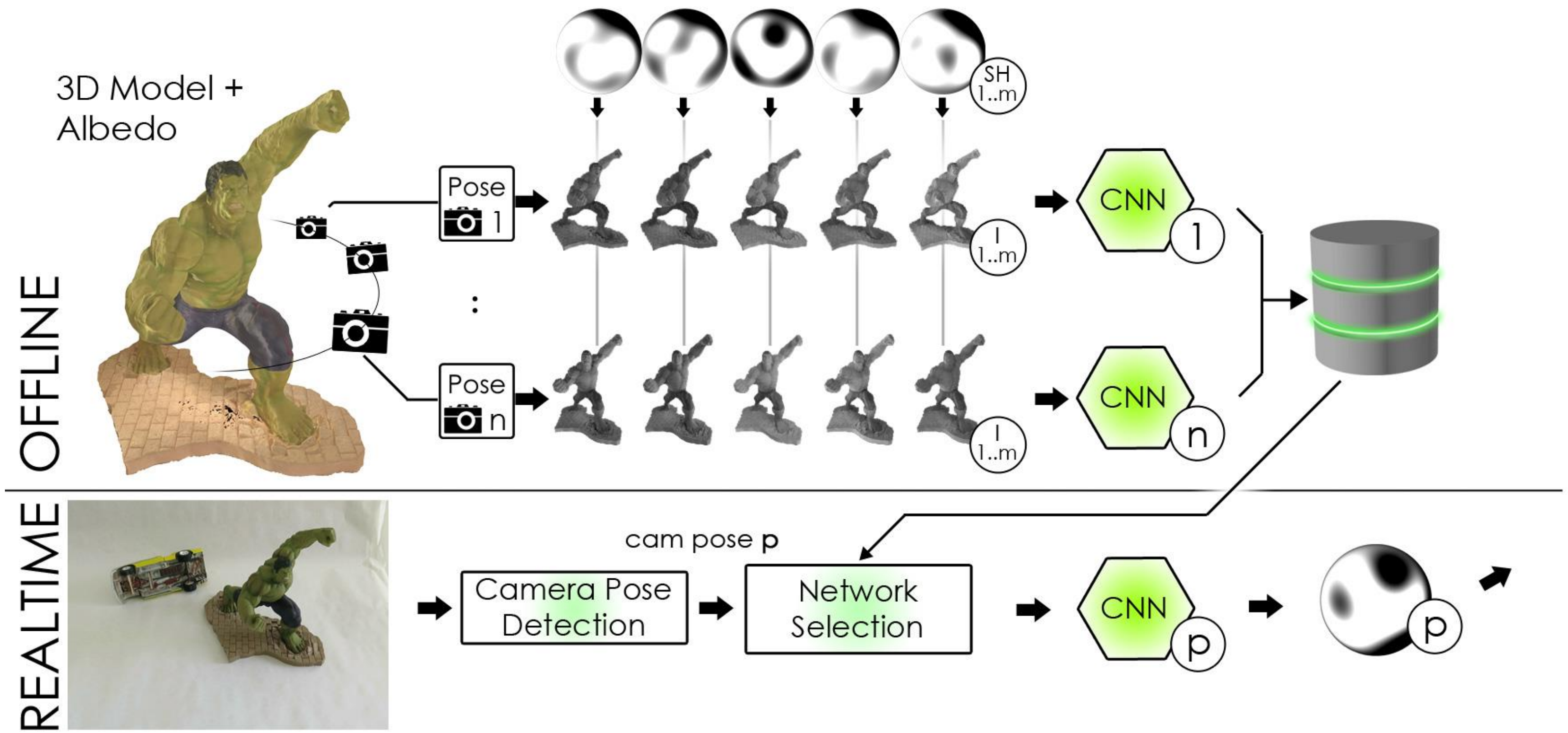
- Create database with different illuminations
- Use spherical harmonics (SH) to represent light sources & transport
- Train CNN to estimate SH coefficients on object



Spherical harmonics

- Functions defined on the surface of a sphere
- Used to approximate diffuse light transport
- Only 9 coefficients needed to represent a lightprobe!

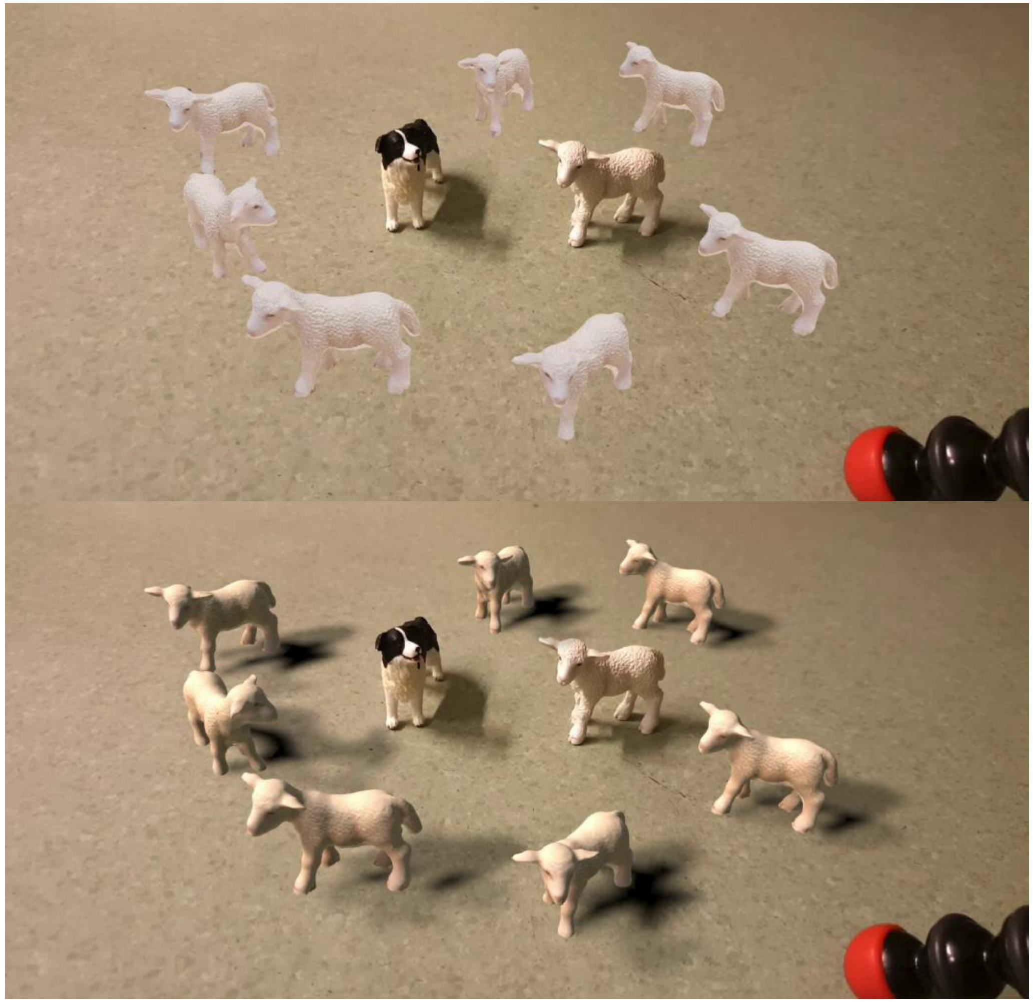
System overview



Results

Shadows

- Greatly improve sense of realism
- Need accurate light sources
- Different types of shadows in AR
  - Real-to-virtual
  - Virtual-to-real
- Need good geometry of the scene!



Direct shadows

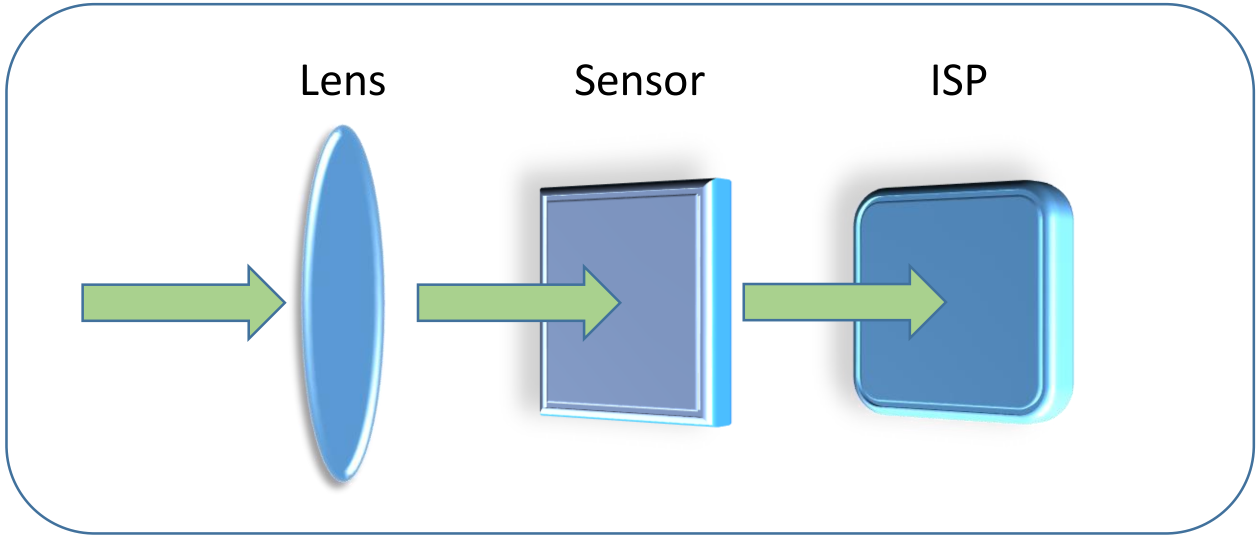
- Estimate dominant light directions in HDR panorama
- Use for shadow mapping

Differential Rendering

- Compute scene radiance with background geometry
- Difference between BG and rendering
- Apply to background
- Combine with rendering

Camera effects

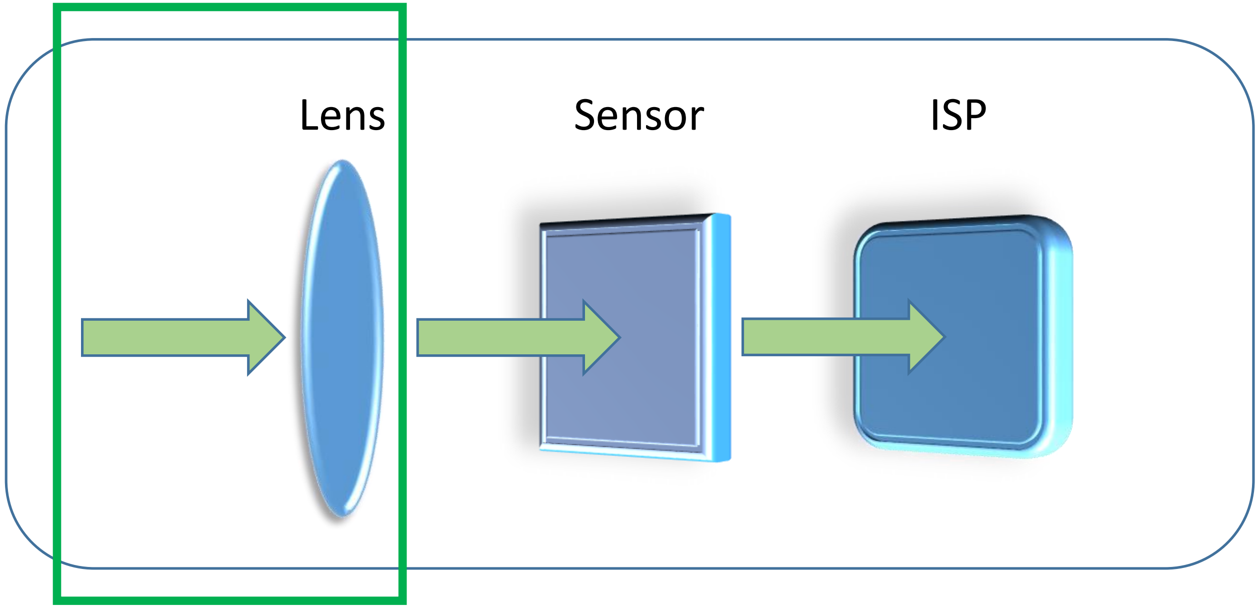
- In Video-See-Through AR there is always a camera!
- Images from a camera are never perfect



Lens effects

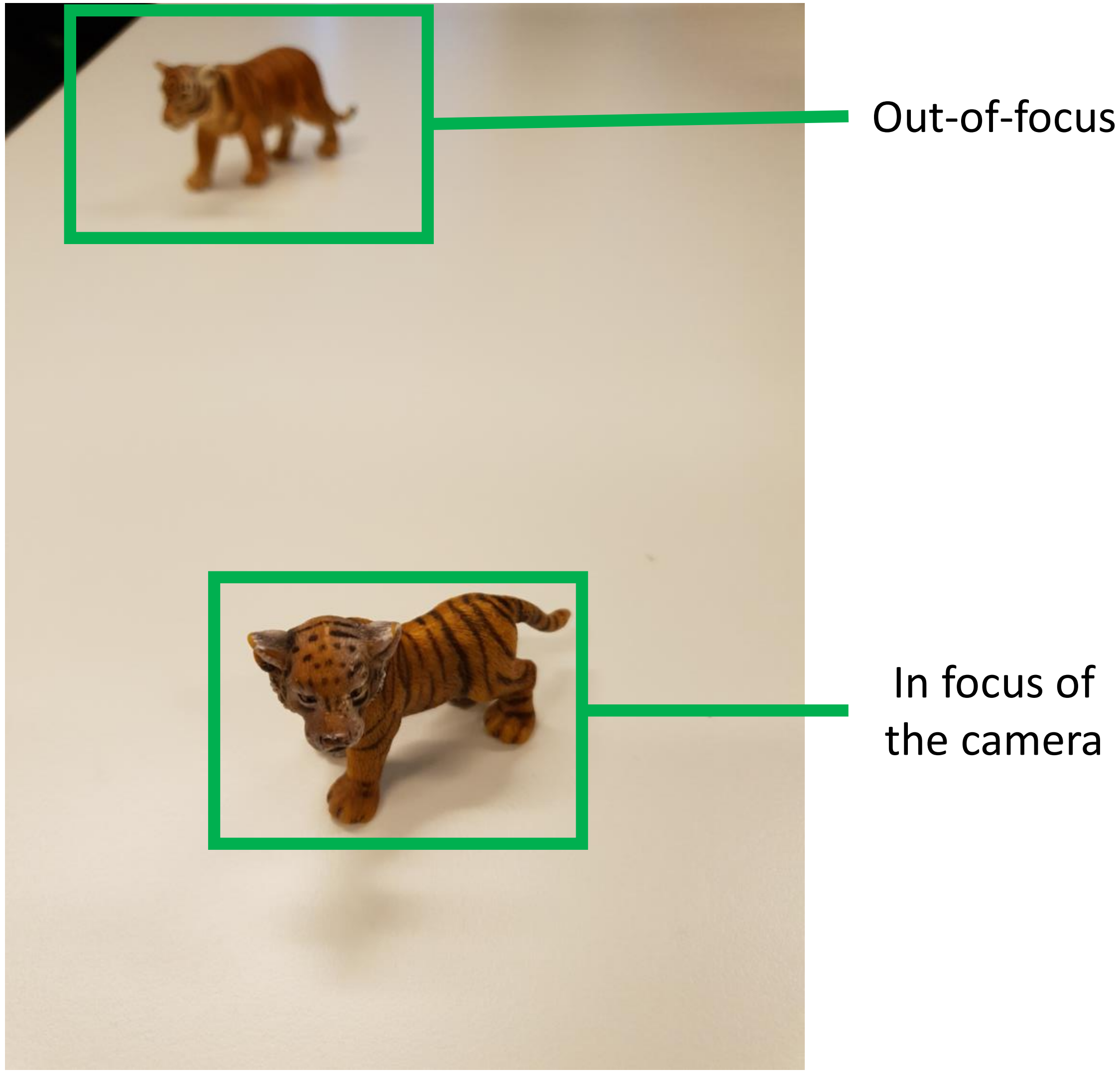
Lens system focuses incoming light onto the sensor

- Depth of Field (DoF)
- Chromatic abberation
- Lens distortion
- Lens vignetting



Depth of Field

- Objects not in the focus plane of the camera appear blurred
- Out-of-focus blur
- Rendered image are usually perfectly sharp!



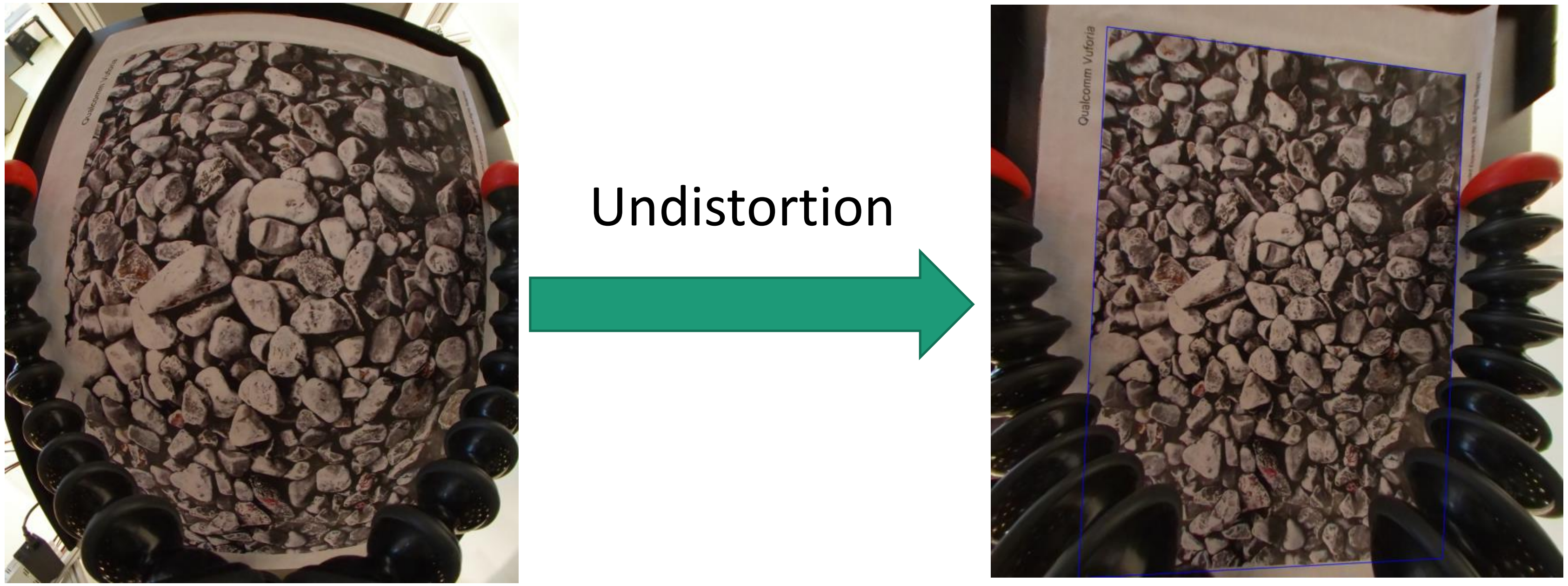
Depth of Field

Post-Process DoF

- Input, rendered RGB image + Depth
- Compute CoC per pixel
- Weighted sum of all neighbouring CoCs

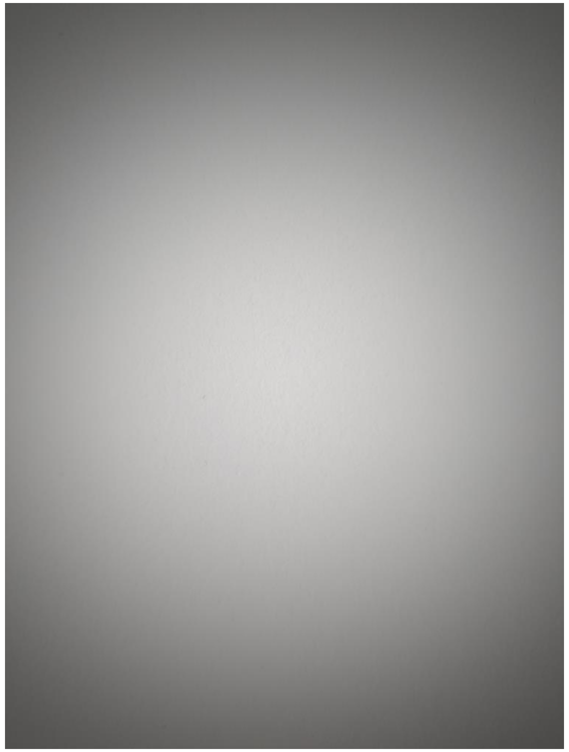
Lens distortion

- Can be measured by intrinsic camera calibration
- Distortion coefficients
- Apply to rendered image



Lens Vignetting

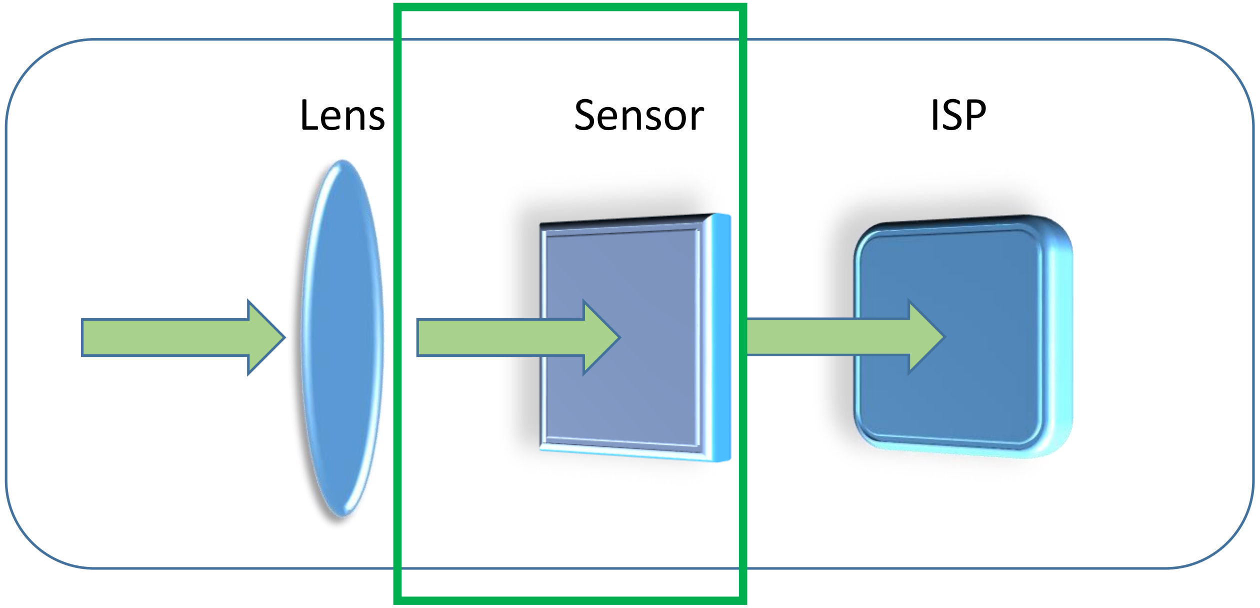
- Darkening in corner of sensor image
- Estimated by taking grayscale images
- Images of uniform white background
- Vignetting texture



Sensor

Sensor imperfections

- Noise
- Motion Blur
- Bayer artifacts

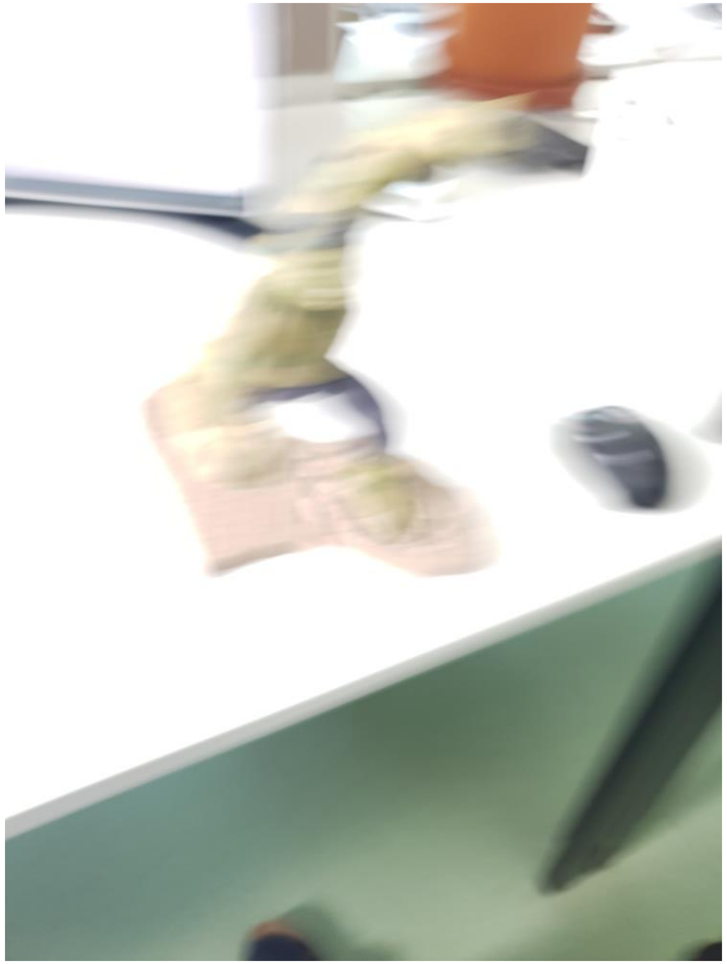


Senor Noise

- Many sources, photon shot noise, readout noise, ...
- Estimate from source images
- Apply to rendering as noise texture

Motion Blur

- Too long exposure time while camera moves
- Colors “bleed” into neighbouring pixels
- Estimate motion model, apply to rendering using directional blur filter



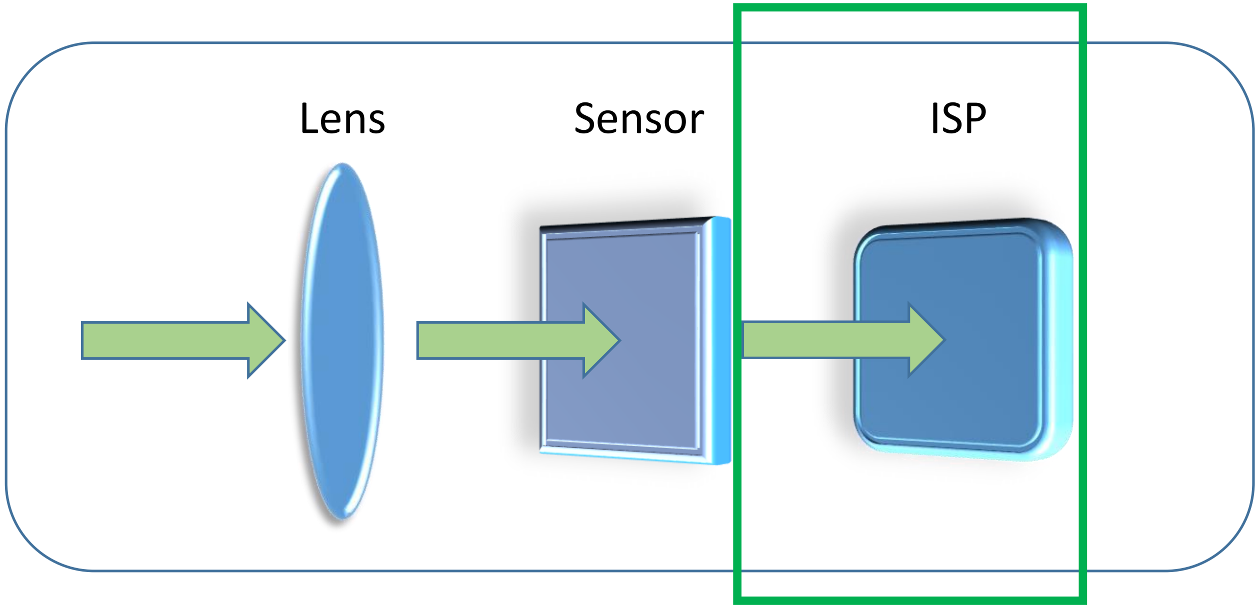
Bayer artifacts

- Come from Bayer CFA on very high frequencies in the image
- Color only covers subpixels of bayer pattern
- Can be applied by
  - Rendering RGB channels individually
  - Shifting them by the CFA pattern
  - Combine channels to image

Image Signal Processor

Post-processing to create final image

- Whitebalance
- Denoising
- Sharpening
- YUV convertion



Thank You!

References

[1] DEBEVEC P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In SIGGRAPH (1998)

[2] DEBEVEC P., TCHOU C., GARDNER A., HAWKINST., POULLIS C., STUMPFEL J., JONES A., YUN N., EINARSSON P., LUNDGREN T., FAJARDO M., MARTINEZ P.: Estimating Surface Reflectance Properties of a Complex Scene under Captured Natural Illumination. Tech. Rep. ICT-TR-06.2004, University of Southern California Institute for Creative Technologies, 2004.

[3] Kronander, J., Banterle, F., Gardner, A., Miandji, E., & Unger, J. (2015). Photorealistic rendering of mixed reality scenes. *Computer Graphics Forum*, *34*(2), 643–665

[4] G. Klein and D. W. Murray. Simulating low-cost cameras for augmented reality compositing. IEEE Transactions on Visualization and Computer Graphics, 16(3):369–380, 2010.

[5] D. Mandl, K. M. Yi, P. Mohr, P. M. Roth, P. Fua, V. Lepetit, D. Schmal-stieg, and D. Kalkofen. Learning lightprobes for mixed reality illumina-tion. InIEEE International Symposium on Mixed and Augmented Reality,pages 82–89,

[6] Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, and Jean-François Lalonde. 2017. Deep Outdoor Illumination Estimation. InIEEE Conferenceon Computer Vision and Pattern Recognition.

[7] Mathieu Garon, Kalyan Sunkavalli, Sunil Hadap, Nathan Carr,and Jean-Franc ̧ois Lalonde. Fast spatially-varying indoorlighting estimation. InProceedings of the IEEE Conferenceon Computer Vision and Pattern Recognition, pages 6908–6917, 2019

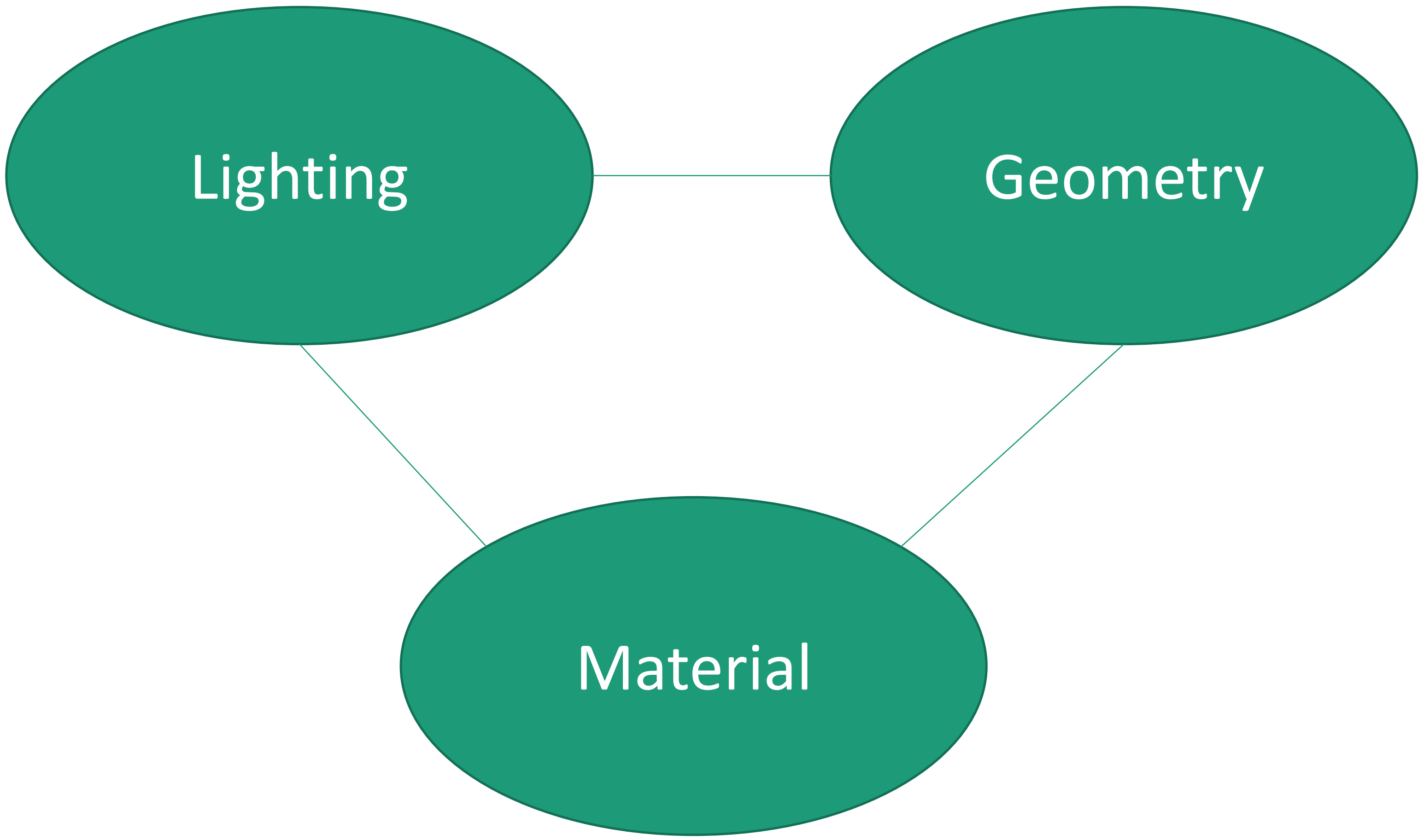
enjoy Vienna

EUROGRAPHICS 2021

Material Estimation

K. S. Ladefoged

Rendering ”Triangel”



Why estimating Materials

- Having known material can be used to estimate lighting conditions
- Digitizing real world objects
  - Re-rendering
  - Analysis (damage analysis etc.)
  - Cultural Heritage preservation
  - Many more.

Describing Materials

Types of Generalized Bidirectional Functions

- 4 Dimensions
  - BRDF
  - BTDF
- 6 Dimensions
  - BTF
  - SVBRDF
  - BSDF
- 8 Dimensions
  - BSSRDF

- Overview paper
  - Guarnera et al. 2016 [1]

Spatially Varying Bidirectional Reflectance Distribution Function

$$f_r(x, \Psi \to \Theta) = \frac{dL(x \to \Theta)}{dE(x \leftarrow \Psi)}$$

- Describes the fraction of incoming light that leaves the point x
- This is a general function
- There exits numerous models to describe BRDF of surfaces
    - Lambertian
    - Phong
    - Ward

Approaches

Measuring Equipment

- Large, one off, builds that are very hard to recreate.
- Some acquire geometry and Spatially Varying Reflectance at once,
- Others are specialized in singular reflectance.

- Some papers using this approach:
  - Köhler et al. 2013 [2]
  - Nöll et al. 2013 [3]
  - Nöll et al. 2015 [4]
  - Tunwattanapong et al. 2013 [5]
  - Chen et al. 2014 [56]

Optimization

- Minimizes some error function in relation to a given BRDF model
- Data amount is dependent on model complexity
- Usually needs to split the object into a given number of materials to have enough data for specularity estimation

- Paper:
  - Nam et al. 2018 [7]

Using Known Lighting

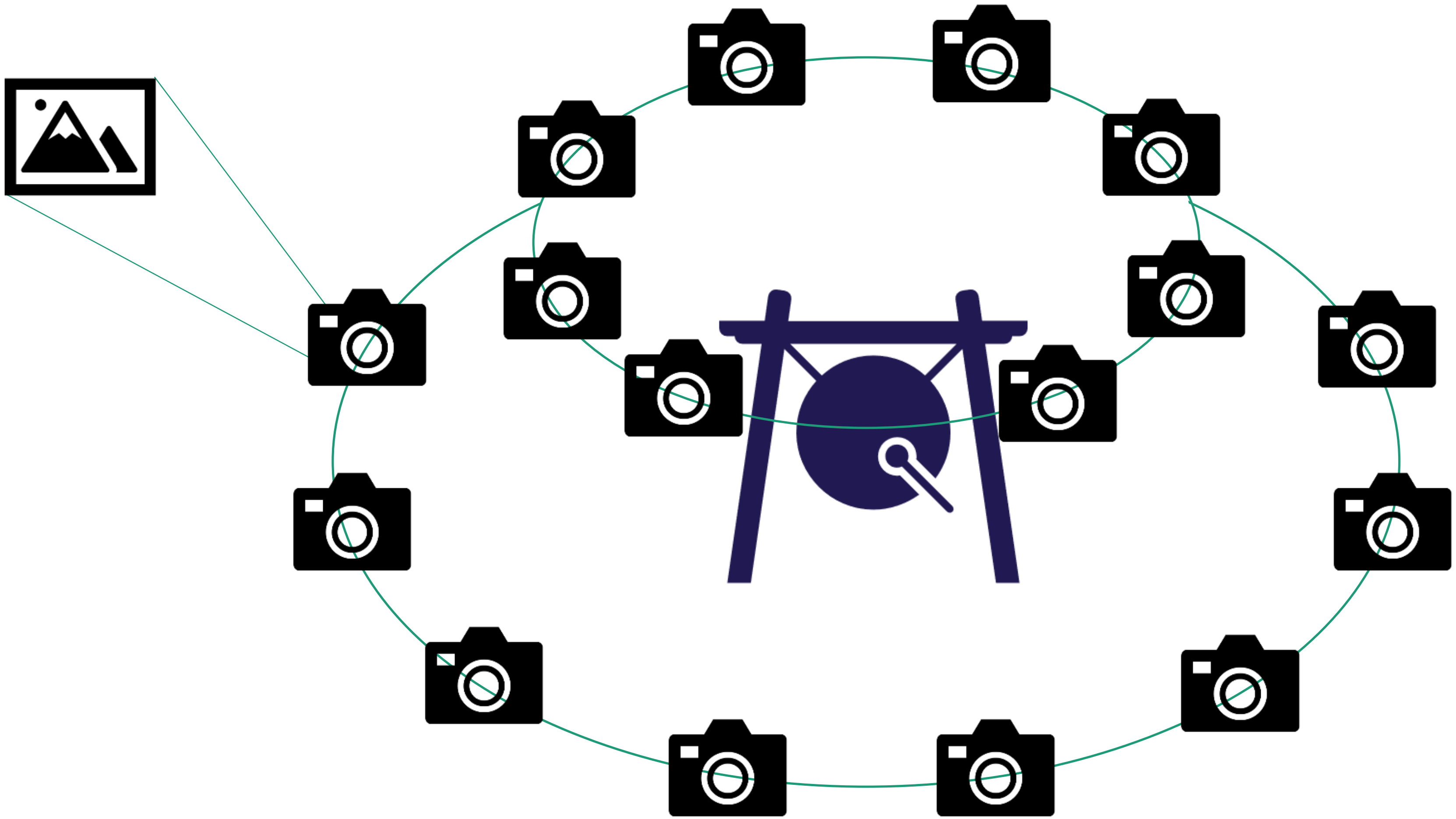
Ladefoged, K. S., & Madsen, C. B. (2020). Spatially-Varying Diffuse Reflectance Capture Using Irradiance Map Rendering for Image-Based Modeling Applications. In *2019 IEEE International Symposium on Mixed and Augmented Reality* (pp. 46-54). [8943701] IEEE Computer Society Press. https://doi.org/10.1109/ISMAR.2019.00-27 [8]

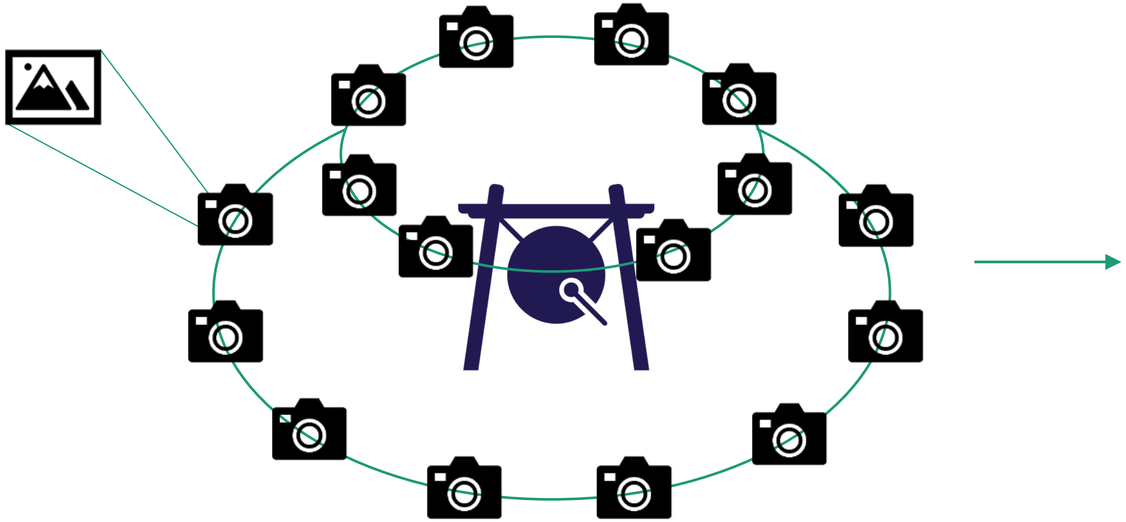
THE PROBLEM

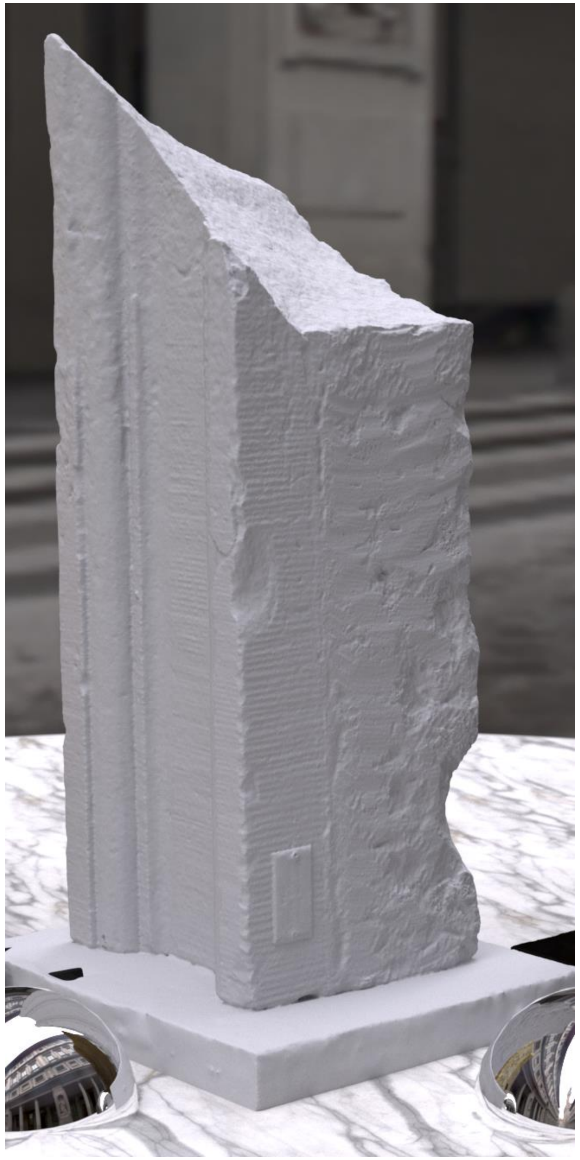
- Need for digitizing an object?
- Does Structure from Motions not produce textures that are usable?

