元宇宙中的交互计算与包容普惠

Computational Interaction for a Universally Accessible Metaverse

Ruofei Du (杜若飞)
Senior Research Scientist
Google Labs, San Francisco
www.ruofeidu.com
me@duruofei.com

Self Intro

www.duruofei.com

Publications
Projects
Software
Videos
Media
Talks
Artsy

Featured Publications

DepthLab: Real-Time 3D Interaction With Depth Maps for Mobile Augmented Reality *13K Downloads*

Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte, Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces, Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, and David Kim

Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST), 2020.

pdf, doi | website, project, video, slides, code, demo, supp | cited by, cite

Geollery: A Mixed Reality Social Media Platform *Online Demo of a Metaverse of Mirrored World*

Ruofei Du, David Li, and Amitabh Varshney

Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI), 2019.

pdf, doi | website, project, video, slides, demo | cited by, cite

Montage4D: Real-Time Seamless Fusion and Stylization of Multiview Video Textures *Microsoft TechFest 2018*

Ruofei Du, Ming Chuang, Wayne Chang, Hugues Hoppe, and Amitabh Varshney

Journal of Computer Graphics Techniques (JCGT), 2019.

pdf, lowres, doi | website, project, video, slides | cited by, cite

Social Street View: Blending Immersive Street Views With Geo-Tagged Social Media *Best Paper Award*

Ruofei Du and Amitabh Varshney

Proceedings of the 21st International Conference on Web3D Technology (Web3D), 2016.

pdf, lowres, doi | website, project, video, slides | cited by, cite

Self Intro

Ruofei Du (杜若飞)

PEOPLE ›

Ruofei Du

About

Ruofei Du is a Senior Research Scientist at Google and works on creating novel interactive technologies for virtual and augmented reality. Du's research covers a wide range of topics in VR and AR, including depth-based interaction (DepthLab), mixed-reality social platforms (Geollery and Social Street View), 4D video-based rendering (Montage4D and VideoFields), foveated rendering (KFR, EFR, Foveated360), and deep learning in graphics (HumanGPS and SketchColorization). Du served as a committee member in CHI, UIST, SIGGRAPH Asia XR, ICMI and an Associate Editor of Frontiers in Virtual Reality and IEEE TCSVT. Du holds a Ph.D. in Computer Science from University of Maryland, College Park. In their own words: I am passionate about inventing interactive technologies with computer graphics, 3D vision, and HCI. Feel free to visit my research, artsy, projects, youtube, talks, github, and shadertoy demos for fun!

Personal website
Google scholar

Research Areas

- Human-Computer Interaction and Visualization →
- Machine Intelligence →
- Machine Perception →

Authored publications

Google publications

Filters

Research areas +

Year +

Sort by: Year ⌄

19 publications

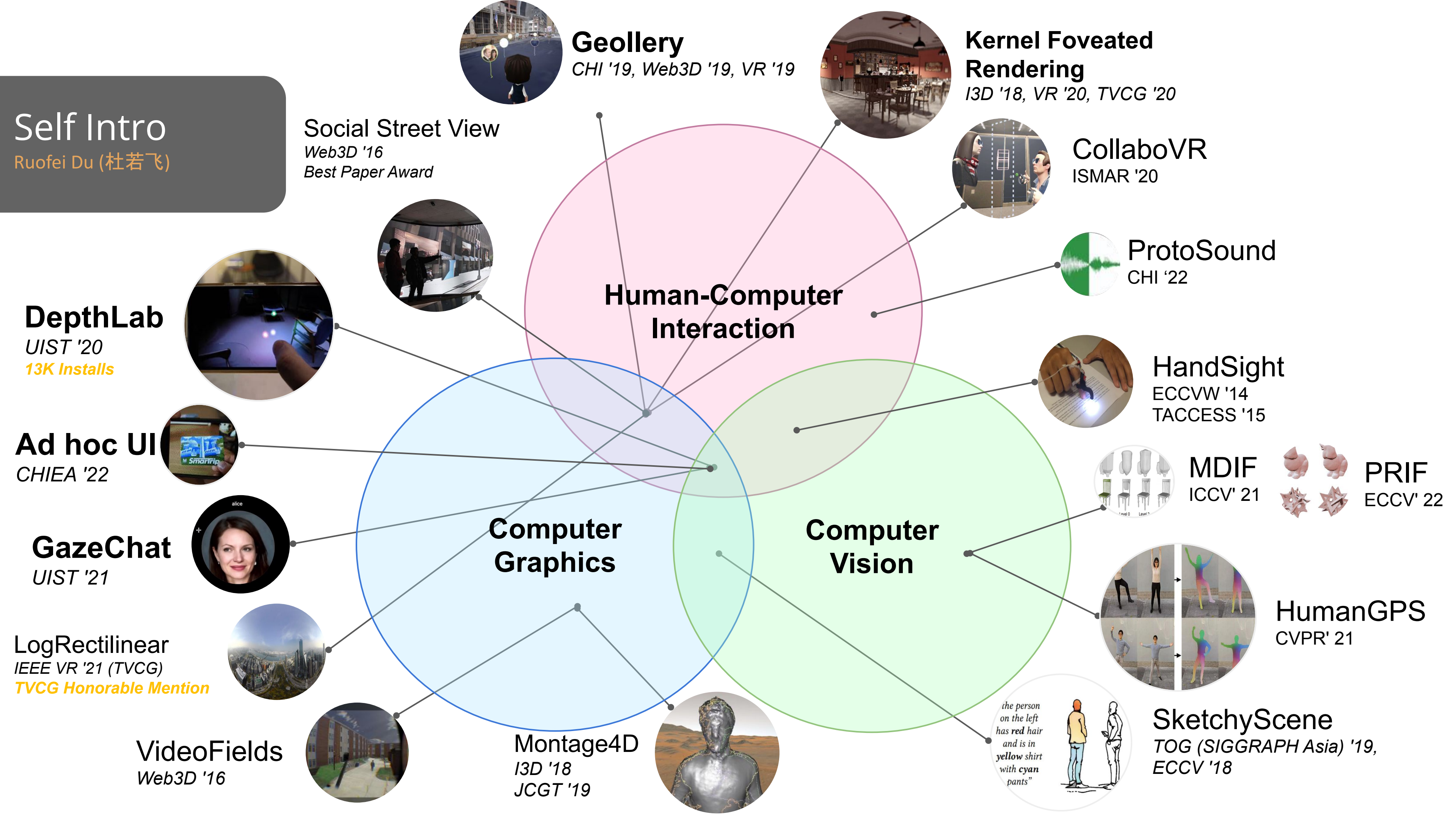
“Slurp” Revisited: Using 'system re-presencing' to look back on, encounter, and design with the history of spatial interactivity and locative media
Shengzhi Wu, Daragh Byrne, Ruofei Du, Molly Steenson • *ACM Conference on Designing Interactive Systems*, ACM (2022)

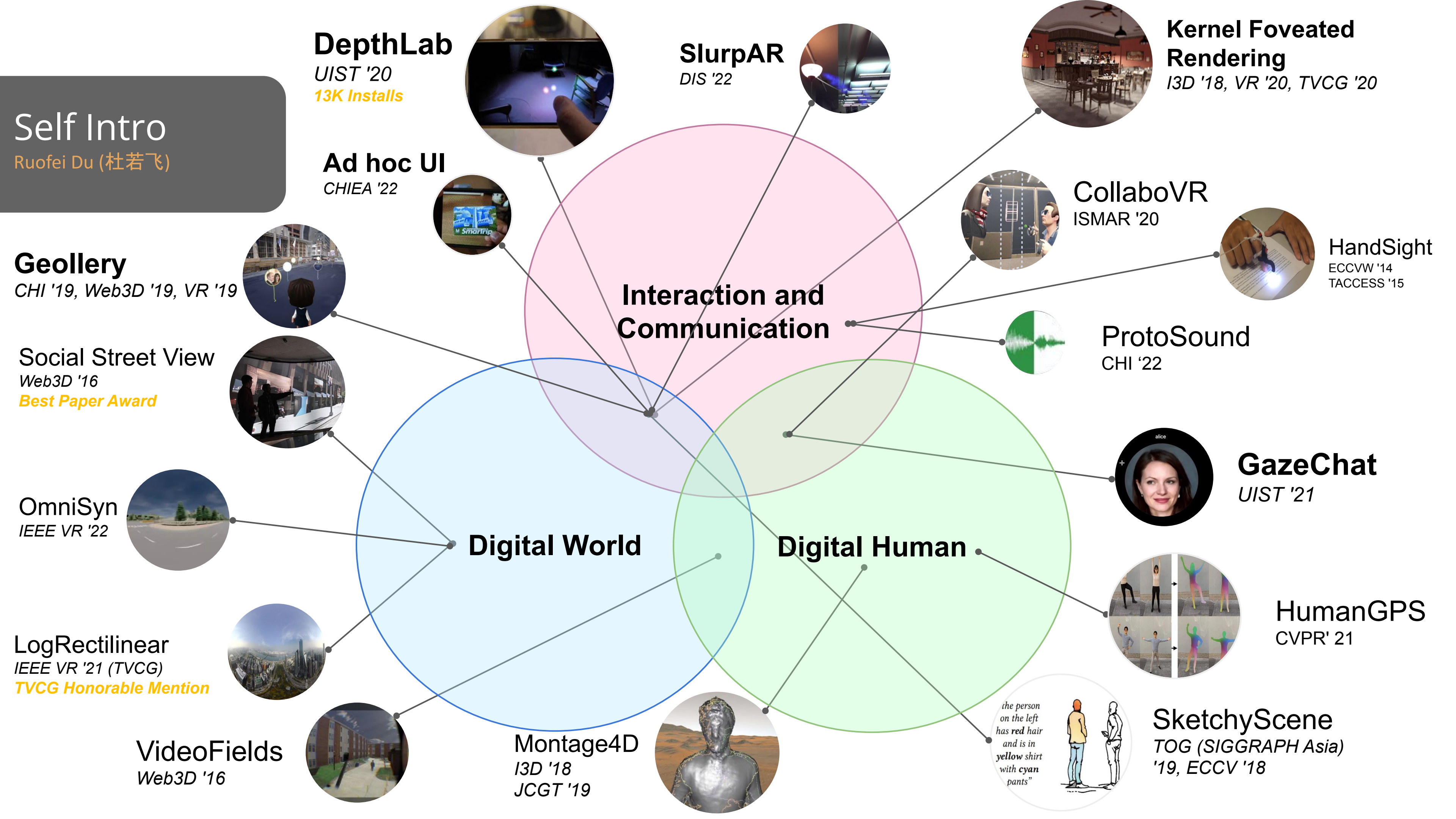
OmniSyn: Synthesizing 360 Videos with Wide-baseline Panoramas
David Li, Yinda Zhang, Christian Haene, Danhang "Danny" Tang, Amitabh Varshney, Ruofei Du • *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE

Opportunistic Interfaces for Augmented Reality: Transforming Everyday Objects into Tangible 6DoF Interfaces Using Ad hoc UI
Ruofei Du, Mathieu Le Goc, Alex Olwal, Shengzhi Wu, Danhang "Danny" Tang, Yinda Zhang, Jun Zhang, David Joseph New Tan, Federico Tombari, David Kim • *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, ACM

PRIF: Primary Ray-based Implicit Function
Brandon Yushan Feng, Yinda Zhang, Danhang "Danny" Tang, Ruofei Du, Amitabh Varshney • *European Conference on Computer Vision (ECCV)* (2022)

ProtoSound: A Personalized and Scalable Sound Recognition System for Deaf and Hard-of-Hearing Users
DJ Jain, Khoa Huynh Anh Nguyen, Steven Goodman, Rachel Grossman-Kahn, Hung Ngo, Aditya Kusupati, Ruofei Du, Alex Olwal, Leah Findlater, Jon E. Froehlich • *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI)*, ACM, pp. 24


Self Intro
Ruofei Du (杜若飞)
Geollery
CHI '19, Web3D '19, VR '19
Kernel Foveated Rendering
I3D '18, VR '20, TVCG '20
Social Street View
Web3D '16
Best Paper Award
CollaboVR
ISMAR '20
ProtoSound
CHI '22
Human-Computer Interaction
DepthLab
UIST '20
13K Installs
HandSight
ECCVW '14
TACCESS '15
Ad hoc UI
CHIEA '22
MDIF
ICCV' 21
PRIF
ECCV' 22
GazeChat
UIST '21
Computer Graphics
Computer Vision
HumanGPS
CVPR' 21
LogRectilinear
IEEE VR '21 (TVCG)
TVCG Honorable Mention
VideoFields
Web3D '16
Montage4D
I3D '18
JCGT '19
SketchyScene
TOG (SIGGRAPH Asia) '19,
ECCV '18



Self Intro
Ruofei Du (杜若飞)
DepthLab
UIST '20
13K Installs
SlurpAR
DIS '22
Kernel Foveated Rendering
I3D '18, VR '20, TVCG '20
Ad hoc UI
CHIEA '22
CollaboVR
ISMAR '20
HandSight
ECCVW '14
TACCESS '15
Geollery
CHI '19, Web3D '19, VR '19
Interaction and Communication
ProtoSound
CHI '22
Social Street View
Web3D '16
Best Paper Award
GazeChat
UIST '21
OmniSyn
IEEE VR '22
Digital World
Digital Human
HumanGPS
CVPR' 21
LogRectilinear
IEEE VR '21 (TVCG)
TVCG Honorable Mention
SketchyScene
TOG (SIGGRAPH Asia) '19, ECCV '18
VideoFields
Web3D '16
Montage4D
I3D '18
JCGT '19
the person on the left has red hair and is in yellow shirt with cyan pants

元宇宙中的交互计算与包容普惠

Computational Interaction for a Universally Accessible Metaverse

Ruofei Du (杜若飞)
Senior Research Scientist
Google Labs, San Francisco
www.ruofeidu.com
me@duruofei.com

Metaverse

Neal Stephenson, 1992.

Metaverse


ROBLOX


horizon
Worlds
Meta

Metaverse

- Future of Internet?
- Internet of Things?
- Virtual Reality?
- Augmented Reality?
- Decentralization?
- Blockchain + NFT?
- Mirrored World?
- Digital Twin?
- VR OS?
- Web 3.0?

Metaverse

Metaverse envisioned a ***persistent*** digital world where people are fully connected as ***virtual representations***,

As a teenager, my dream was to live in a metaverse...

However, today I wish metaverse is only a tool to **make information more useful and accessible** and help people to **live a better physical life**.

元宇宙中的交互计算与包容普惠

Computational Interaction for a Universally Accessible Metaverse

第一章 · 镜像世界与实时渲染

Ruofei Du (杜若飞)
Senior Research Scientist
Google Labs, San Francisco
www.ruofeidu.com
me@duruofei.com

Project Geollery.com: Reconstructing a Live Mirrored World With Geotagged Social Media

Digital twin?
Metaverse?

Ruofei Du[†], David Li[†], and Amitabh Varshney

{ruofei, dli7319, varshney}@umiacs.umd.edu | www.Geollery.com | ACM CHI 2019 & Web3D 2016 Best Paper Award & 2019

UNIVERSITY OF MARYLAND

UMIACS

THE AUGMENTARIUM
VIRTUAL AND AUGMENTED REALITY LAB
AT THE UNIVERSITY OF MARYLAND

COMPUTER SCIENCE
UNIVERSITY OF MARYLAND, COLLEGE PARK

Introduction
Social Media

image courtesy: plannedparenthood.org

Introduction

Social Media + Topics

image courtesy: huffingtonpost.com

Motivation

Social Media + XR

Motivation

Social Media + XR

image courtesy:
instagram.com,
facebook.com,
twitter.com

Motivation

2D layout

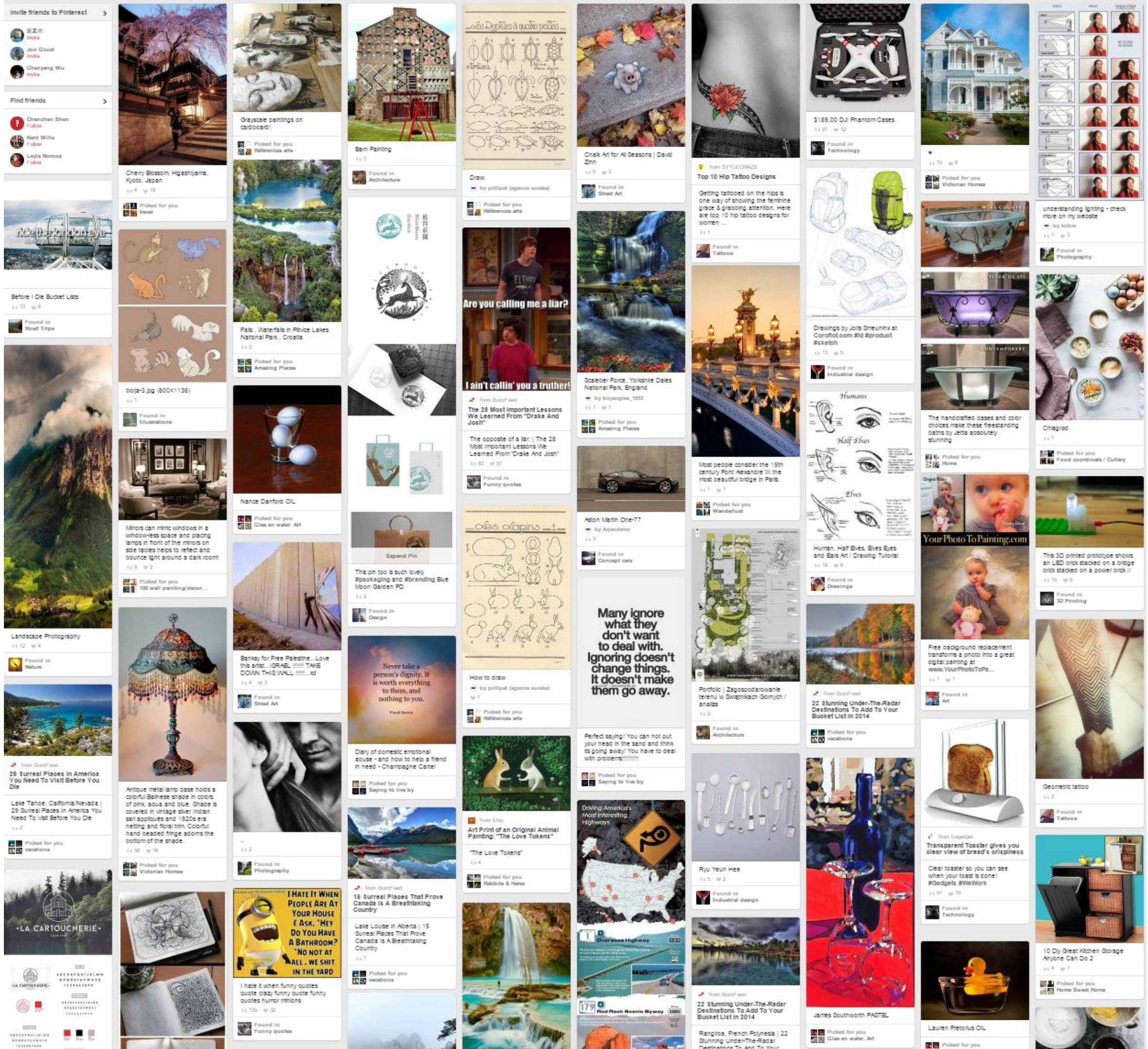


image courtesy: pinterest.com

Motivation

Immersive Mixed Reality?