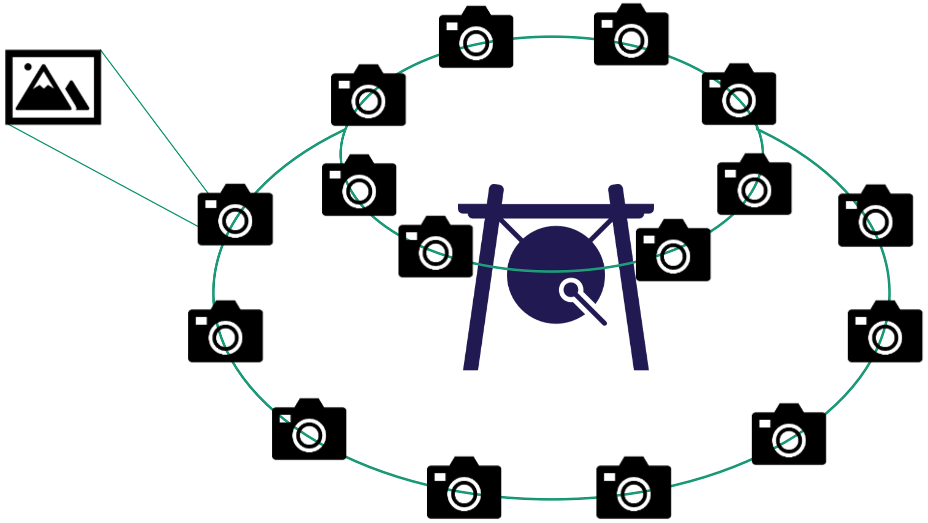
BEST CASE RECONSTRUCTION

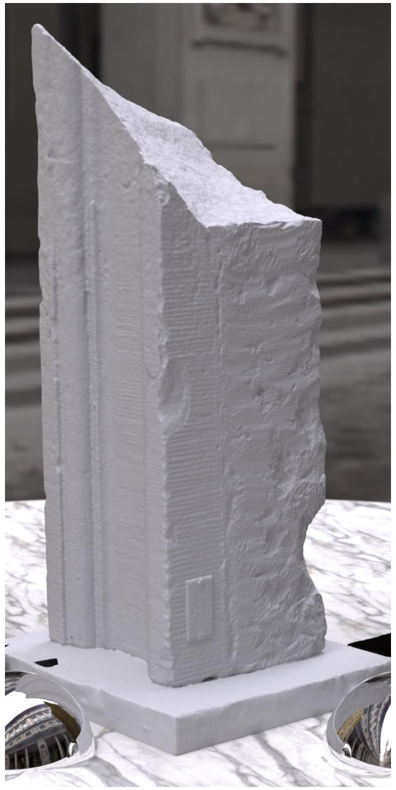
RECONSTRUCTED REFLECTANCE

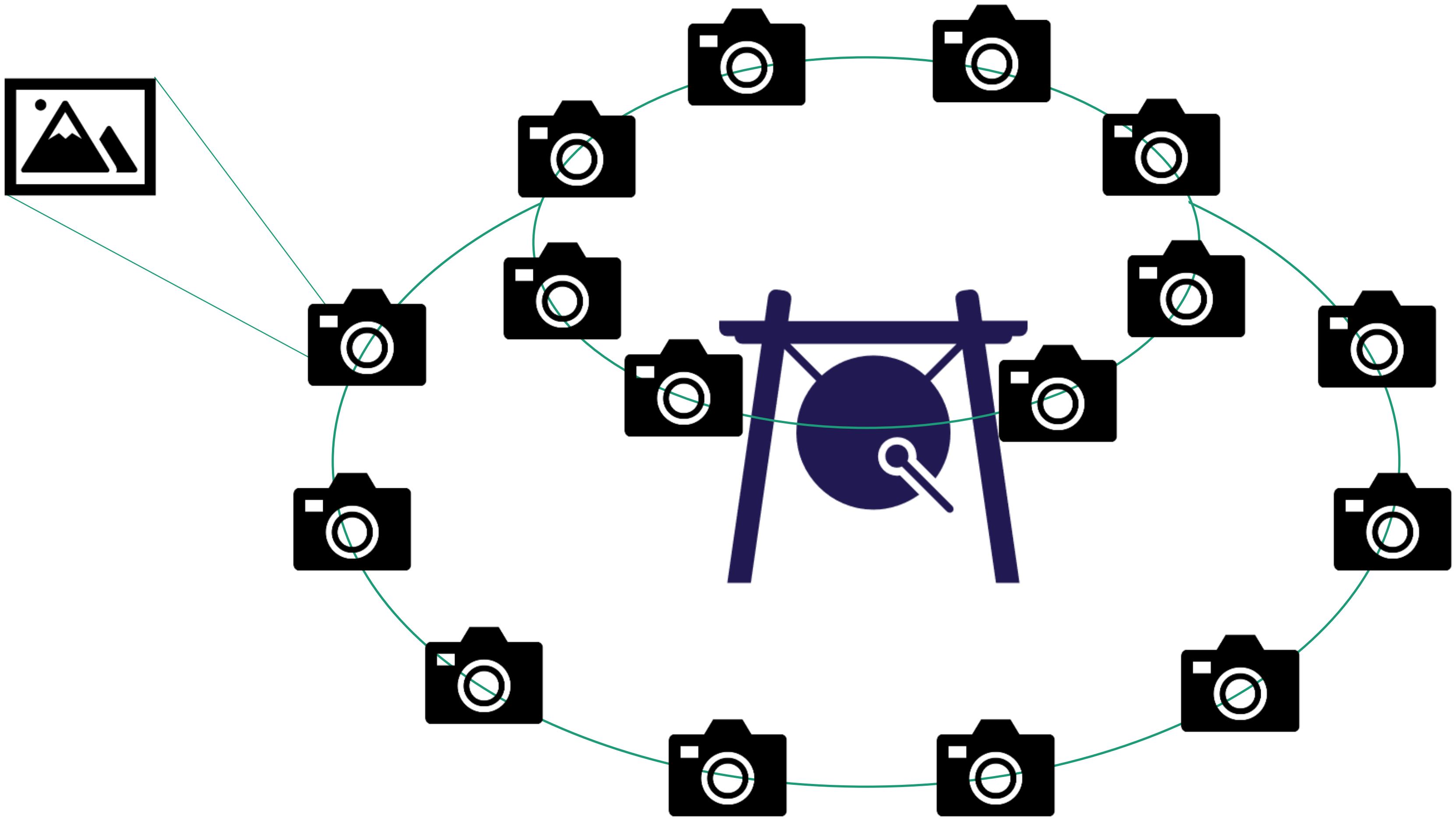


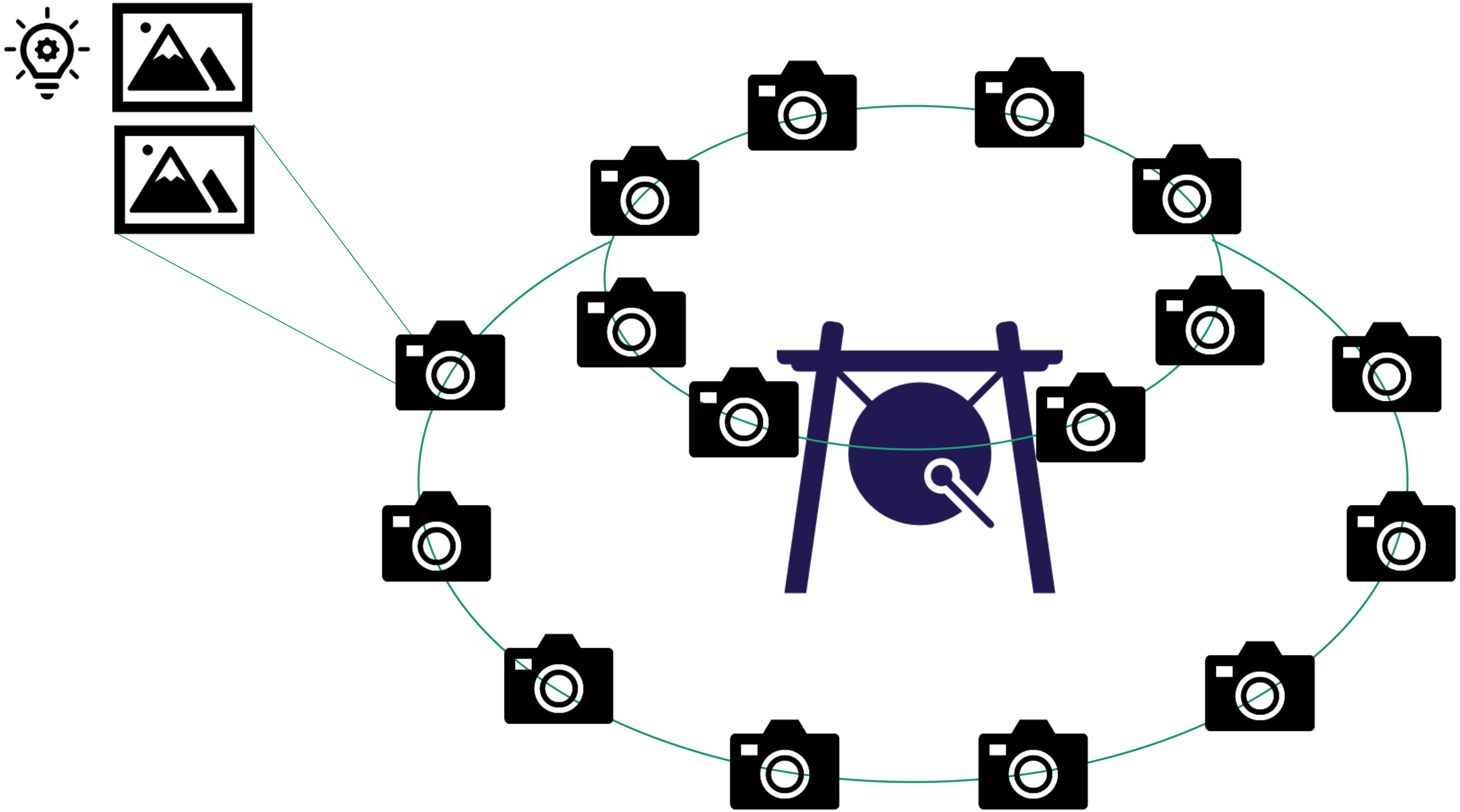


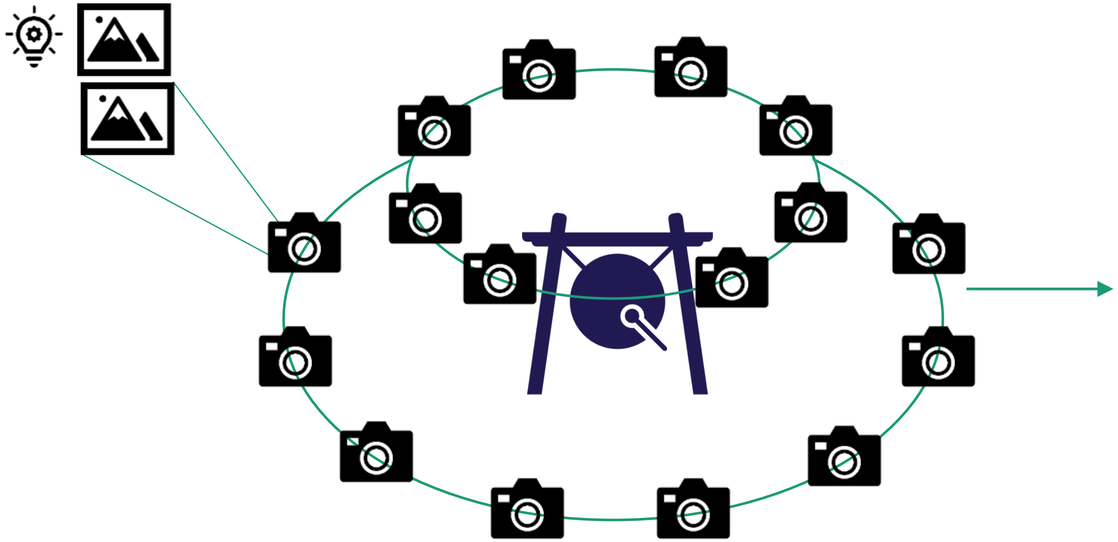


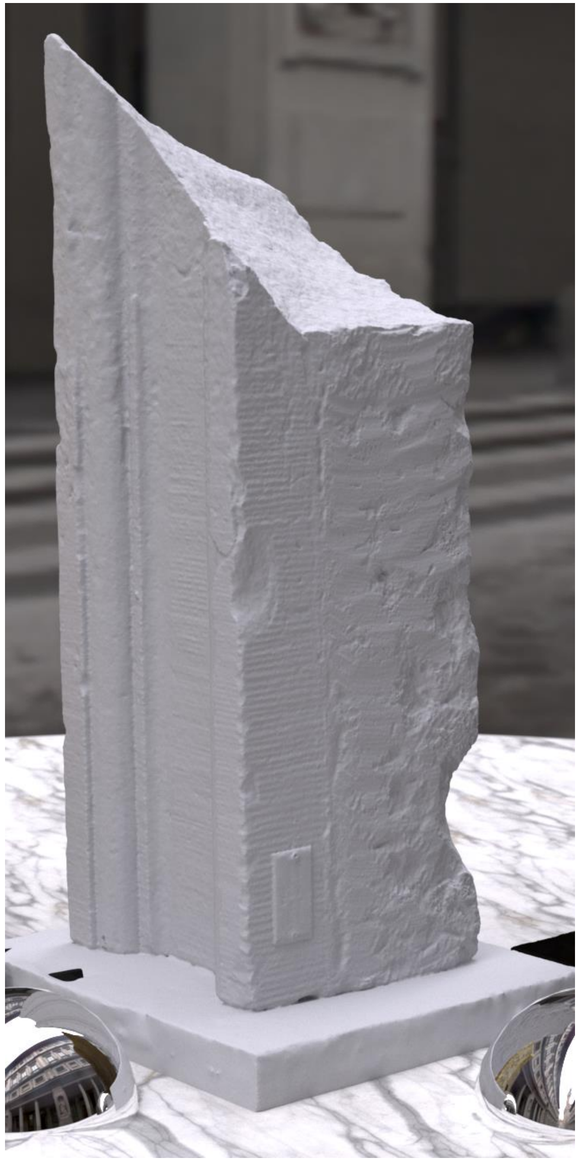


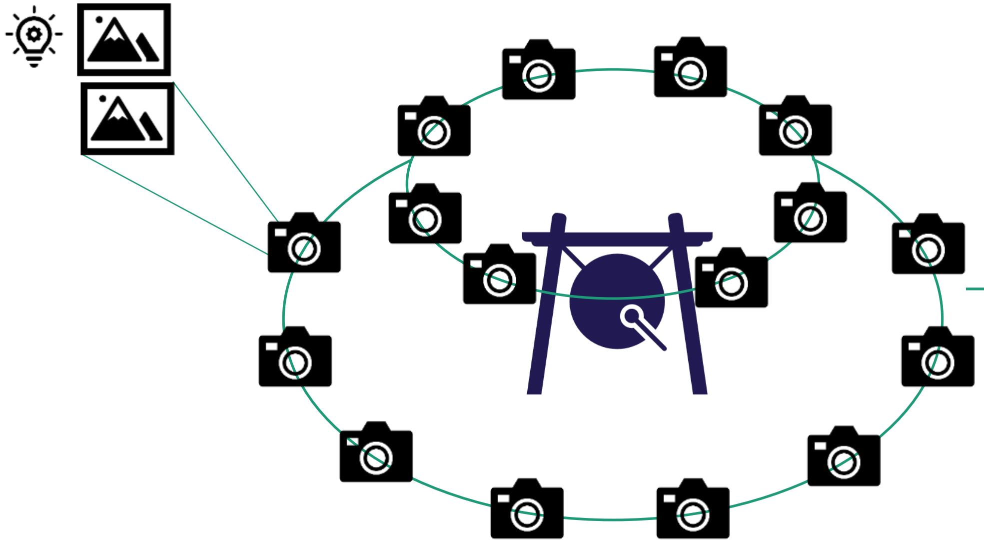


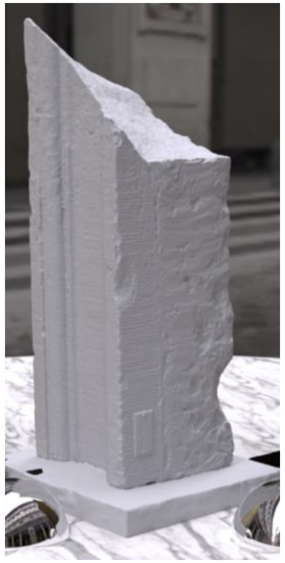


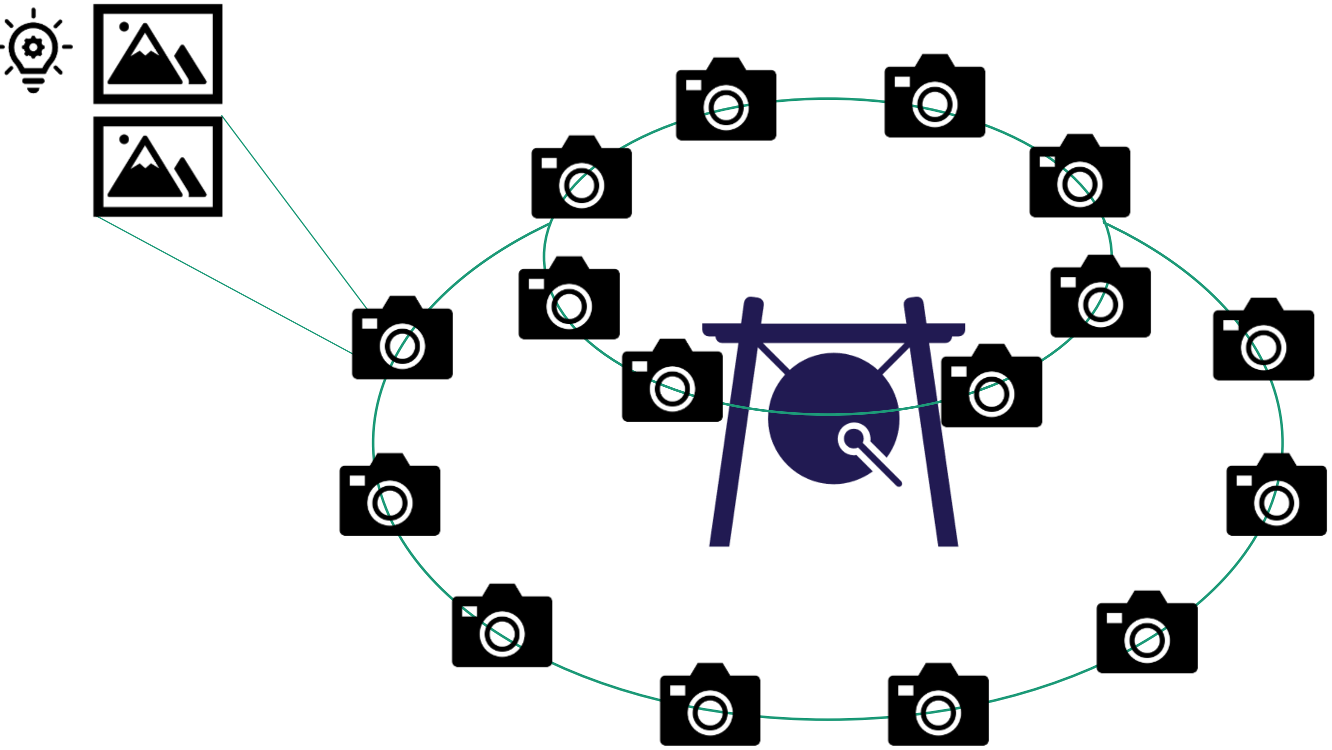


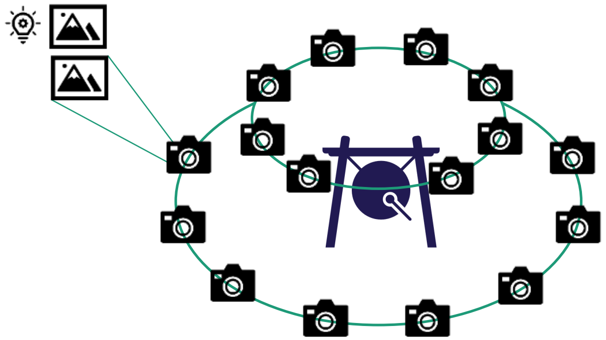


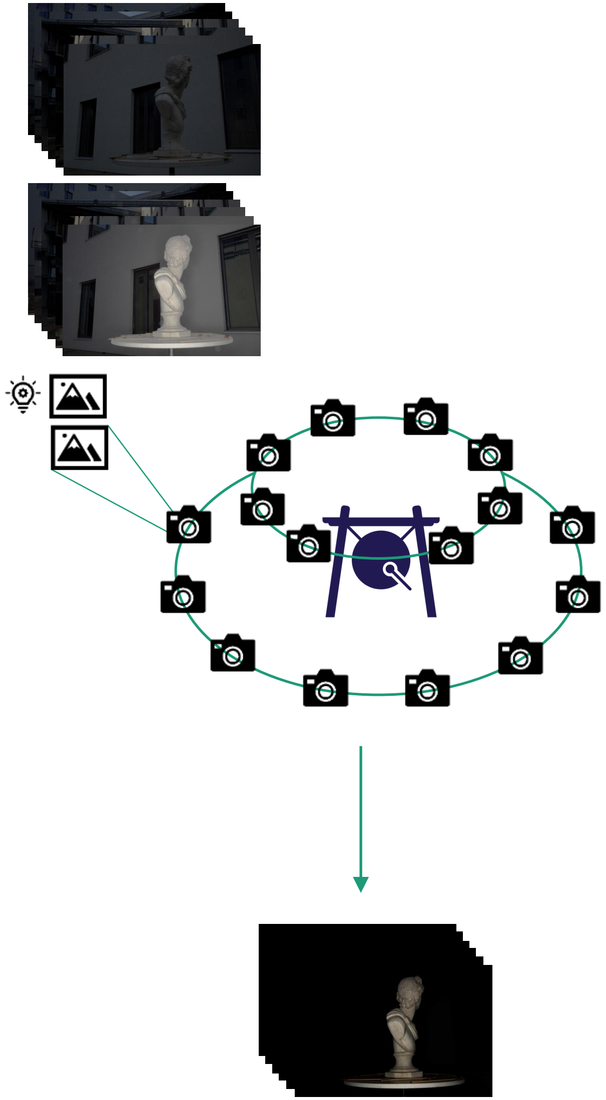


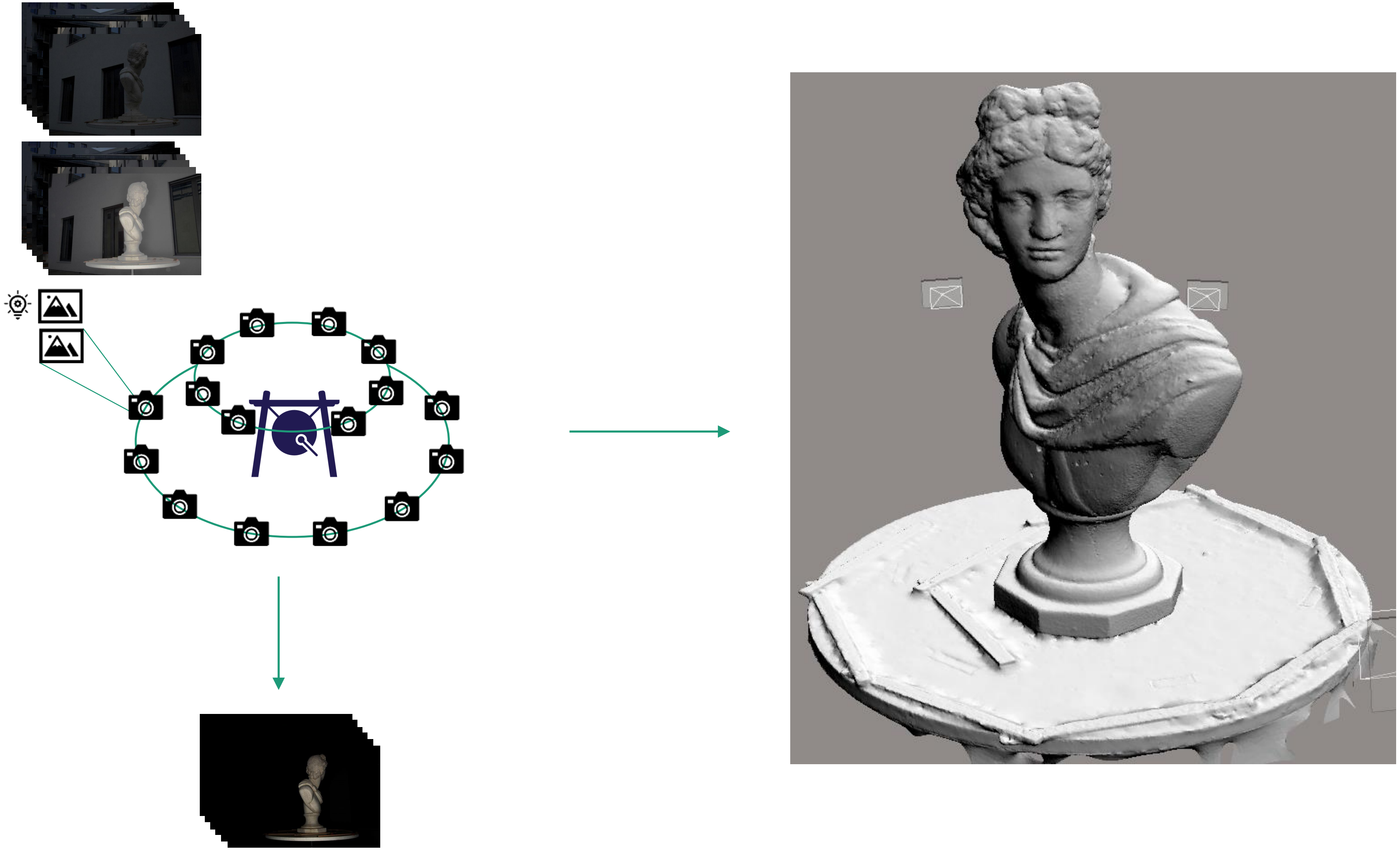


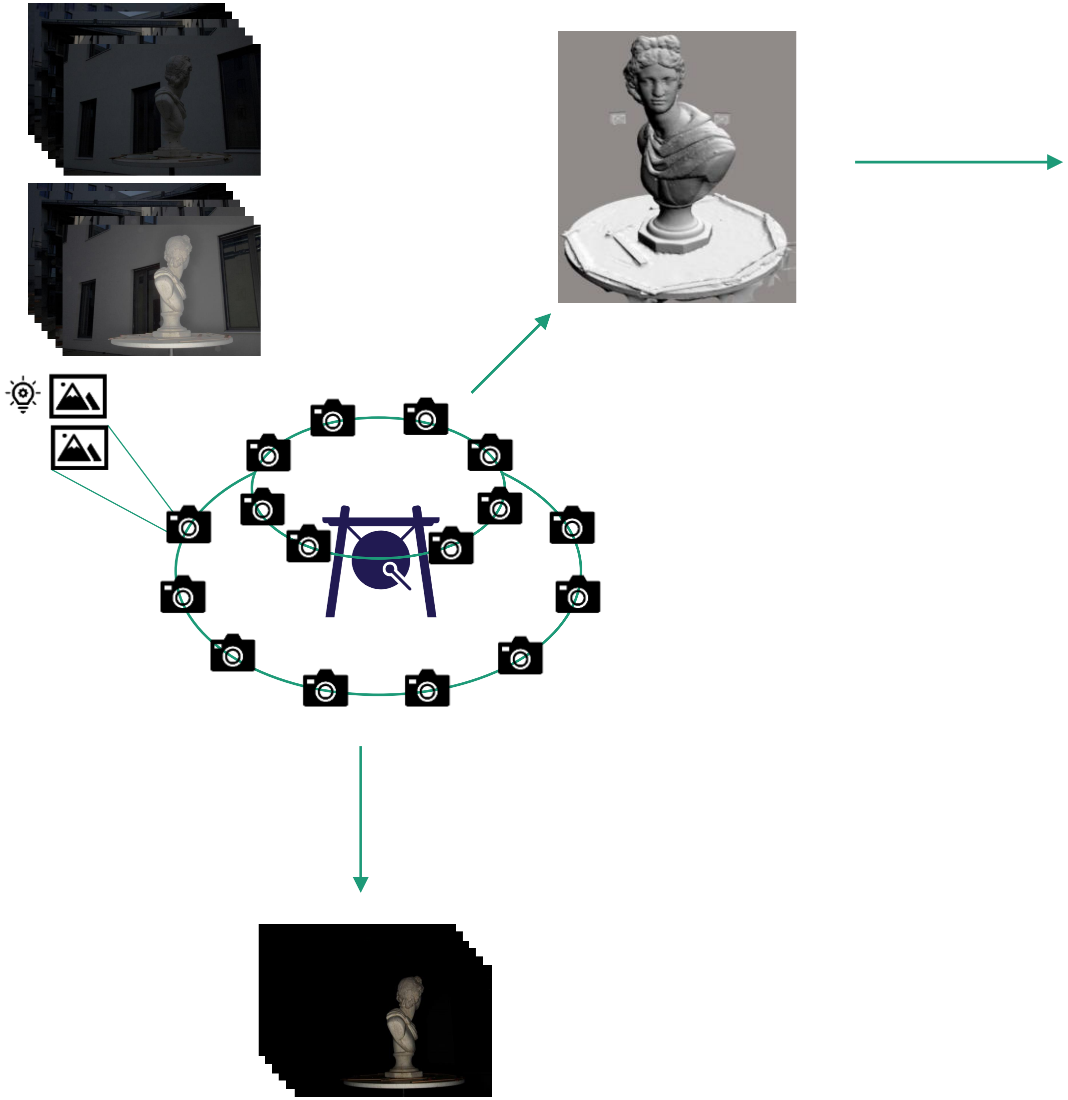


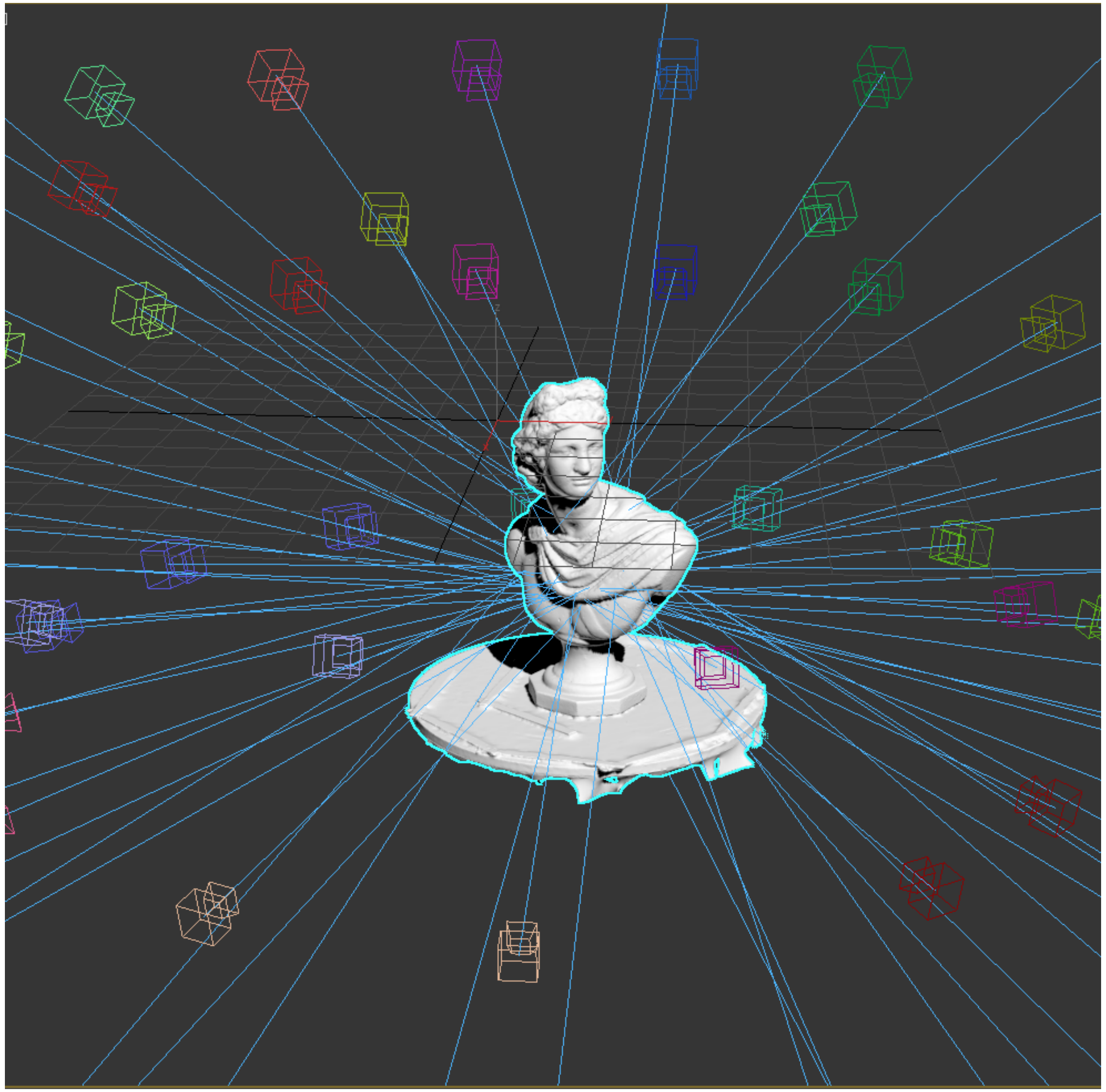


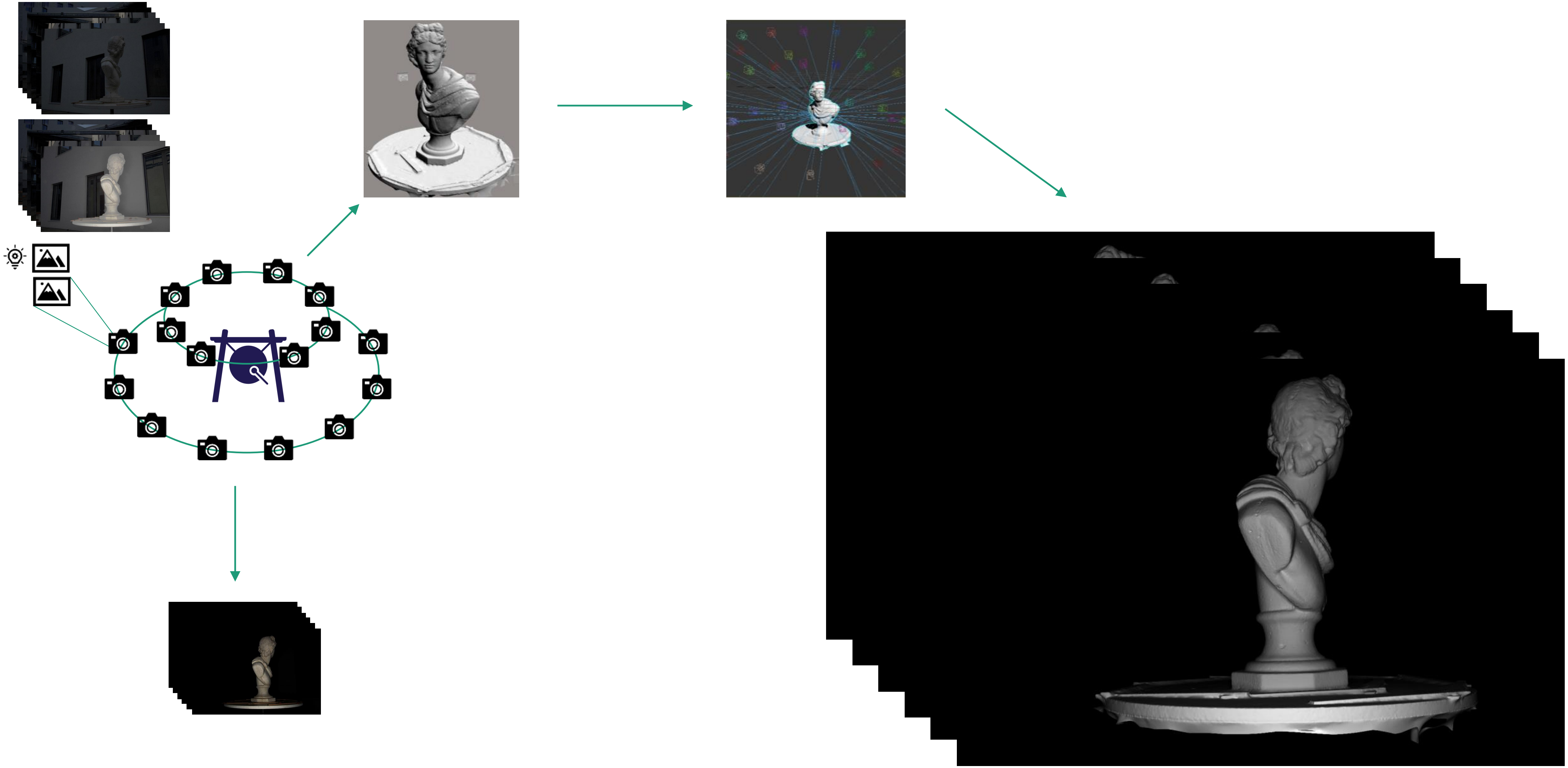


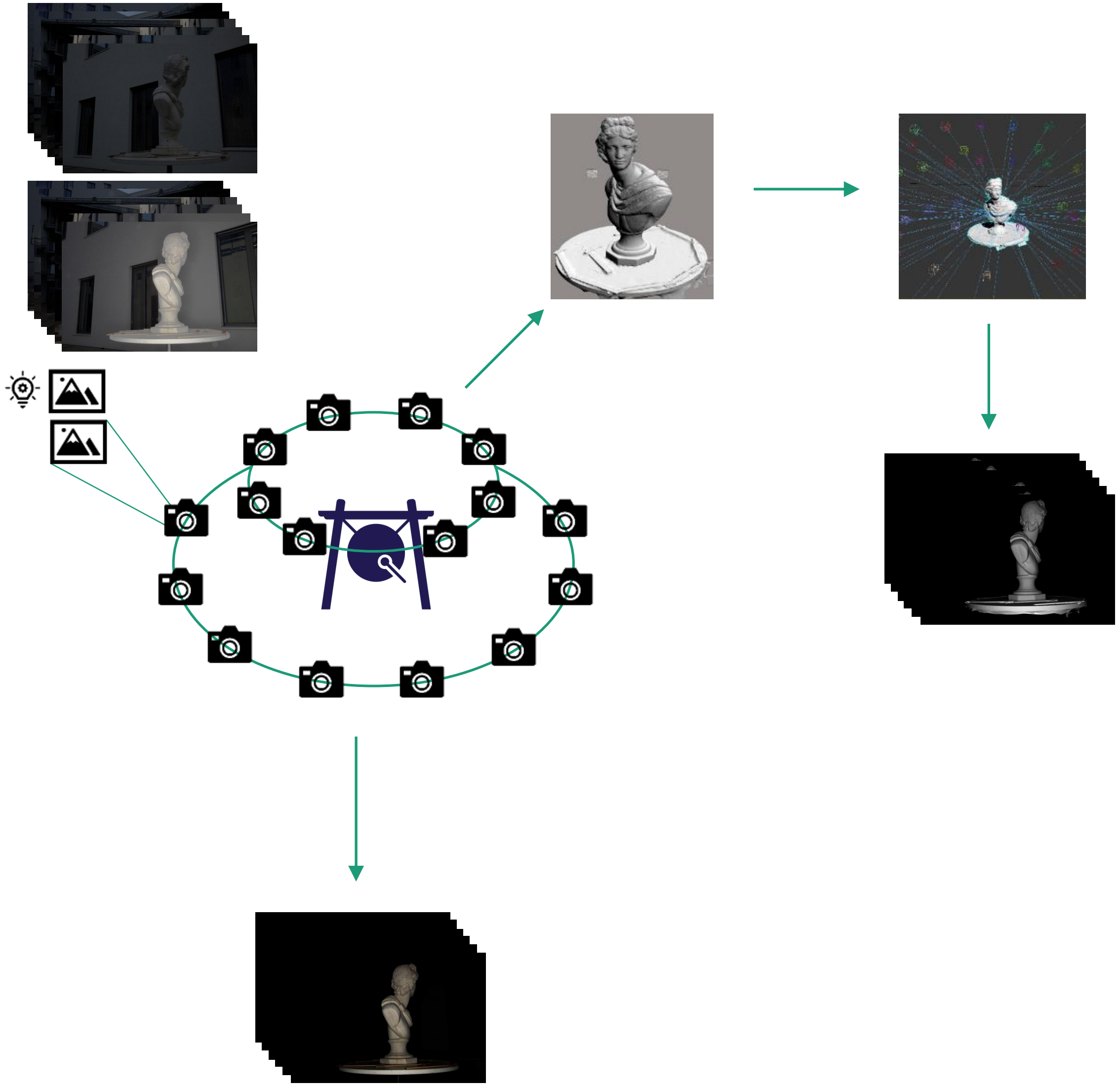


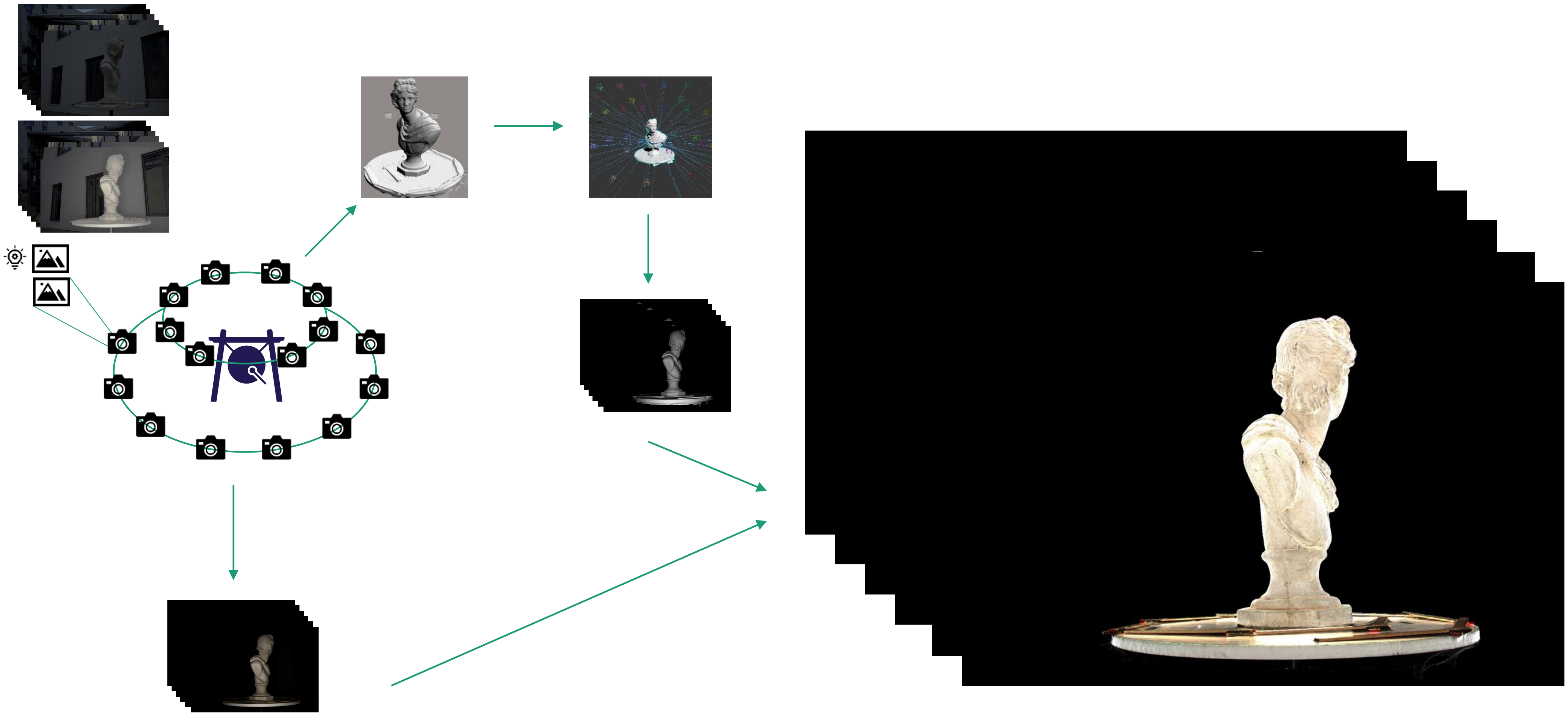






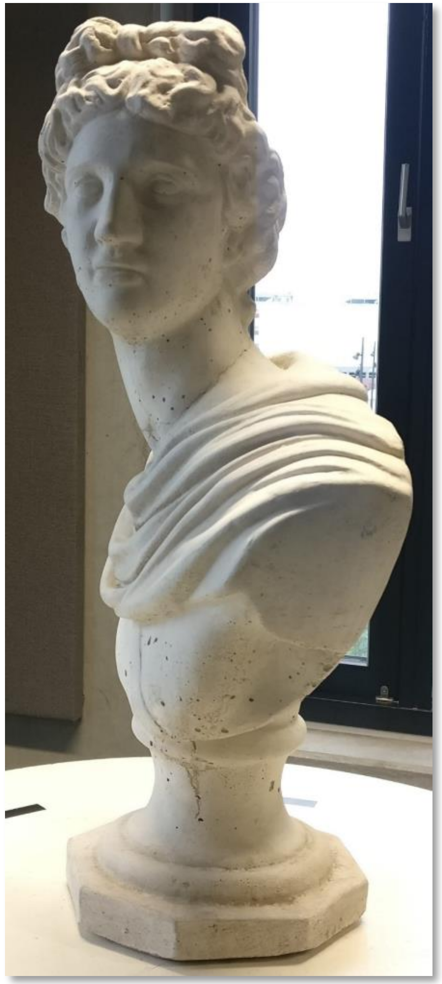






LIGHTING INVARIANCE

OUTSIDE



INSIDE

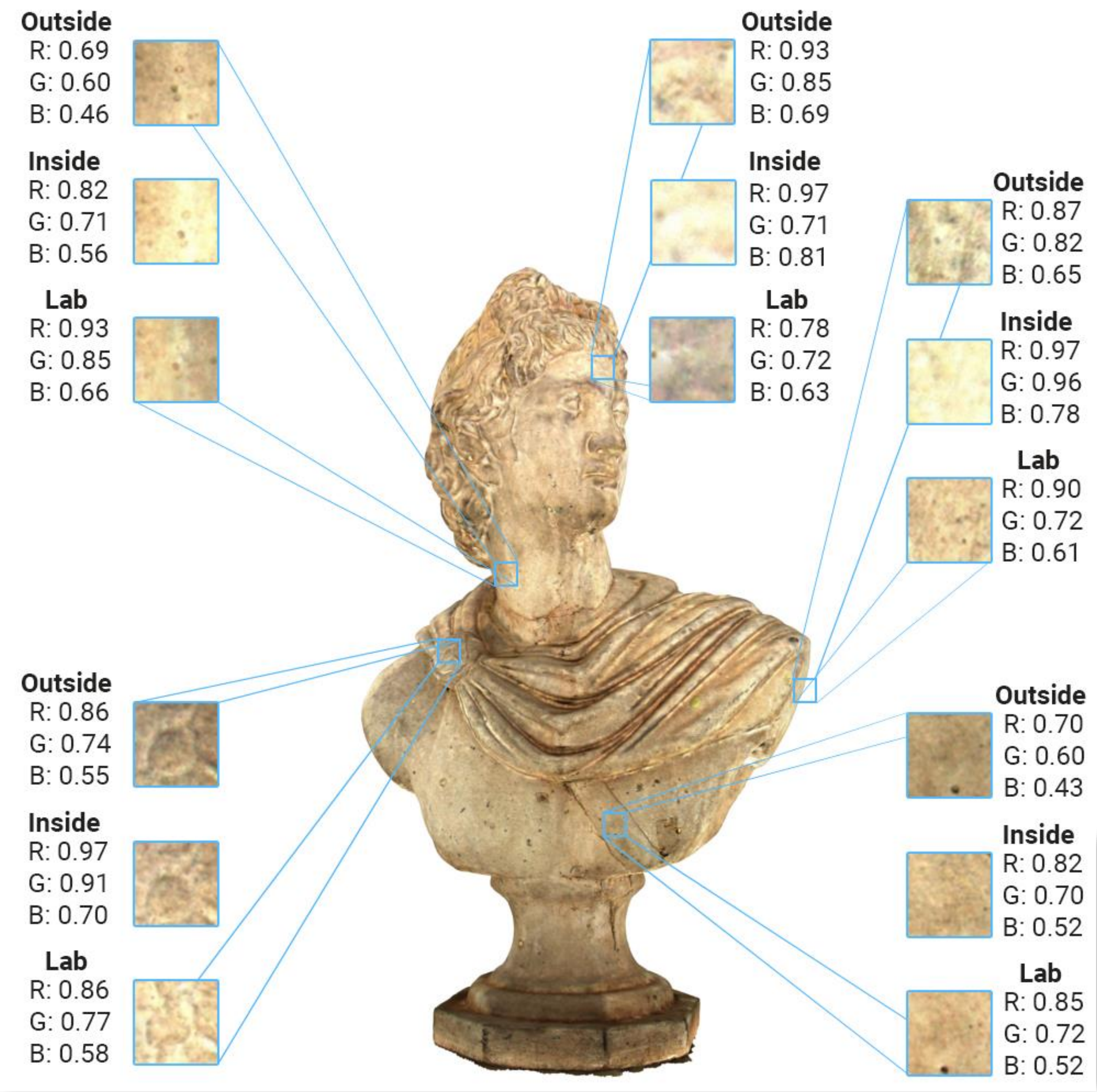
LAB

LIGHTING INVARIANCE

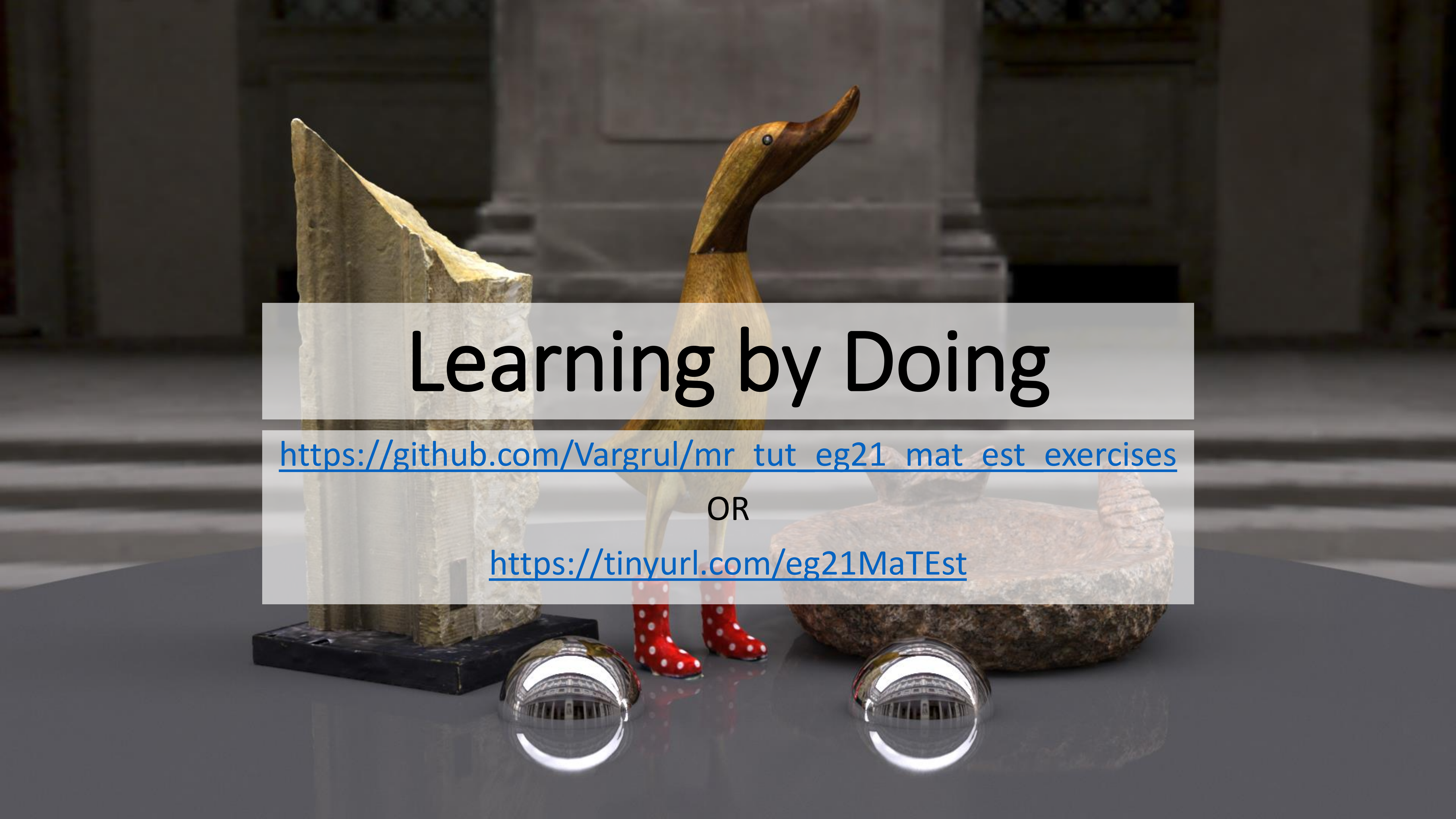
LIGHTING INVARIANCE

- Took patches to compare quantitatively
- Five (5) locations on the bust
- Compared between environments

- *For specific results, please reference the paper.*



QUALITATIVE RESULTS



Learning by Doing

https://github.com/Vargrul/mr_tut_eg21_mat_est_exercises

OR

https://tinyurl.com/eg21MaTEst

ISOLATING KNOWN LIGHT

Exploiting the fact that light is additive

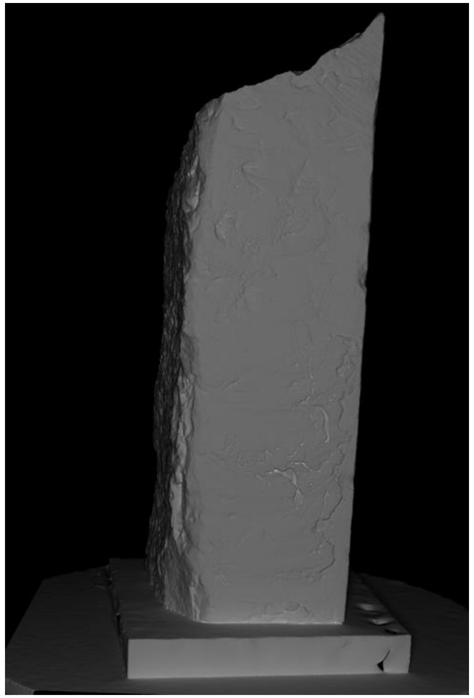
$$P_k(x) = P_{k+u}(x) - P_u(x)$$

Resulting in an image only containing light originating from the known light source

Reflectance Calculation

$$\frac{\rho_d(x)}{\pi} = \frac{\frac{S_u}{S_{u+k}} \cdot P_k(x)}{P_i(x)}$$

For this exercises the $\frac{S_u}{S_{u+k}}$ are assumed to be 1 hence can be ignored.



EXERCISE

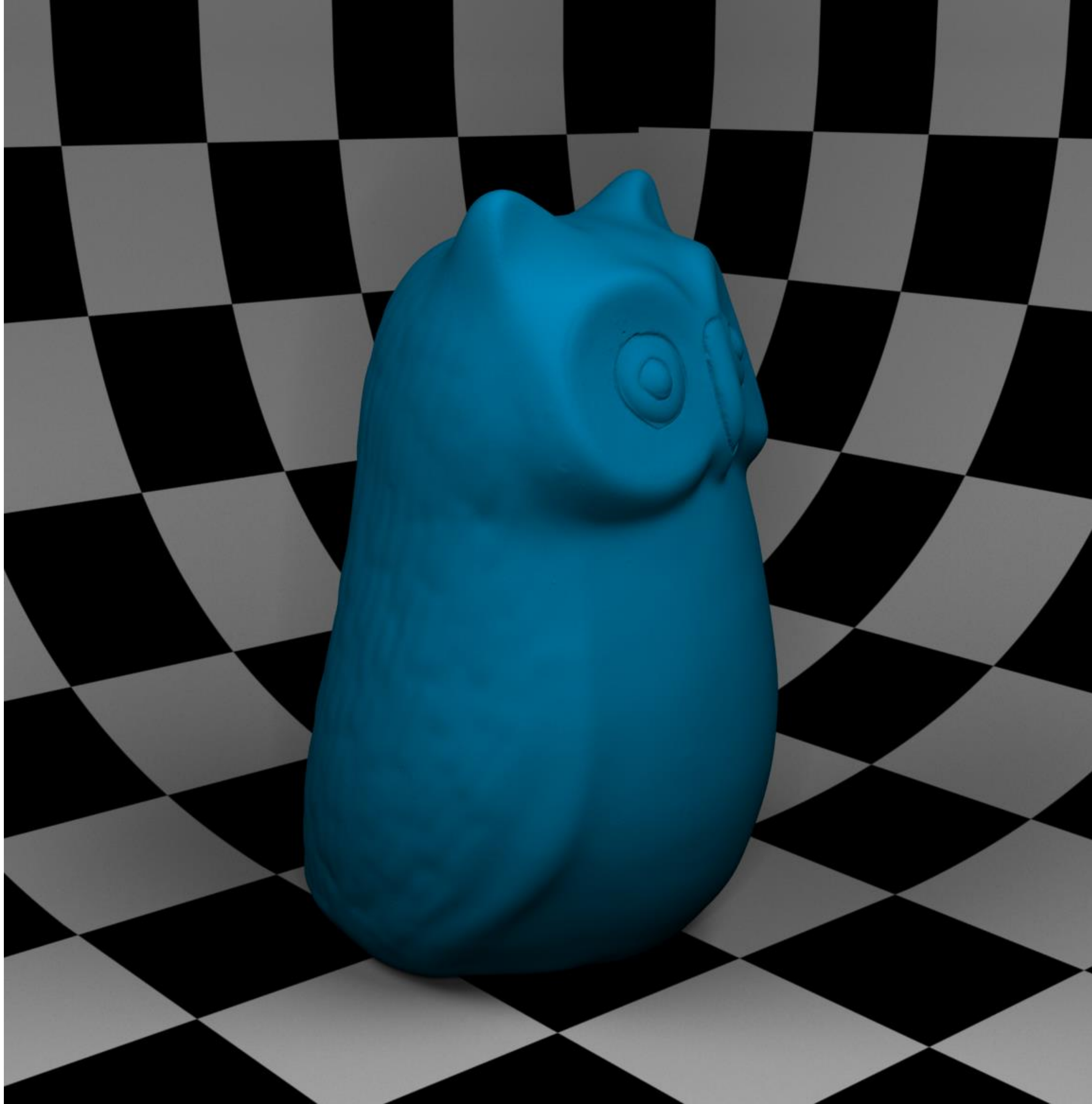
1: Calculate reflectance image from the given data images

2: Calculate Accuracy and Precision (aka error) in pixel value

3: Calculate a, per pixel, error map (image)

4: Examine the reason for the error (hint: there are some interesting information/patterns in the error map, and the intermediate calculated images could also be of interest ;))

5: Riminess a little about the sources of error and how these could be decreased



Hope you learned something!

Email: ksla@create.aau.dk

Literature

1. Guarnera, D., Guarnera, G., Ghosh, A., Denk, C., & Glencross, M. (2016). BRDF Representation and Acquisition. *Computer Graphics Forum*, *35*(2), 625–650. https://doi.org/10.1111/cgf.12867
2. J. Köhler, T. Nöll, G. Reis, and D. Stricker. A full-spherical device for simultaneous geometry and reflectance acquisition. In *2013 IEEE Workshop on Applications of Computer Vision (WACV)*, pp. 355–362. IEEE, 2013.
3. T. Nöll, J. Köhler, G. Reis, and D. Stricker. Faithful, compact and complete digitization of cultural heritage using a full-spherical scanner. In *Digital Heritage International Congress (DigitalHeritage), 2013,* vol. 1, pp. 15–22. IEEE, 2013.
4. T. Nöll, J. Köhler, G. Reis, and D. Stricker. Fully automatic, omnidirectional acquisition of geometry and appearance in the context of cultural heritage preservation. *Journal on Computing and Cultural Heritage (JOCCH)*, 8(1):2, 2015.
5. B. Tunwattanapong, G. Fyffe, P. Graham, J. Busch, X. Yu, A. Ghosh, and P. Debevec. Acquiring reflectance and shape from continuous spherical harmonic illumination. *ACM Transactions on graphics (TOG)*, 32(4):109, 2013.
6. G. Chen, Y. Dong, P. Peers, J. Zhang, and X. Tong. Reflectance scanning: estimating shading frame and brdf with generalized linear light sources. *ACM Transactions on Graphics (TOG)*, 33(4):117, 2014
7. G. Nam, J. H. Lee, D. Gutierrez, and M. H. Kim. Practical svbrdf acquisition of 3d objects with unstructured flash photography. In *SIGGRAPH Asia 2018 Technical Papers*, p. 267. ACM, 2018.
8. Ladefoged, K. S., & Madsen, C. B. Spatially-Varying Diffuse Reflectance Capture Using Irradiance Map Rendering for Image-Based Modeling Applications. In *2019 IEEE International Symposium on Mixed and Augmented Reality* (pp. 46-54). 2020. [8943701] IEEE Computer Society Press. https://doi.org/10.1109/ISMAR.2019.00-27

enjoy Vienna

EUROGRAPHICS 2021

Diminished Reality (DR)

Shohei Mori

Diminished Reality (DR)

“While most applications of AR are concerned with the addition of virtual objects to a real scene, ***diminished reality describes the conceptual opposite*** *— namely, the seamless removal of real objects from a real scene.”*