image courtesy: viralized.com

Motivation

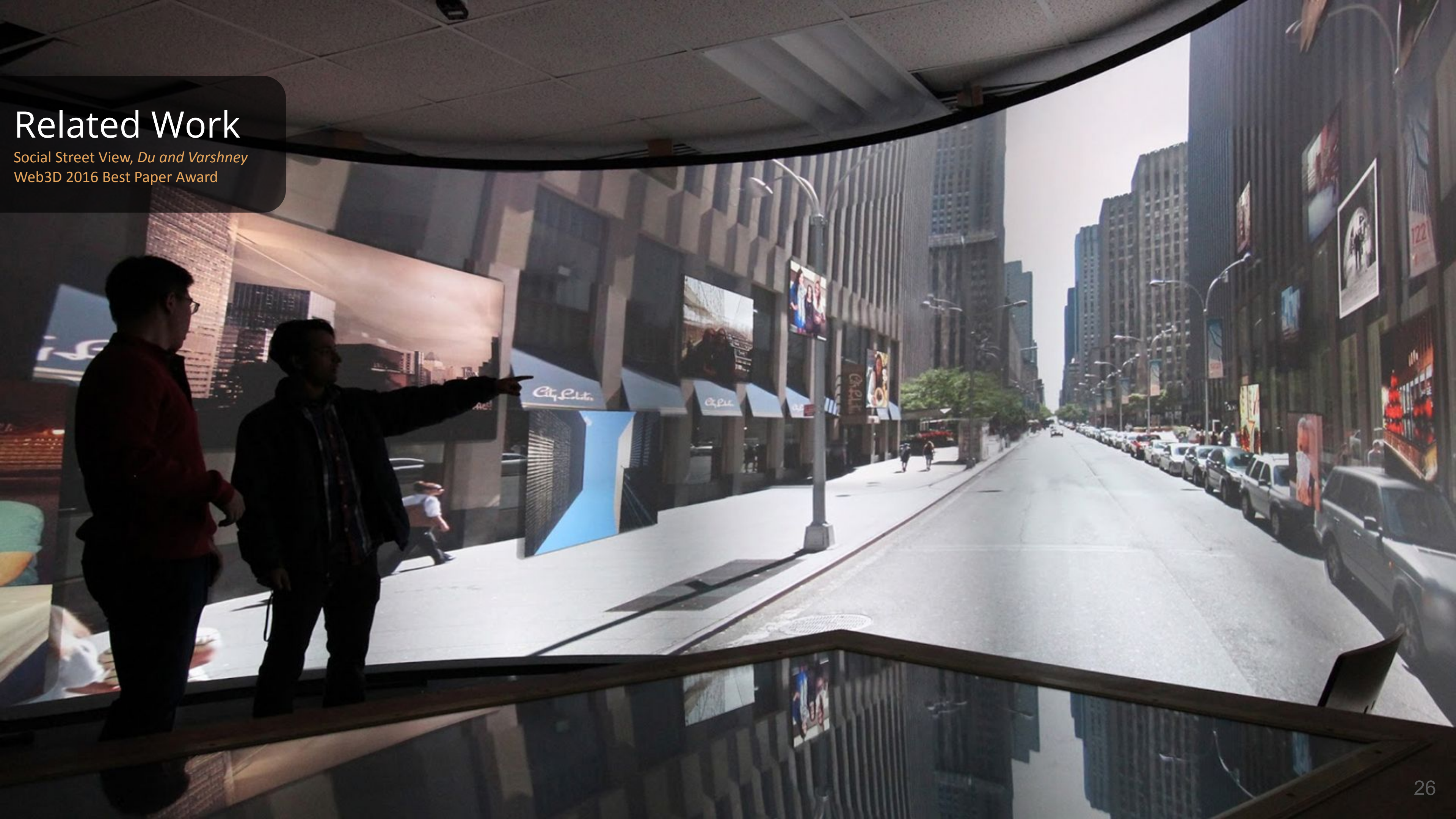
Pros and cons of the classic

Motivation

Pros and cons of the classic

Related Work

Social Street View, *Du and Varshney*
Web3D 2016 Best Paper Award



Related Work

Social Street View, *Du* and *Varshney*
Web3D 2016 Best Paper Award

Related Work

Social Street View, *Du* and *Varshney*
Web3D 2016 Best Paper Award

Related Work

3D Visual Popularity

Bulbul and *Dahyot, 2017*

Related Work

Virtual Oulu, *Kukka et al.*
CSCW 2017

Related Work

Immersive Trip Reports
Brejcha et al. UIST 2018

Related Work

High Fidelity, Inc.

Related Work

Facebook Spaces, 2017

What's Next?

Research Question 1/3

What may a social media platform look like in mixed reality?

What's Next?

Research Question 2/3

What if we could allow social media sharing in a live mirrored world?

What's Next?

Research Question 3/3

What use cases can we benefit from social media platform in XR?

Geollery.com

A Mixed-Reality Social Media Platform


Real-time Texturing

Conception, architecting & implementation

1

Geollery

A mixed reality system that can depict geotagged social media and online avatars with 3D textured buildings.

Extending the design space of

2

3D Social Media Platform

Progressive streaming, aggregation approaches, virtual representation of social media, co-presence with virtual avatars, and collaboration modes.

Conducting a user study of

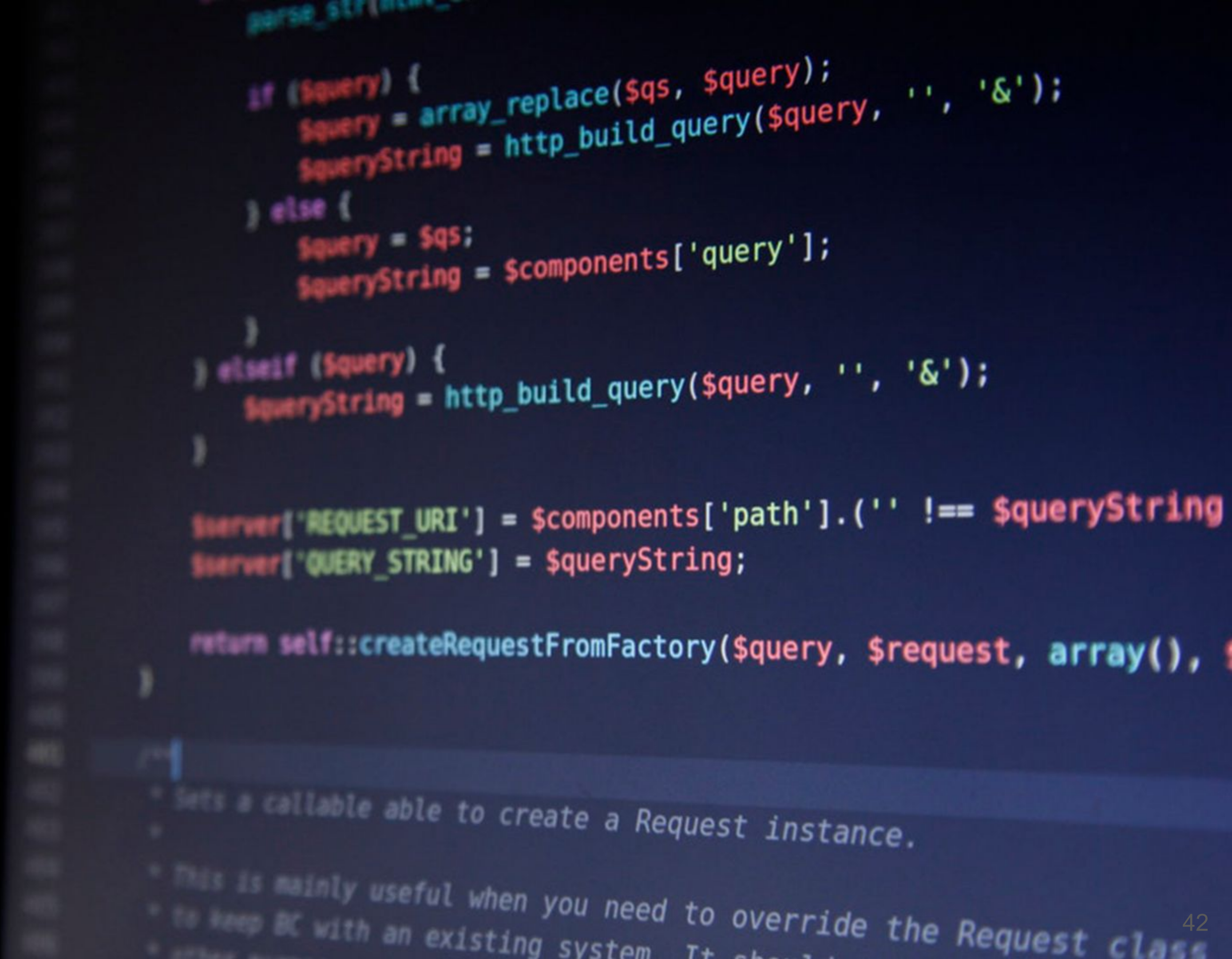
3

Geollery vs. Social Street View

by discussing their benefits, limitations, and potential impacts to future 3D social media platforms.

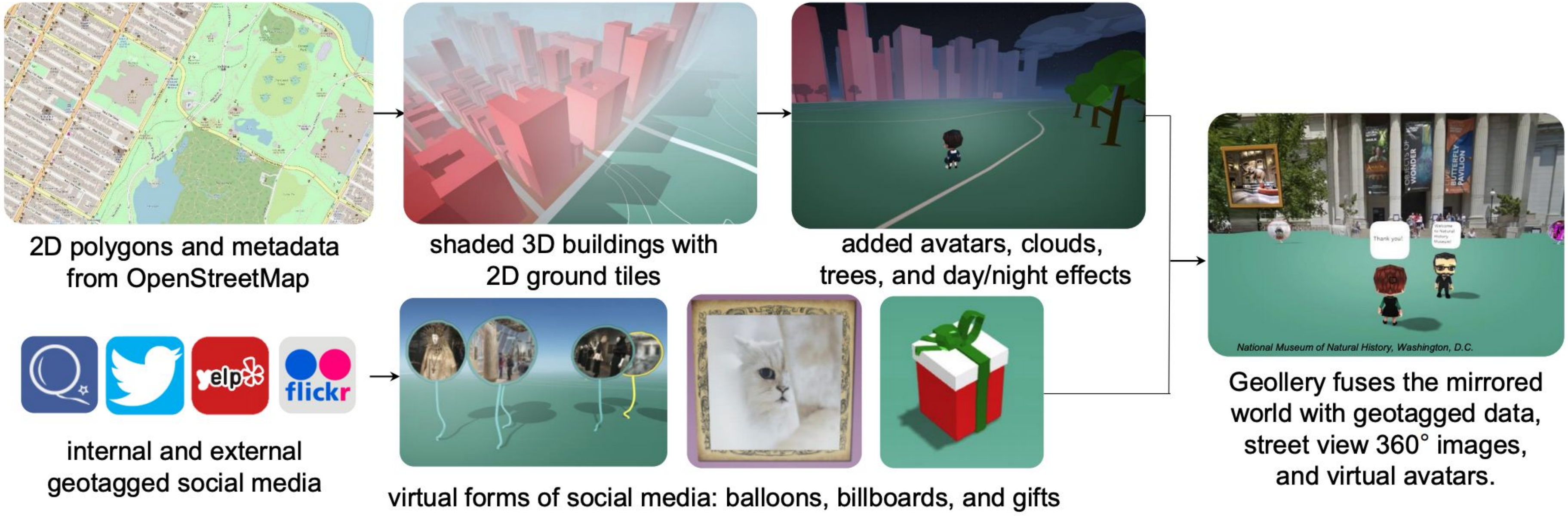
System Overview

Geollery Workflow



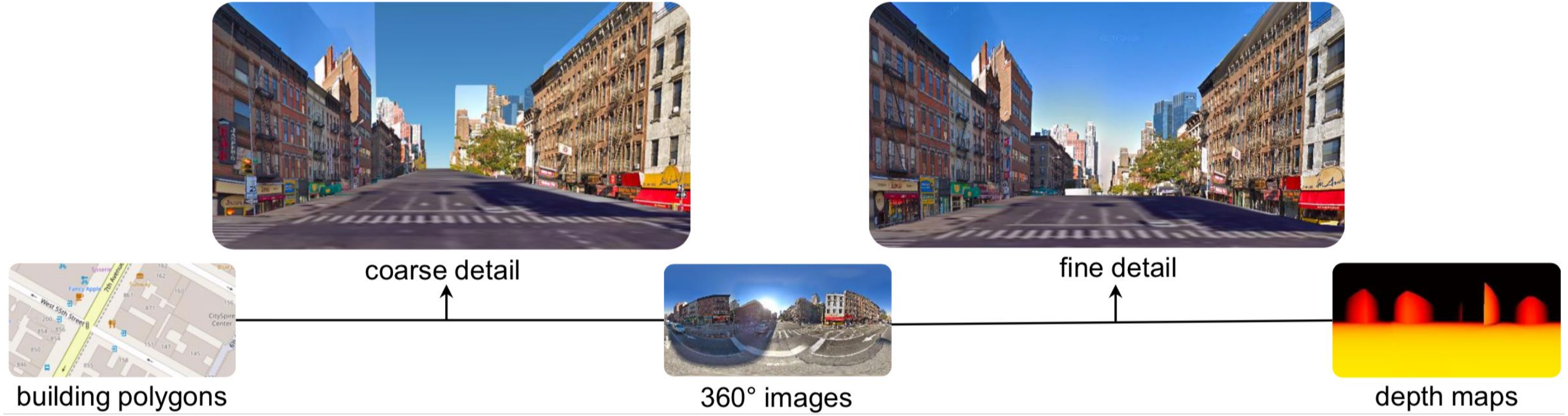
System Overview

Geollery Workflow



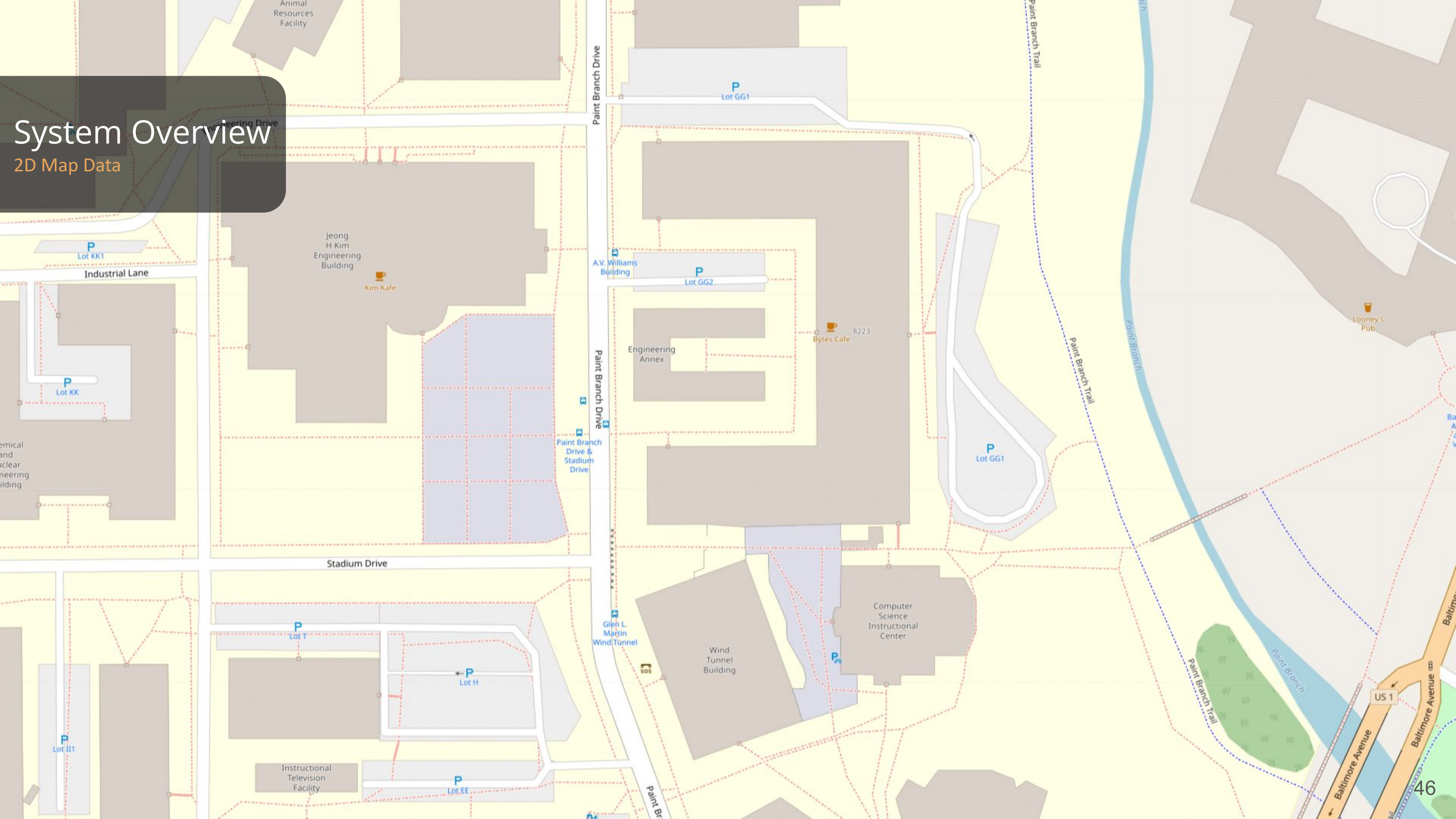
Geollery.com

v2: a major leap



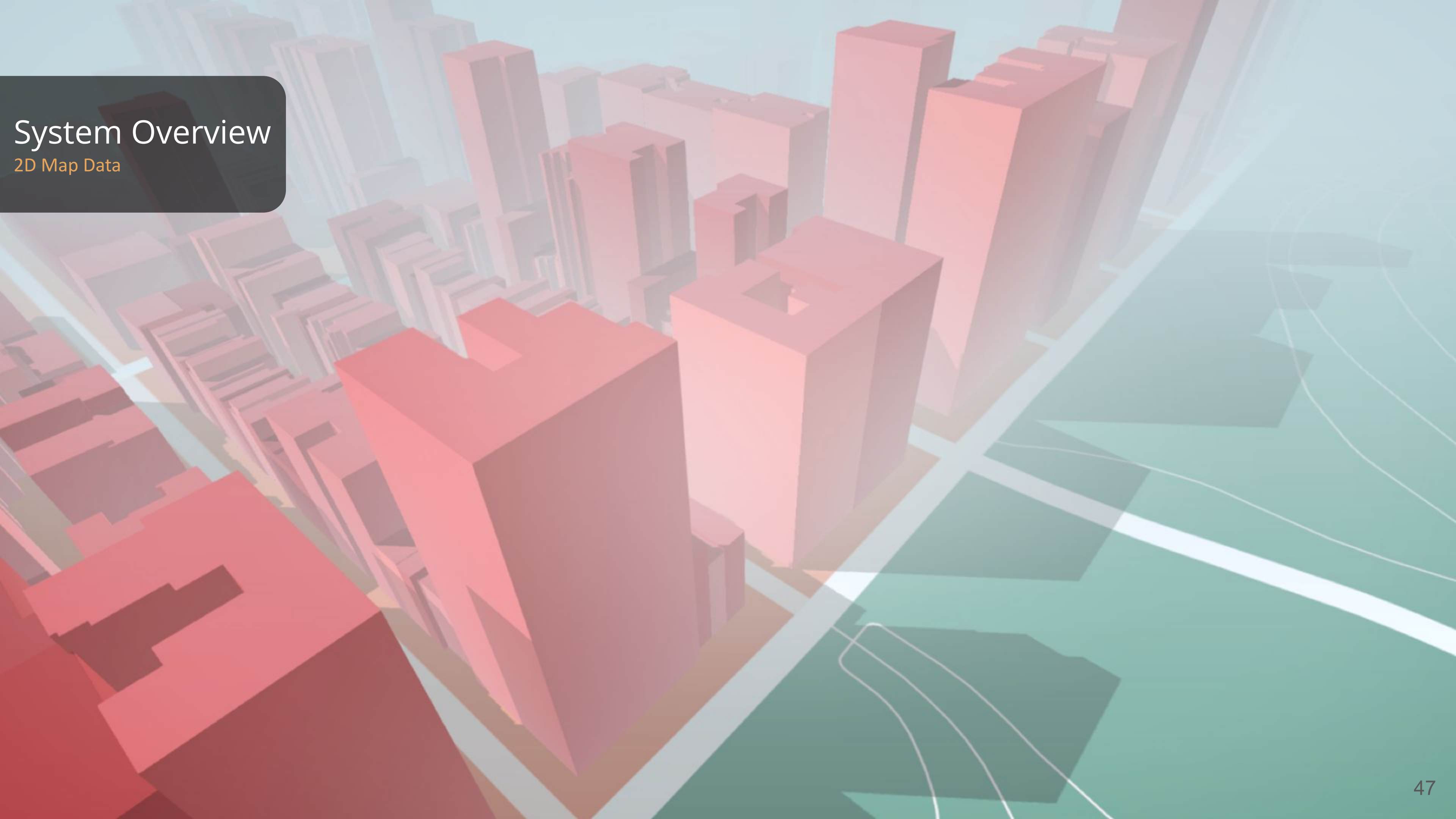
System Overview

2D Map Data



System Overview

2D Map Data



System Overview

+Avatar +Trees +Clouds