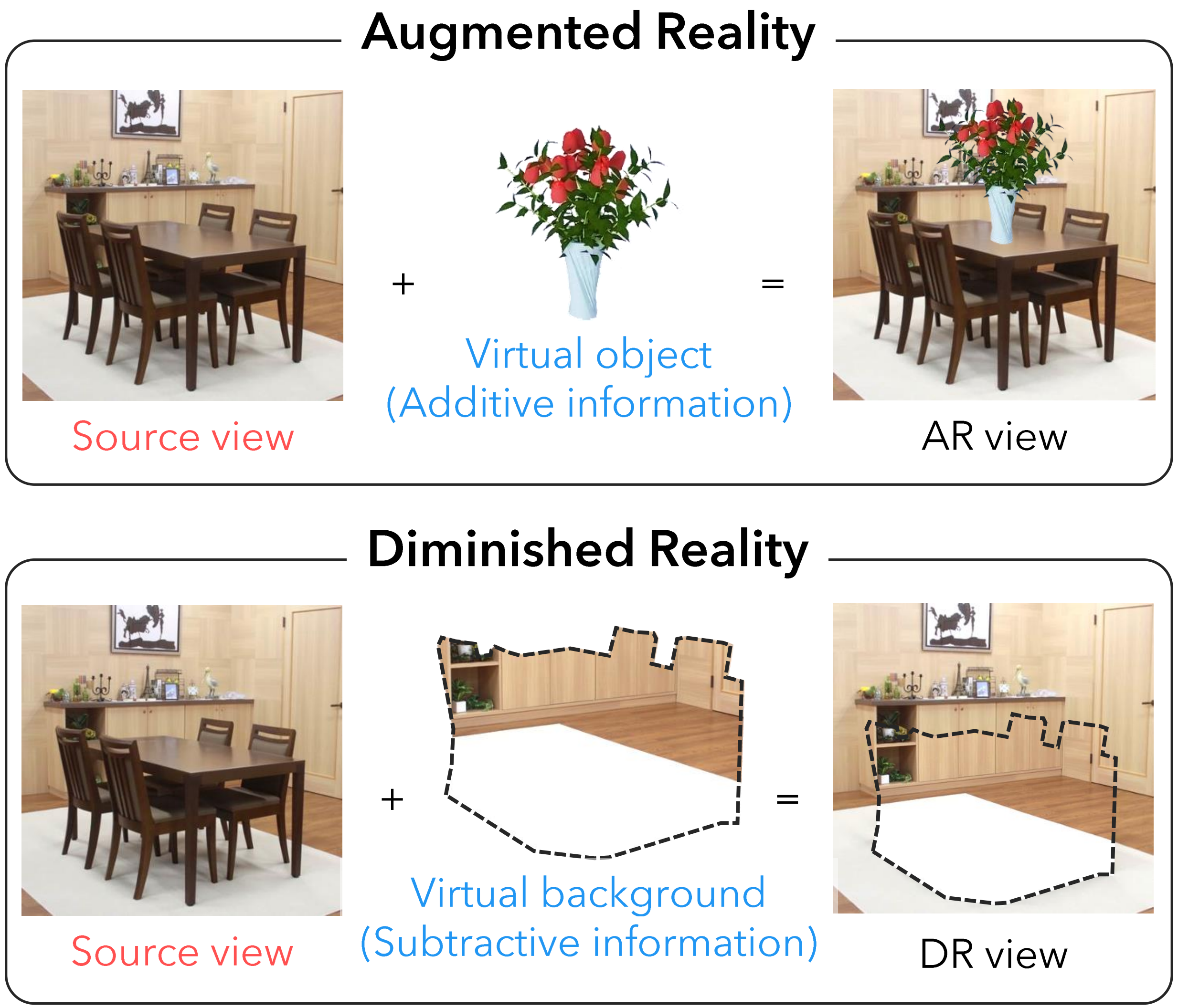
D. Schmalstieg and T. Hollerer (2016) Augmented Reality: Principles and Practice, Addison-Wesley Professional

DR is a set of methodologies for diminishing the reality, and concealing, eliminating, and seeing through objects in a perceived environment in real time.

S. Mori, S. Ikeda, and H. Saito: A Survey of Diminished Reality: Techniques for Visually Concealing, Eliminating, and Seeing Through Real Objects, IPSJ Trans. on Computer Vision and Applications (CVA), Vol. 9, No. 17, SpringerOpen, DOI: 10.1186/s41074-017-0028-1 (2017.6)

AR vs. DR

Conceptually
⇔



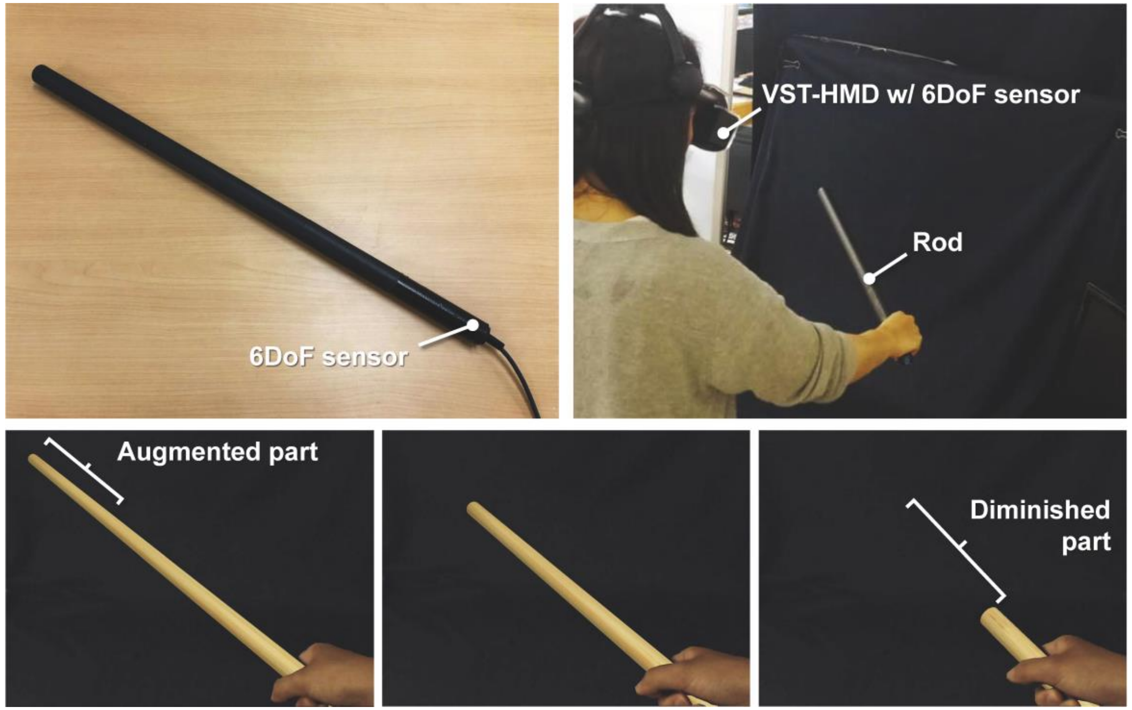
Technically
≒

Figures based on
S. Mori, S. Ikeda, and H. Saito: A Survey of Diminished Reality: Techniques for Visually Concealing, Eliminating, and Seeing Through Real Objects, IPSJ Trans. on Computer Vision and Applications (CVA), Vol. 9, No. 17, SpringerOpen, DOI: 10.1186/s41074-017-0028-1 (2017.6)

Real-time Capability Matters!

A DR system must present an “experience” through multi-modal displays

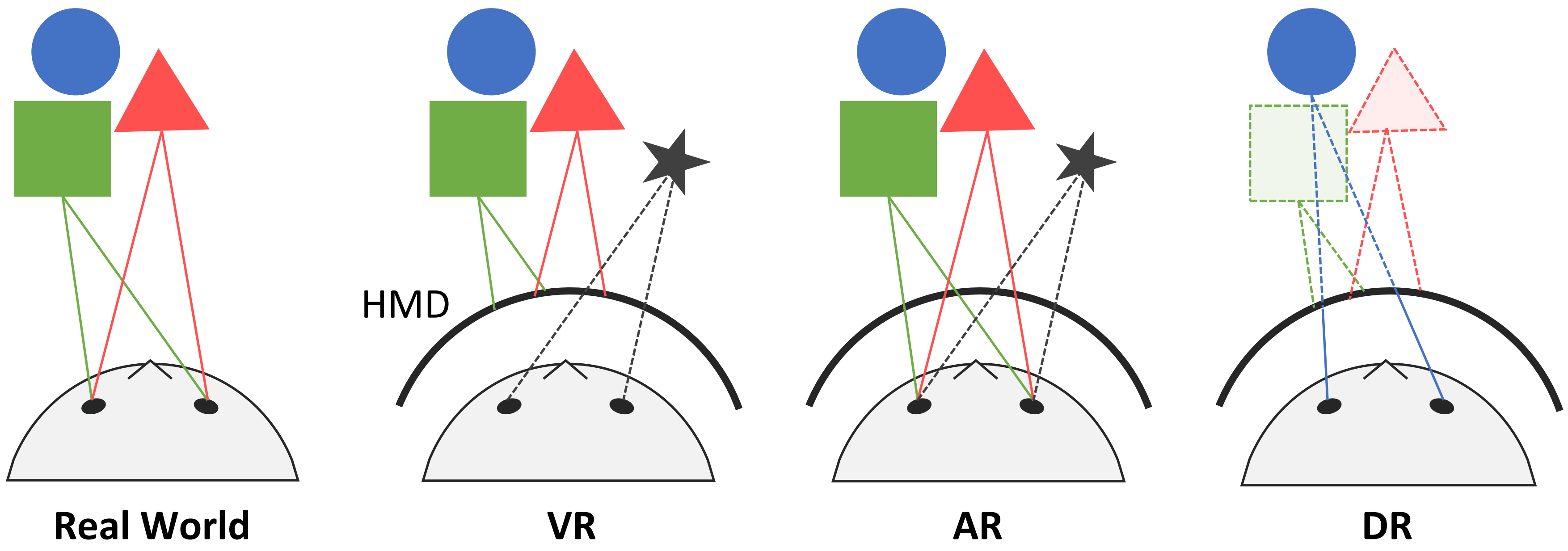
- Usually targeting to **30Hz** refresh rate at **640** × **480** pixel resolution



Figures based on
S. Hashiguchi, S. Mori, M. Tanaka, F. Shibata, and A. Kimura, “Perceived Weight of a Rod under Augmented and Diminished Reality Visual Effects”,
Proc. The ACM Symp. on Virtual Reality Software and Technology (VRST) (2018.11)

Displays for DR

- DR displays are capable of selectively occluding real light rays
- Light rays occluded by frontal objects need to be recovered virtually



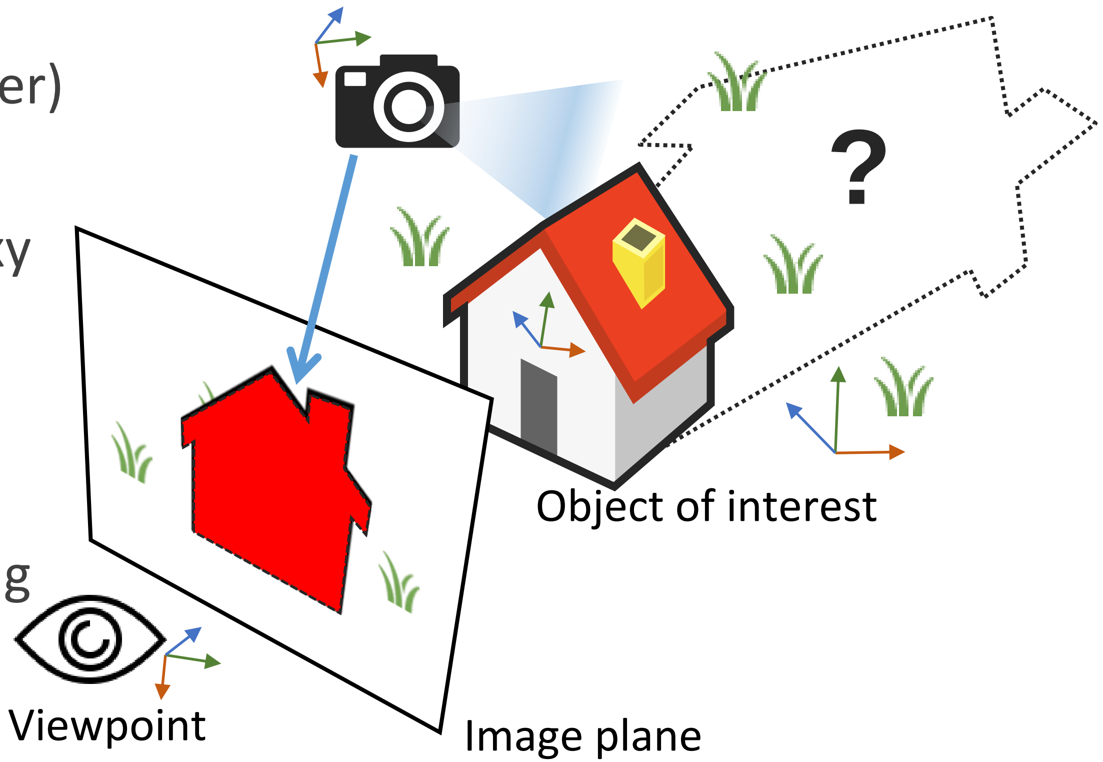
Figures based on S. Mori and H. Saito, "An Overview of Augmented Visualization: Observing the Real World as Desired"
APSIPA Trans. on Signal and Information Processing, Vol. 7, pages E12 (2018.10)

Non-video-based Displays are not ready for DR

Yet

Implementing a DR System

1. Tracking
   - Camera or scene tracking (e.g., vSLAM / marker)
2. Background proxy modeling
   - Planar proxy / multi-plane proxy / full 3D proxy
3. ROI detection
   - User annotation / semantic-segmentation
4. Background synthesis
   - Image-based rendering / Homography warping
5. Composition
   - Intensity interpolation / seamless cloning / smooth alpha masking / lighting estimation



Figures based on
S. Mori and H. Saito, "An Overview of Augmented Visualization: Observing the Real World as Desired" APSIPA Trans. on Signal and Information Processing, Vol. 7, pages E12 (2018.10)

Background Resources

a) Multi-viewpoint images

- (+) Resources from *observations*
- (−) Hardware sync., calibration, color compensation, etc.

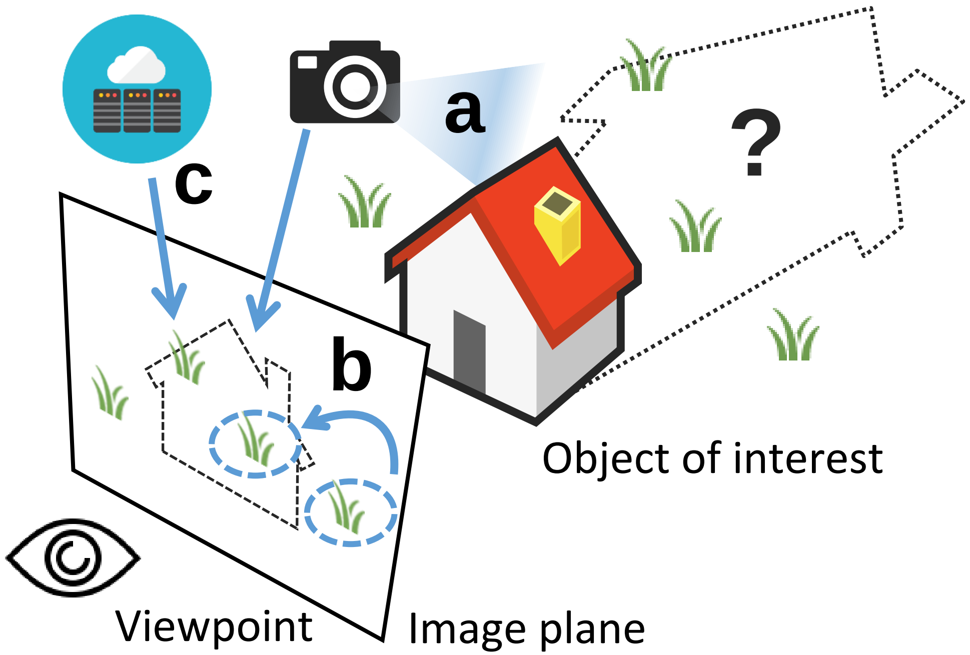
b) Pixels within the FoV (Inpainting)

- (+) No additional hardware, thus, portable
- (−) *Hallucinated* background
- (−) Fast (multi-view) inpainting is hard

c) Dataset (Photo collection / Features)

- (+) On-demand resource
- (+) Well-prepared resources
- (−) Large memory or network connection
- (−) Day/time compensation

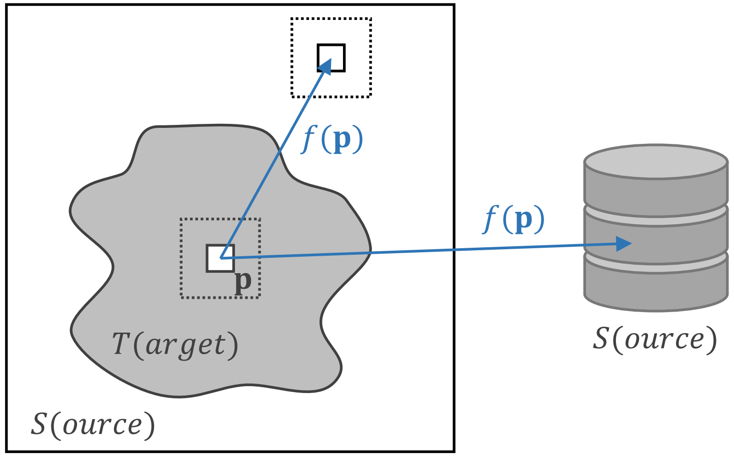
d) Combinations of the above



Figures based on
S. Mori and H. Saito, "An Overview of Augmented Visualization: Observing the Real World as Desired" APSIPA Trans. on Signal and Information Processing, Vol. 7, pages E12 (2018.10)

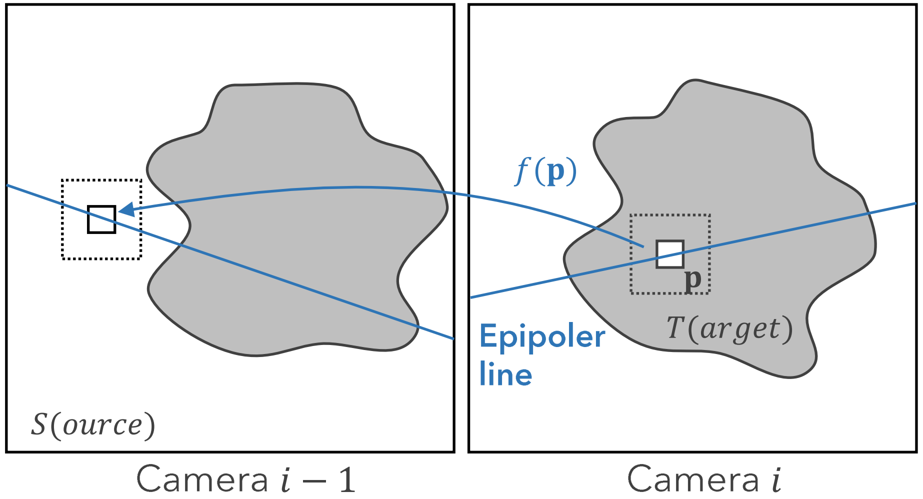
Formulating DR Problems

Goal Find $f: T \rightarrow S$



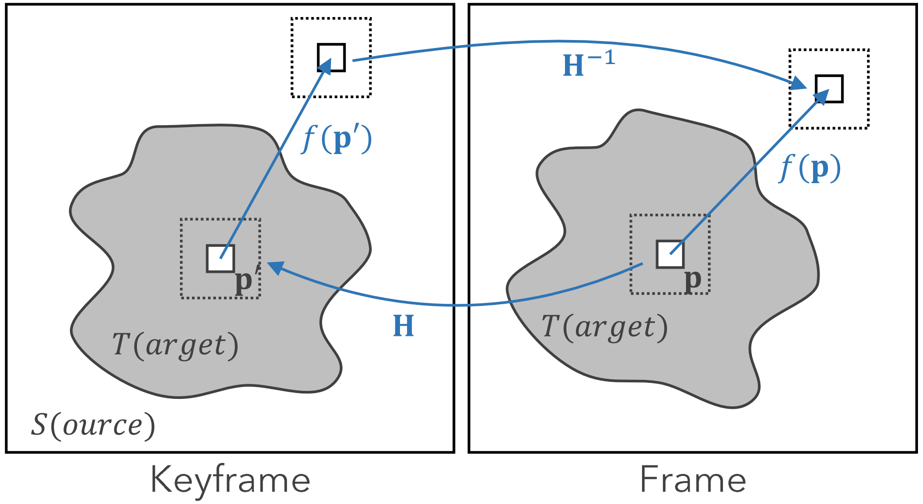
Multi-view approach

[Zokai+, ISMAR03][Rameau+, TVCG16]



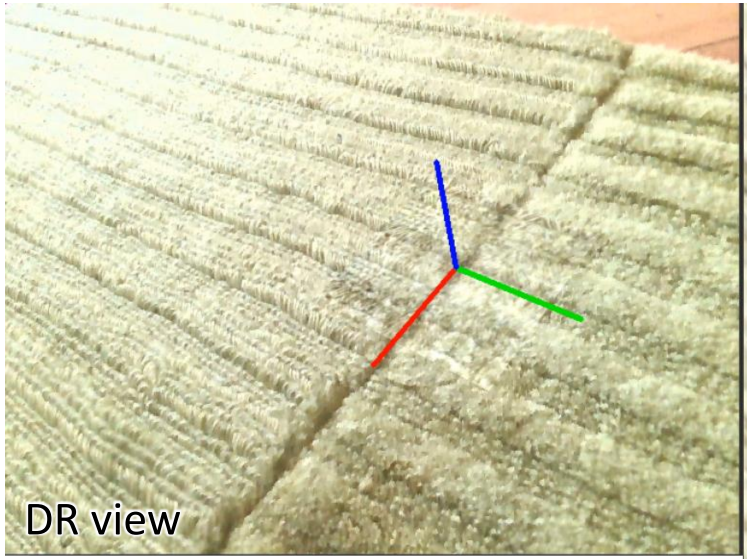
Inpainting approach

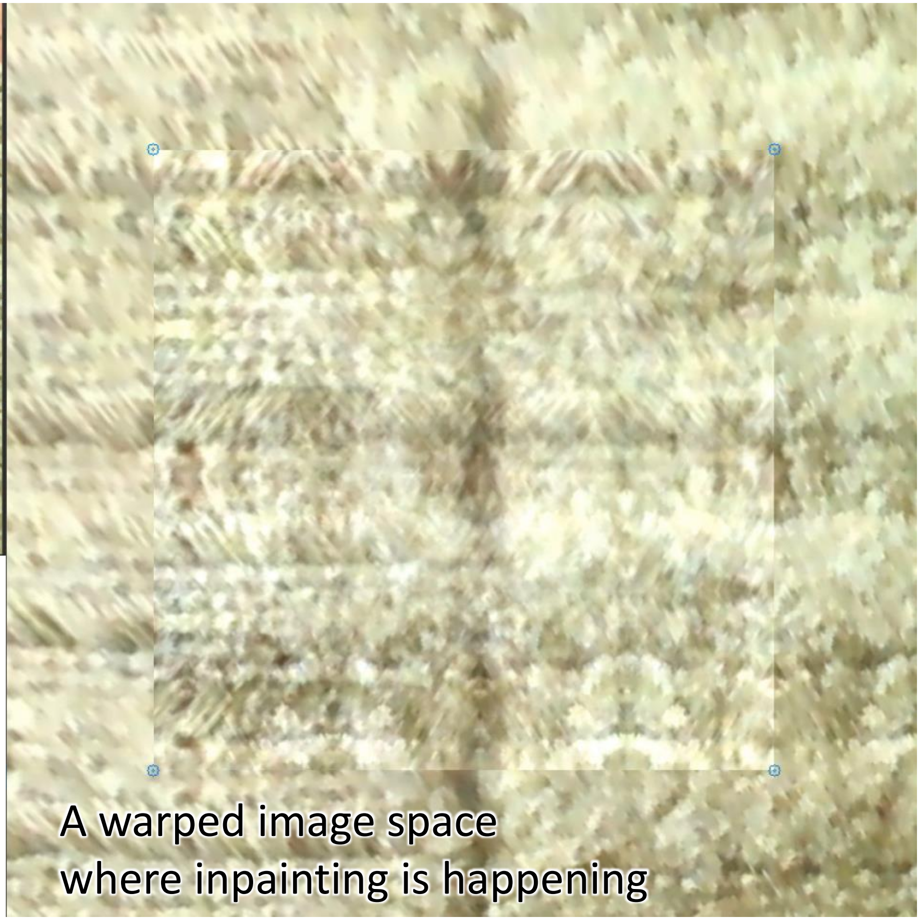
[Korkalo+, ISMAR10][Herling+, TVCG14][Kawai+, TVCG16, 17]



Fast Inpainting for Marker Hiding

S. Siltanen, “Texture Generation over the Marker Area”, Proc. ISMAR, 2006.



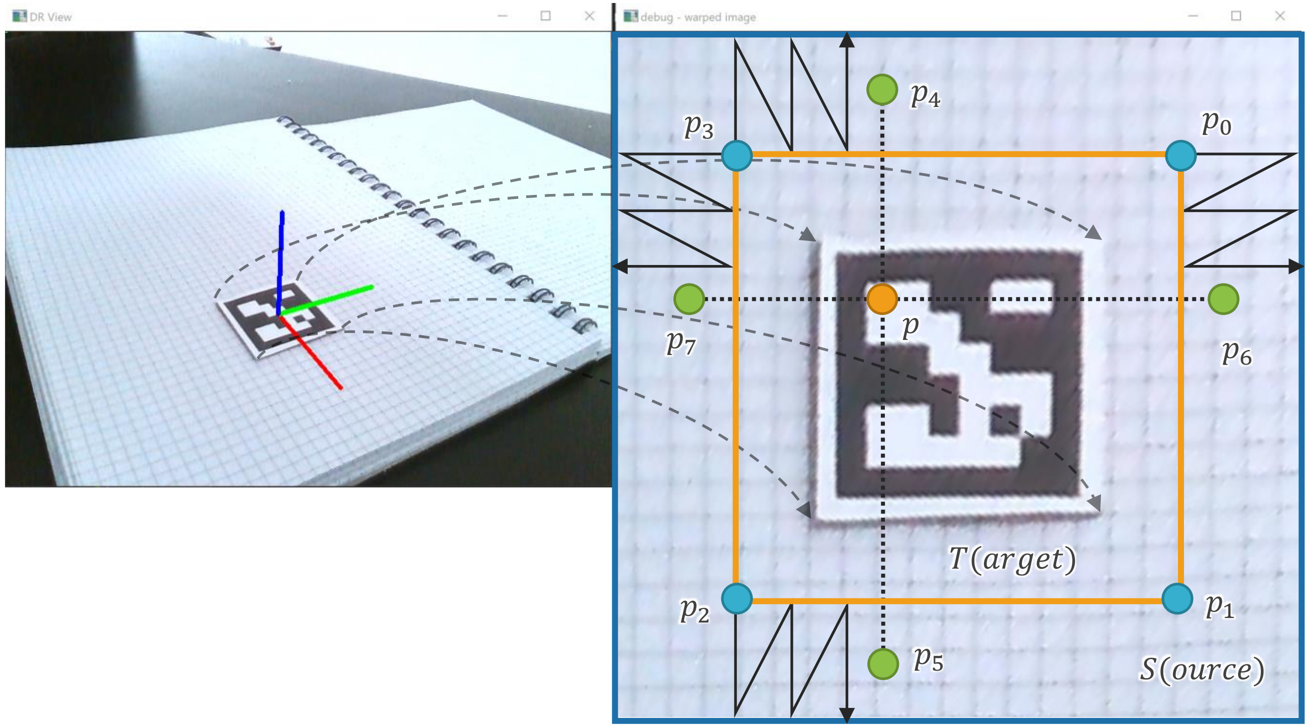


Fast Inpainting for Marker Hiding

S. Siltanen, "Texture Generation over the Marker Area", Proc. ISMAR, 2006.

A pioneering marker hiding method

Mirroring and mixing the vicinity pixels towards the marker region

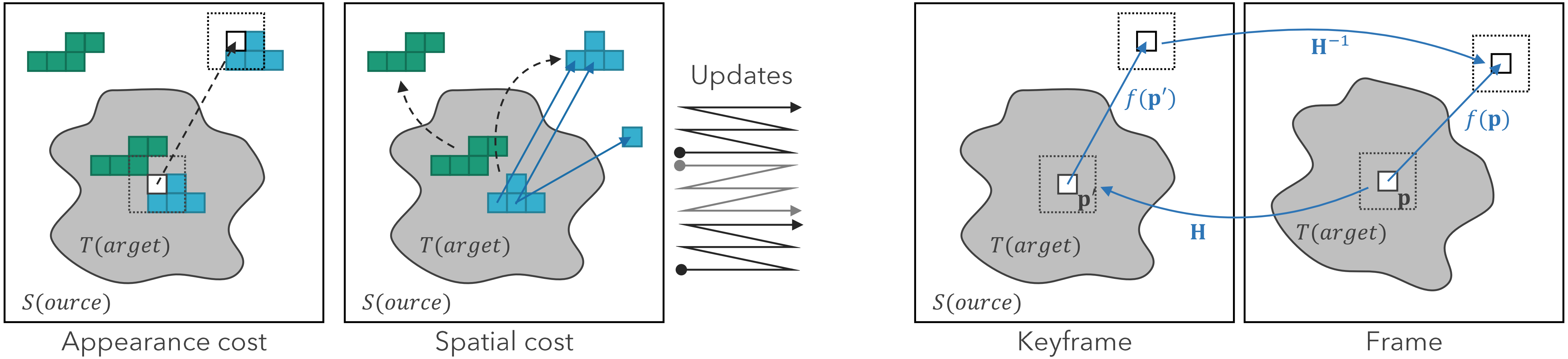


$$p = \sum_{i=\{0,\dots,7\}} w_i p_i$$

PixMix – A Keyframe-based Approach

J. Herling and W. Broll, "High-Quality Real-Time Video Inpainting with PixMix," IEEE TVCG, Vol. 20, Issue 6, pp. 866 - 879, 2014.