System Overview

+Avatar +Trees +Clouds +Night

System Overview

Street View Panoramas

System Overview

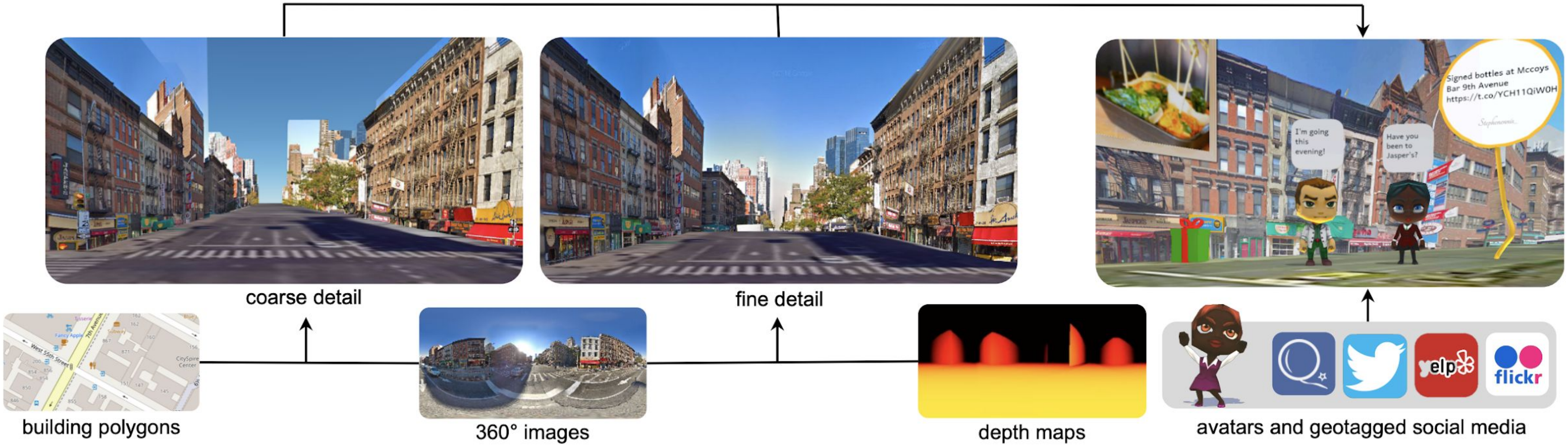
Street View Panoramas

System Overview

Street View Panoramas

System Overview

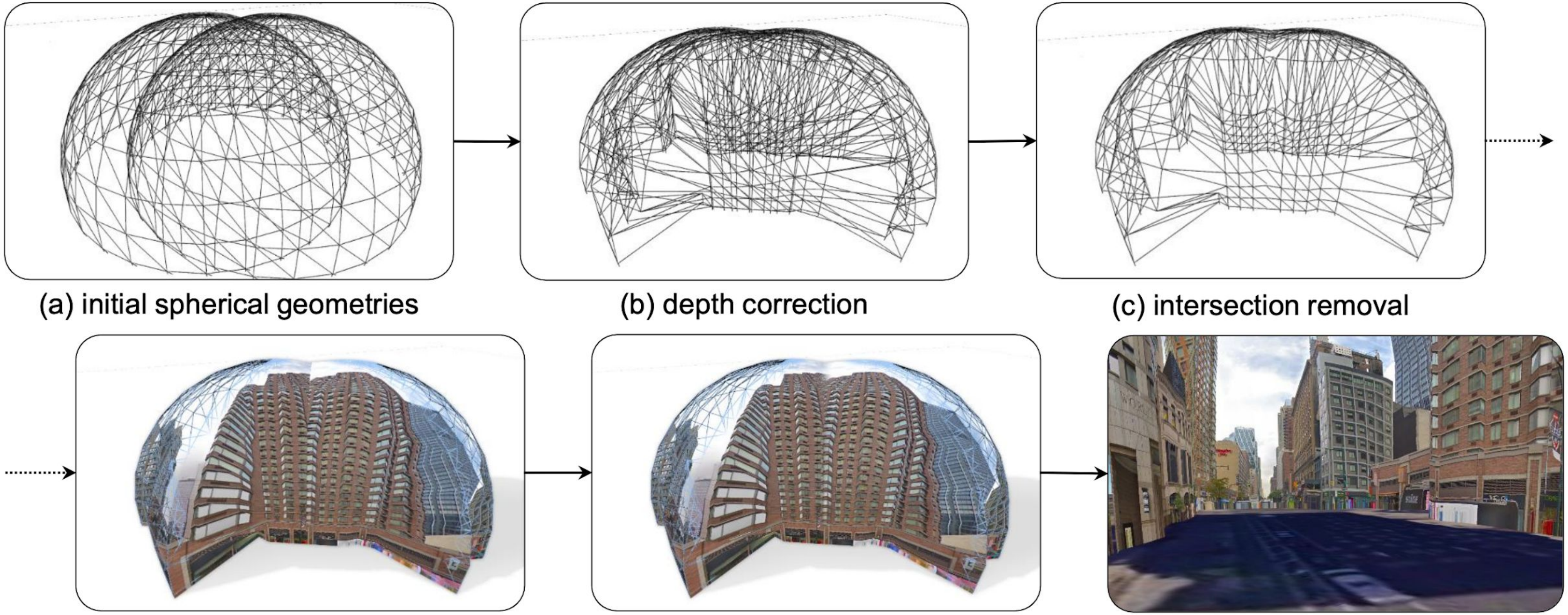
Geollery Workflow



All data we used is publicly and widely available on the Internet.