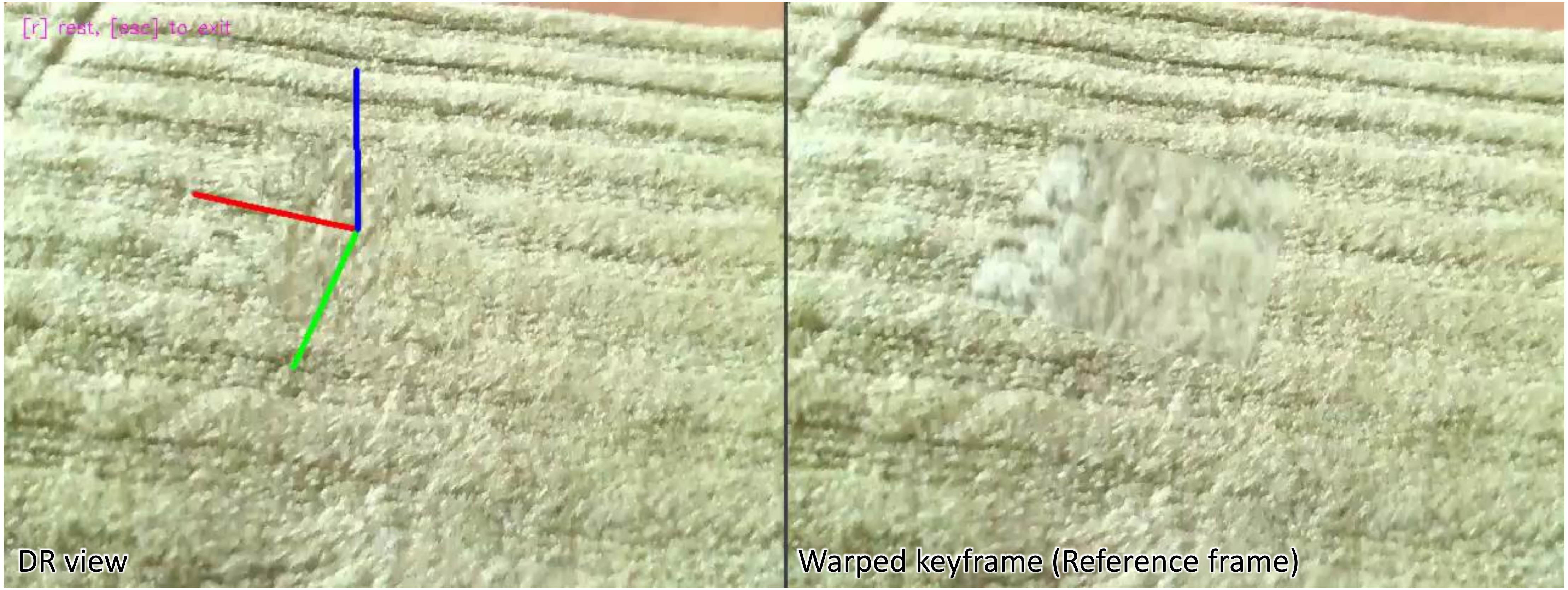
- Inpaint a frame (Keyframe) and **warp** it to the current frame as a reference
- Keep copying adjacent pixels when good pixels are not found
- Region-wise parallel pixel updates in an image



$$\min_f \sum_{p \in T} cost_\alpha(p) = \alpha\ cost_{appearance}(p)\ + (\alpha - 1)\ cost_{spatial}(p)$$

For more details: https://github.com/Mugichoko445/PixMix-Inpainting

Marker Hiding Using PixMix



Multi-plane Inpainting

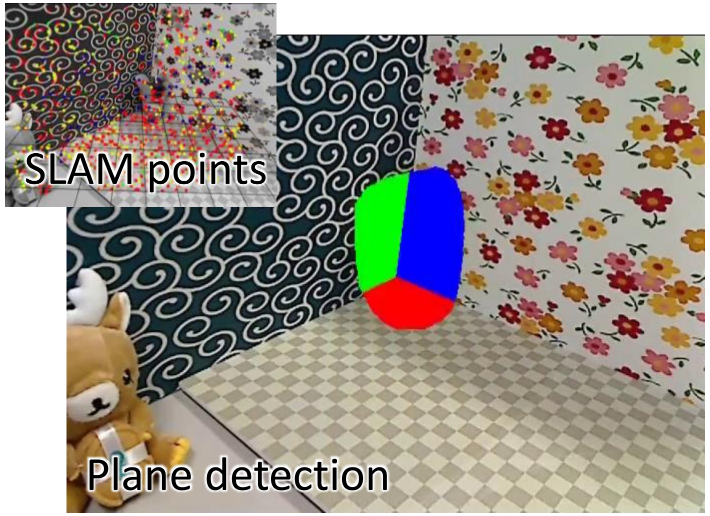
N. Kawai, T. Sato, and N. Yokoya. "Diminished Reality based on Image Inpainting Considering Background Geometry", IEEE TVCG, Vol. 22 Issue 3, pp. 1236 - 1247, 2015.

Video: Courtesy of Dr. N. Kawai

Multi-plane Inpainting

N. Kawai, T. Sato, and N. Yokoya. "Diminished Reality based on Image Inpainting Considering Background Geometry", IEEE TVCG, Vol. 22 Issue 3, pp. 1236 - 1247, 2015.

- Inpaint the ROI on **independent plains** in a **keyframe**
- Tracking & inpainting on different **threads**
- Show **on-going** inpainting results

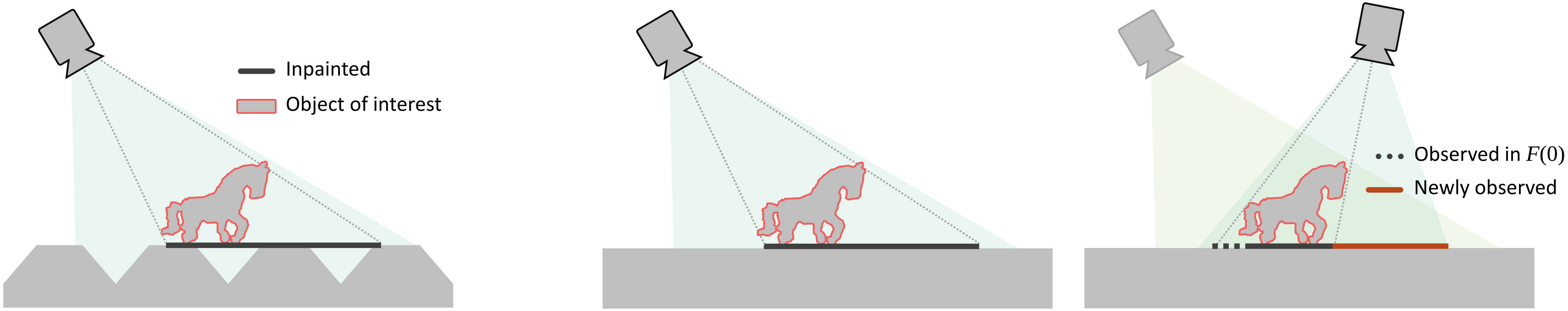


Figures: Courtesy of N. Kawai

Marker Hiding Using Multi-threading

Plane(s) as Background Geometry Proxy?

- Image-inpainting works in an image-space
- Limitations to AR/DR
  - No interaction with the background after a DR method is applied
  - No automatic updates when new real object pixels are observed
- How can we extend inpainting for 3D AR scenes?



Keyframe inpainting | First keyframe $F(0)$ | Upcoming frames

InpaintFusion – 3D Inpainting for AR Scenes

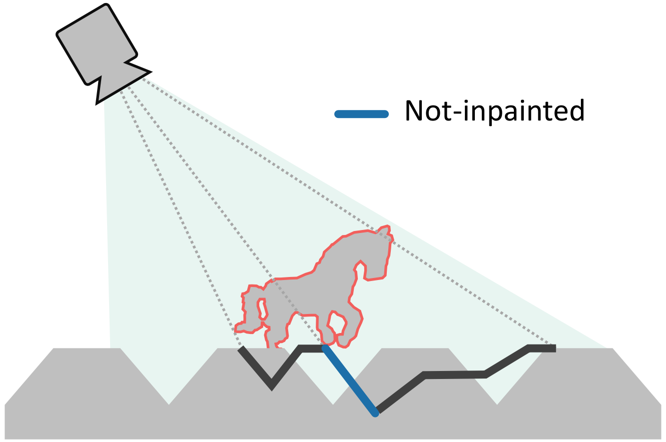
S. Mori, O. Erat, W. Broll, H. Saito, D. Schmalstieg, and D. Kalkofen, "InpaintFusion: Incremental RGB-D Inpainting for 3D Scenes", IEEE TVCG, Vol. 26, Issue 10, 2020.

Ideas

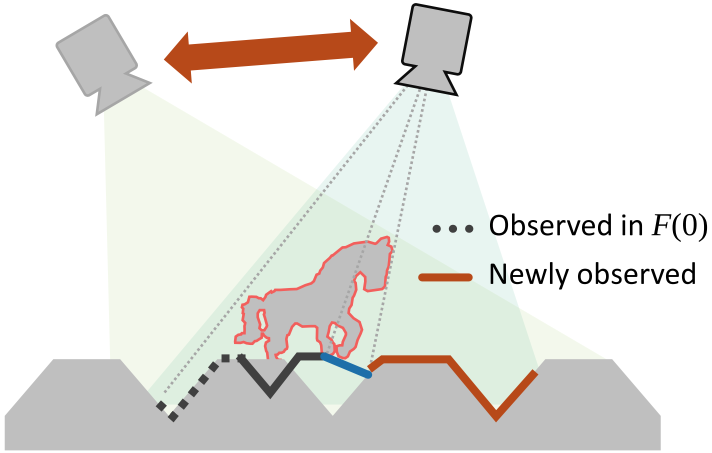
S. Mori, O. Erat, W. Broll, H. Saito, D. Schmalstieg, and D. Kalkofen, “InpaintFusion: Incremental RGB-D Inpainting for 3D Scenes”, IEEE TVCG, Vol. 26, Issue 10, 2020.

Multi-keyframe inpainting with **RGBD fusion** and an **IBR** technique
Image-Based Rendering

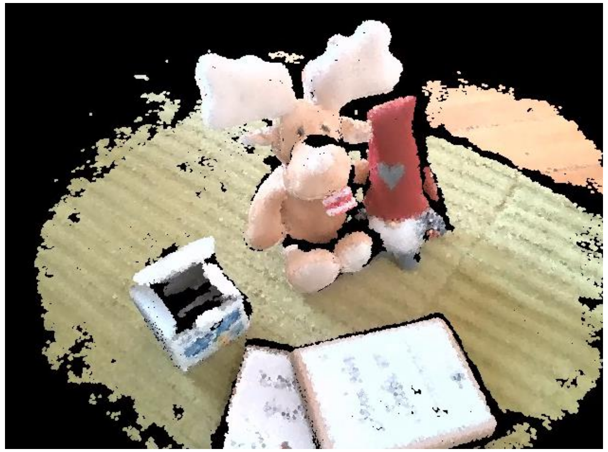
- *RGBD inpainting* per keyframe
- Filling in missing pixels in the *next keyframes* and *fuse* them
- Pixel *blending* based on view-/surfel-priorities



First keyframe $F(0)$



Upcoming frames

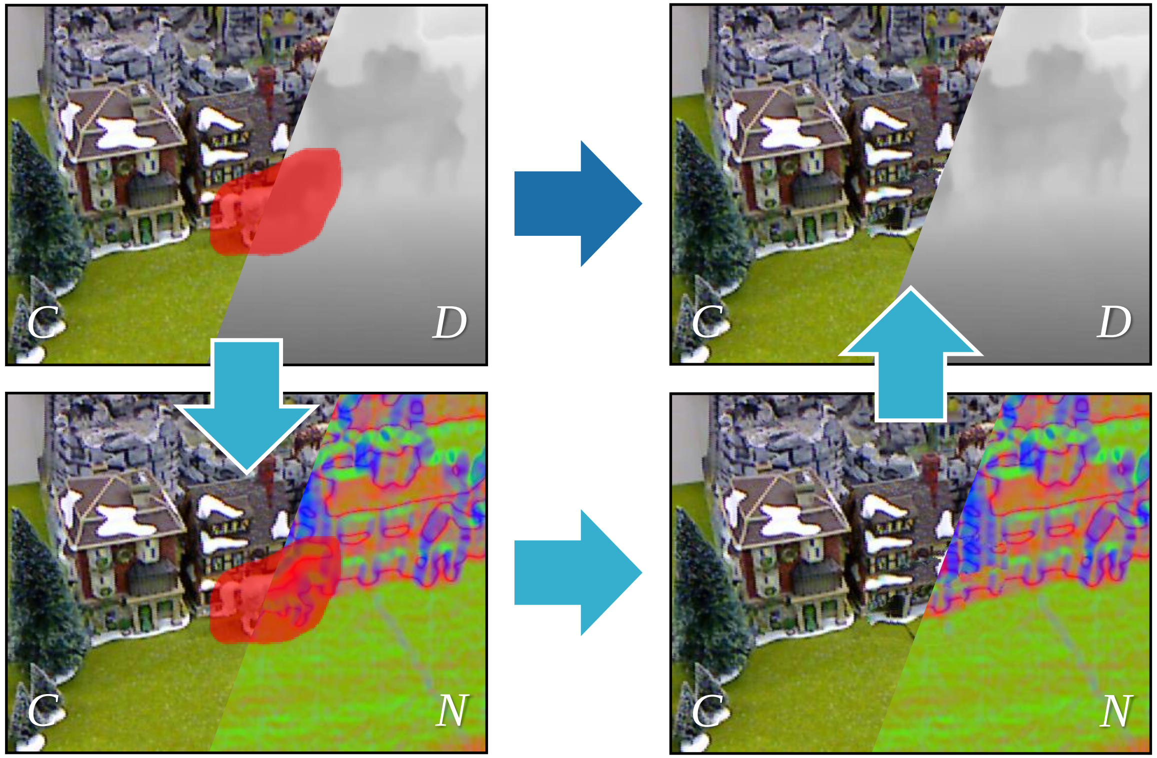


Surfel fusion

RGBD Keyframe Inpainting

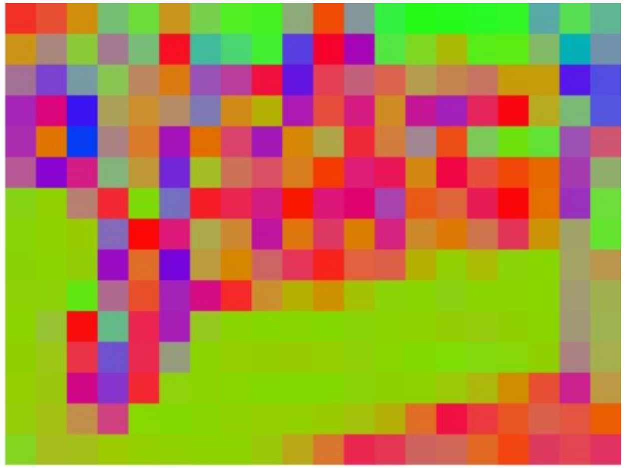
S. Mori, O. Erat, W. Broll, H. Saito, D. Schmalstieg, and D. Kalkofen, “InpaintFusion: Incremental RGB-D Inpainting for 3D Scenes”, IEEE TVCG, Vol. 26, Issue 10, 2020.

- RGBD inpainting via RGB-Normal inpainting
- Depth from depth gradient samples from f^*



$$f^* = \arg\min_f \sum_{\mathbf{u} \in T} \left(w\rho_t(f, \mathbf{u})\rho_g(f, \mathbf{u}) + (1 - w)\rho_s(f, \mathbf{u}) \right)$$

Texture term — Geometry term — Spatial term



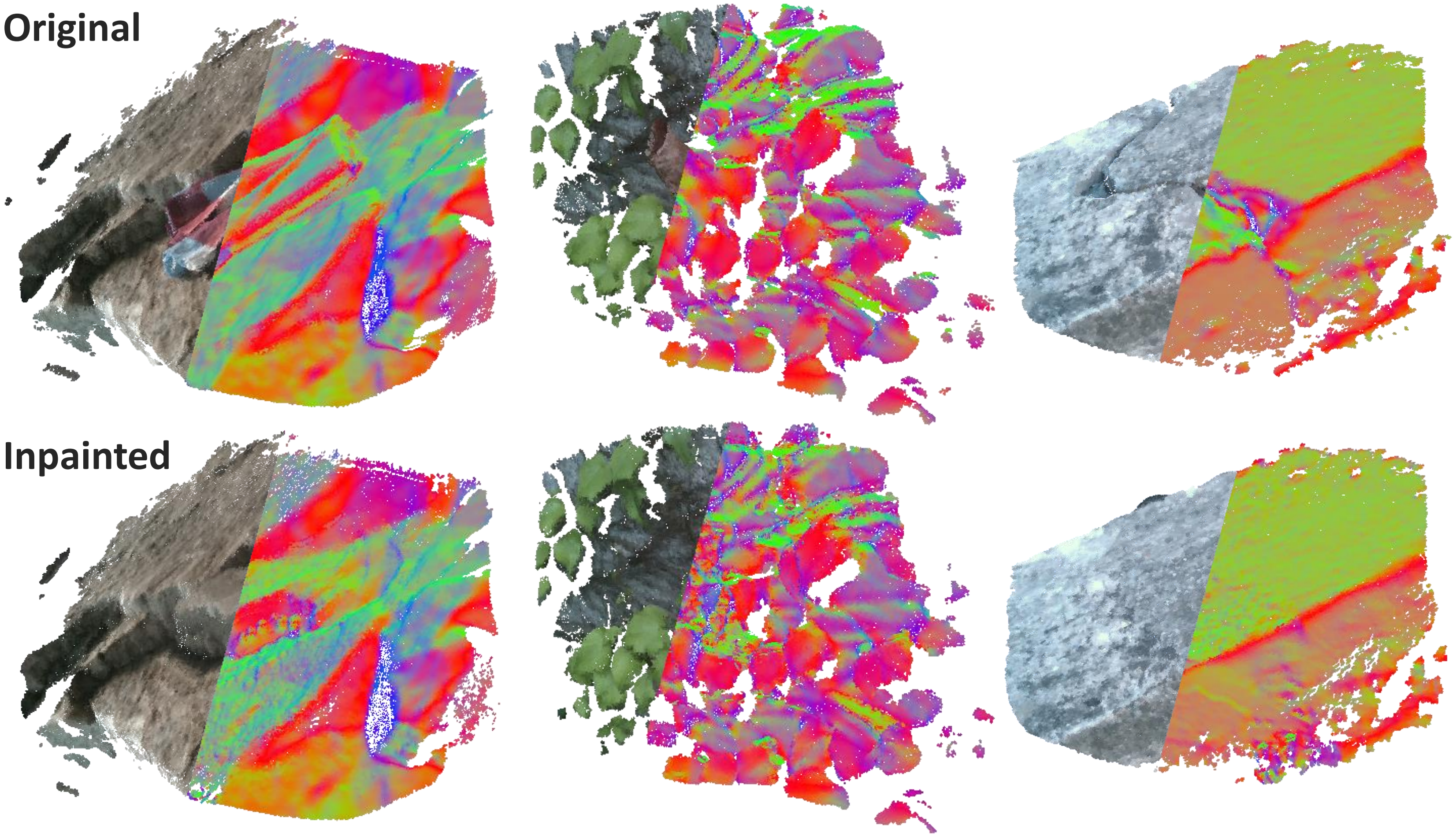
Keyframe (KF) Propagation and Blending

S. Mori, O. Erat, W. Broll, H. Saito, D. Schmalstieg, and D. Kalkofen, "InpaintFusion: Incremental RGB-D Inpainting for 3D Scenes", IEEE TVCG, Vol. 26, Issue 10, 2020.

- KF is inserted when the sensor gets away from the closest KF
- KF's transformation map f is transformed to a new KF
- Multiple KFs are **blended** over the inpainted global surfel map

Fusion Results

S. Mori, O. Erat, W. Broll, H. Saito, D. Schmalstieg, and D. Kalkofen, “InpaintFusion: Incremental RGB-D Inpainting for 3D Scenes”, IEEE TVCG, Vol. 26, Issue 10, 2020.



← For **tracking** and **user-annotation** (ROI)

← For **3D AR interactions**

Summary

- Diminished Reality (DR)
  - DR is a **conceptual opposite** to AR, while they are **technically similar**
  - The majority of DR systems are **video-based**
  - Multi-view cameras, inpainting, and dataset

- (Semi-)Real-time inpainting for DR experiences
  - Mirroring & mixing, keyframe, multi-plane approaches
  - InpaintFusion for full 3D DR and AR

Take-home Message

- DR is a missing piece that compensates AR
- Real-time 3D inpainting is still challenging
- All inpainting-based DR systems rely on exemplar-based approaches
- Multi-modal DR is an un-touched research area

enjoy Vienna

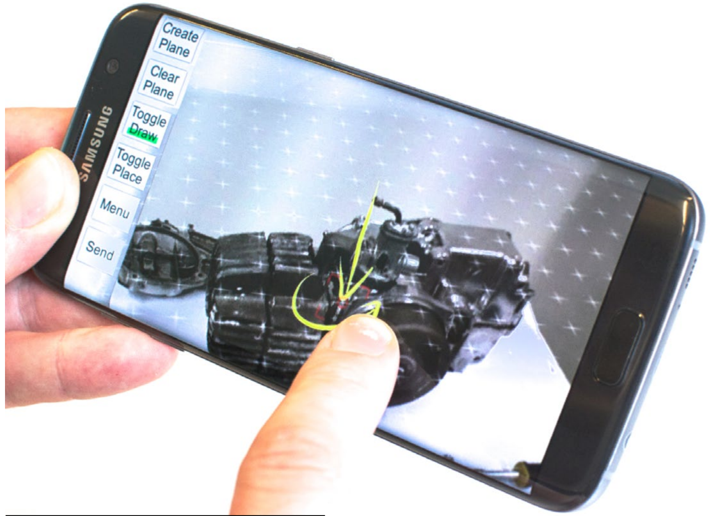
EUROGRAPHICS 2021

Perceptual Issues of Augmented Reality Visualization

Markus Tatzgern, Salzburg University of Applied Sciences

Perceptual Issues

- A short overview of perceptual issues of AR visualization with a focus on issues that AR visualizations and applications typically face
- Visual clutter
- Temporal coherence
- Registration errors
- Visual interference
- Viewport of scene

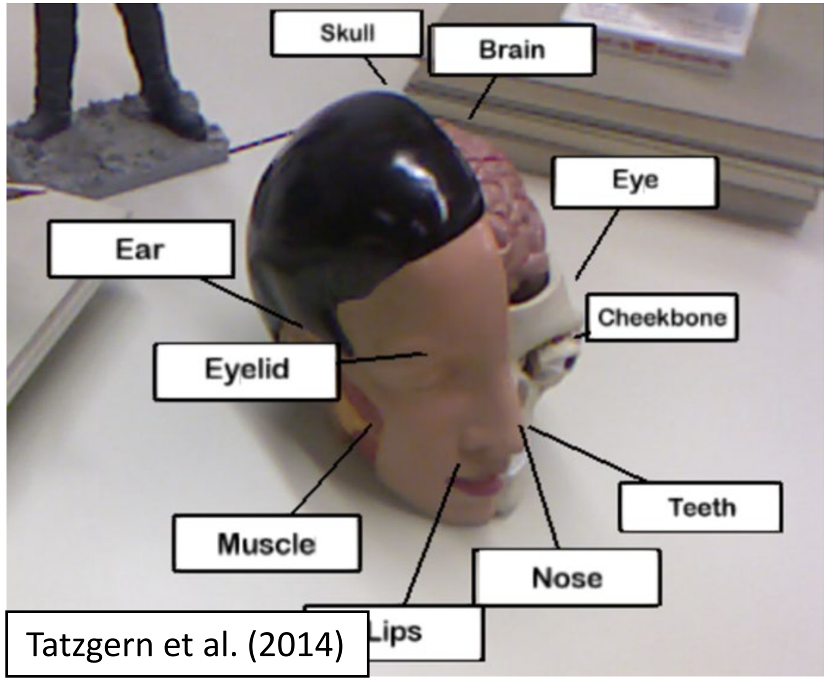


Mohr et al. (2020)

Typical AR Visualizations

Object Annotations

- Annotated an object in the view
- Update layout at run time



Tatzgern et al. (2014)

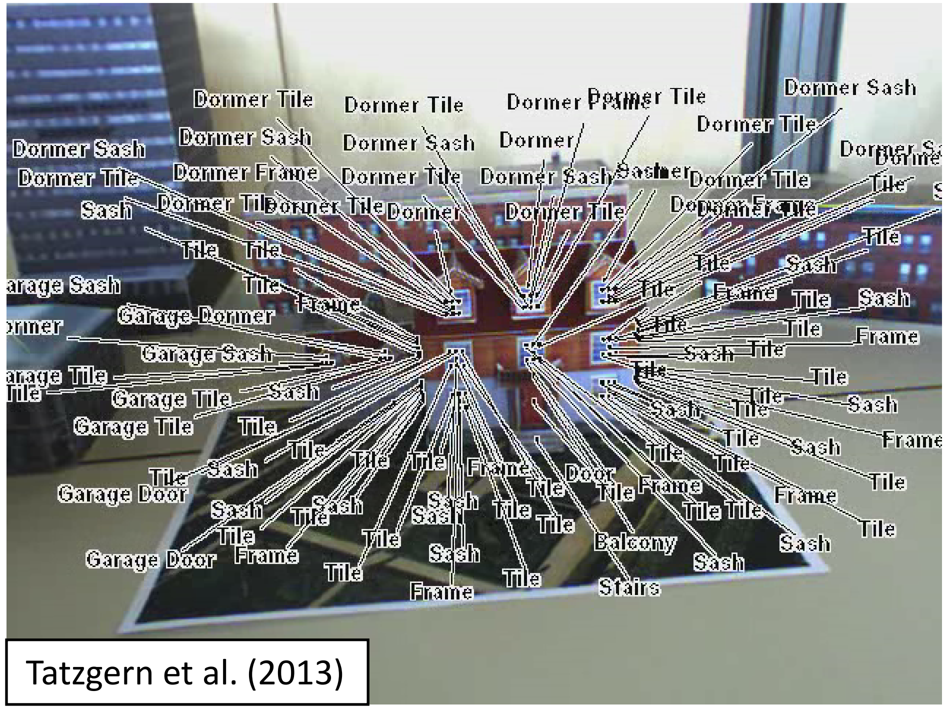
X-ray Vision

- See through structures
- Typically uses part of the video as context

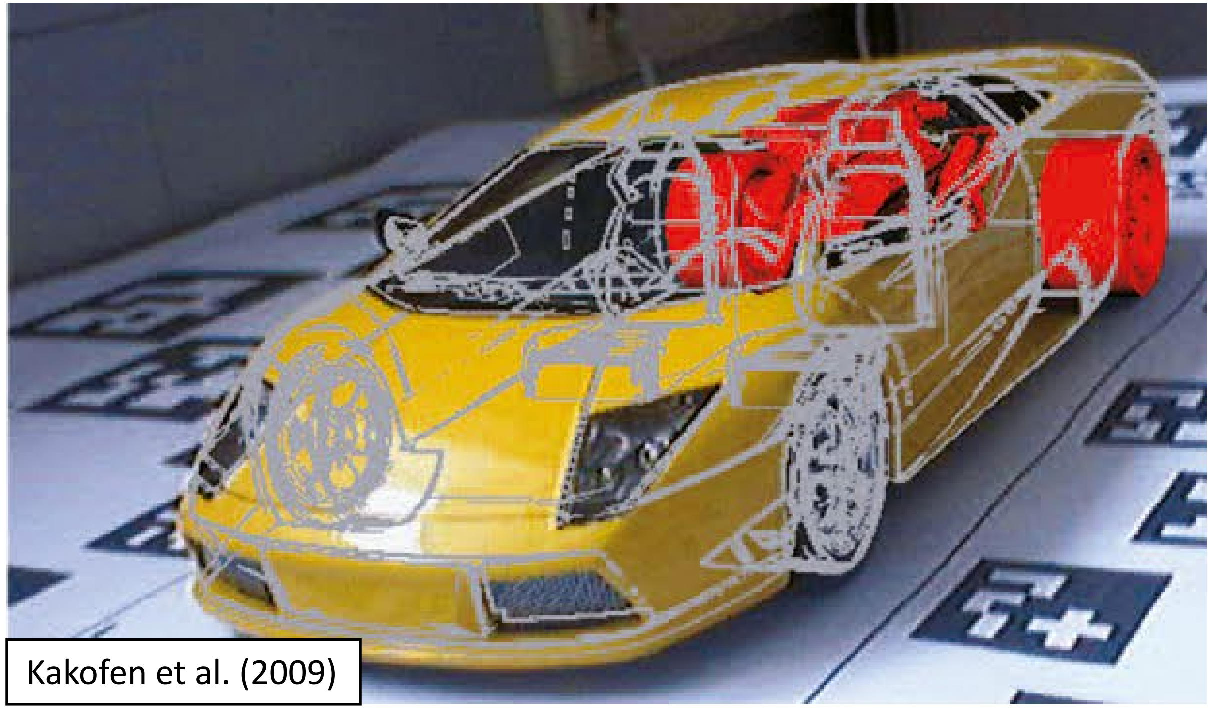
Kakofen et al. (2009)

Visual Clutter

- Data overload can easily lead to visual clutter and an unreadable visualization



Tatzgern et al. (2013)



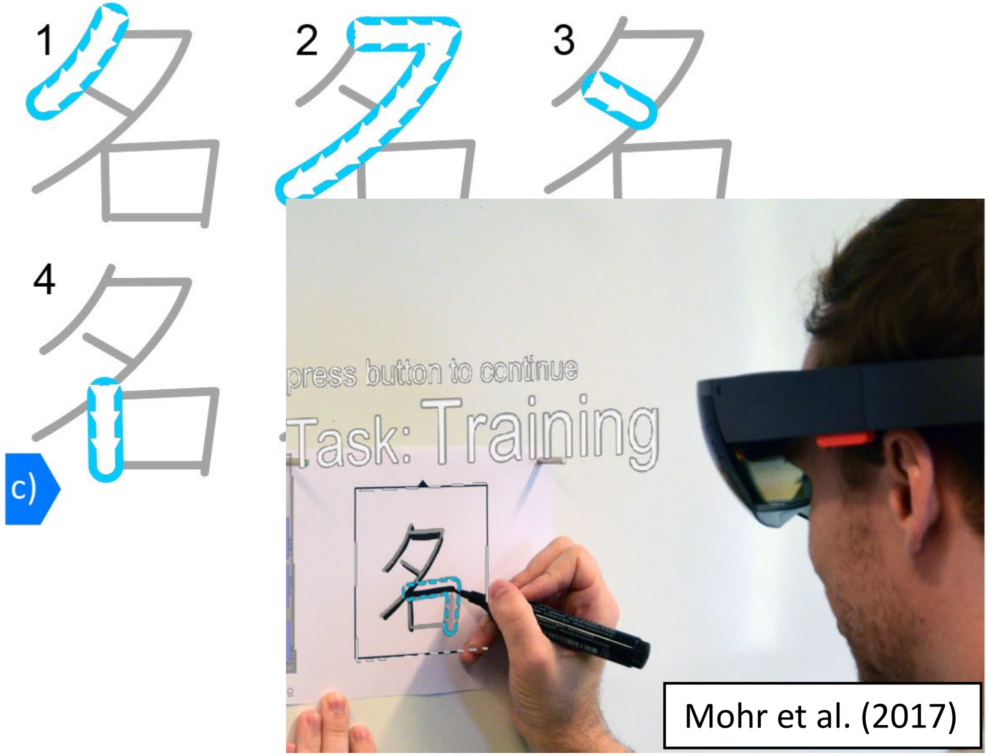
Kakofen et al. (2009)

Filtering

- Knowledge-based Filter
- Spatial Filter
- Hybrid Filter

Knowledge-based Filter

- Use knowledge about data, such as tasks/subtasks, prioritized search criteria or similarities in the data to filter

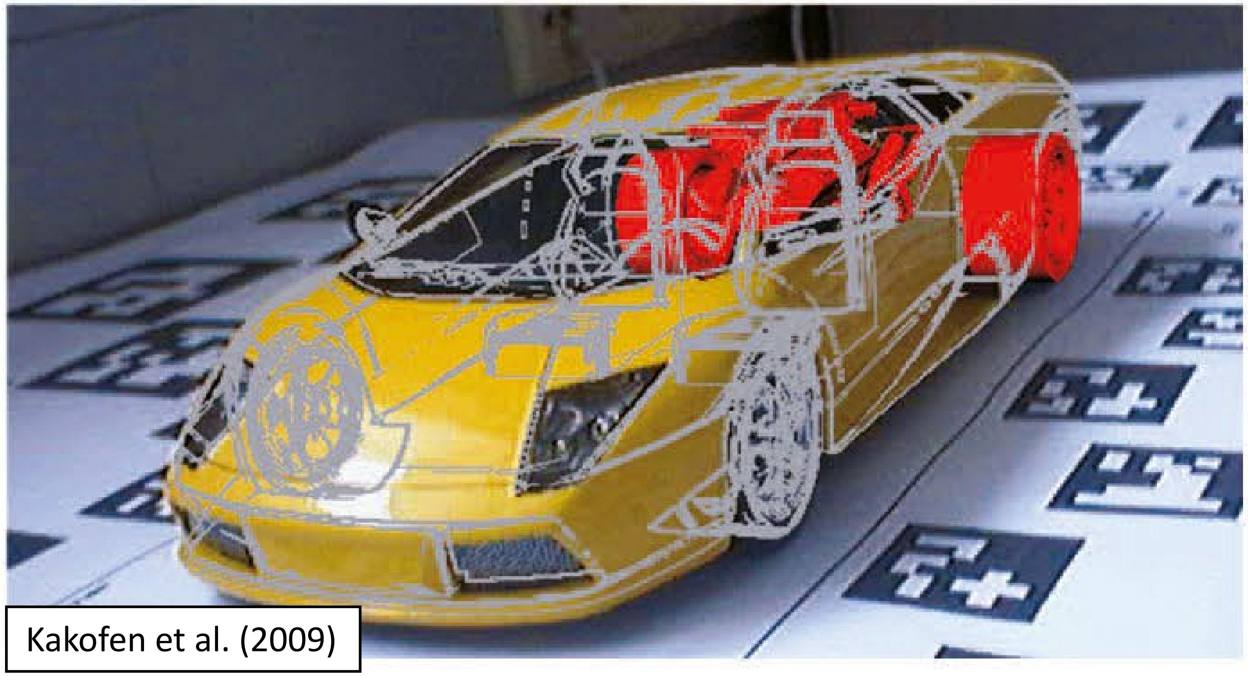


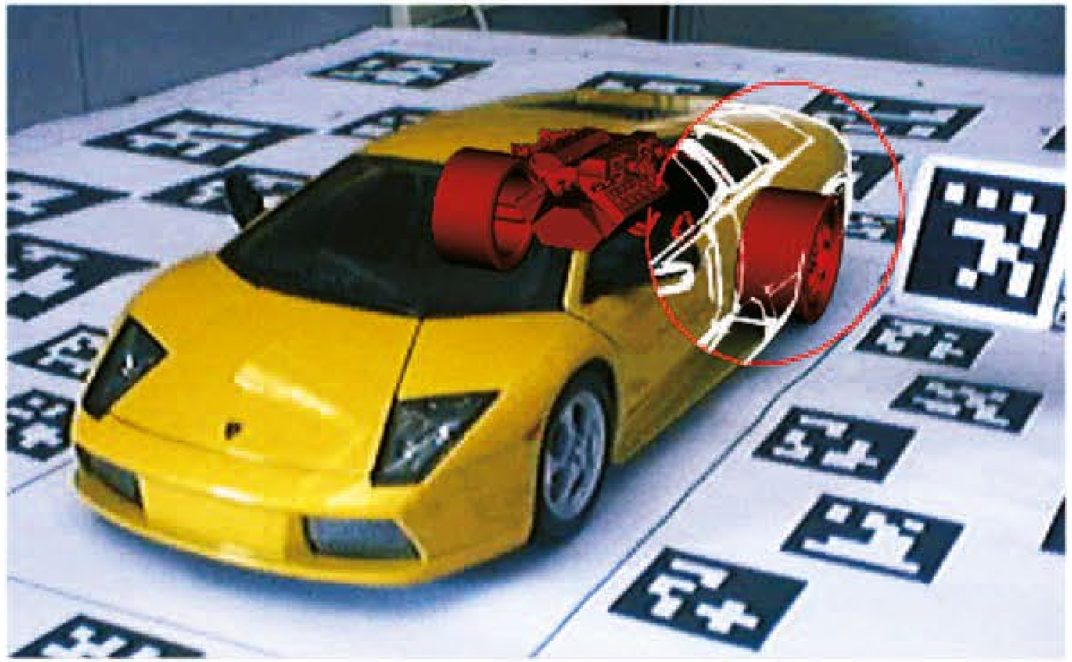
Mohr et al. (2017)

Tatzgern et al. (2016)

Spatial Filter

- Spatial filter filters data based on distance, or a region specified by a magic lens

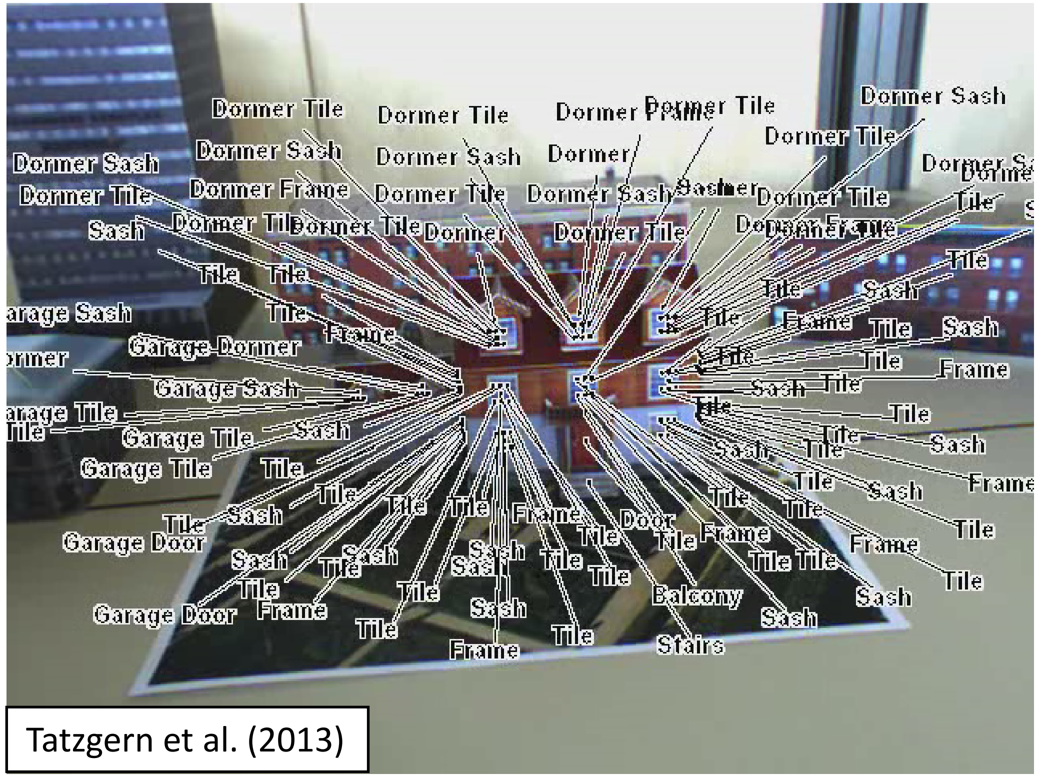




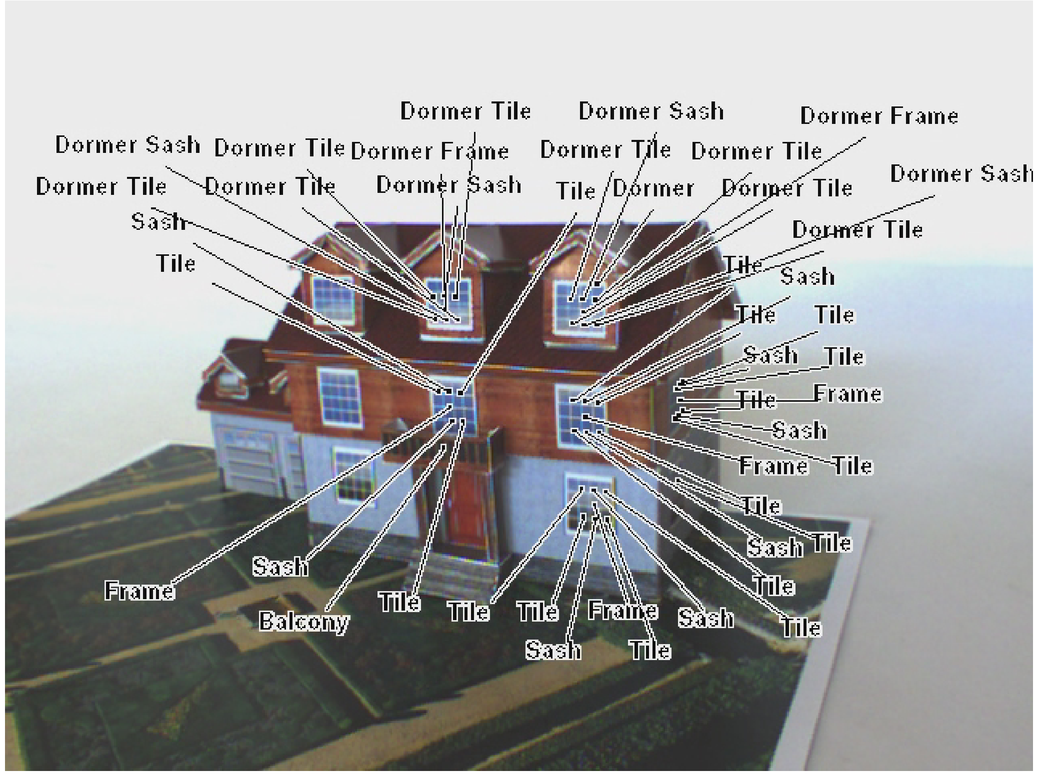
Kakofen et al. (2009)

Spatial Filter Issues

- Localized filter can lead to unbalanced amount of data due to missing data or regional data overload and data can group in a single region

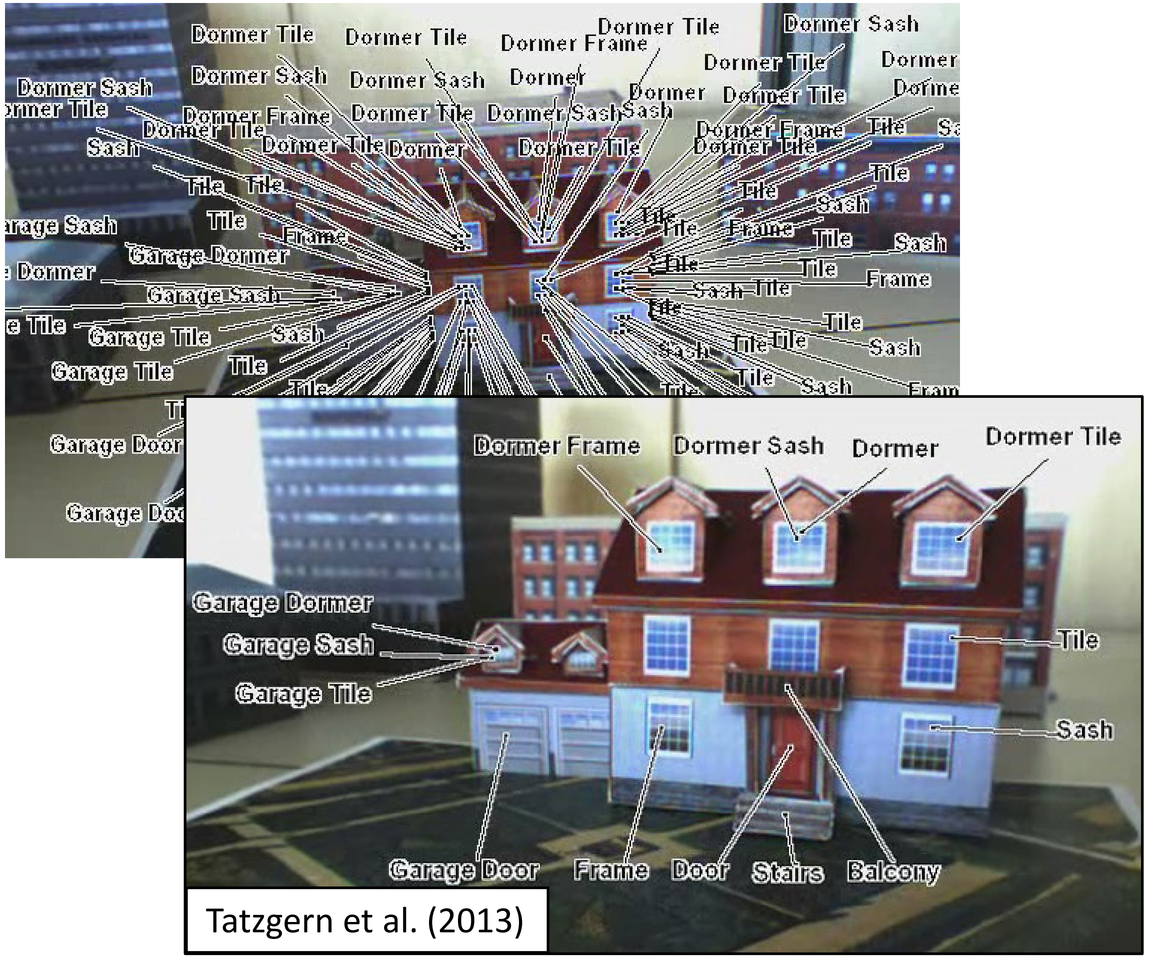


Tatzgern et al. (2013)



Hybrid Filter – Compact Visualization

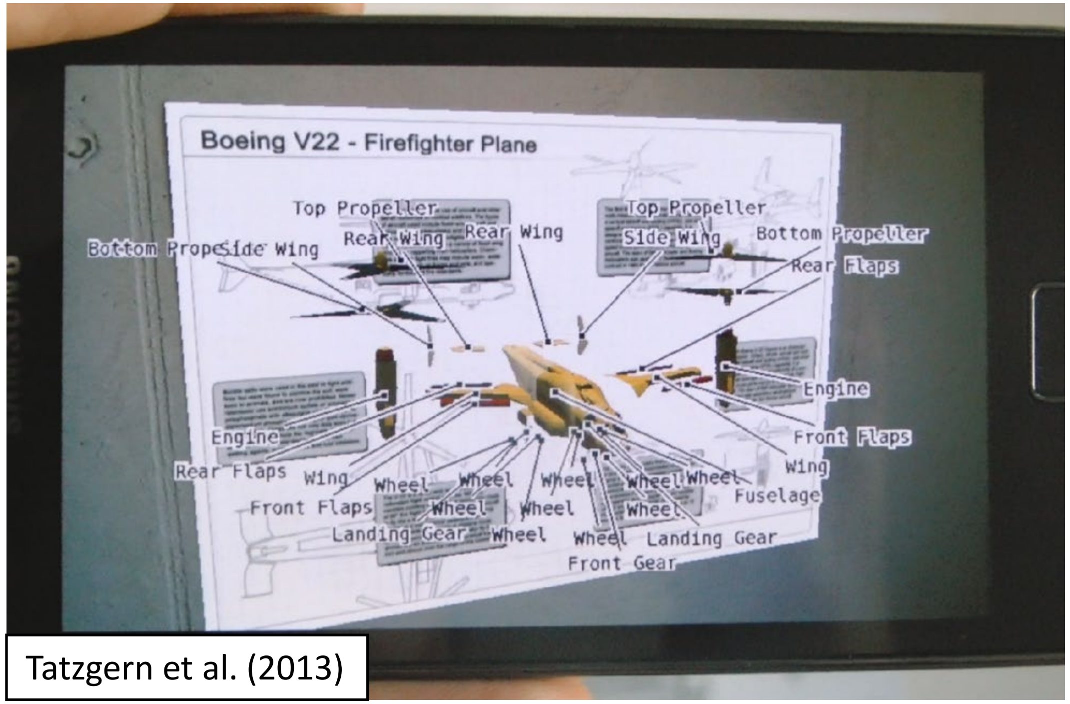
- Analyse data for similarities, e.g., underlying 3D shape, similarities in labels, etc. and create clusters
- Optimize selection of representative elements based on criteria such as the available screen estate and the current viewpoint of the user



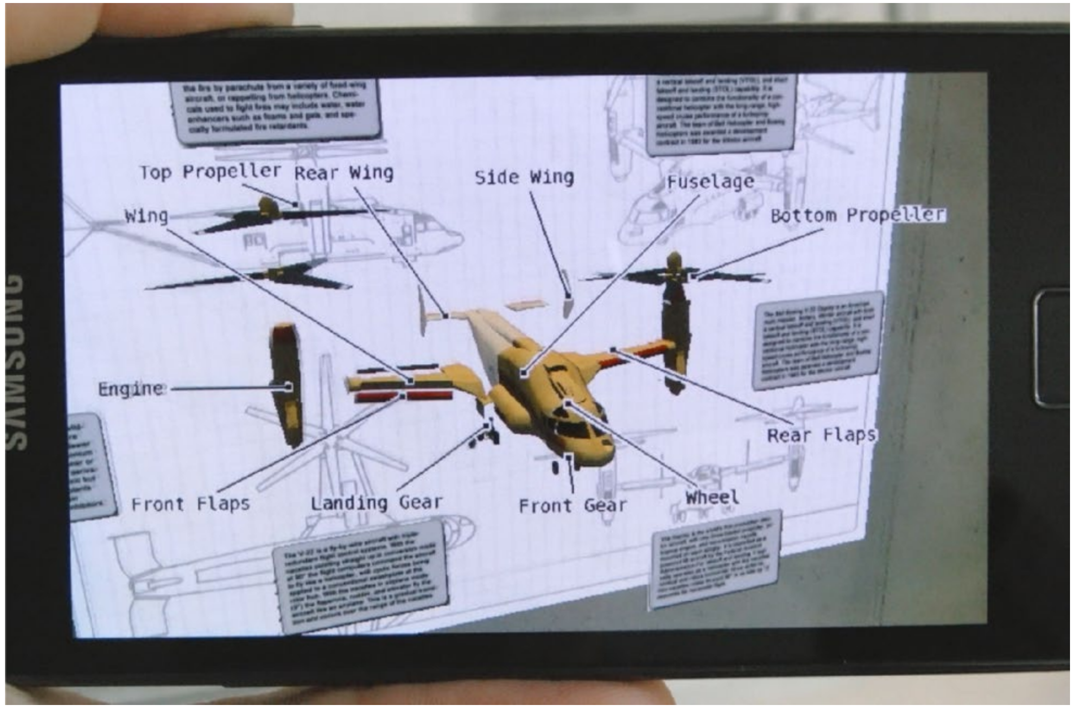
Tatzgern et al. (2013)

Hybrid Filter – Compact Visualization

- The method can also be applied to other visualizations, such as explosion diagrams



Tatzgern et al. (2013)



Perceptual Issues

- A short overview of perceptual issues of visualization with a focus on issues that AR visualizations and applications typically face
- Visual clutter
- **Temporal coherence**
- Registration errors
- Visual interference
- Viewport of scene

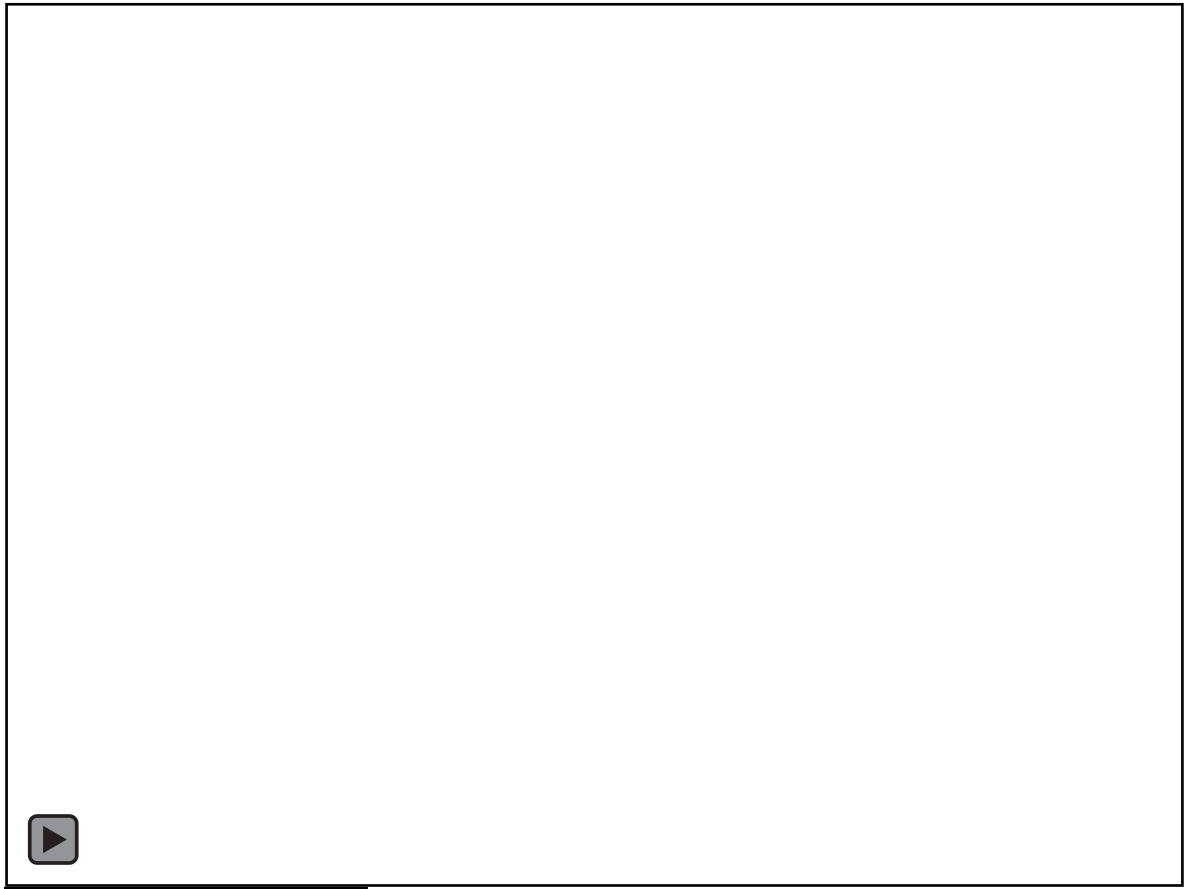
Temporal Coherence

- Visualizations must be stable and avoid undesired distractions of the viewer
- In AR, camera is always in motion
  - Viewpoint changes
  - Shaky hand / head
- Scene analysis can also cause distracting visual artifacts, e.g.,
  - when extracting occluding features from the video
  - when annotating objects that are not continuously recognized in the view

Tatzgern et al. (2014)

Strategy: Animation

- Animate changes so that users can follow
- Very straightforward, but animations can be distracting when they are too frequent



Tatzgern et al. (2014)

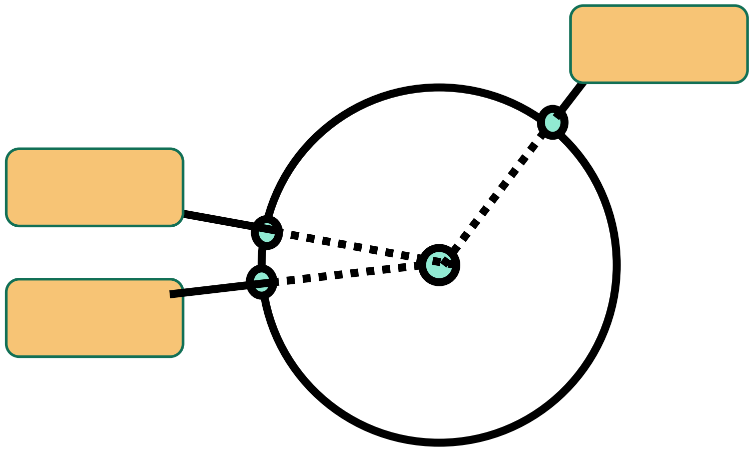
Strategy: Hysteresis

- Hysteresis delays updates to avoid high frequent changes
- Perform changes to the visualization only when they are stable for several frames, e.g.,
  - when a better layout has been found
  - Object has been safely detected / lost

Tatzgern et al. (2014)

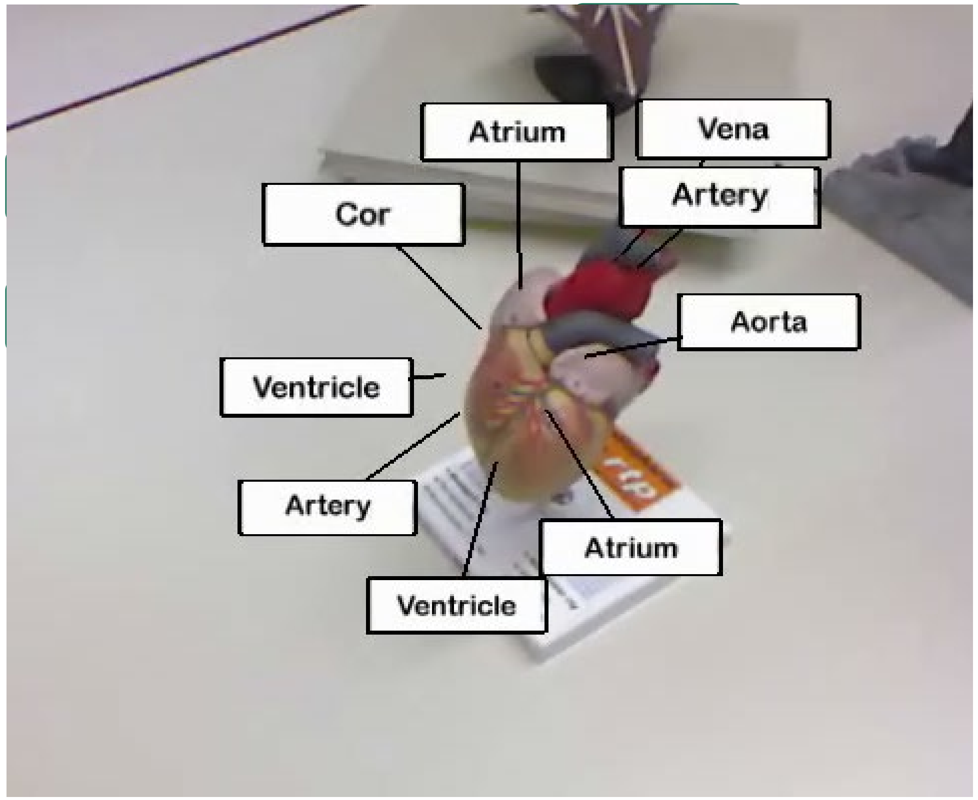
Strategy: Hedgehog Labeling

- We redesigned the labeling algorithm to use radial 3d labeling
- Avoid changes due to crossing lines by using radial layout
  - → No crossing lines also during viewpoint changes
- Reduces degrees-of-freedom by moving only along “poles”



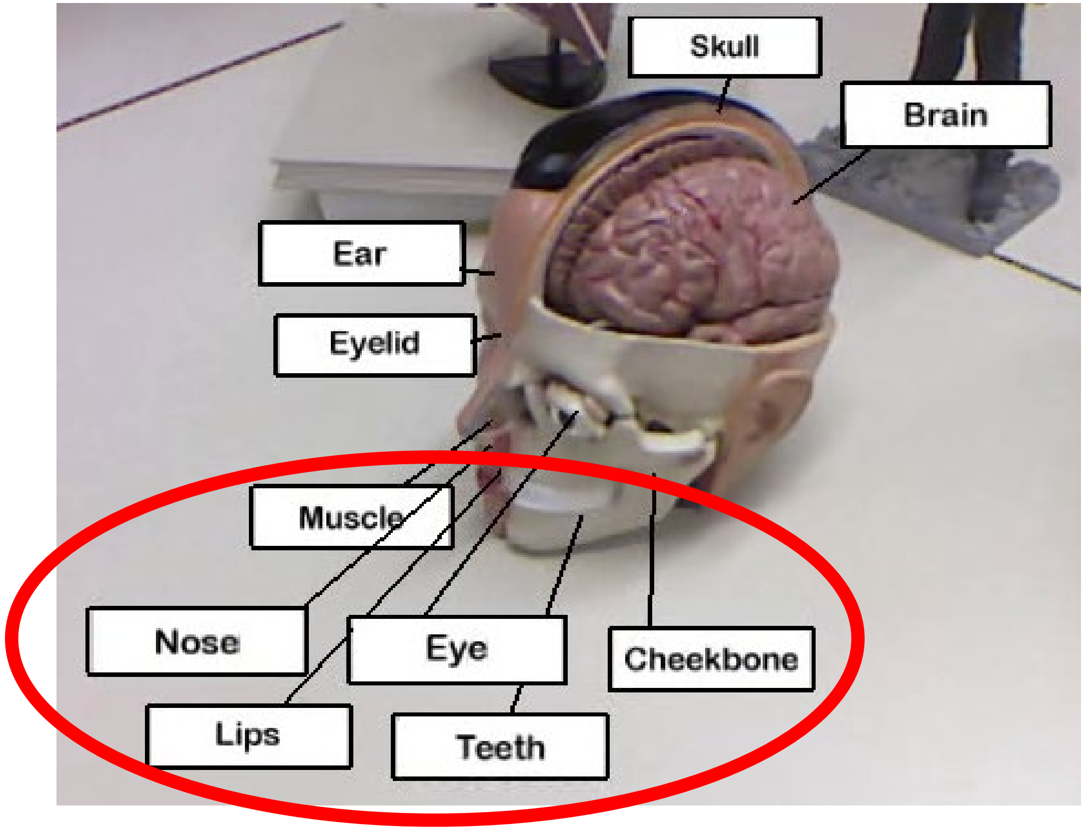
Strategy: Radial 3D Labelling

- We redesigned the labeling algorithm to use radial 3d labeling
- Avoid changes due to crossing lines by using radial layout
  - → No crossing lines also during viewpoint changes
- Reduces degrees-of-freedom by moving only along “poles”



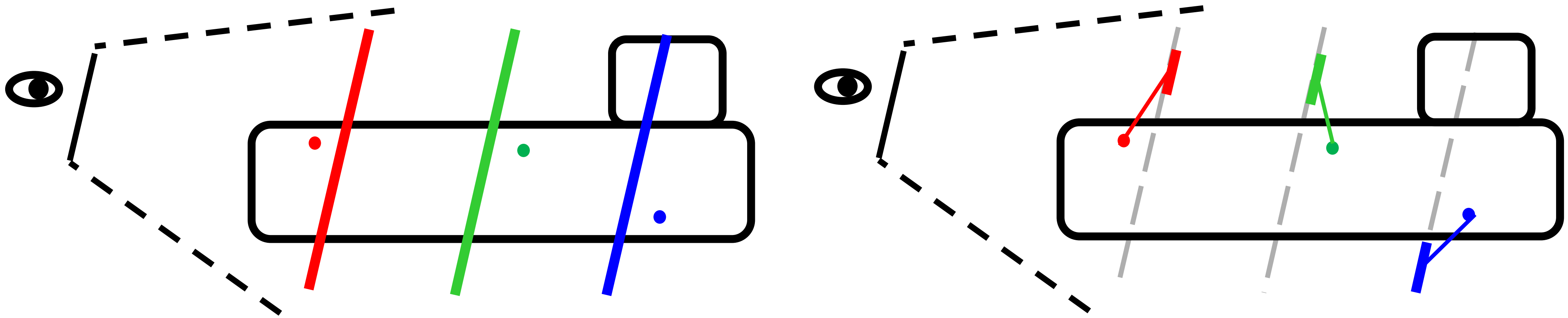
Strategy: Radial 3D Labeling

- We redesigned the labeling algorithm to use radial 3d labeling
- Avoid changes due to crossing lines by using radial layout
  - → No crossing lines also during viewpoint changes
- Reduces degrees-of-freedom by moving only along "poles"
  - Can lead to clustering / stacking of labels



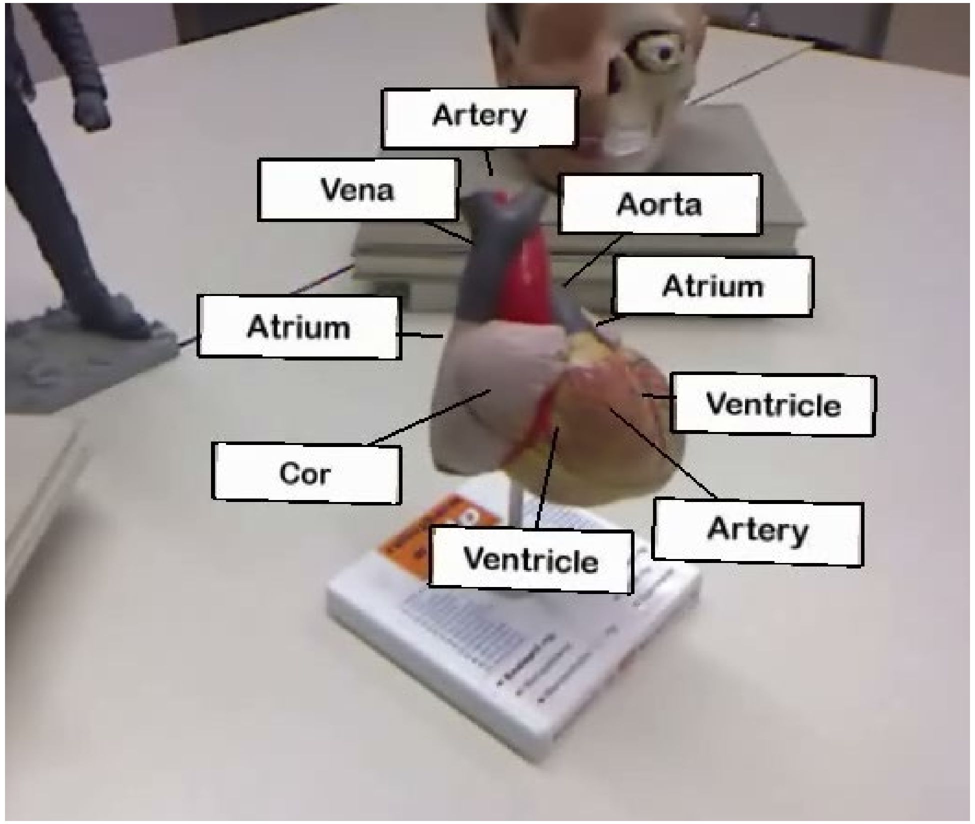
Strategy: Plane-based 3D Labeling

- Planes are defined parallel to viewing plane
- Each label is assigned to the closest plane
- Labels move only in their plane
- Temporal coherence: freeze planes and labels once optimized



Strategy: Plane-based 3D Labeling

- Labels frozen in planes are prone to occlusions, but
  - Depth ordering provides additional depth cues
  - Occlusions can easily be resolved via viewpoint changes
- Once the view of the layout degrades (e.g., angle too large) switch to new layout

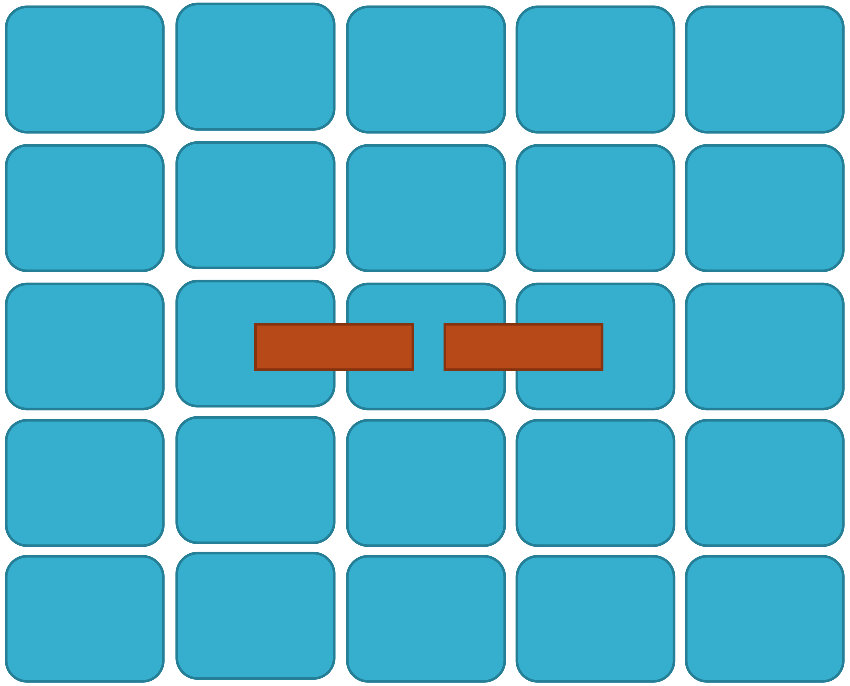


Perceptual Issues

- A short overview of perceptual issues of visualization with a focus on issues that AR visualizations and applications typically face
- Visual clutter
- Temporal coherence
- **Registration errors**
- Visual interference
- Viewport of scene

Registration Errors

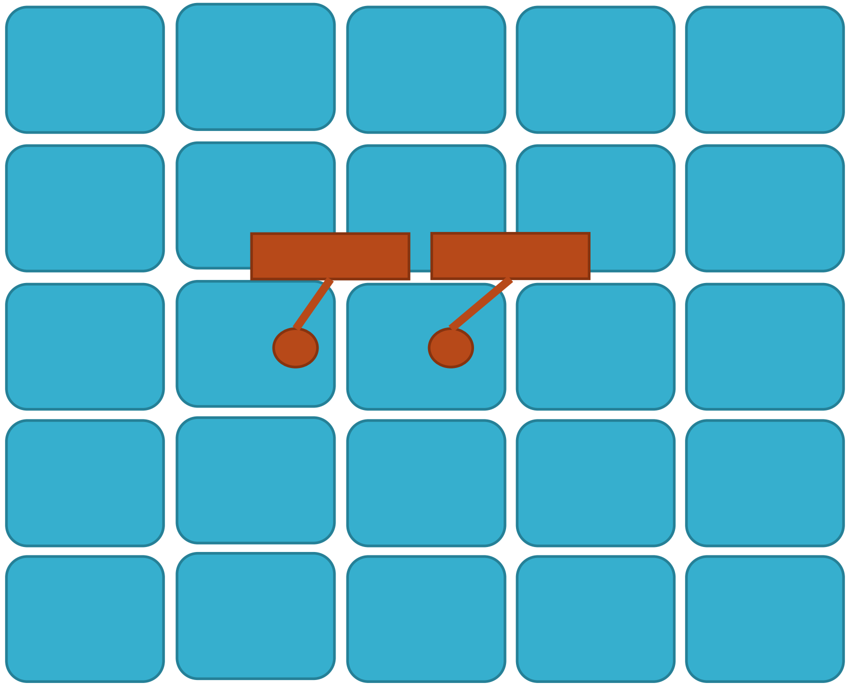
- Registration errors lead to misalignment of AR visualizations with the real world
- Internal labels annotating objects become ambiguous
- Solution: switch to external labels as the anchor point has a smaller footprint that may be more tolerant towards errors



Based on Coelho et al. (2004)

Solution: Change Representation

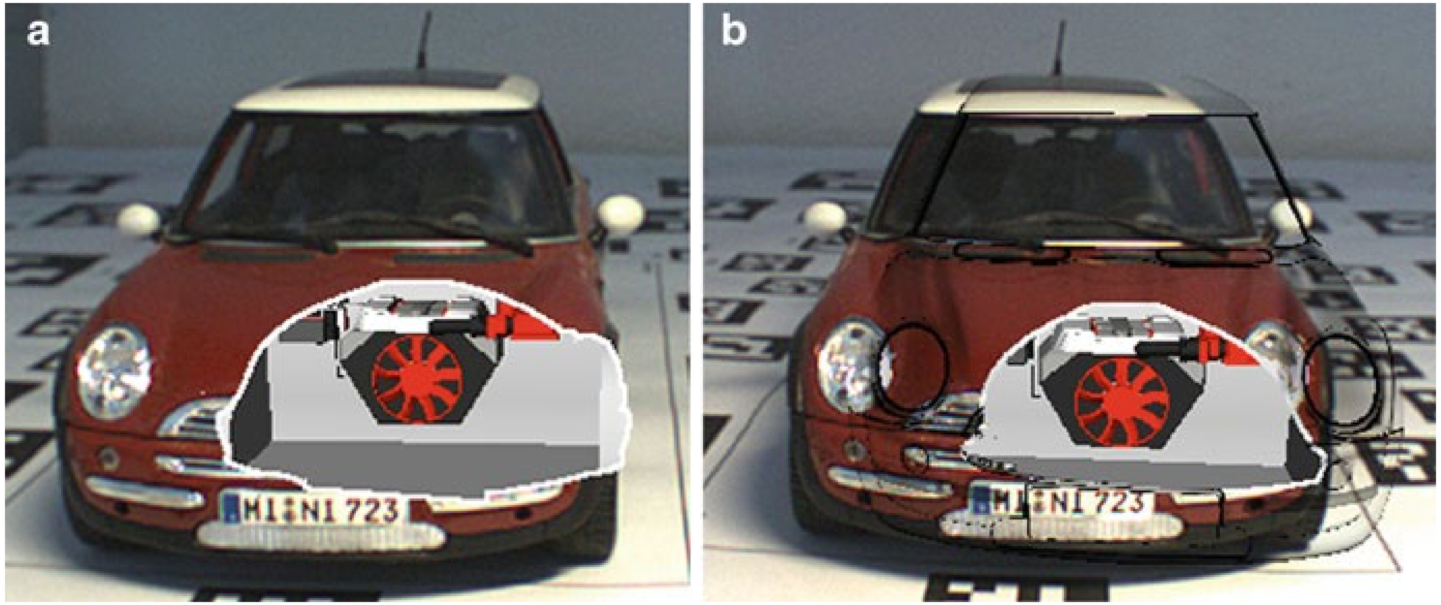
- Switch to external labels as the anchor point has a smaller footprint that may be more tolerant towards errors



Based on Coelho et al. (2004)

Solution: Provide additional Context

- A visualization enhances a real-world object and provides additional information, but is not registered correctly
- Providing additional context may help users to understand the spatial relationships



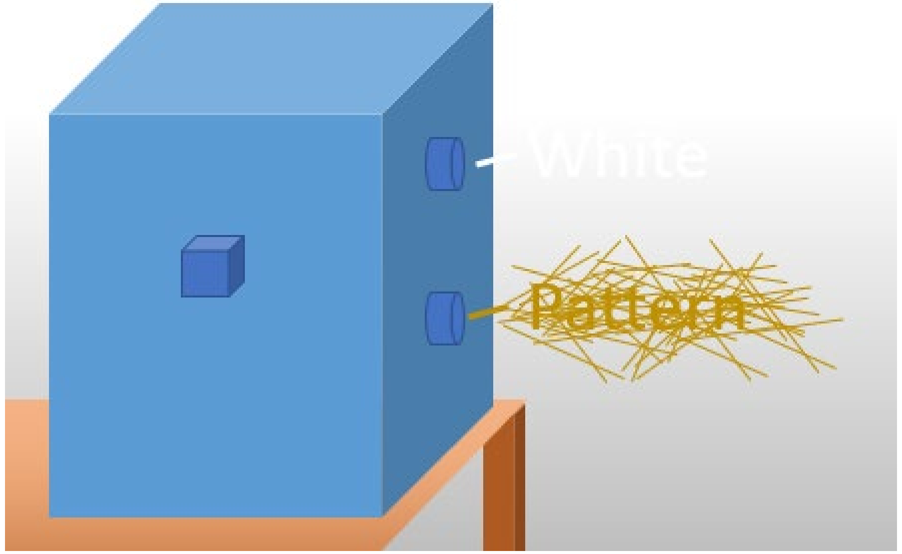
Kalkofen et al. (2011)

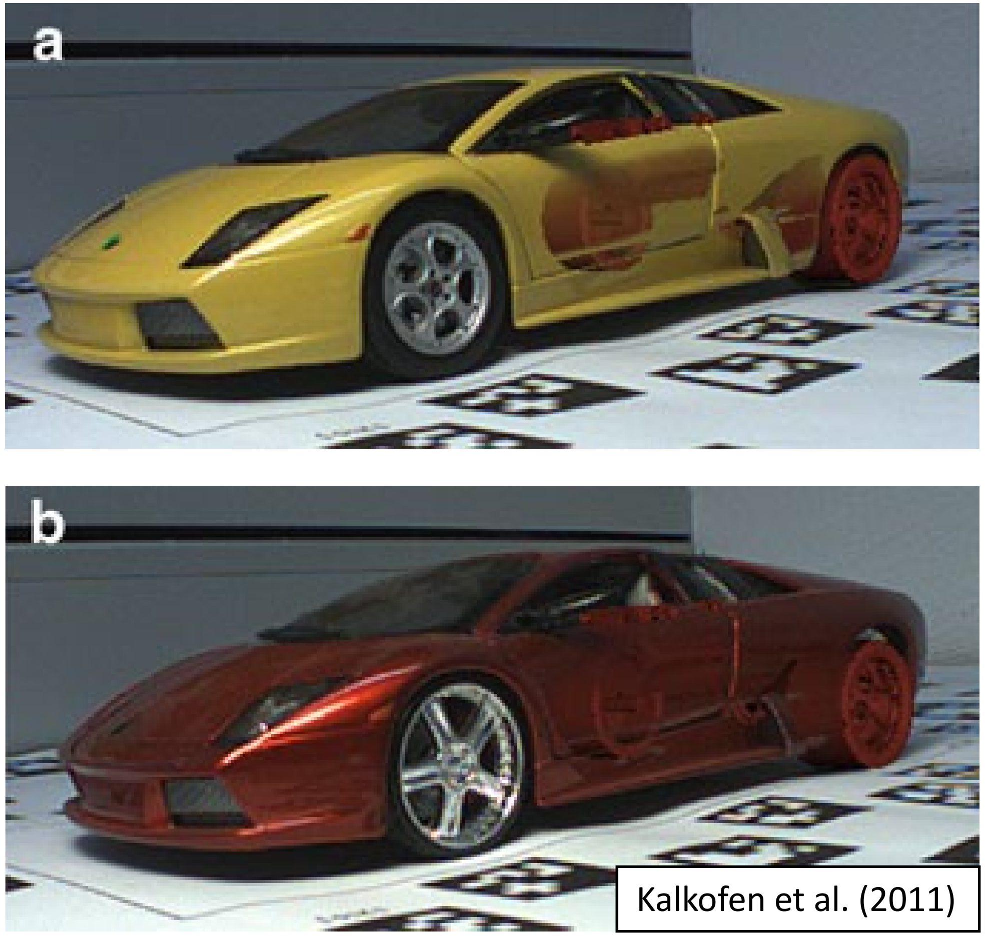
Perceptual Issues

- A short overview of perceptual issues of visualization with a focus on issues that AR visualizations and applications typically face
- Visual clutter
- Temporal coherence
- Registration errors
- **Visual interference**
- Viewport of scene

Visual Interference

- AR augmentations interfere with the real-world background
- Contrast problem that reduces legibility and comprehensibility

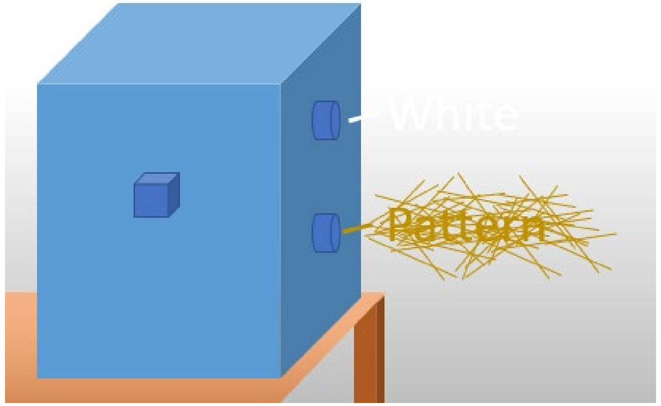




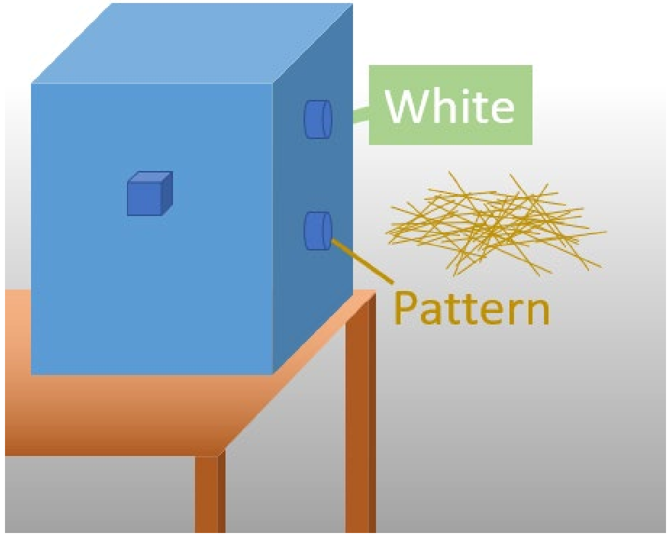
Kalkofen et al. (2011)

Solution: Adaptive Visualization

- Adapt visualization to scene background
  - Adapt contrast by changing appearance
  - Avoid placement in regions of low contrast by moving visualizations



- Style adaptations only work well for video see-through devices
- Issues are aggravated by optical see-through devices due to transparent display



Optical See-through HMDs

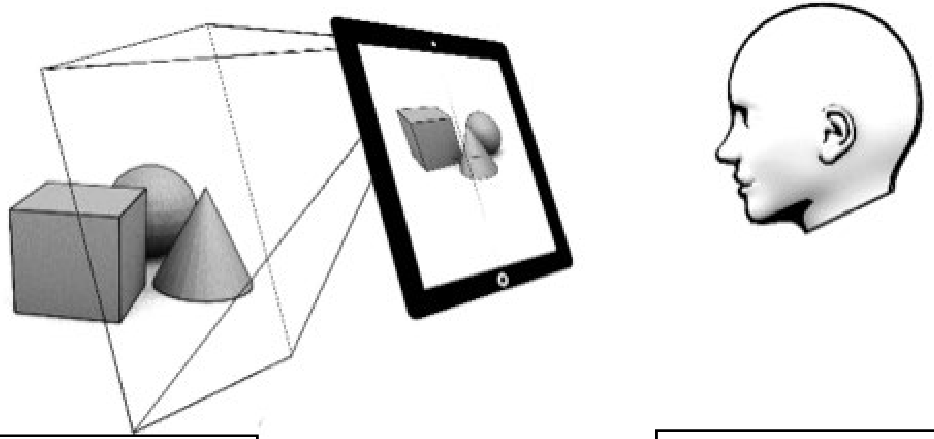
- Adaptation only works to a certain degree due to
  - additive color generation behavior
  - Inability of displays to occlude real-world
  - Inability of displays to render black
- Issues lead to
  - Users seeing uintended colors on display
  - washed out colors due to background illumination
  - Lack of contrast
- Ongoing research topic
  - Alternative display designs
  - Color calibration methods to optimize presented colors