Rendering Pipeline

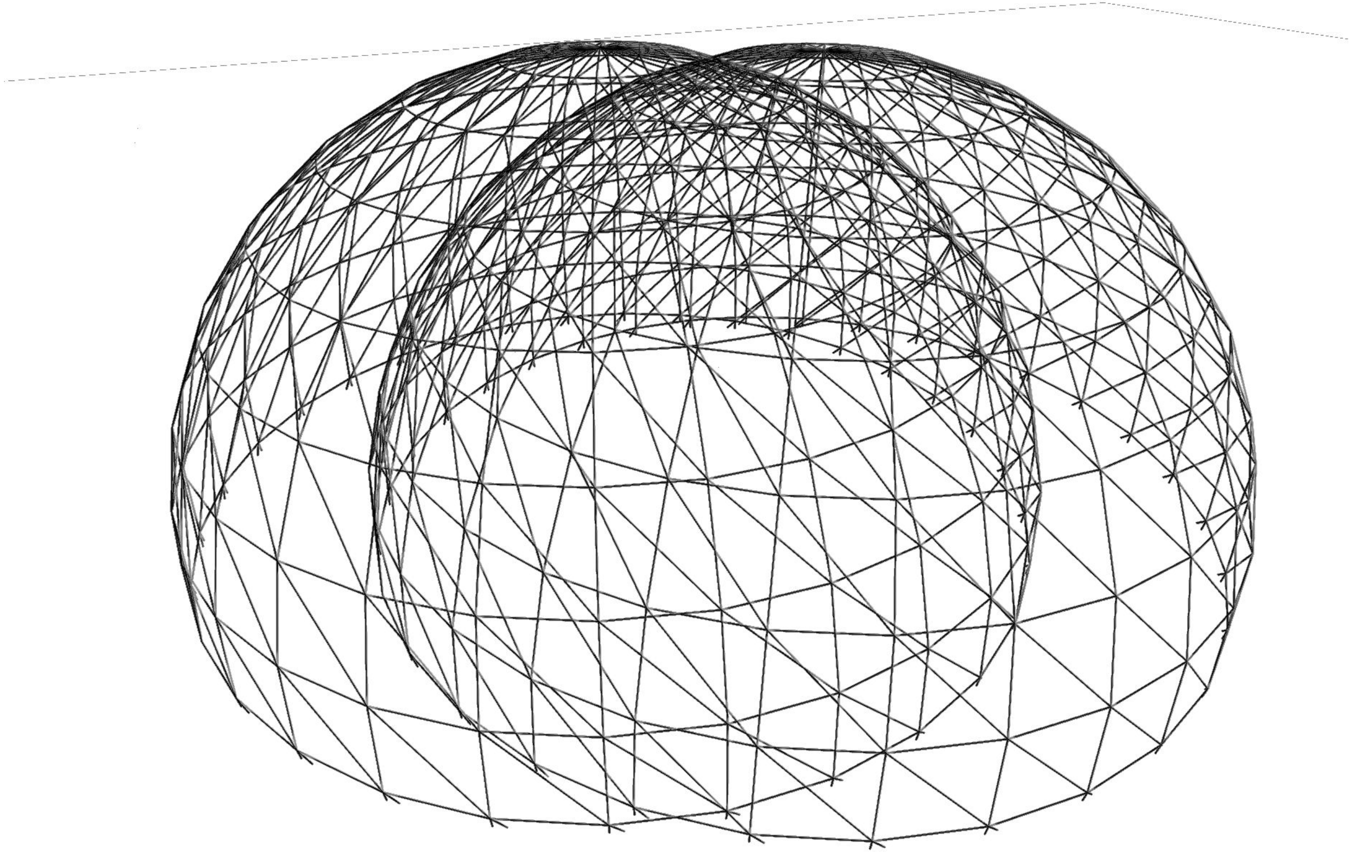
Close-view Rendering



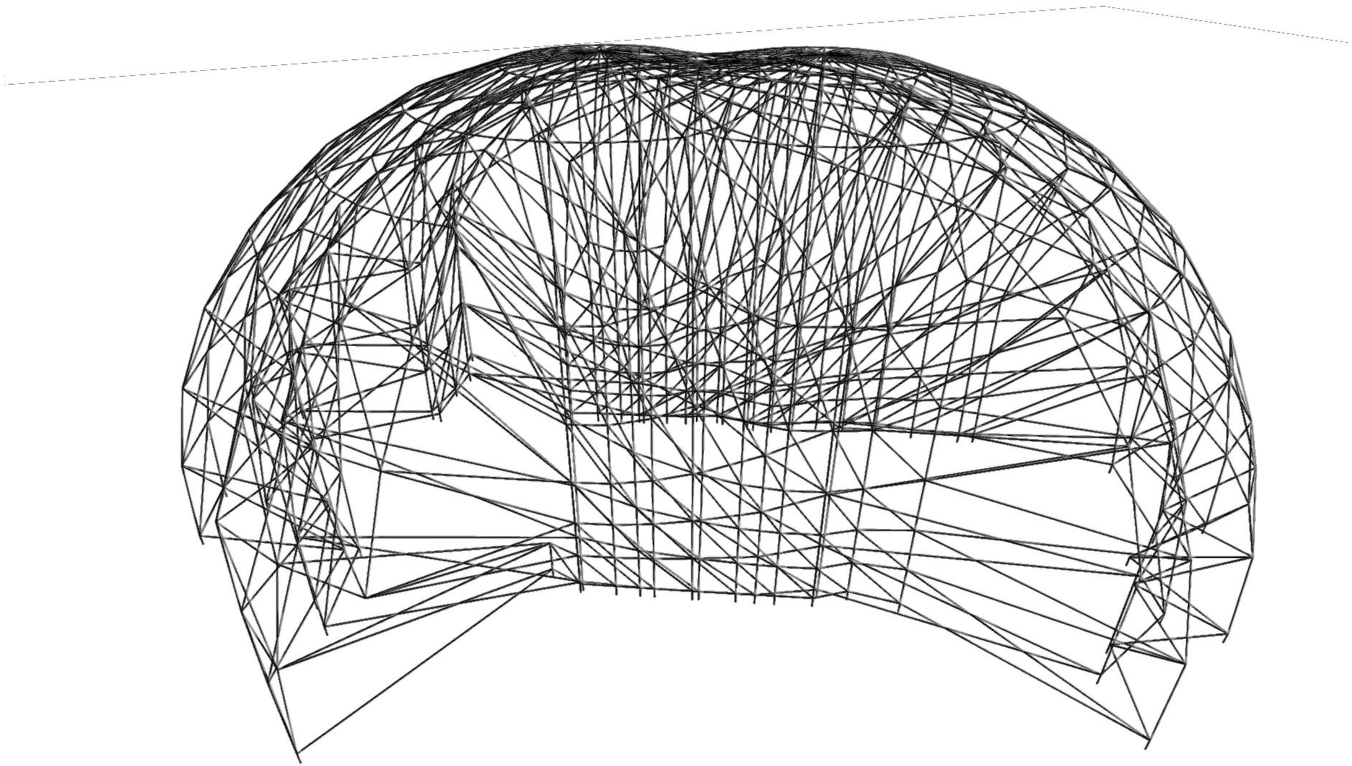
(a) initial spherical geometries
(b) depth correction
(c) intersection removal
(d) texturing individual geometry
(e) texturing with alpha blending
(f) rendering results in fine detail

Rendering Pipeline

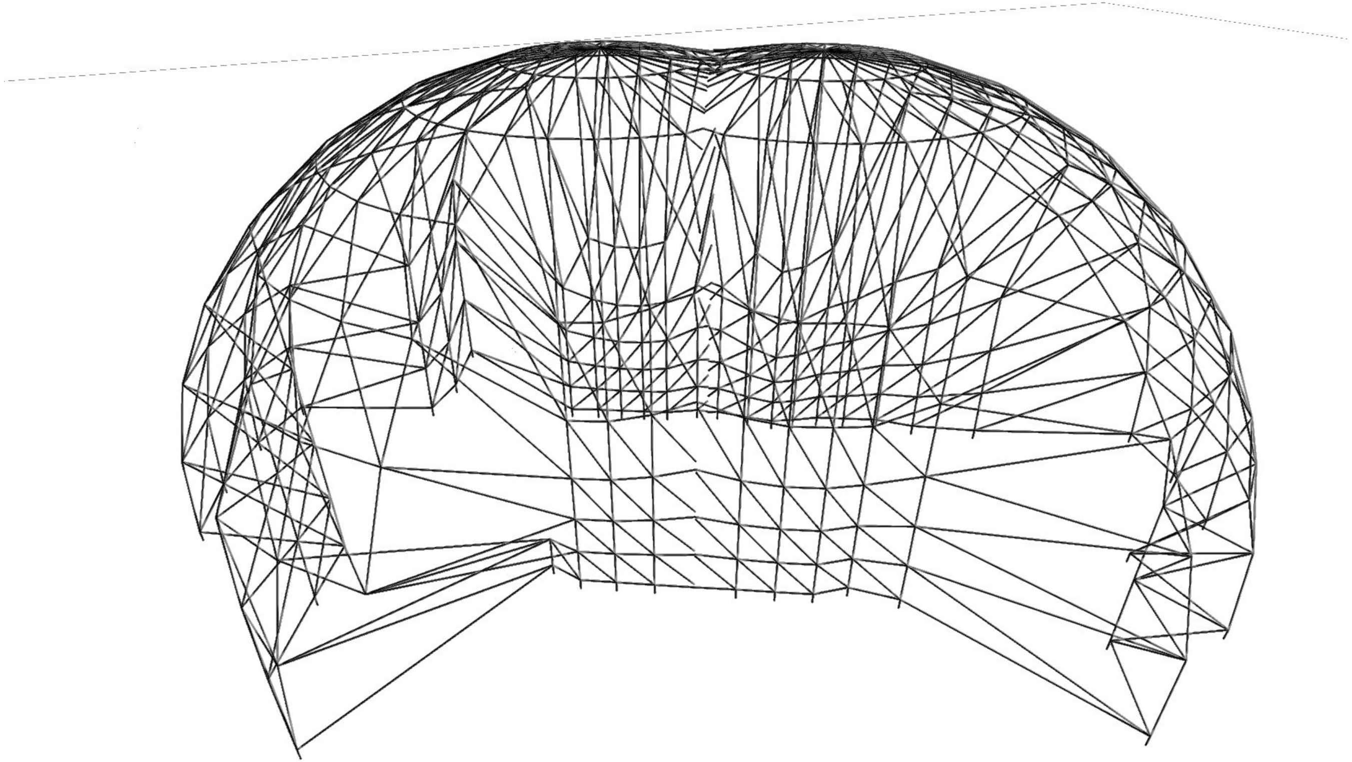
Initial spherical geometries



Rendering Pipeline
Depth correction

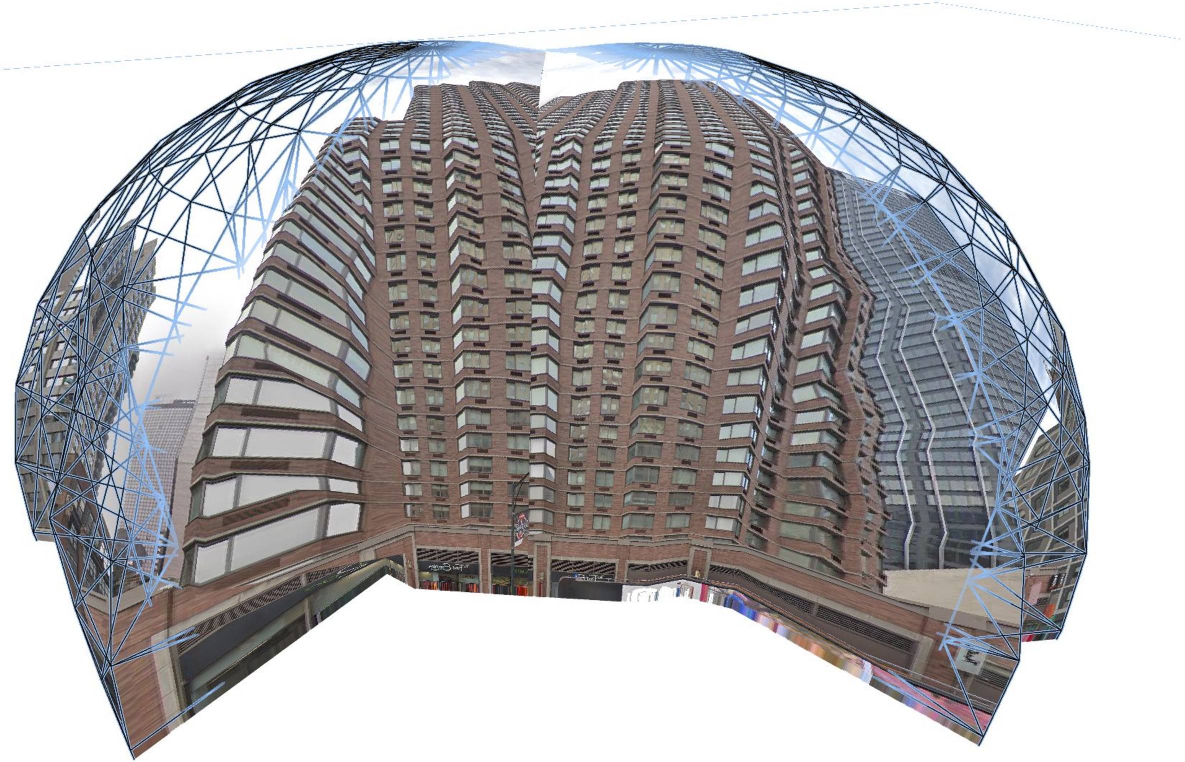


Rendering Pipeline
Intersection removal

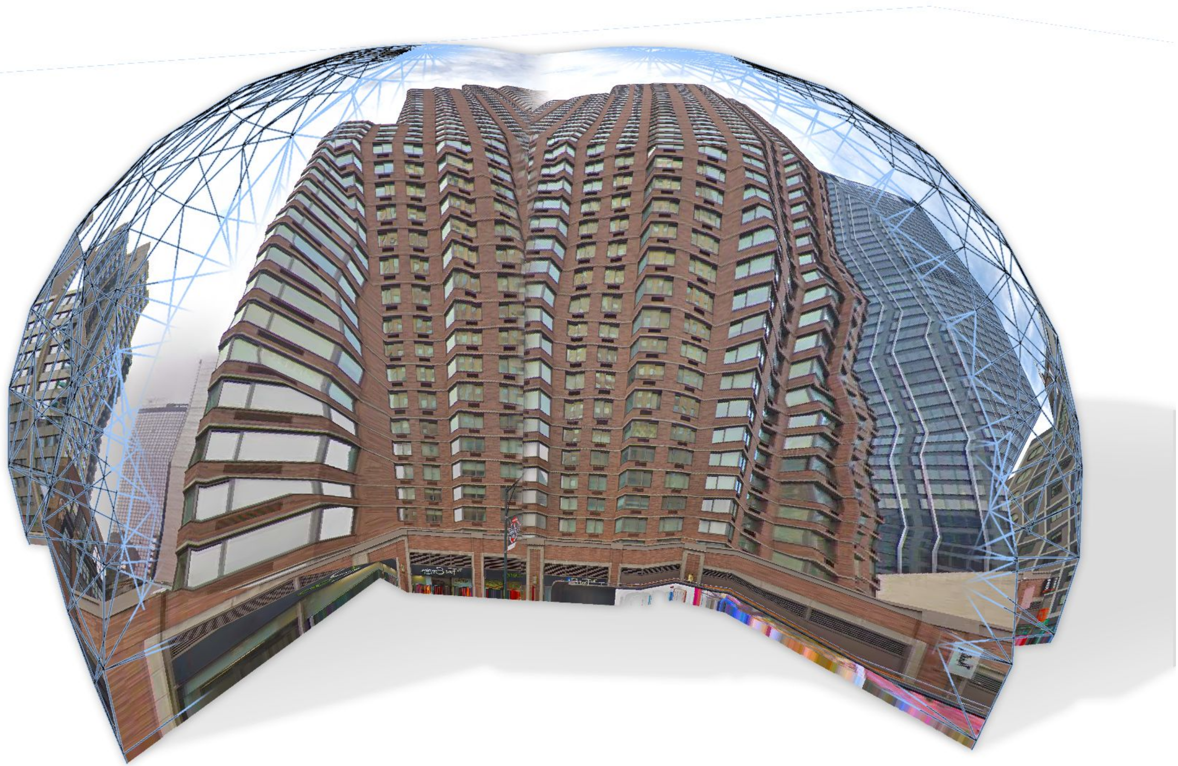


Rendering Pipeline

Texturing individual geometry

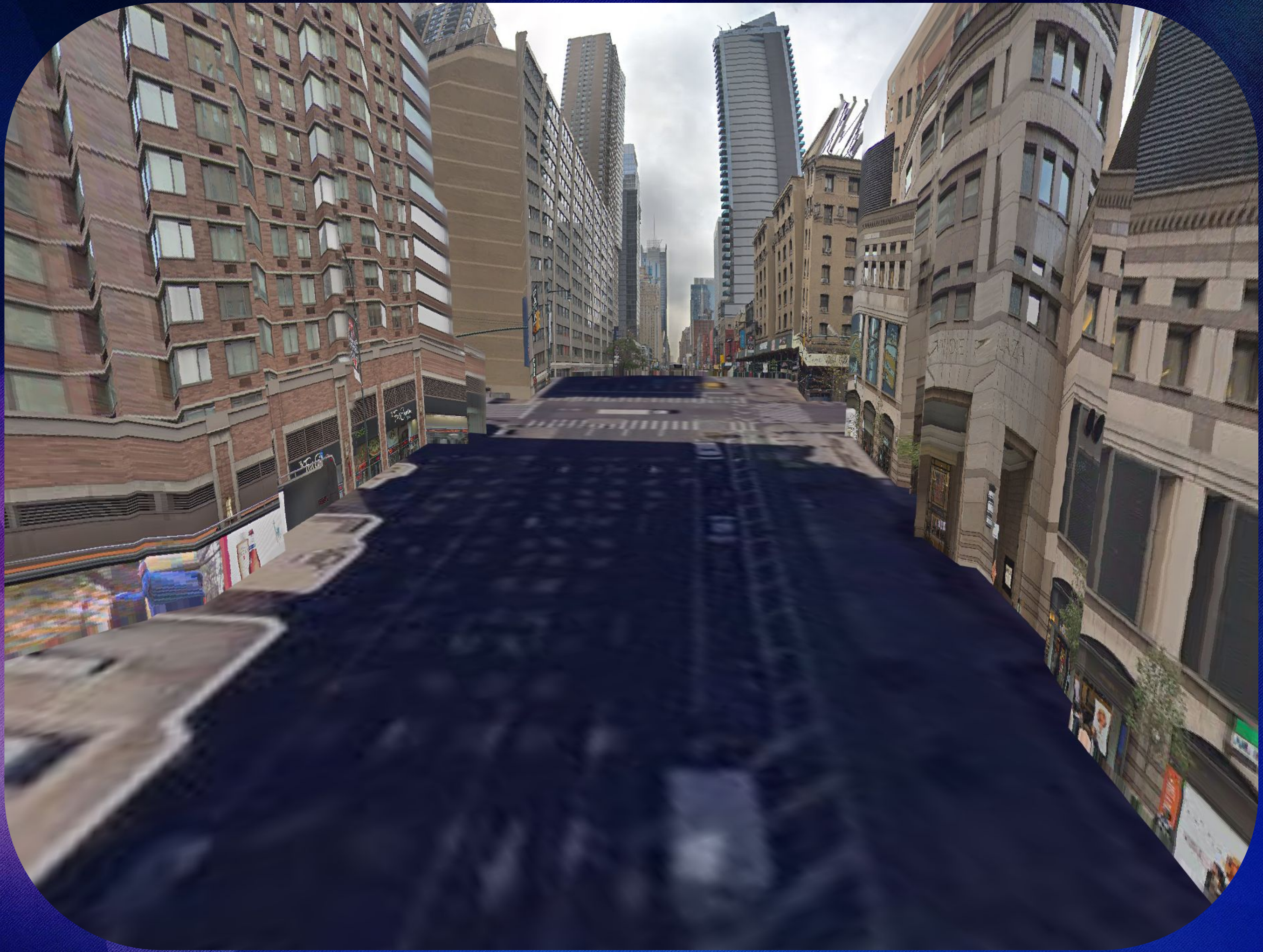


Rendering Pipeline
Texturing with alpha blending



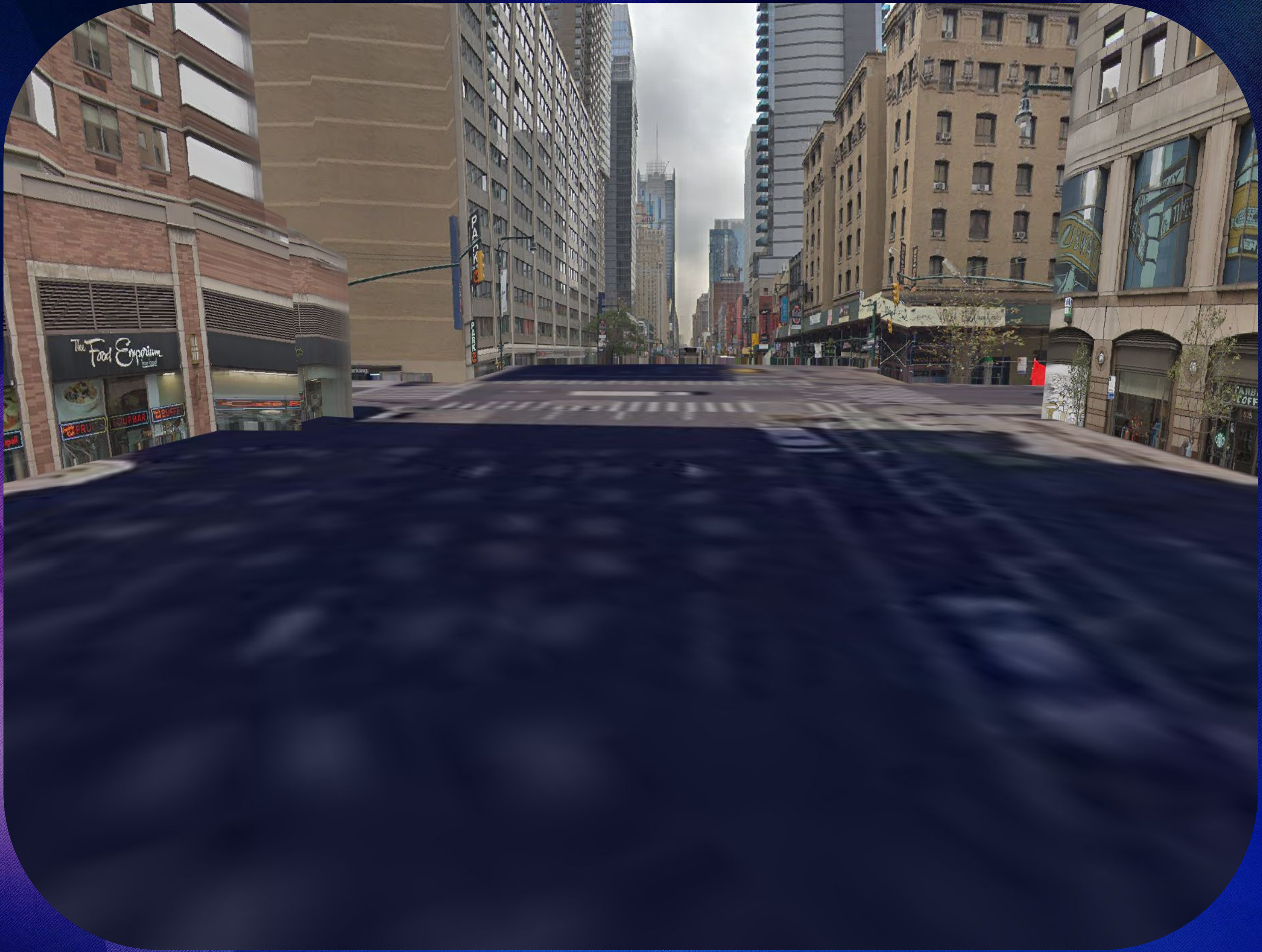
Rendering Pipeline

Rendering result in the fine detail



Rendering Pipeline

Rendering result in the fine detail

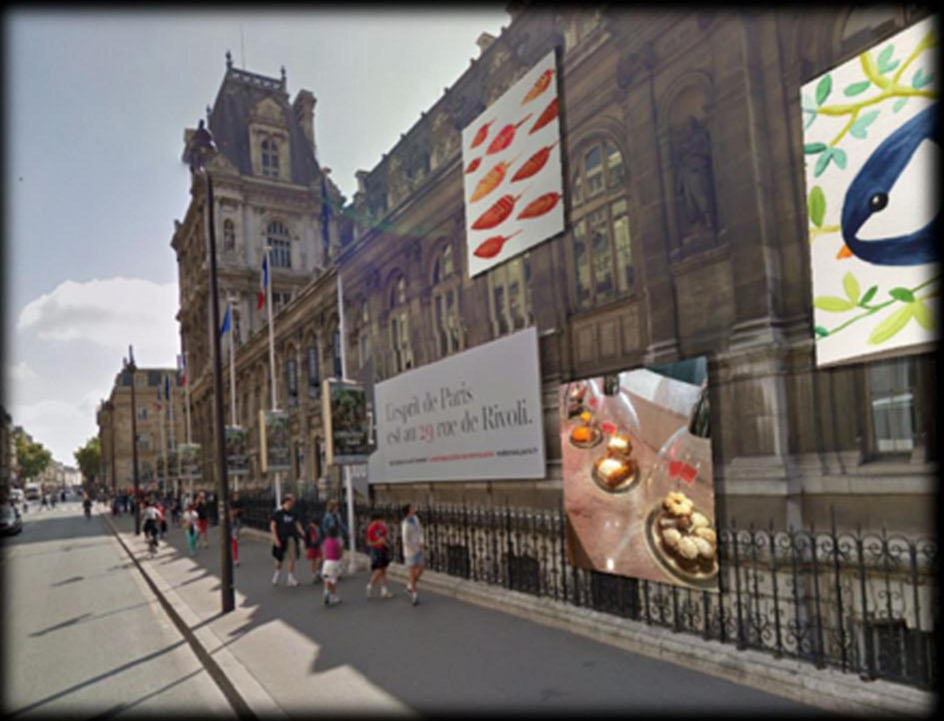


Rendering Pipeline

Rendering result in the fine detail

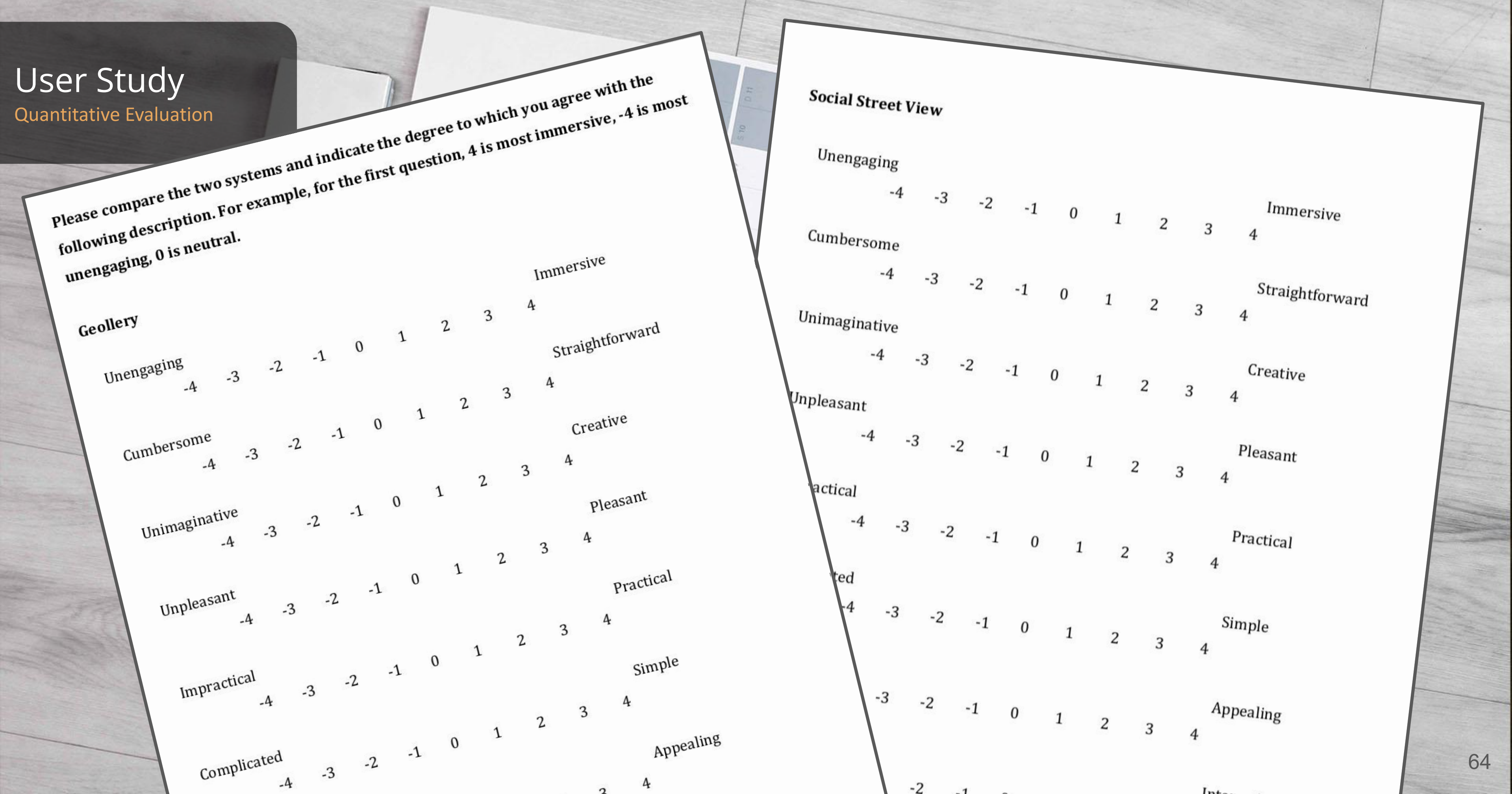
User Study

Social Street View vs. Geollery