Perceptual Issues

- A short overview of perceptual issues of visualization with a focus on issues that AR visualizations and applications typically face
- Visual clutter
- Temporal coherence
- Registration errors
- Visual interference
- **Viewport of scene**

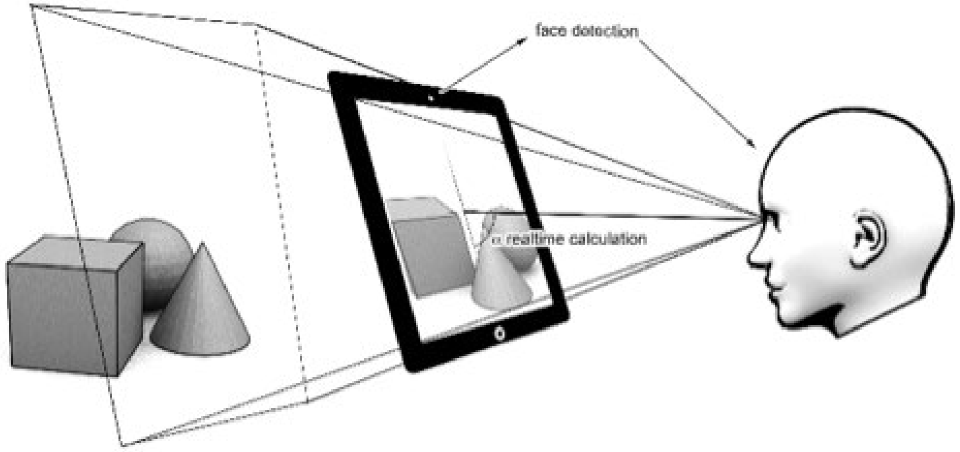
Viewport of Scene

- When using AR on handheld devices, there is a viewport mismatch between the user‘s eyes and the device camera
- The render technique „user-perspective rendering“ solves this issue by rendering the real-world view as if the display was transparent



Mohr et al. (2017)

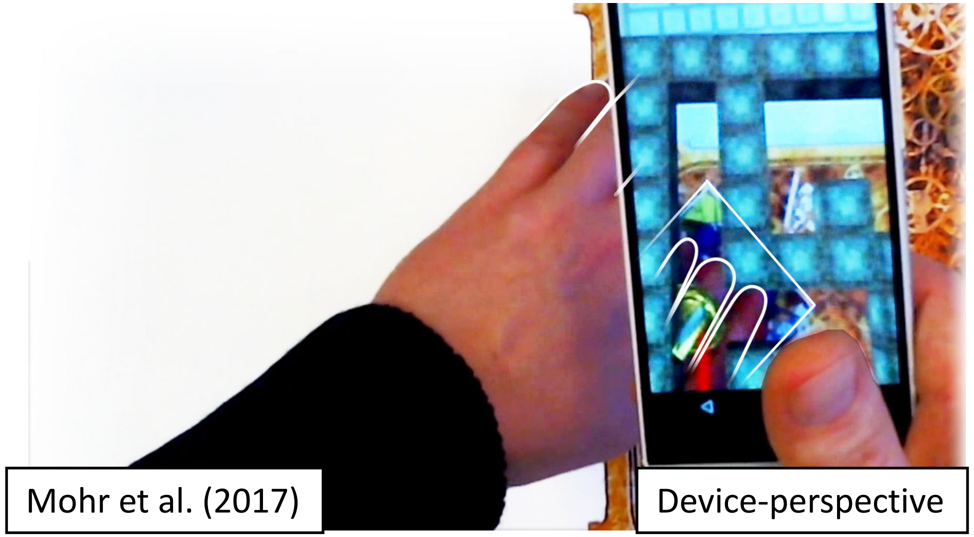
Device-perspective



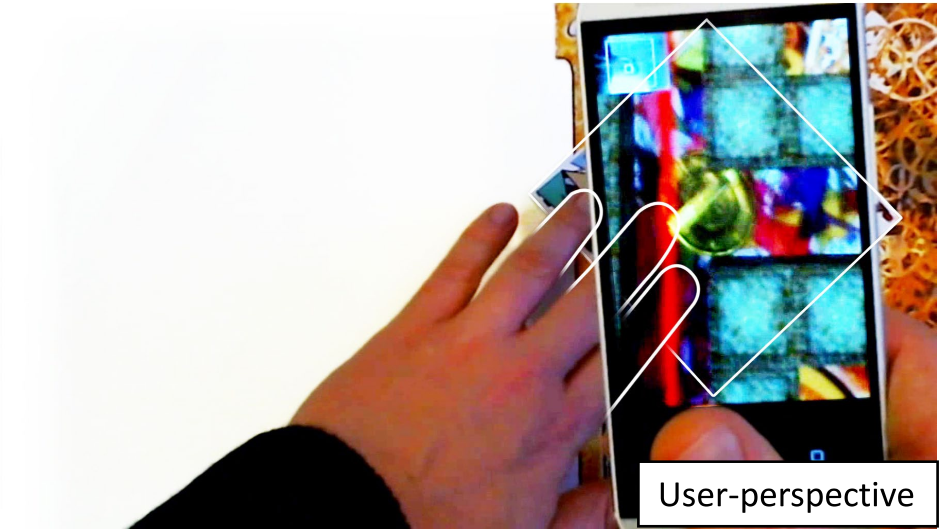
User-perspective

Viewport of Scene

- When using AR on handheld devices, there is a viewport mismatch between the user‘s eyes and the device camera
- The render technique „user-perspective rendering“ solves this issue by rendering the real-world view as if the display was transparent



Mohr et al. (2017)

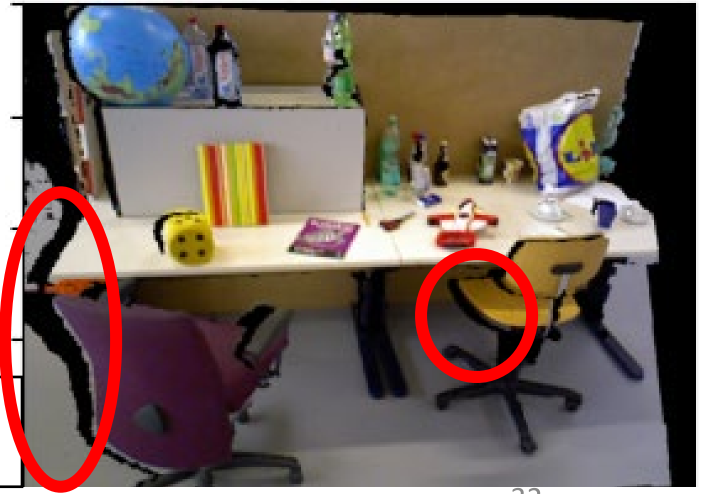


Optimal User-perspective Rendering

- Track eye position / head of user relative to display
- Create a novel view of the scene by using a textured 3d model of the real-world scene
  - Reconstructed or image-based rendered
  - Render a novel viewpoint based on tracked head position of user
- Best result, but computationally expensive and requires access to front and backfacing camera at the same time
- Not feasible for most mobile devices due to
  - Limited camera access and costly 3D reconstruction to fill in missing data

ground truth

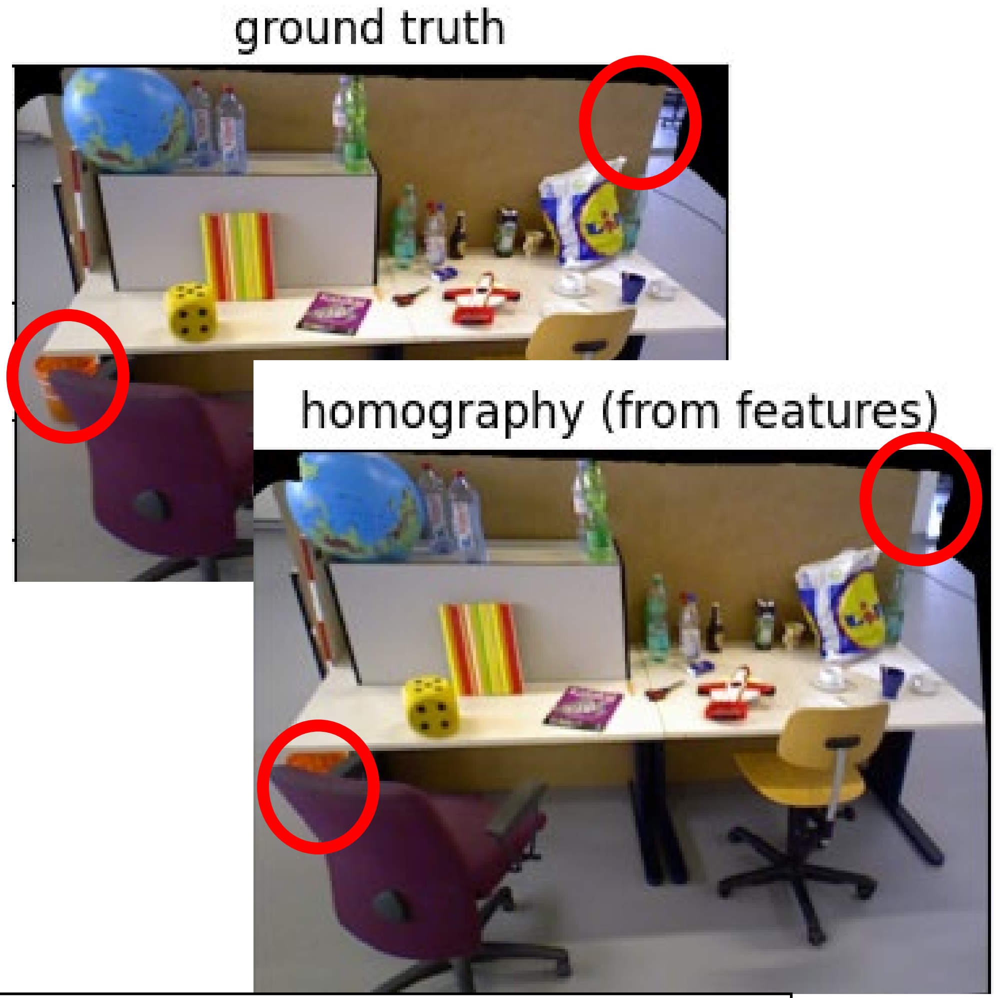
reprojection with holes



Implementation of Salzburg University of Applied Sciences using the TUM RGBD data set (Sturm et al. 2012).

Approximate User-perspective Rendering

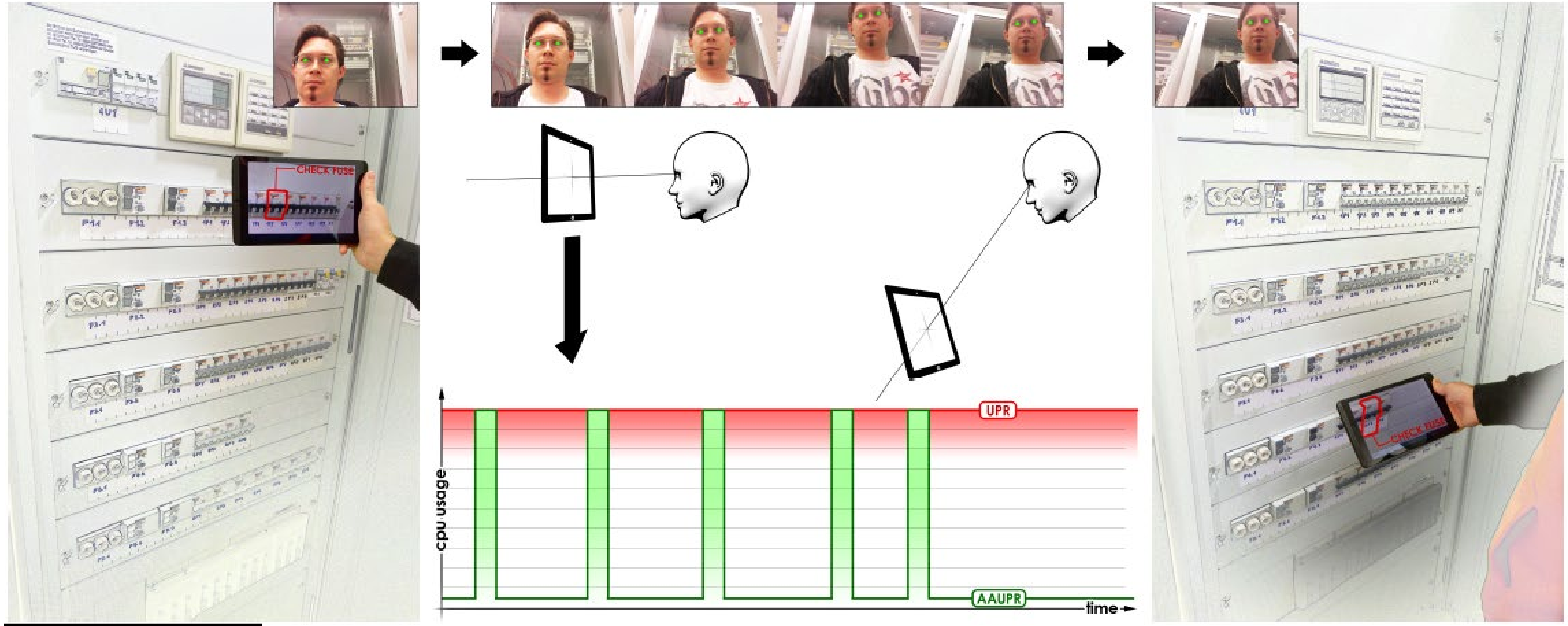
- Calibrate a fixed position for users relative to the display to avoid expensive head tracking
- Distort video of world using a homography
  - Good results for scenes consisting of a (approx. planar geometry)
  - Otherwise occlusion artifacts/distortions
- Fast, but fixed viewpoint calibration is a severe limitation of the approach



Implementation of Salzburg University of Applied Sciences using the TUM RGBD data set (Sturm et al. 2012).

Adaptive User-perspective Rendering

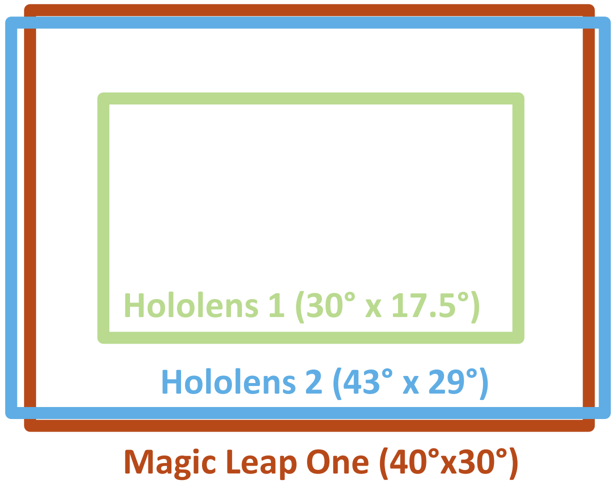
- Adaptive: use expensive head tracking only when users head moves beyond a threshold relative to the device



Mohr et al. (2017)

HMD: Limited Field of View

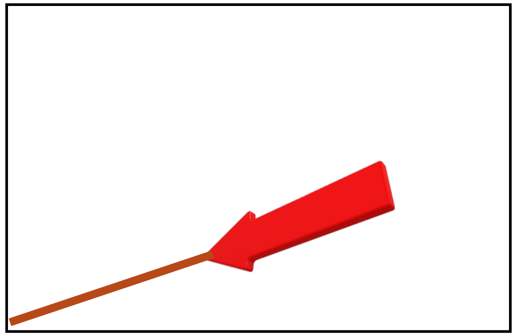
- HMDs do not suffer from the viewpoint mismatch, but a generally very small field of view
- Only a small portion of the human field of view is covered, which leads to human users having to search for the virtual augmentations
- Attention guidance / offscreen visualizations are used to compensate for these issues

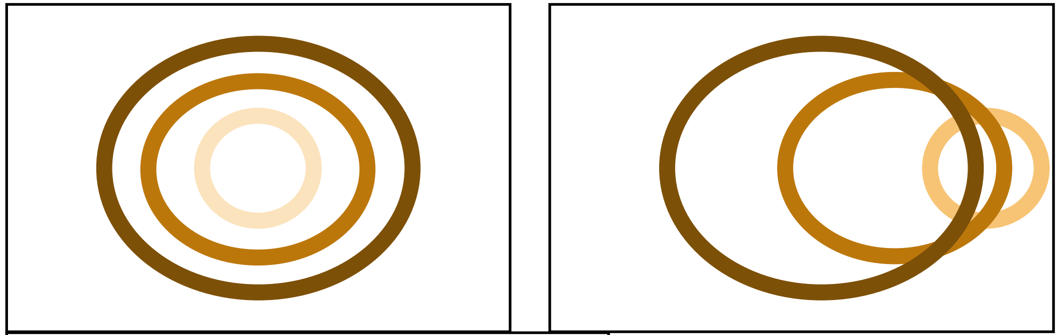


Attention Guidance

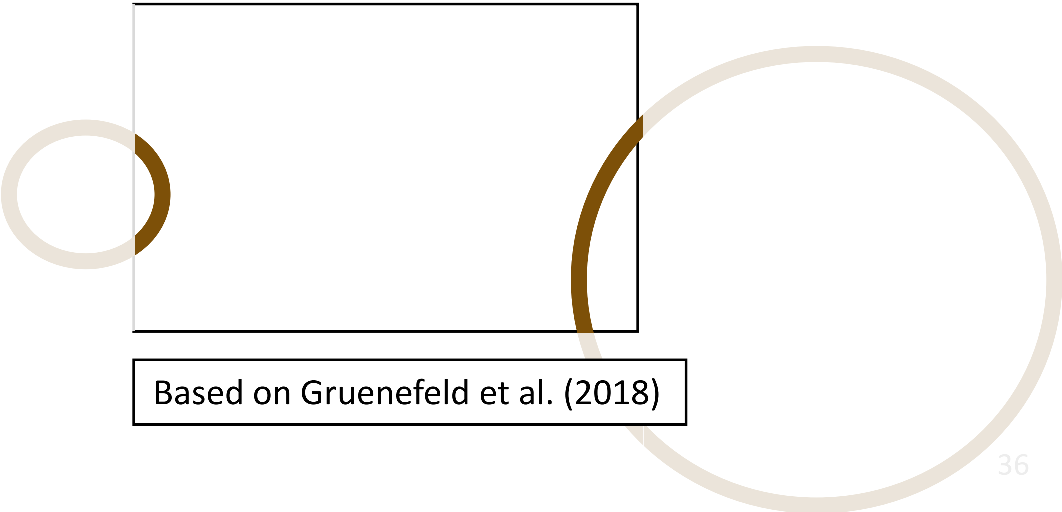
- Visual representations
  - Arrow (+ rubber band line)
  - Attention funnel
  - Halos
  - Radar-like visualizations

- Audio
- Haptic feedback
  - Requires additional hardware for vibrotactile feedback





Based on Schwerdtfeger et al. (2011)



Based on Gruenefeld et al. (2018)

Publications

Peter Mohr, Shohei Mori, Tobias Langlotz, Bruce Thomas, Dieter Schmalstieg, and Denis Kalkofen. Mixed Reality Light Fields for Interactive Remote Assistance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20). https://doi.org/10.1145/3313831.3376289

Tatzgern, M., Kalkofen, D., Grasset, R., & Schmalstieg, D. (2014). Hedgehog Labeling: View Management Techniques for External Labels in 3D Space. *IEEE Virtual Reality*. 10.1109/VR.2014.6802046

Kalkofen, D., Mendez, E., & Schmalstieg, D. (2009). Comprehensible visualization for augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, *15*(2), 193–204. https://doi.org/10.1109/TVCG.2008.96

Tatzgern, M., Kalkofen, D., & Schmalstieg, D. (2013). Dynamic compact visualizations for augmented reality. *IEEE Virtual Reality (VR)*, 3–6. https://doi.org/10.1109/VR.2013.6549347

Mohr, P., Mandl, D., Tatzgern, M., Veas, E., Schmalstieg, D., & Kalkofen, D. (2017). Retargeting Video Tutorials Showing Tools With Surface Contact to Augmented Reality. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17, July*, 6547–6558. https://doi.org/10.1145/3025453.3025688

Tatzgern, M., Orso, V., Kalkofen, D., Jacucci, G., Gamberini, L., & Schmalstieg, D. (2016). Adaptive information density for augmented reality displays. *Proceedings - IEEE Virtual Reality*, 83–92. https://doi.org/10.1109/VR.2016.7504691

Coelho, E. M., MacIntyre, B., & Julier, S. J. (2004). OSGAR: A Scene Graph with Uncertain Transformations. *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, 6–15. https://doi.org/10.1109/ISMAR.2004.44

Kalkofen, D., Sandor, C., White, S., & Dieter, S. (2011). Visualization Techniques for Augmented Reality. In *Handbook of Augmented Reality* (pp. 65–98). http://link.springer.com/chapter/10.1007/978-1-4614-0064-6_3

Sturm, J., Engelhard, N., Endres, F., Burgard, W., & Cremers, D. (2012, October). A benchmark for the evaluation of RGB-D SLAM systems. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 573-580). IEEE.

Mohr, P., Tatzgern, M., Grubert, J., Schmalstieg, D., & Kalkofen, D. (2017). Adaptive user perspective rendering for Handheld Augmented Reality. *2017 IEEE Symposium on 3D User Interfaces, 3DUI 2017 - Proceedings*. https://doi.org/10.1109/3DUI.2017.7893336

Schwerdtfeger, B., Reif, R., Günthner, W. a., & Klinker, G. (2011). Pick-by-vision: There is something to pick at the end of the augmented tunnel. *Virtual Reality*, *15*(2–3), 213–223. https://doi.org/10.1007/s10055-011-0187-9

Gruenefeld, U., El Ali, A., Boll, S., & Heuten, W. (2018). Beyond Halo and Wedge: Visualizing Out-of-View Objects on Head-mounted Virtual and Augmented Reality Devices. *MobileHCI 2018*, *September*. https://doi.org/10.1145/3229434.3229437

Thank you!

enjoy Vienna

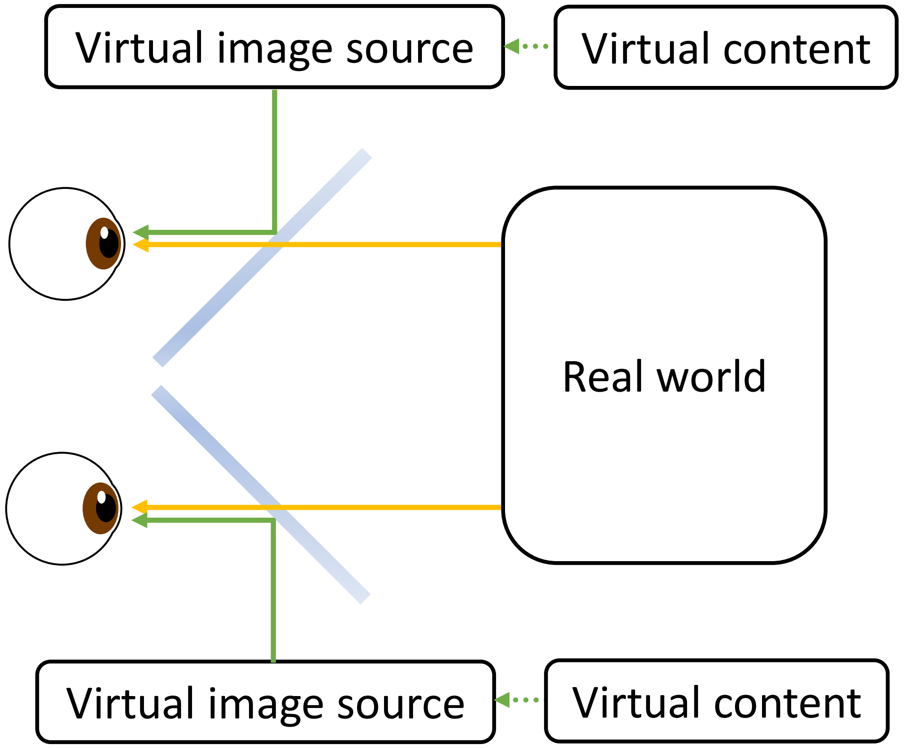
EUROGRAPHICS 2021

Displaying MR Environments

Christoph Ebner

Optical See-Through Displays

Basic working principle



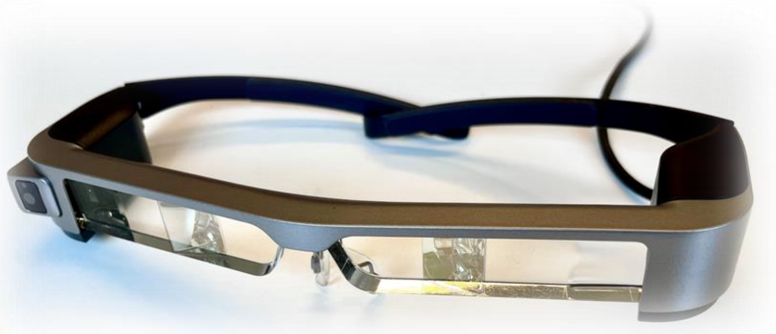
Example of a view through an OST display

Optical See-Through Displays

Examples

Microsoft HoloLens

Vuzix Blade



Epson Moverio BT-300

Optical See-Through Displays

Pros and Cons

Light from the real world is more or less unchanged

- Contrast and brightness
- Dynamic range
- Focus
- Latency

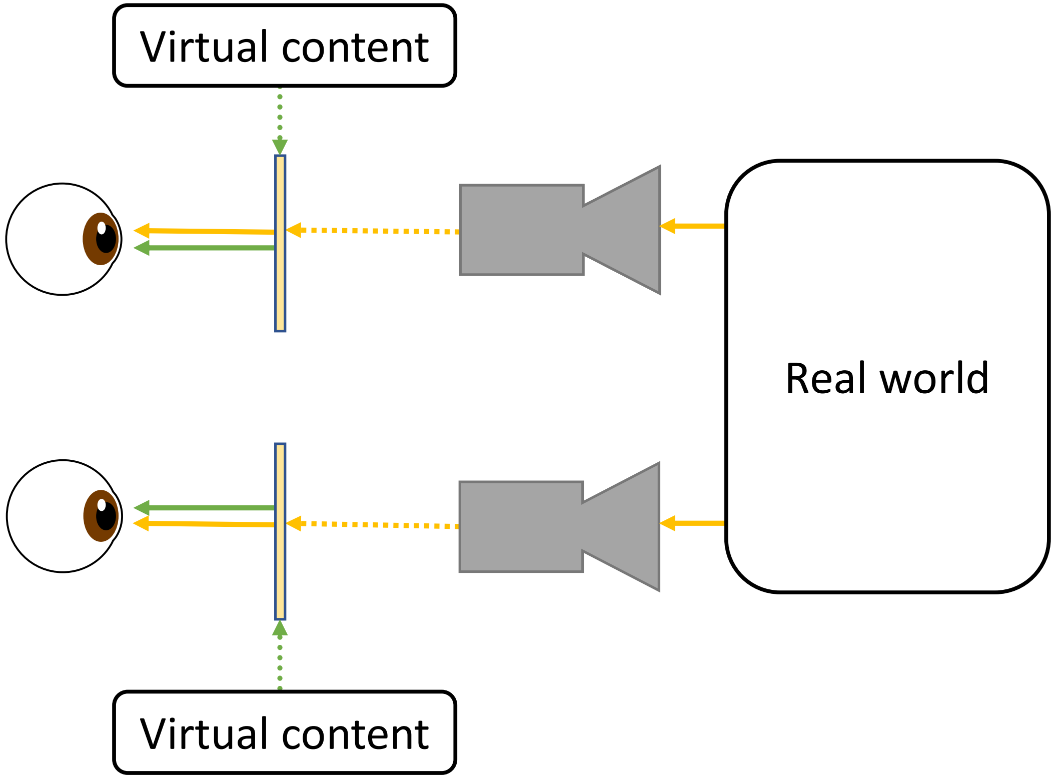
Unable to control real world light

- No occlusions of real objects

Small Field of View

Video See-Through Displays

Basic working principle

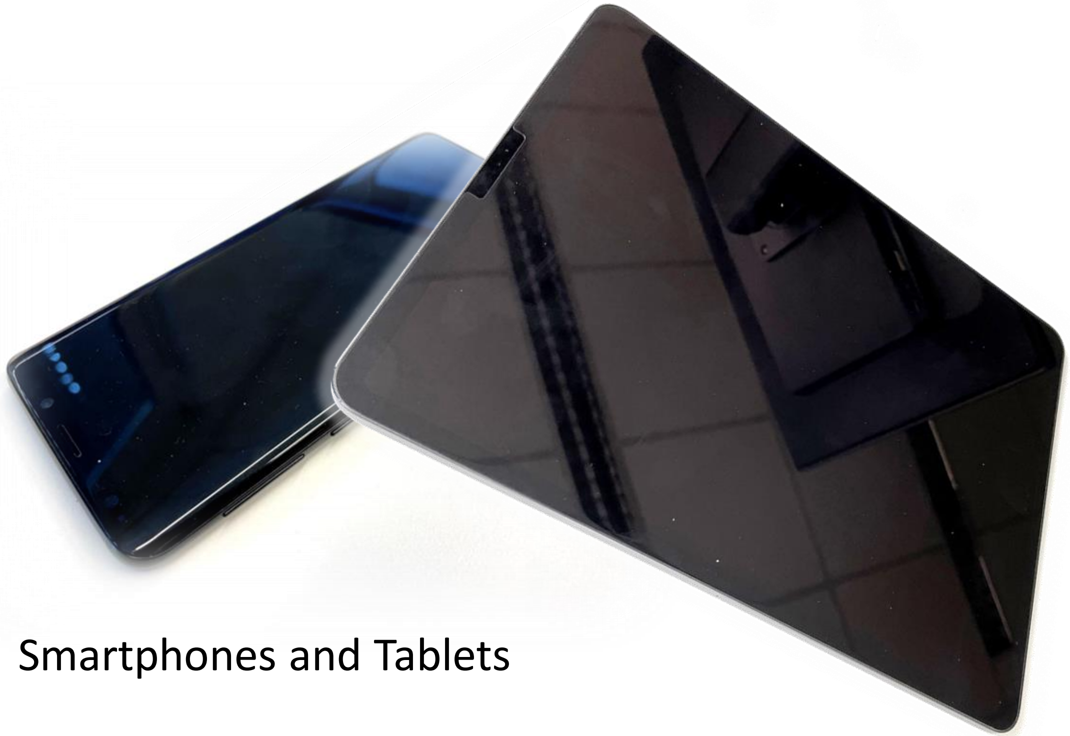


Example of a view through a VST display

Video See-Through Displays

Examples

HTC Vive Pro Eye



Smartphones and Tablets

Video See-Through Displays

Pros and Cons

Real world is perceived through camera stream

- Per-pixel occlusions
- Control of brightness and contrast

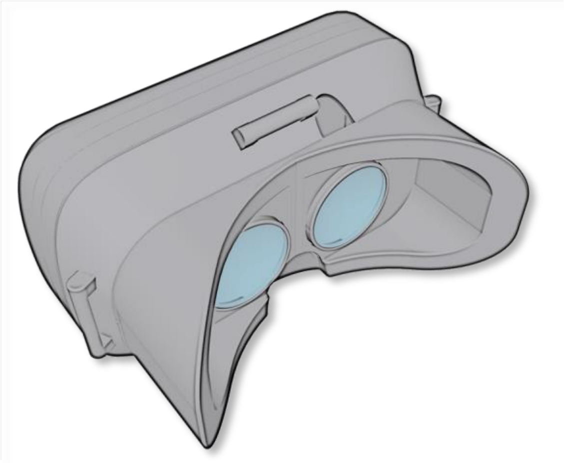
Larger Field of View

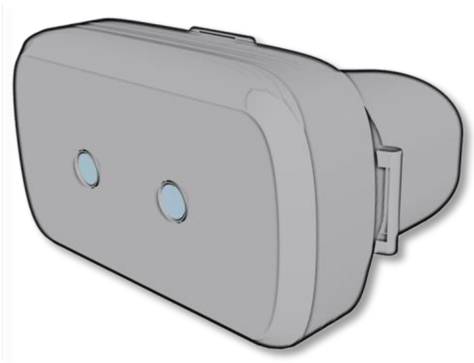
Real world displayed on screen

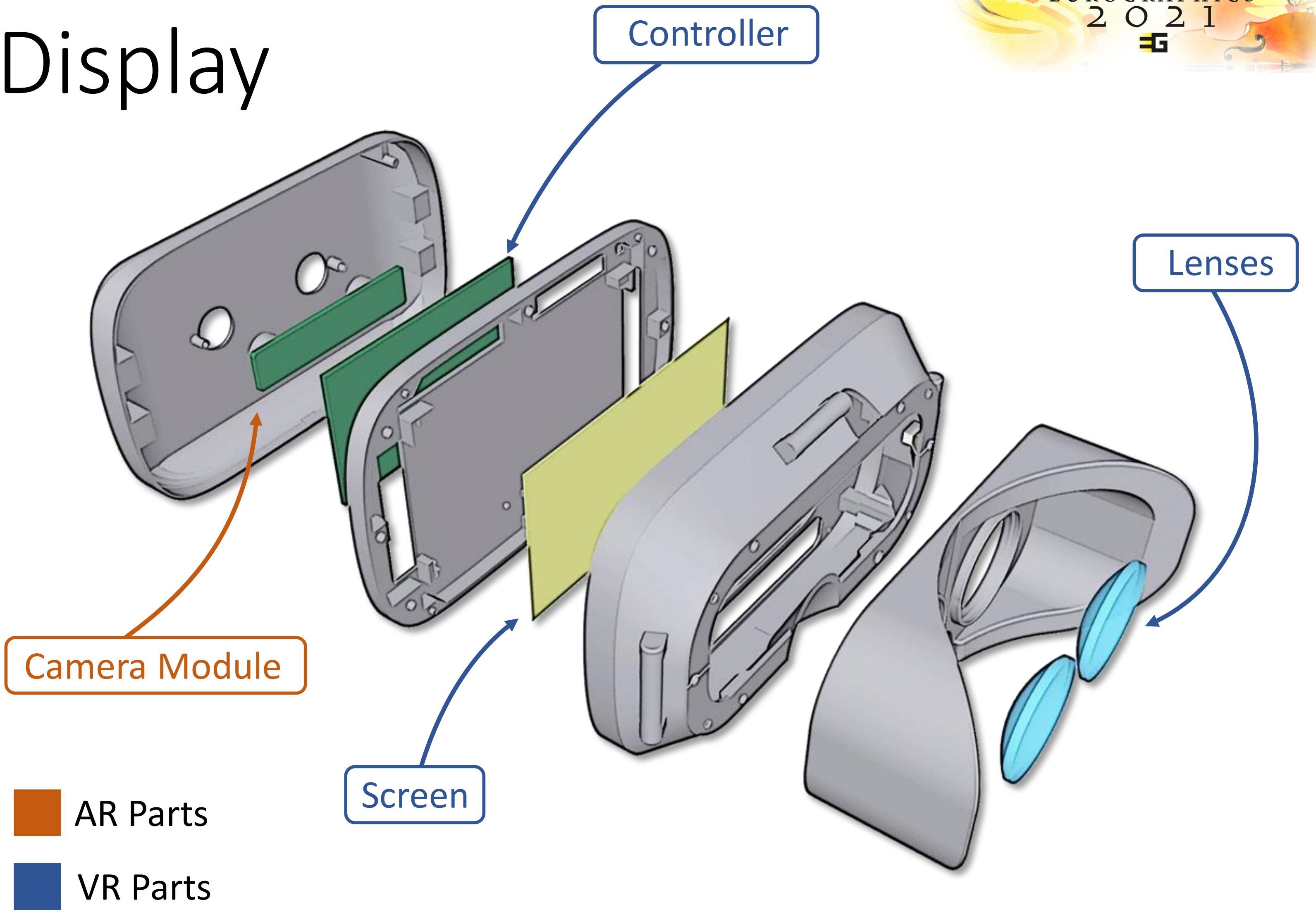
- Limited dynamic range
- Vergence-Accommodation Conflict
- Latency

Building a VST Display

Components

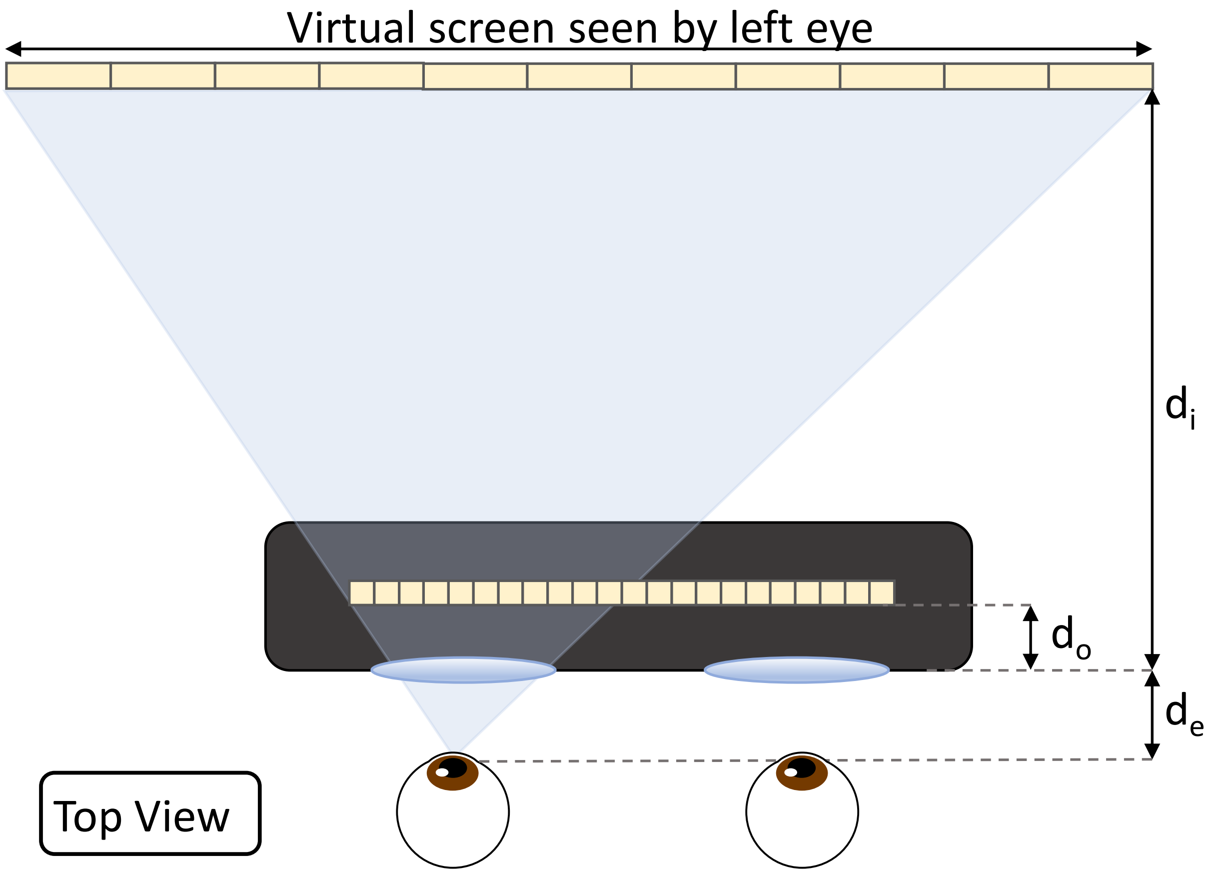






Screen Calibration

Computing virtual image of screen



- Thin lens equation to get **screen distance**:

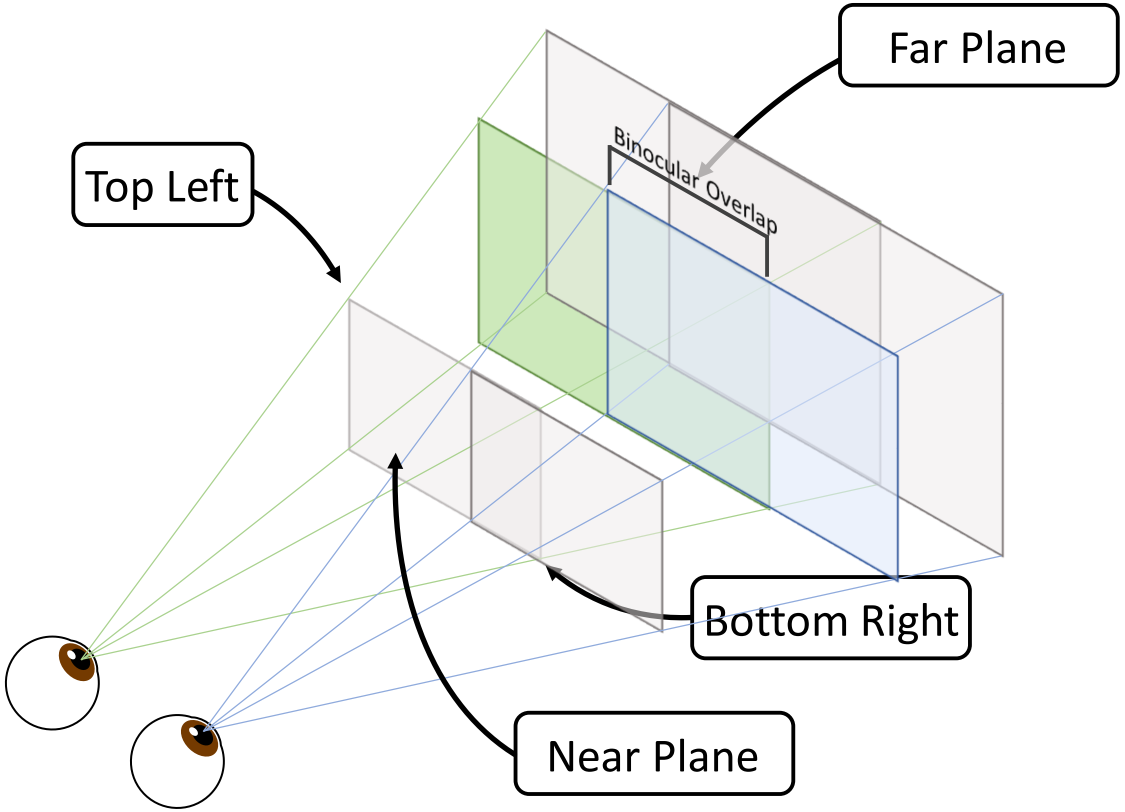
$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i}$$

- **Screen magnification**:

$$M = -\frac{d_i}{d_o}$$

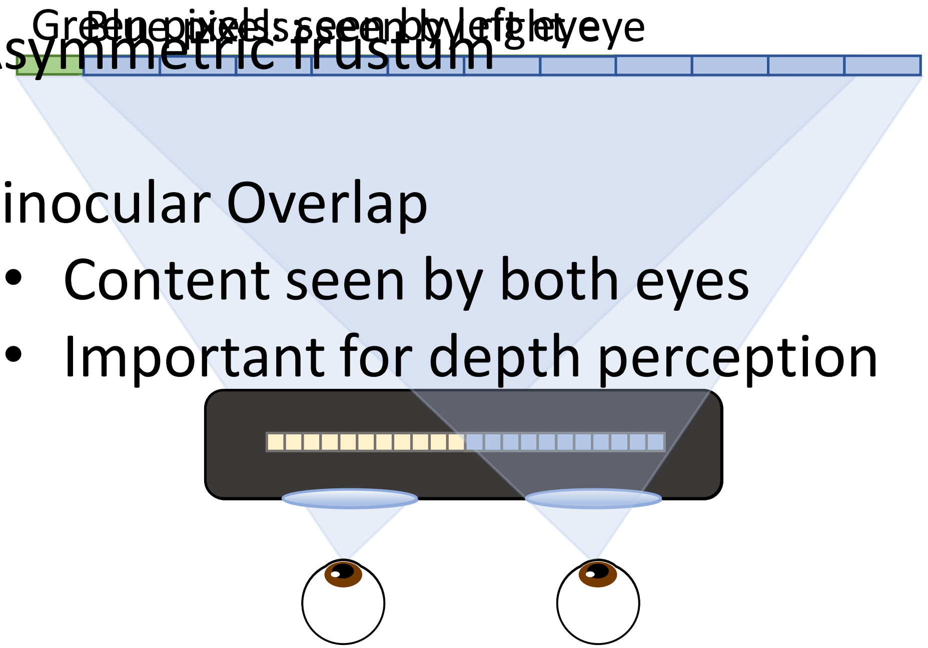
Stereo Rendering

Setting up the asymmetric Frustum



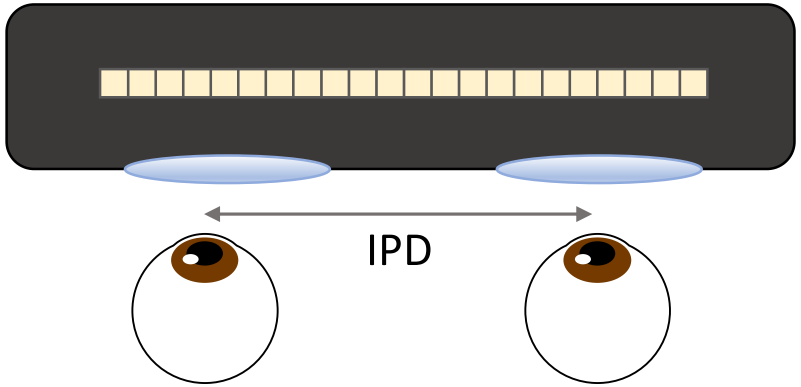
- Asymmetric frustum

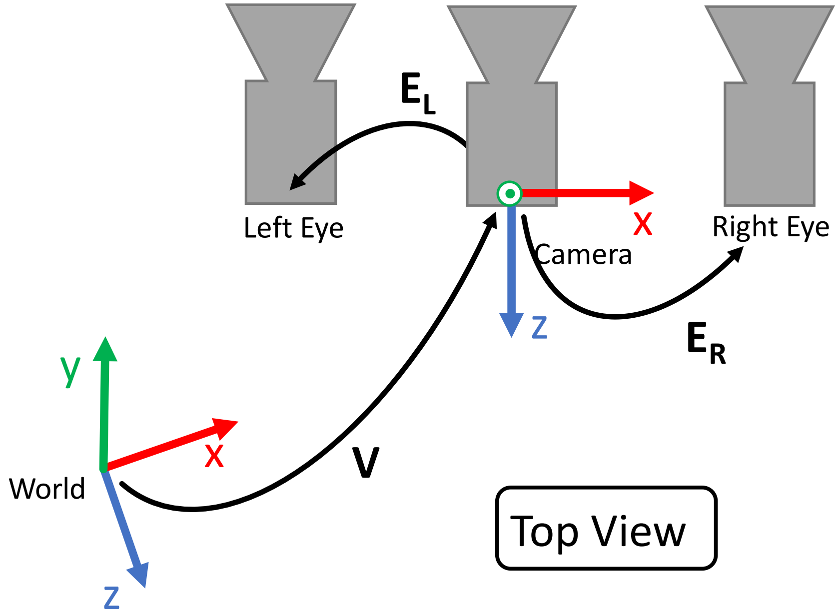
- Binocular Overlap
  - Content seen by both eyes
  - Important for depth perception



Stereo Rendering

Adjusting the View Matrix according to the interpupillary distance (IPD)





- Need to consider IPD in stereo rendering
- Essentially: Additional offset in x after view transform

- Right view transform: $\mathbf{T_R} = \mathbf{E_R} \cdot \mathbf{V}$
- Left view transform: $\mathbf{T_L} = \mathbf{E_L} \cdot \mathbf{V}$

Lens Undistortion

Correcting pincushion distortion in software

- Pincushion distortion is corrected by applying Barrel distortion in software

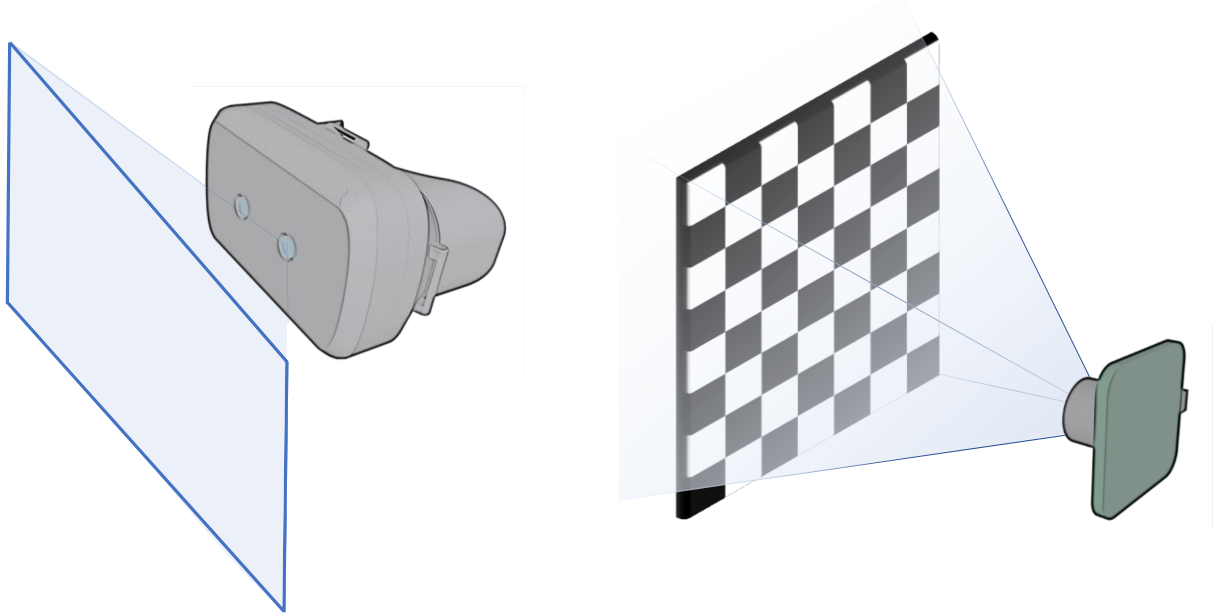
$$x_d = (x_u - x_c)(1 + K_1 * r^2 + K_2 * r^4 + \ldots) + x_c$$
$$y_d = (y_u - y_c)(1 + K_1 * r^2 + K_2 * r^4 + \ldots) + y_c$$
$$r^2 = (x_u - x_c)^2 + (y_u - y_c)^2$$

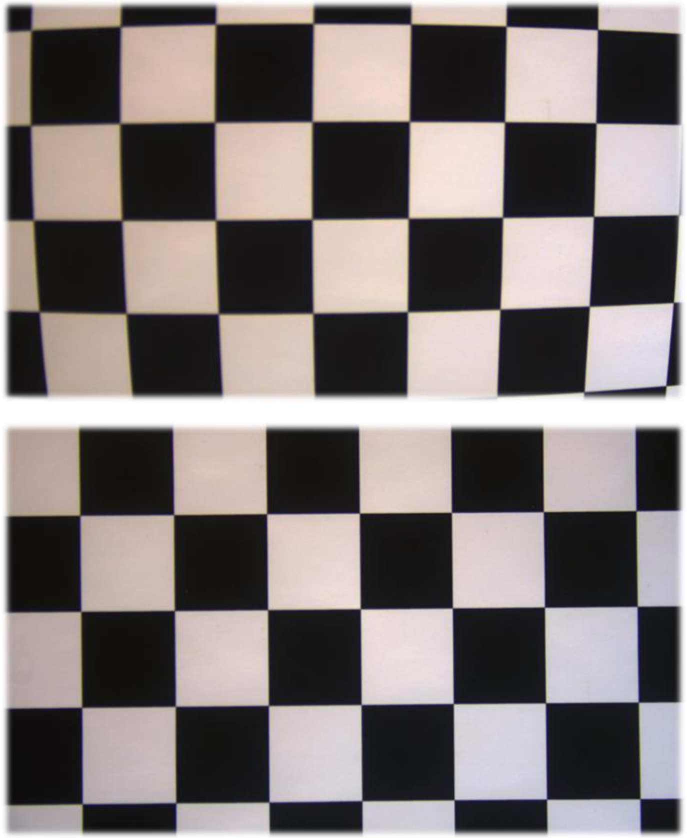
Example of barrel-distorted HMD stereo rendering

Camera Calibration

Obtaining camera matrix and correcting camera distortion

- Need to obtain camera intrinsics and distortion parameters
- A lot of software available for camera calibration
  - OpenCV
  - Matlab (Computer Vision Toolbox)
  - vicalib (https://github.com/arpg/vicalib)



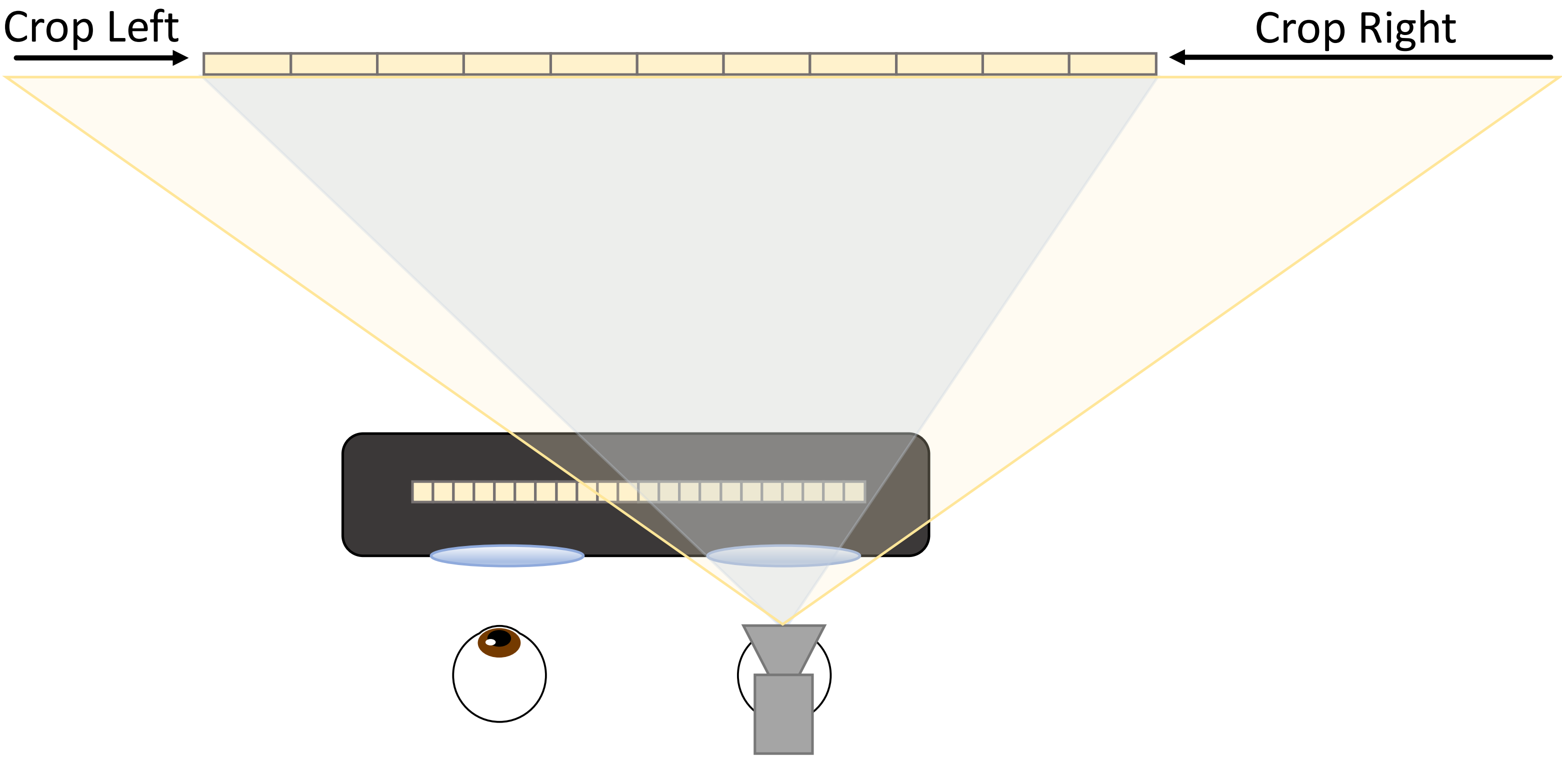


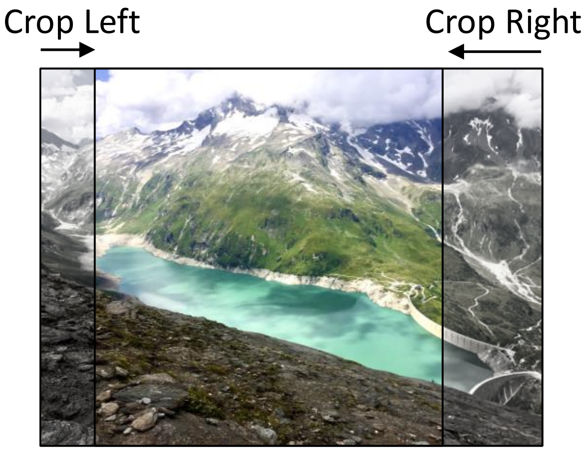
Example of a camera image before (top) and after undistortion (bottom)

Camera Calibration

Adjusting the camera field of view to the display

- Adjust optics as best as possible
- Crop the remaining FOV accordingly

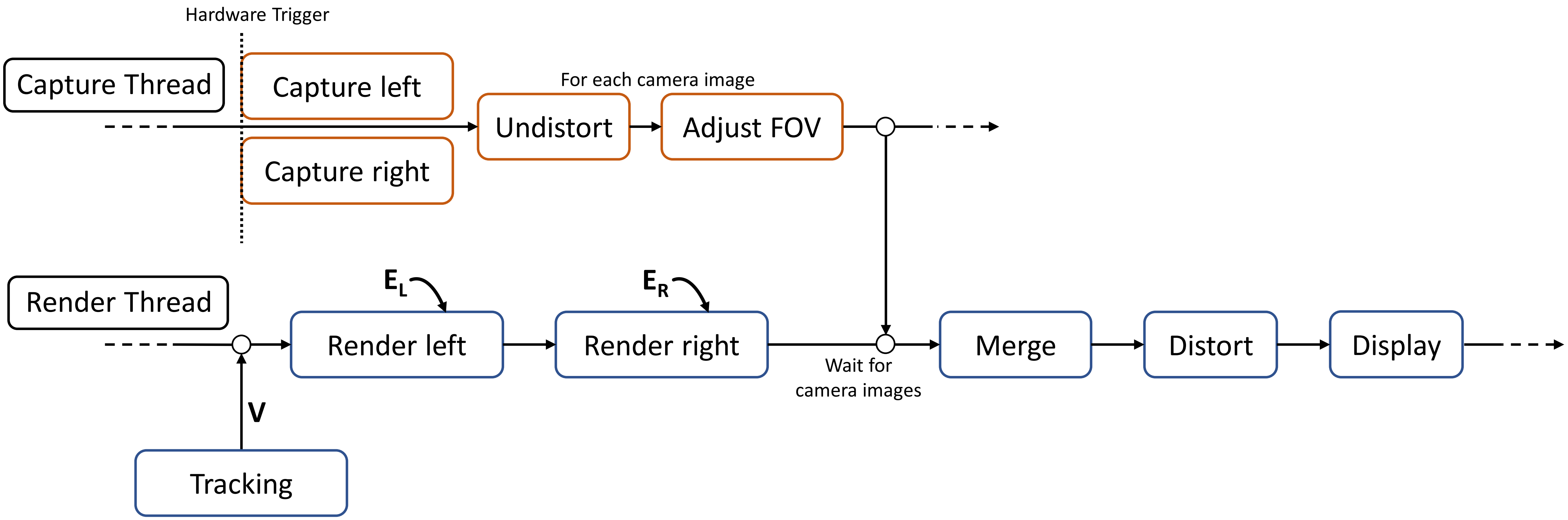




Camera Image

Putting it all Together

Software pipeline example

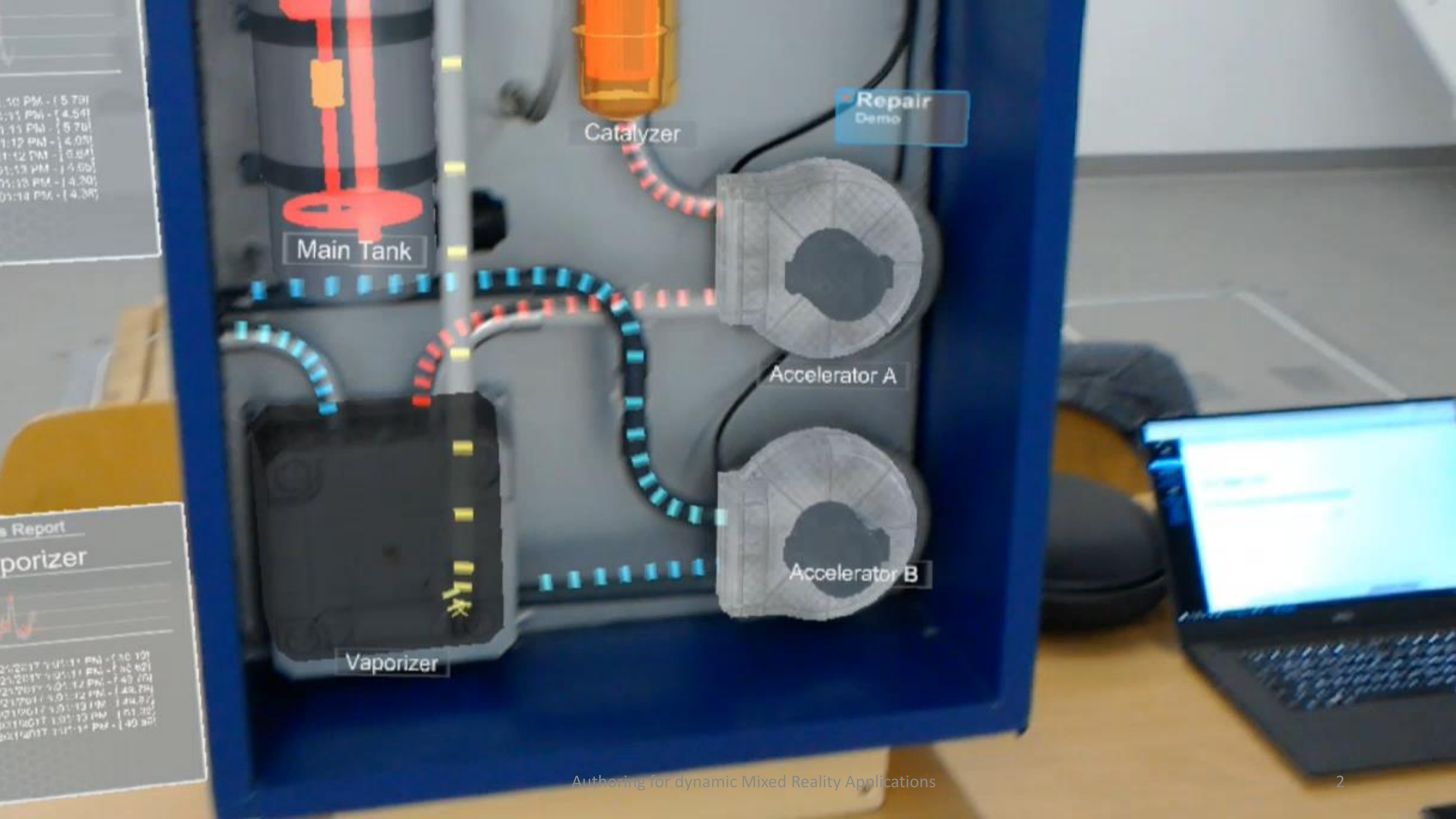


enjoy Vienna

EUROGRAPHICS 2021

Authoring for dynamic Mixed Reality Applications

Peter Mohr


Catalyzer
Repair
Demo
Main Tank
Accelerator A
Accelerator B
Vaporizer
porizer

Problems of traditional content creation for AR

- Conventional Content Creation Tools
  - 3D modelling software
    (e.g. Blender, 3DS Max, Unity)
  - Animations, Path & Label placement by hand

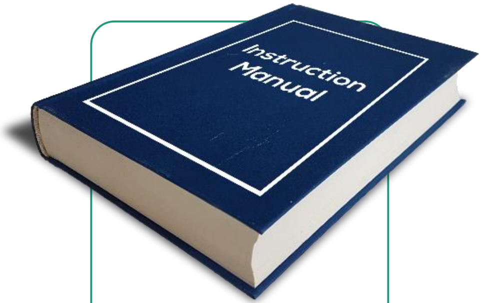
- Drawbacks
  - Requires 3D expert & technical specialist
  - Time consuming
  - Not scalable

➡ **EXPENSIVE**

Retargeting Instructions to Augmented Reality

- Efficient Authoring of Instructions
  - Retargeting from Images
  - Retargeting from Video
  - Authoring using Light Fields (ad hoc)

Retargeting



Traditional Media
(for Instructions)

3D Pose/Path from 2D

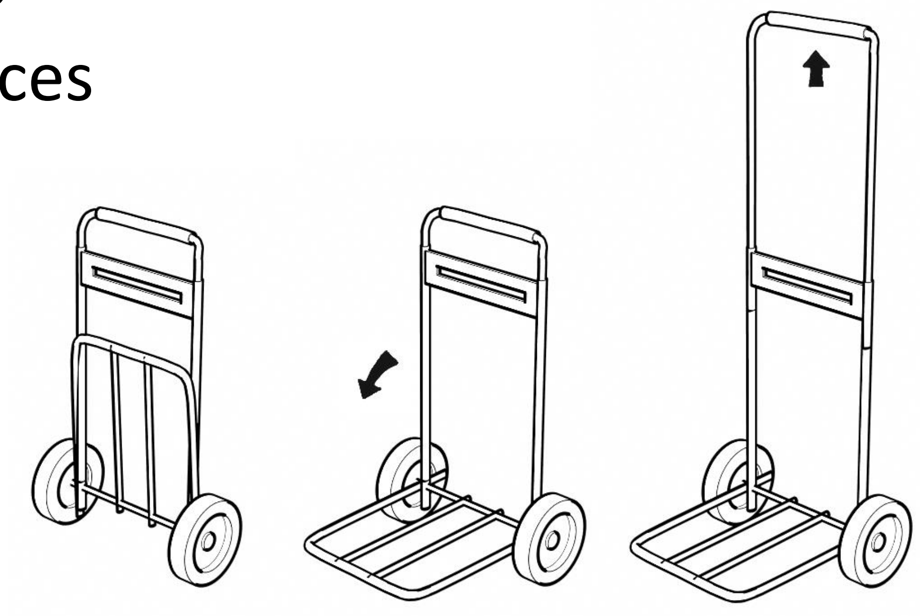
Retargeting

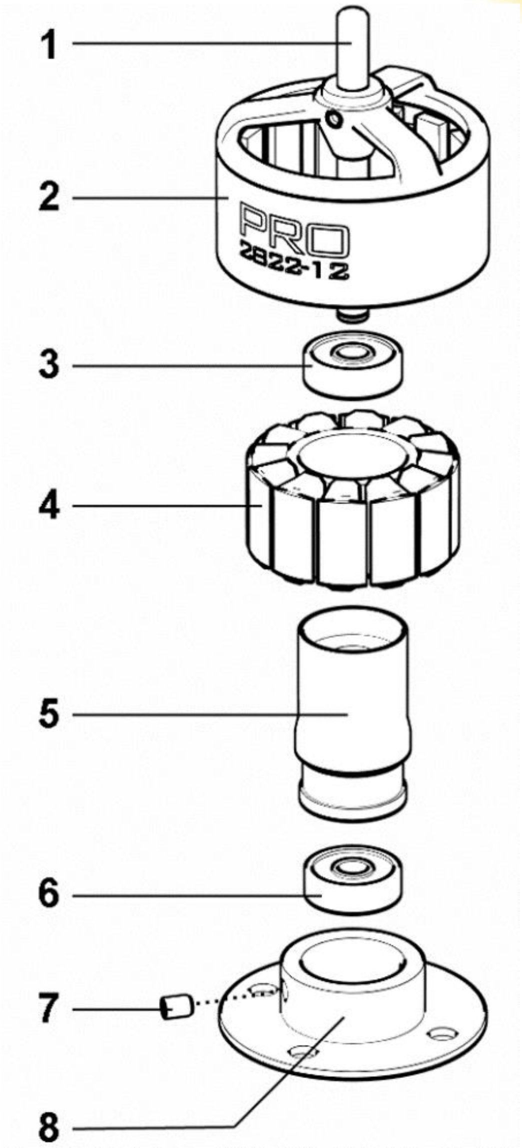
Temporal Segmentation

Interactive 3D Media

Elements of Manuals

- Annotations
- Arrows
- Explosion diagrams
- Image sequences
- Combinations

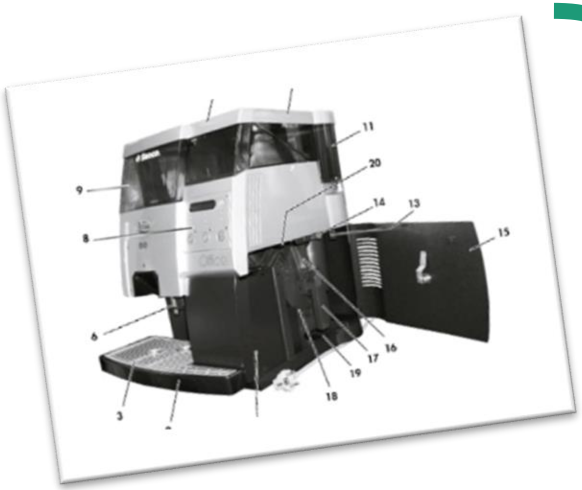


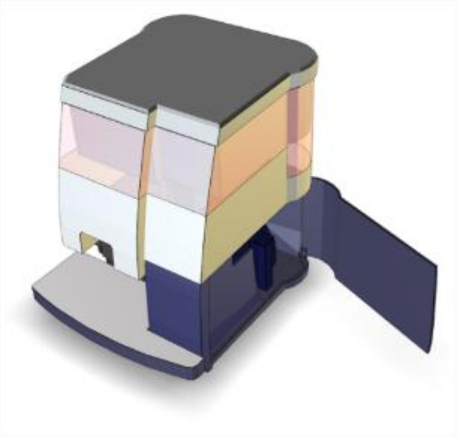


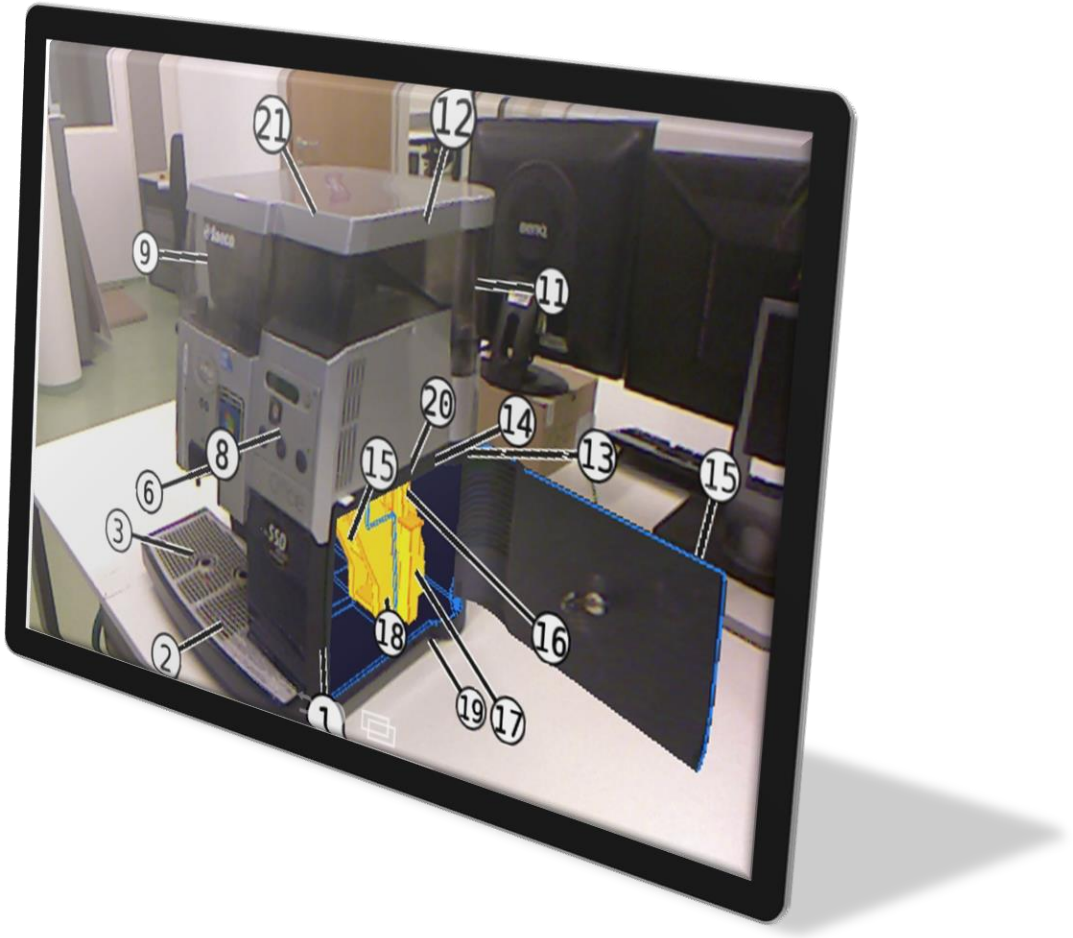
Method

Input Data

- 2D Image(s)
- 3D model







Method

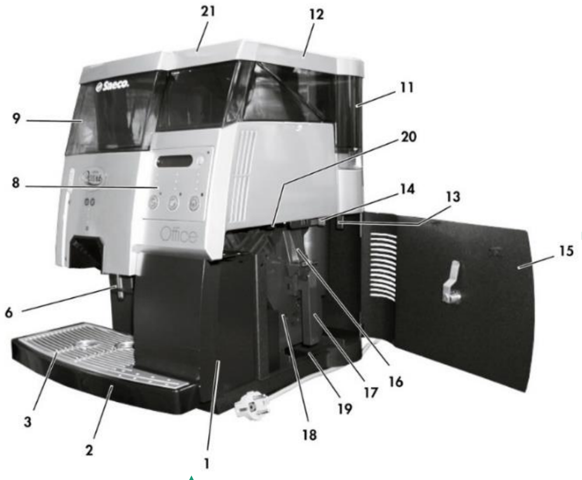
Input Data

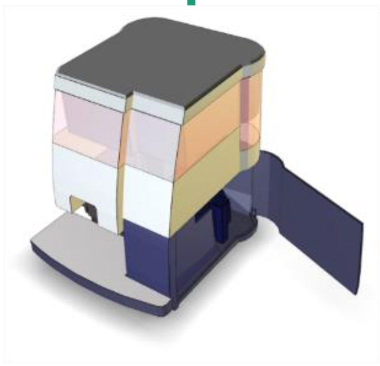
- 2D Image / image sequence
- 3D model (does not need to be exact)

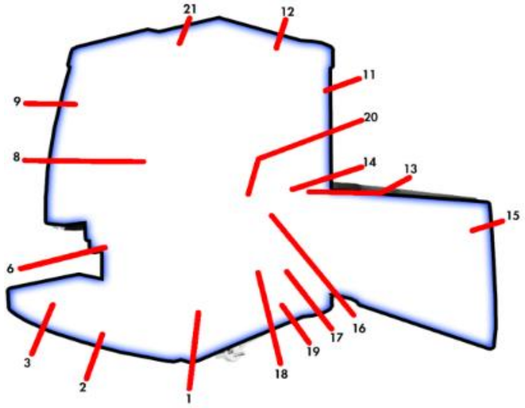
Workflow

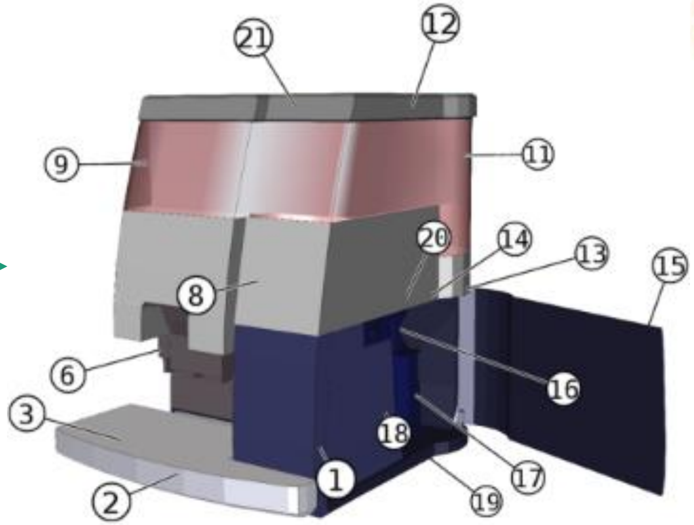
- Align 3D model to image (e.g. PosIt)
- Extract labels
- Extract glyphs (arrows)
- Identify spatial arrangement of parts → infer movement

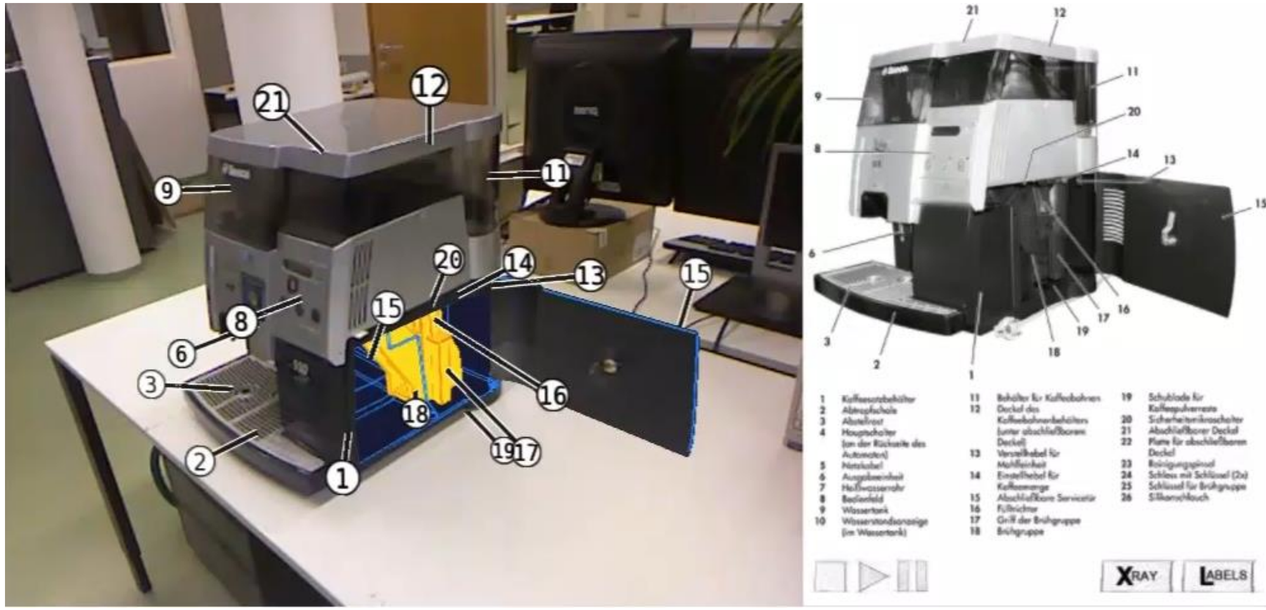
Label Extraction



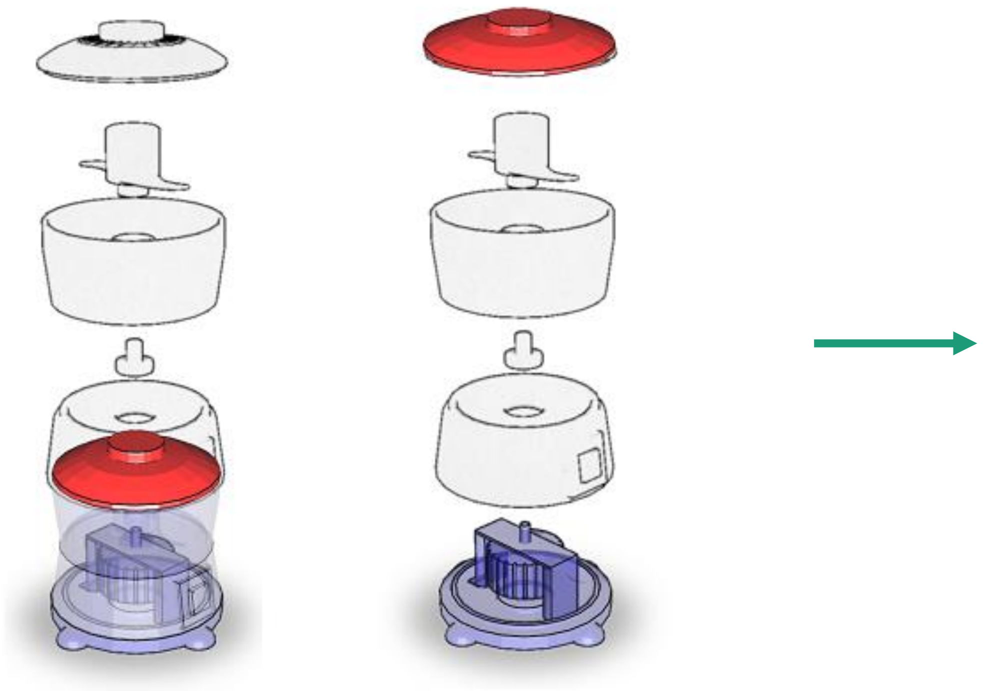


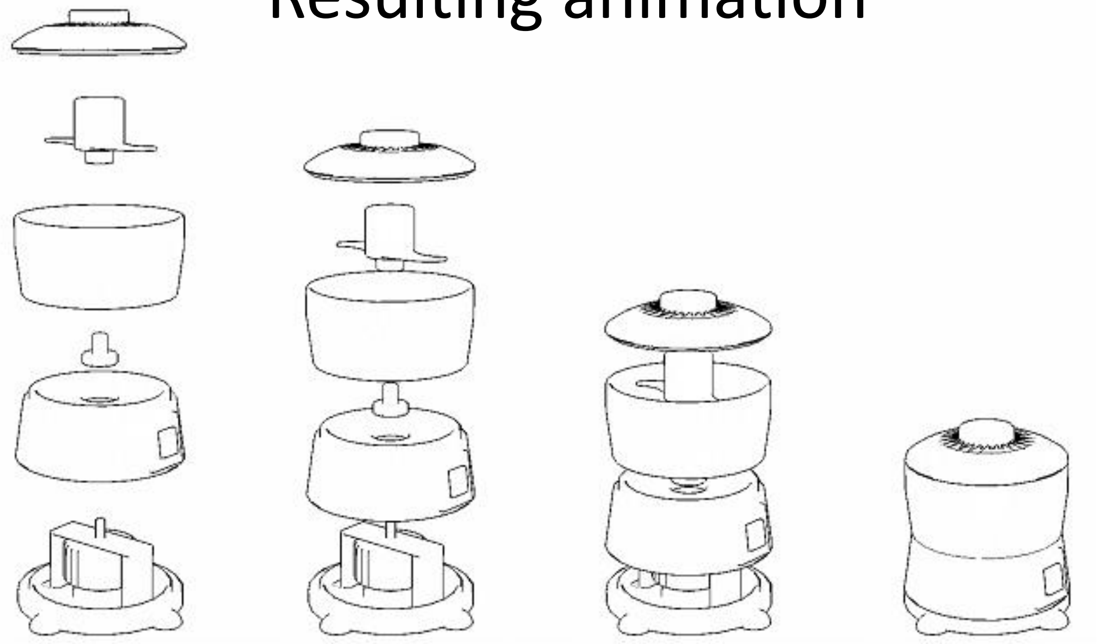






Explosion Diagrams

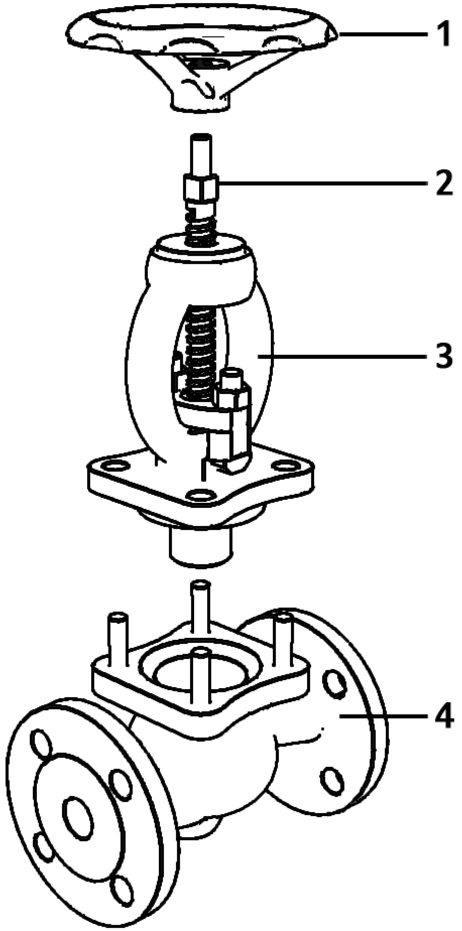


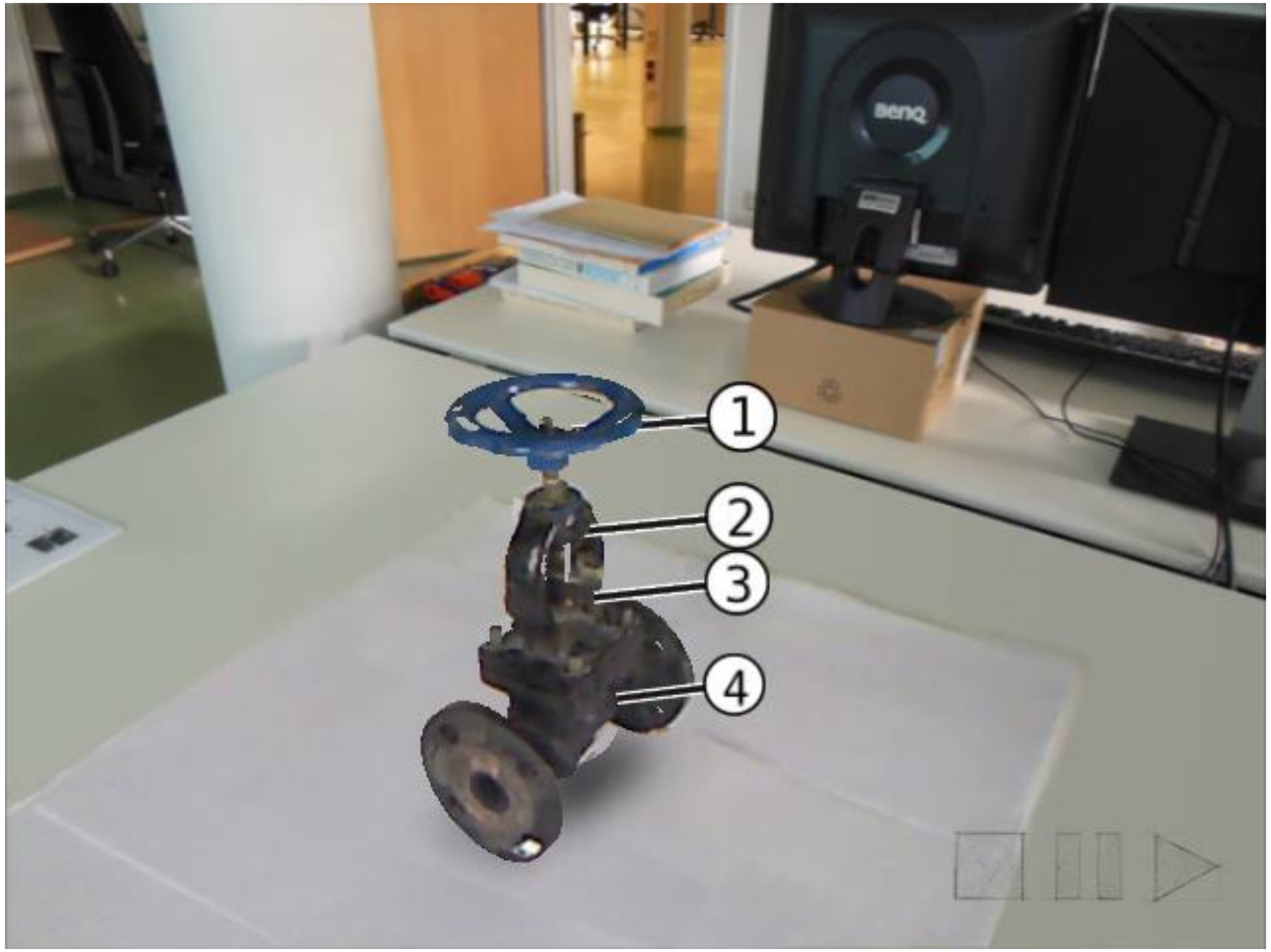


Resulting animation

For every movable part in the 3D model, we find the best configuration to match the input image

Explosion Diagrams





Sequences

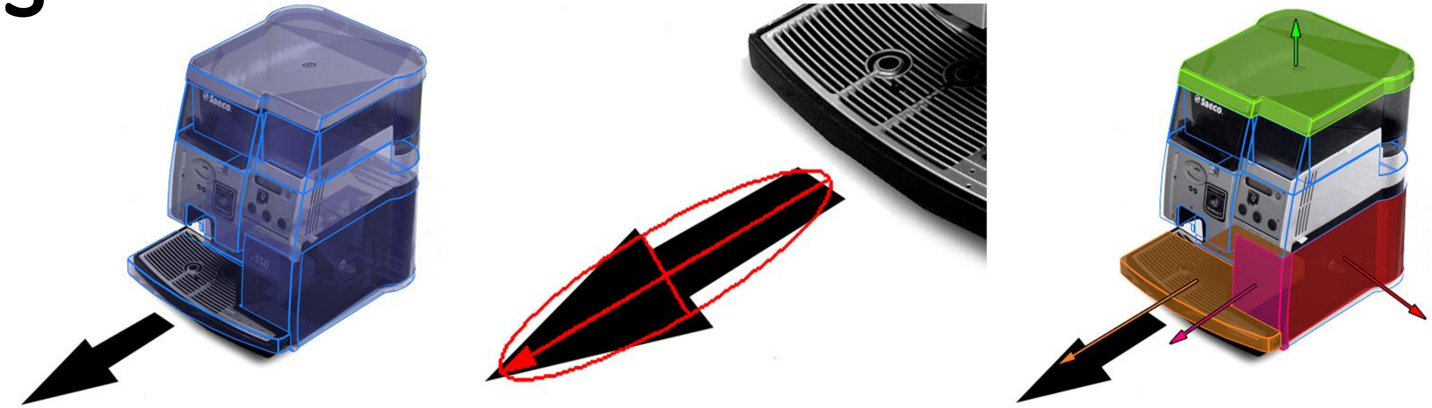
- Identify region of change
- Find best fit

Input images

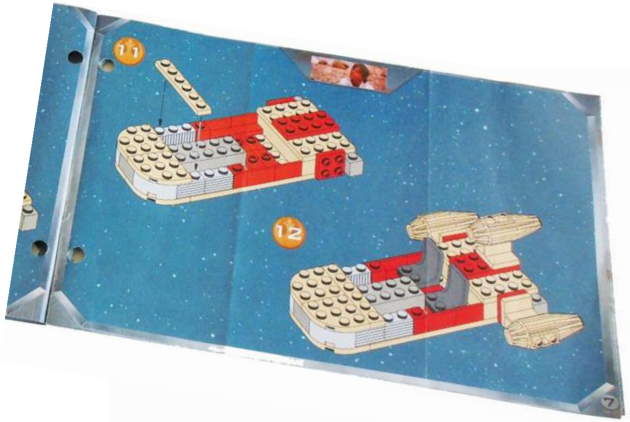
Region of change

Interpreting Arrows

- Indicate important parts
- Convey movement/action

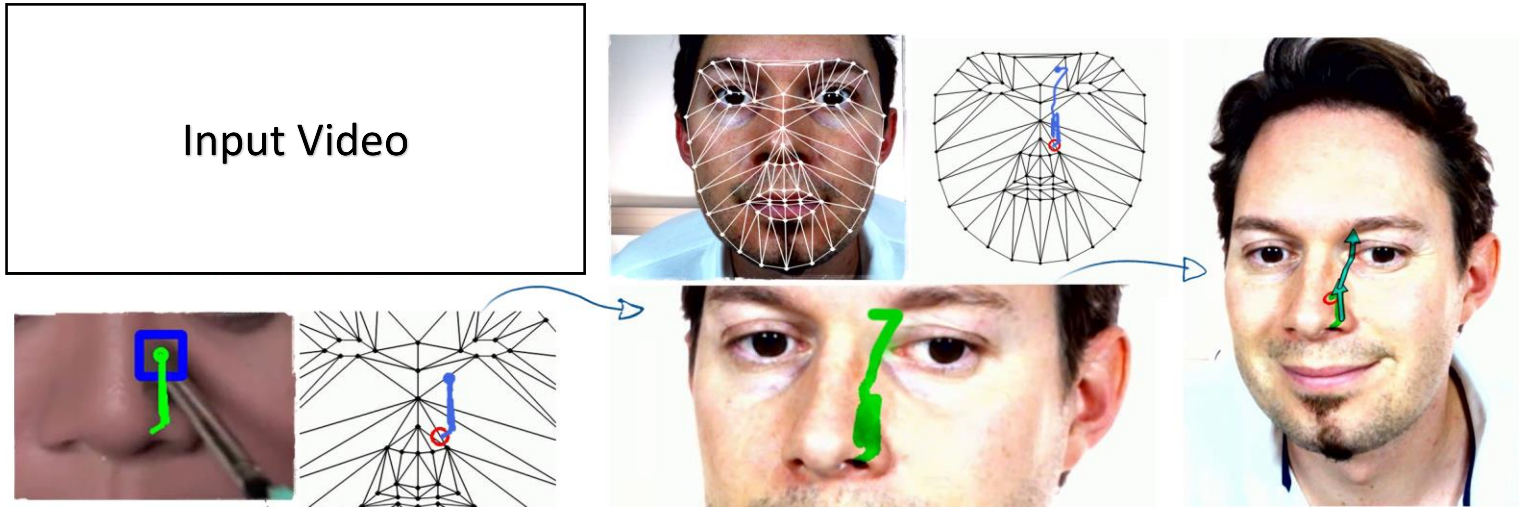


1. Find isolated region
2. Check concavities
3. Fit ellipse
4. Match projected direction & distance with candidate parts



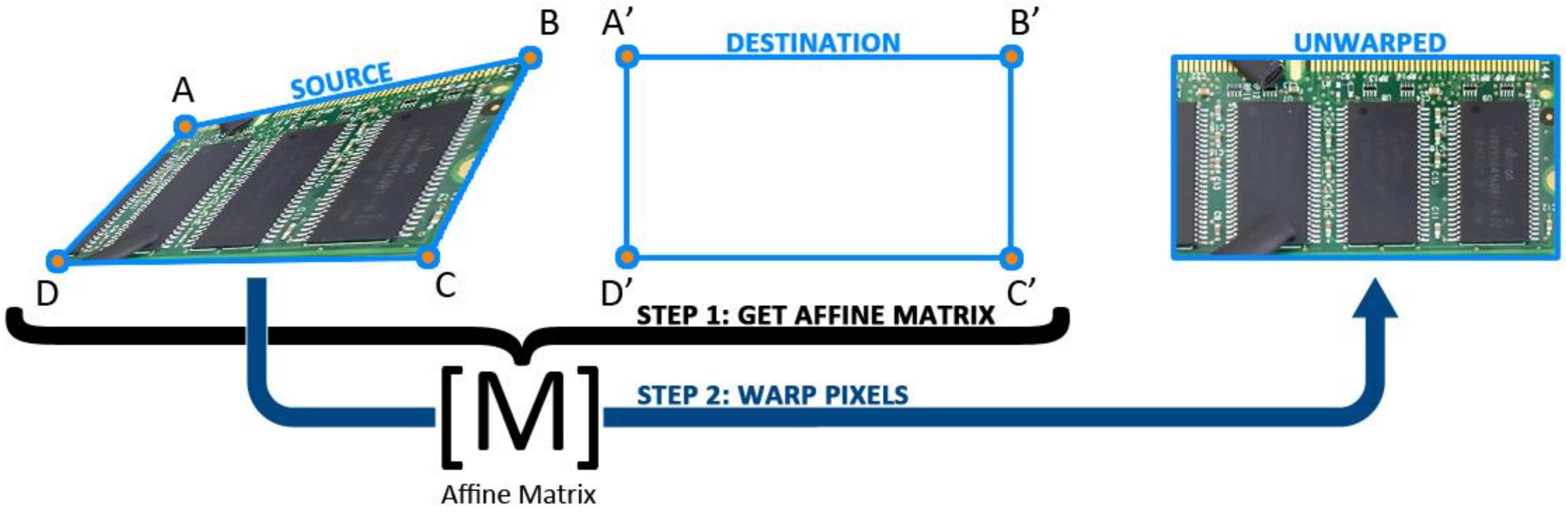
Retargeting: Videos

- Object/user motions are extracted by tracking known model features in the 2D video
- The extracted (path) data is processed and can be edited by the author (path and sequence order)
- The tutorial is retargeted to a (different) live scene



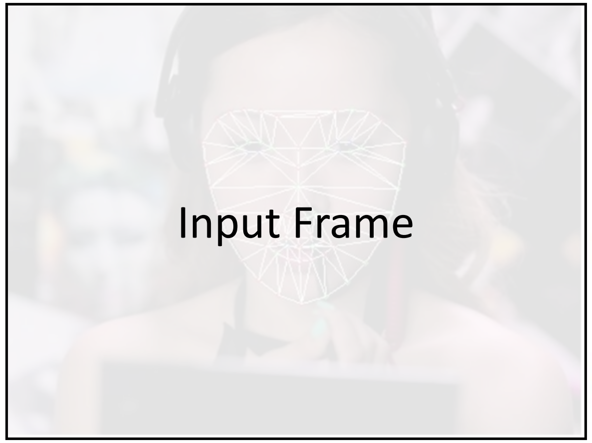
Surface Tracking (planar)

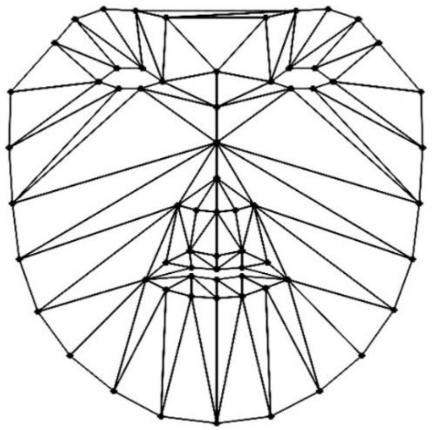
- Author selects surface in input video (reference frame)
- Surface pixels are automatically unwarped for all frames

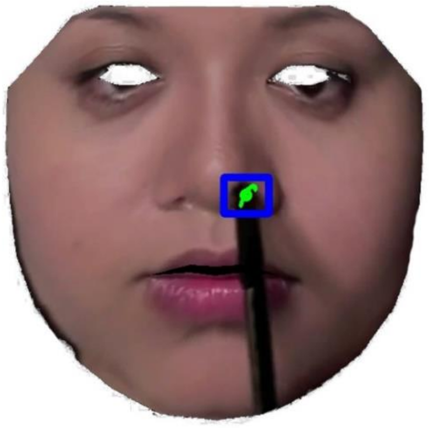


Surface Tracking (3D)

- Faces are automatically detected and tracked using a deformable 3D face model (CLM)
- Surface pixels are unwarped to the UV map of the model for every frame

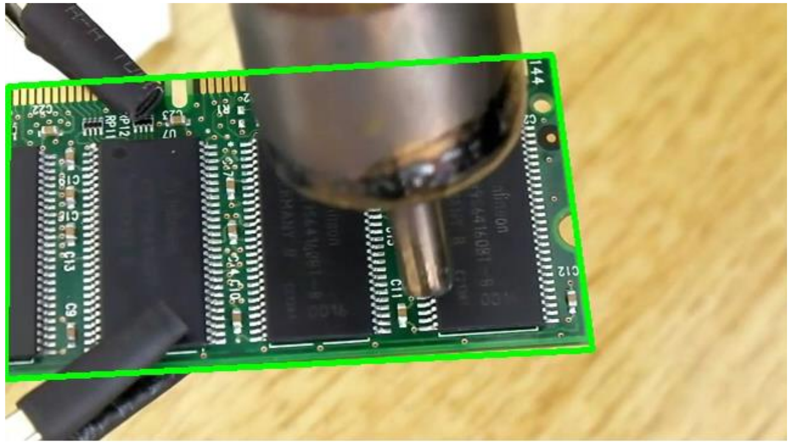


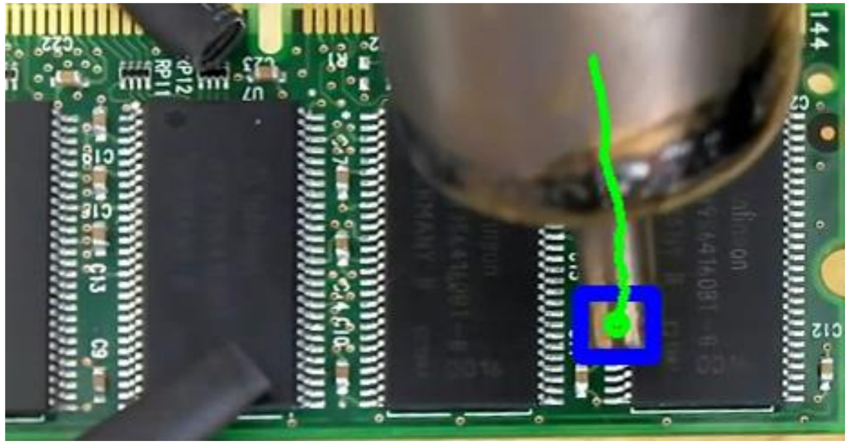




Tool Path Extraction

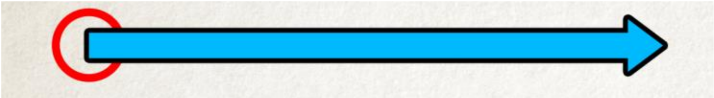
- Author selects the tool
- System tracks the tool path in the **unwarped** video
  - Using TLD (Tracking-Learning-Detection)
- Tool path is stored in surface coordinates





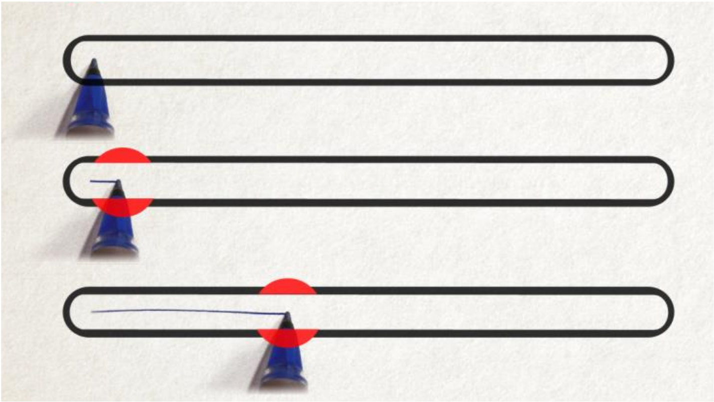
Tool Path Visualization

- Initial Design: Dynamic Glyphs



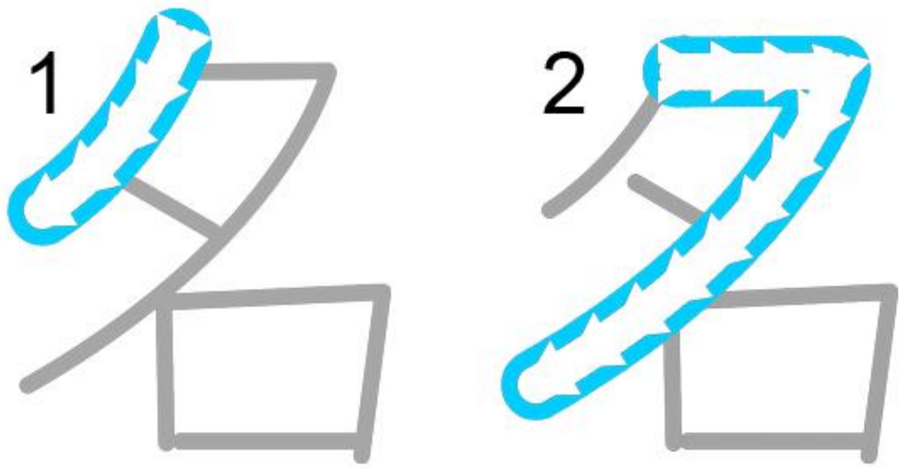
Tool Path Visualization

- Initial Design: Dynamic Glyphs

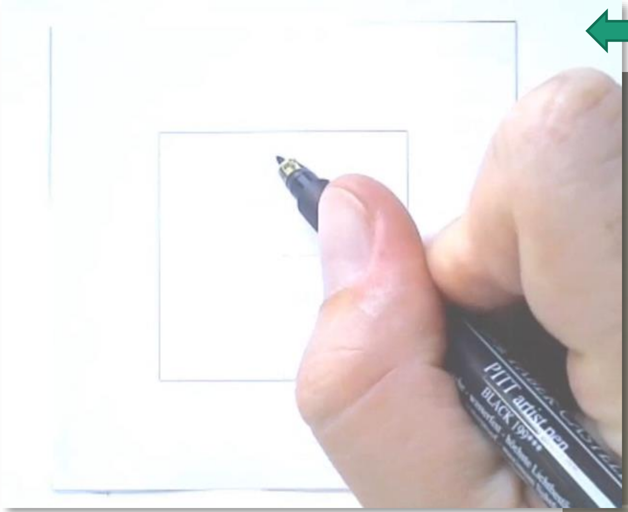


Tool Path Visualization

- Final Design: directional outline



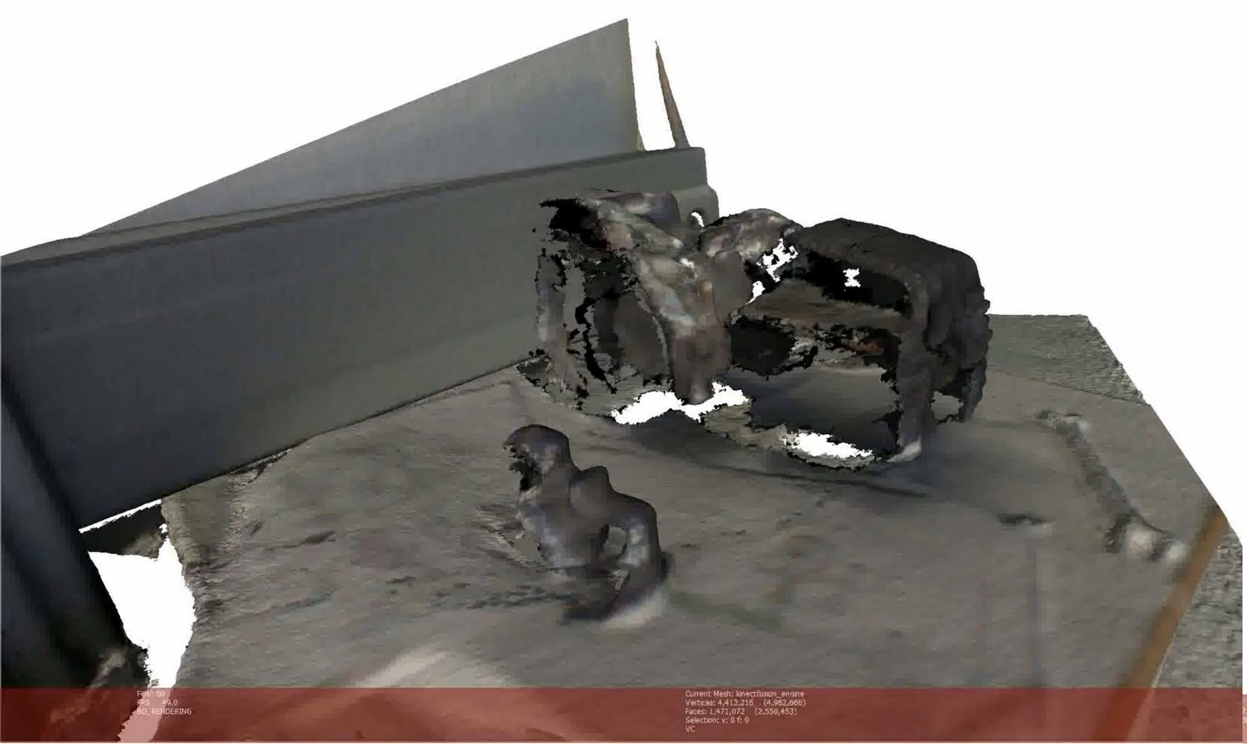
Path Visualization (Study)



Input video tutorial

Authoring using Lightfields

Traditional Approach

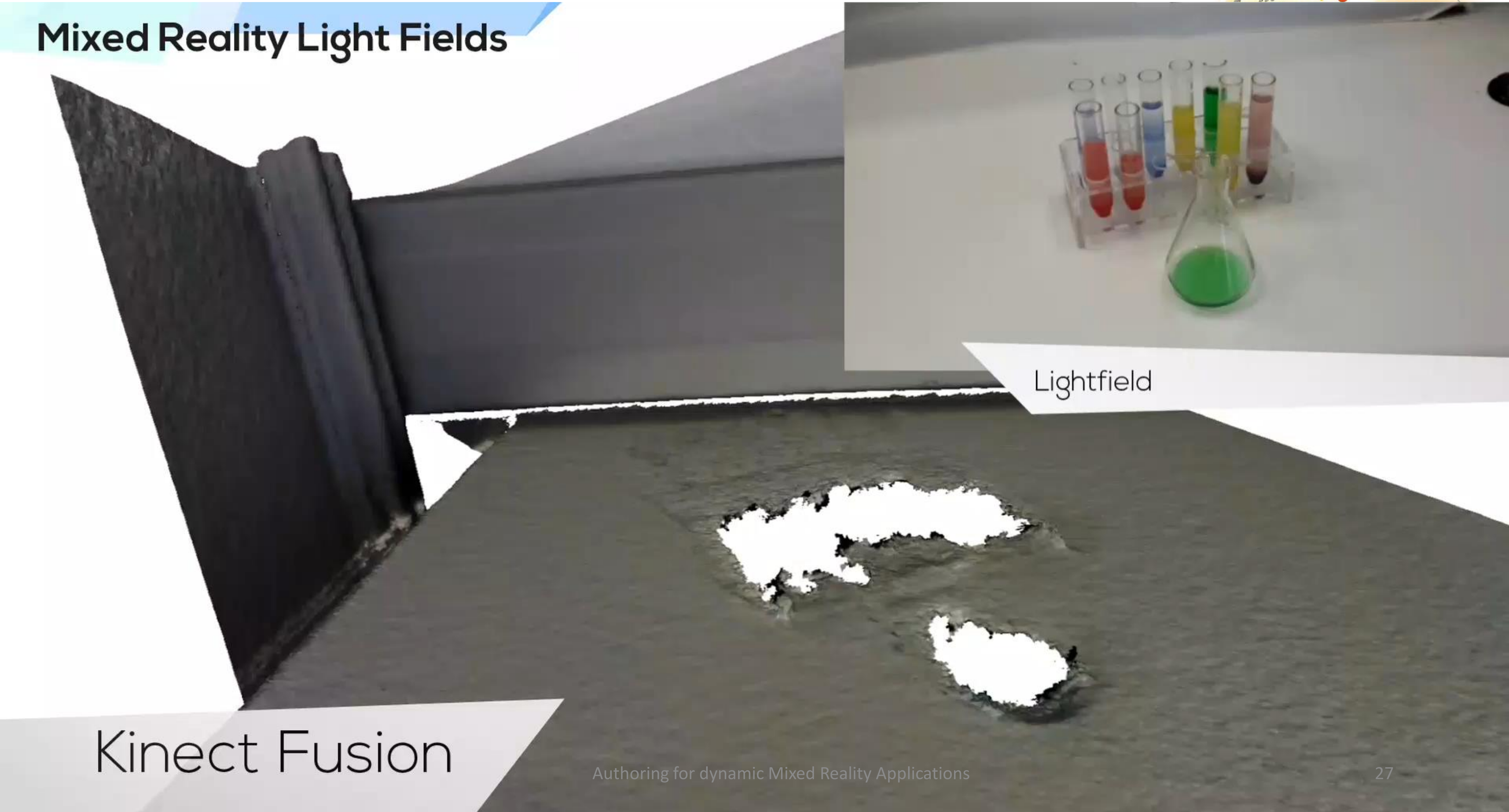


Kinect Fusion (v1 Sensor)

Structure from Motion (187 images, ~59 min)

Problems: **Reflections** **Transparency** **Sunlight(IR)** **Textureless Objects**

Mixed Reality Light Fields

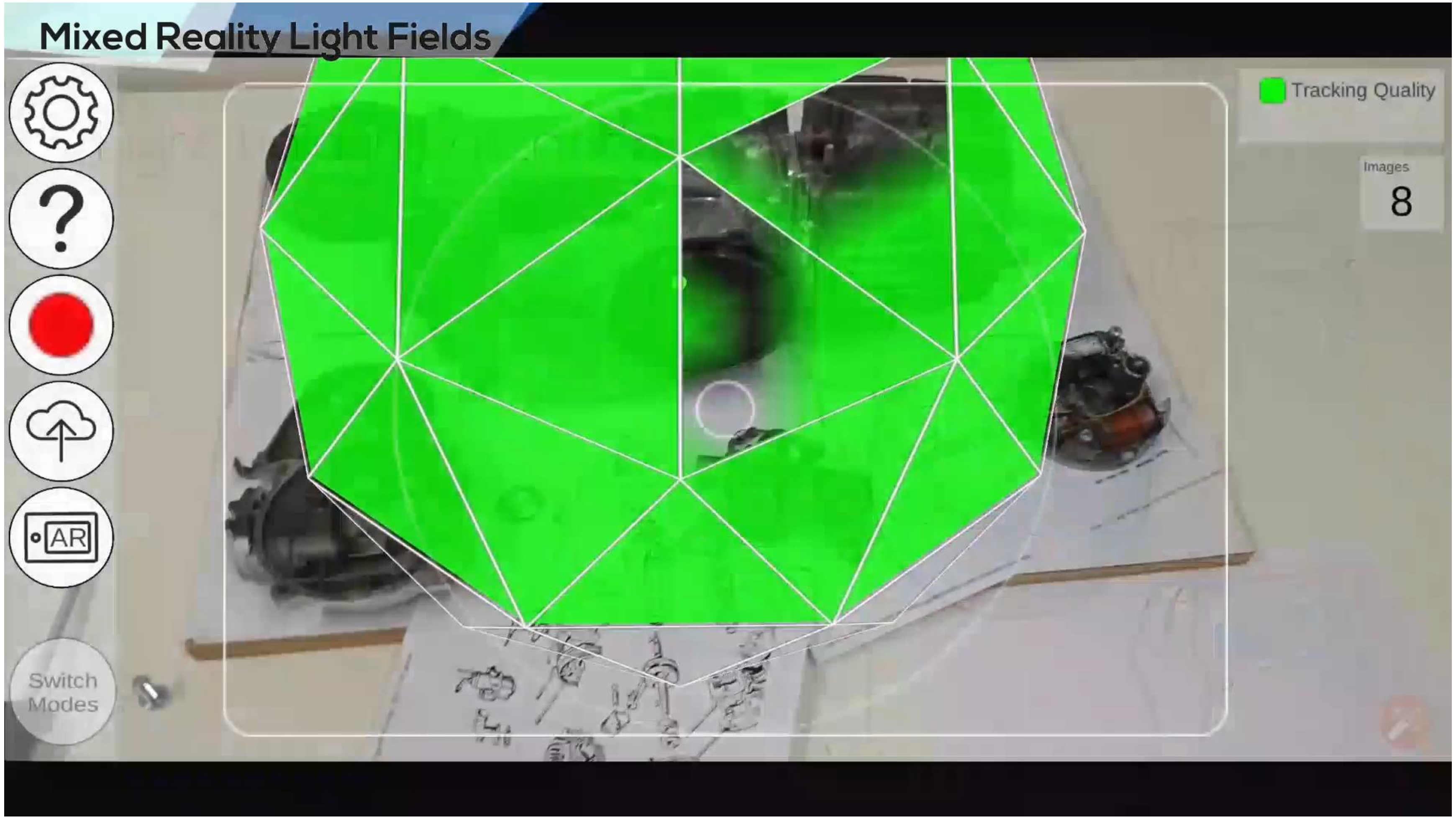


Mixed Reality Light Fields

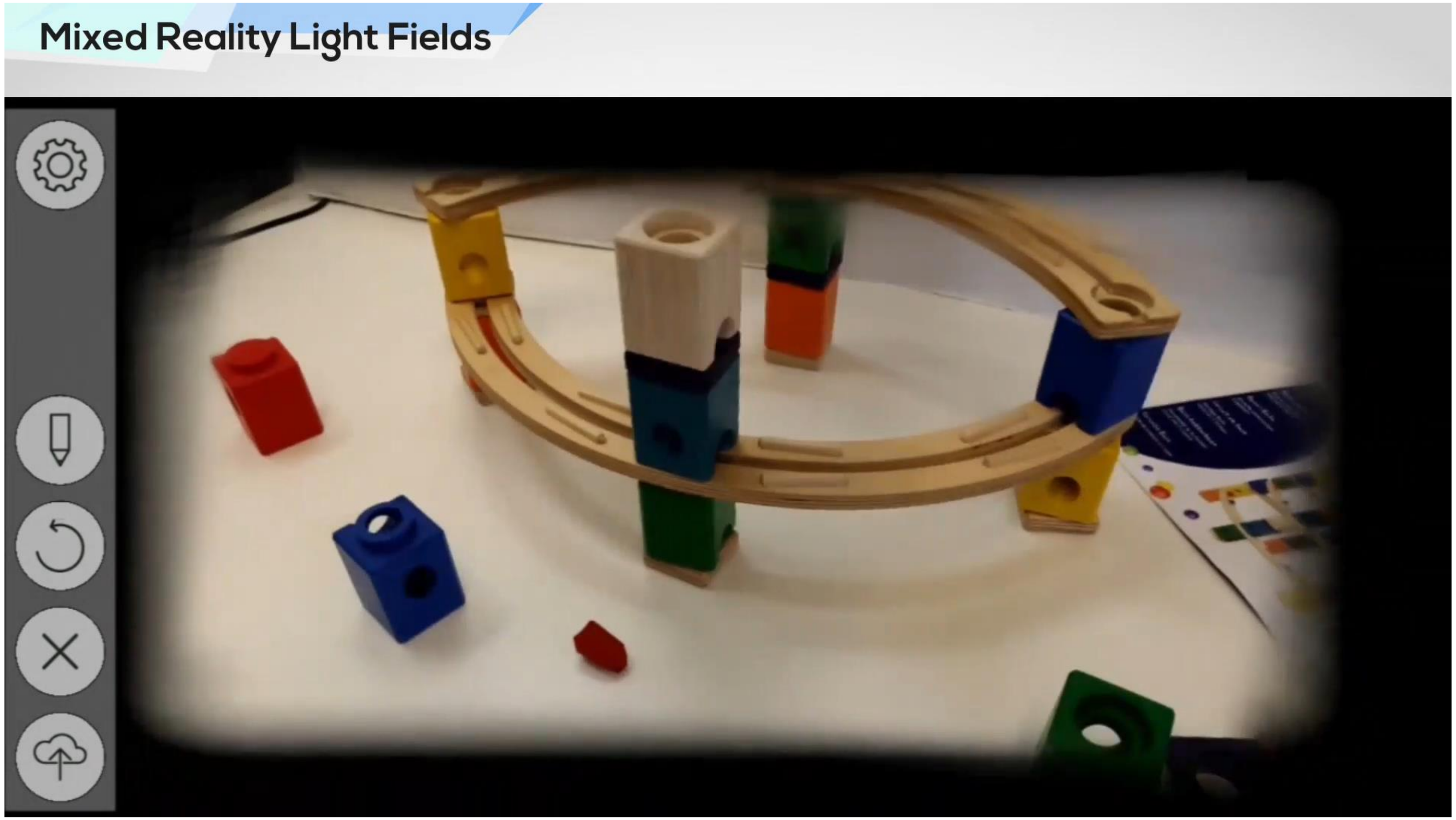
- No 3D explicit model
- Spatially registered image database

- Advantages
  - Scene independent (e.g. reflective or transparent objects, daylight, scale)
  - High quality visual representation
  - No pre-processing necessary
- Disadvantages
  - **No depth data for anchoring annotations**
  - Data size, how to capture

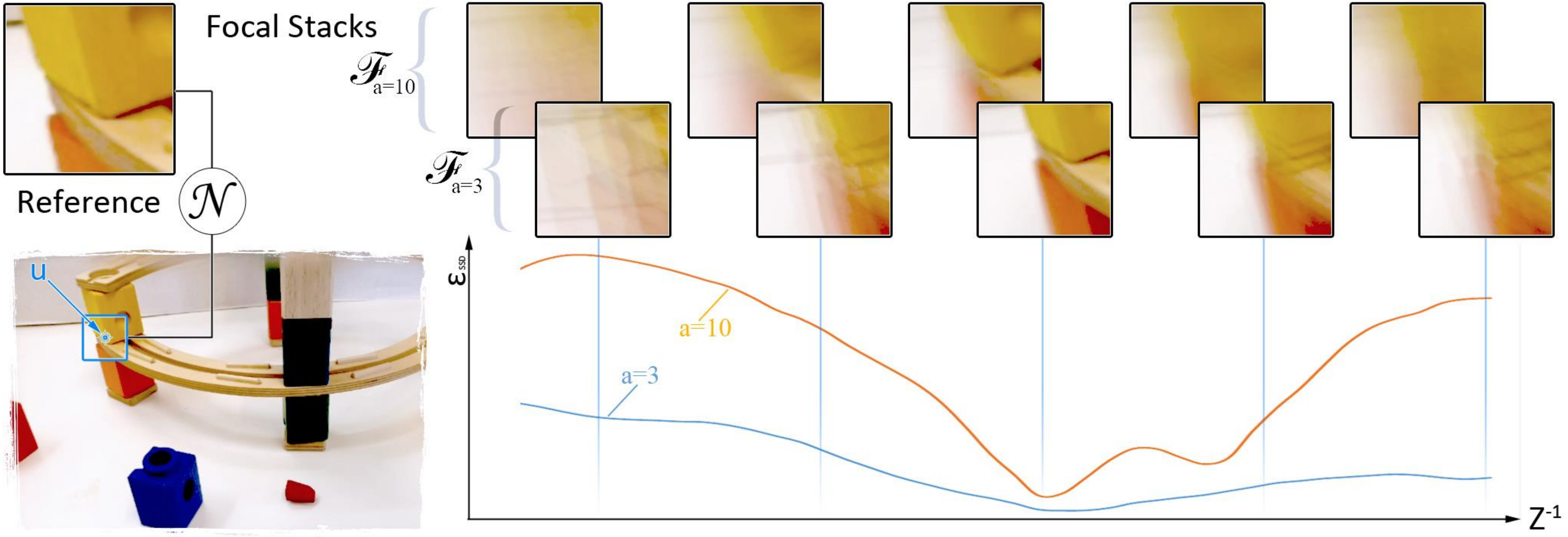
Lightfield Capturing



Lightfield Annotation

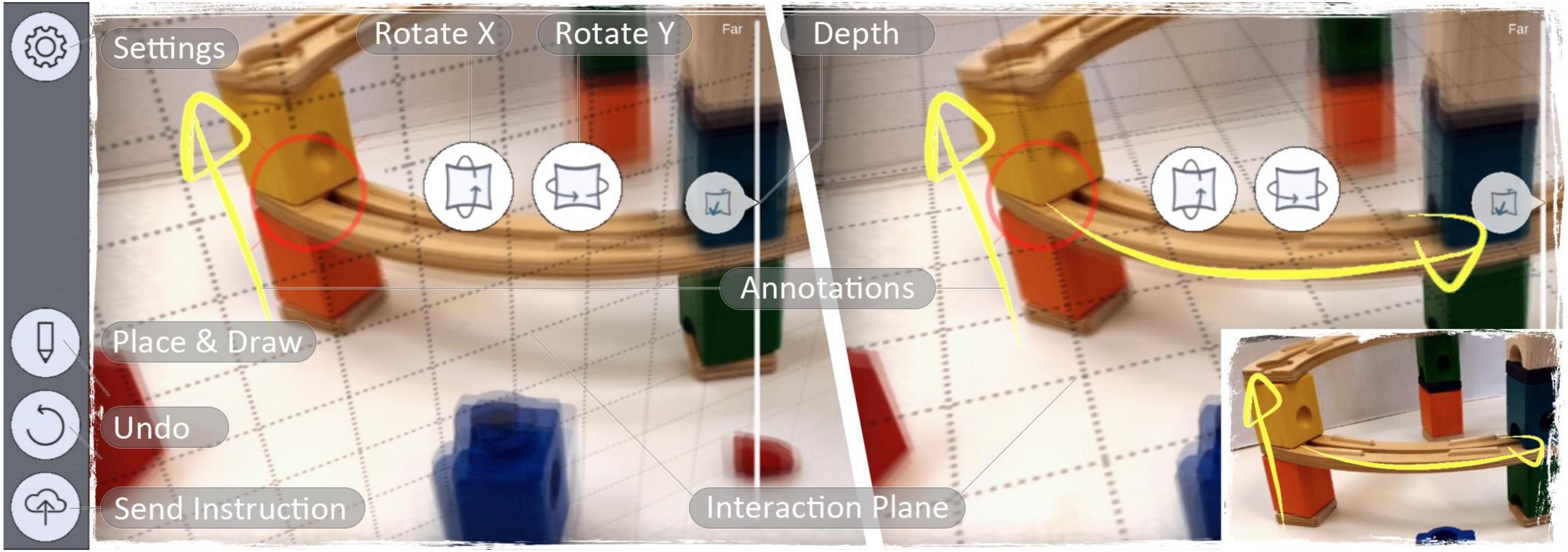


Lightfield Annotation

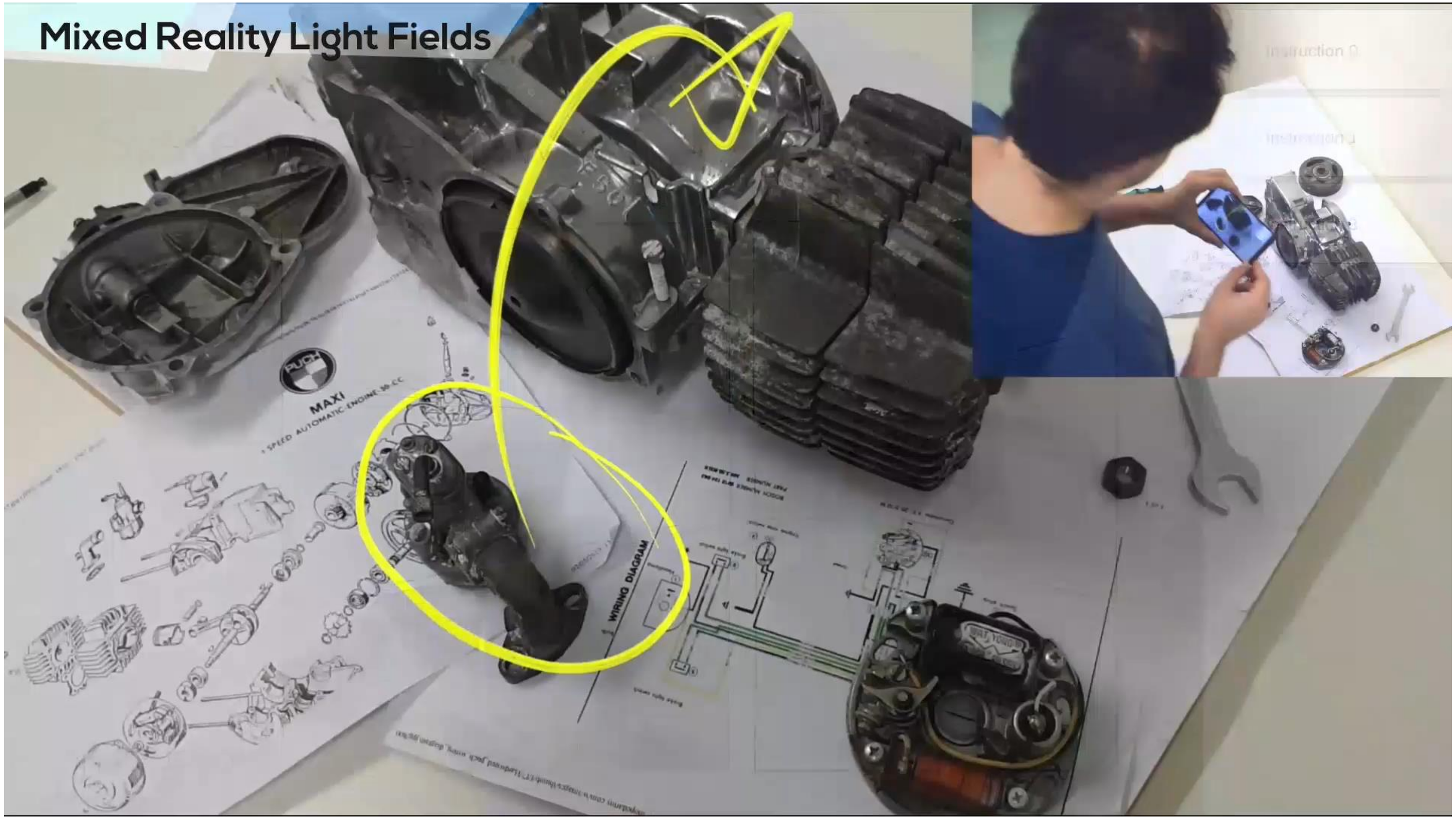


$$\varepsilon(I_{\mathrm{KF}}, I_f) = \sum_{\mathbf{u} \in \mathcal{N}} (I_{\mathrm{KF}}(\mathbf{u}) - I_f(\mathbf{u}))^2$$

Annotation Interface



AR Instruction Display



Publications

Peter Mohr, Shohei Mori, Tobias Langlotz, Bruce Thomas, Dieter Schmalstieg, and Denis Kalkofen. **Mixed Reality Light Fields for Interactive Remote Assistance.** In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20)

Peter Mohr, David Mandl, Markus Tatzgern, Eduardo Veas, Dieter Schmalstieg, and Denis Kalkofen. **Retargeting Video Tutorials Showing Tools With Surface Contact to Augmented Reality**. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI'17)*

Peter Mohr, Bernhard Kerbl, Denis Kalkofen, and Dieter Schmalstieg. **Retargeting Technical Documentation to Augmented Reality**. In *Proc. ACM Conference on Human-Computer Interaction (CHI'15)*, Seoul, South Korea, April 2015.

Institute of Computer Graphics and Vision, Graz University of Technology

Wearable Computer Lab, University of South Australia

Salzburg University of Applied Sciences

HCI Group, University of Otago

zentrum für virtual reality und visualisierung forschungs-gmbh

Graz University of Technology