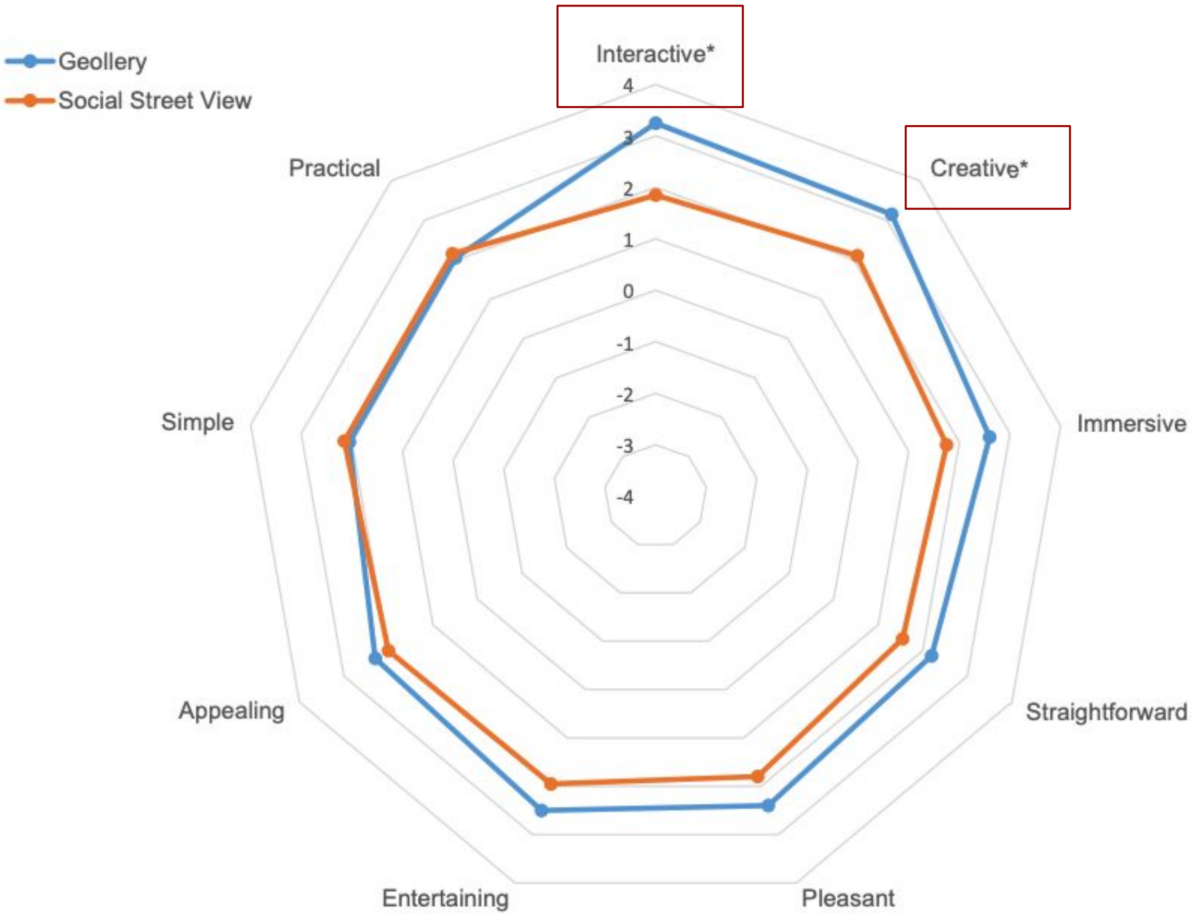


User Study

Quantitative Evaluation



User Study
Quantitative Evaluation



I would like to use it for the food in different restaurants. I am always hesitating of different restaurants. It will be very easy to *see all restaurants with street views*. In Yelp, I can only see one restaurant at a time.

P6 / F

[I will use it for] exploring *new places*. If I am going on vacation somewhere, I could *immerse myself* into the location. If there are avatars around that area, I could *ask questions*.

P1 / M

I think it (Geollery) will be useful for families. I just taught my grandpa how to use Facetime last week and it would great if I could teleport to their house and meet with them, then we could chat and share photos with our avatars.

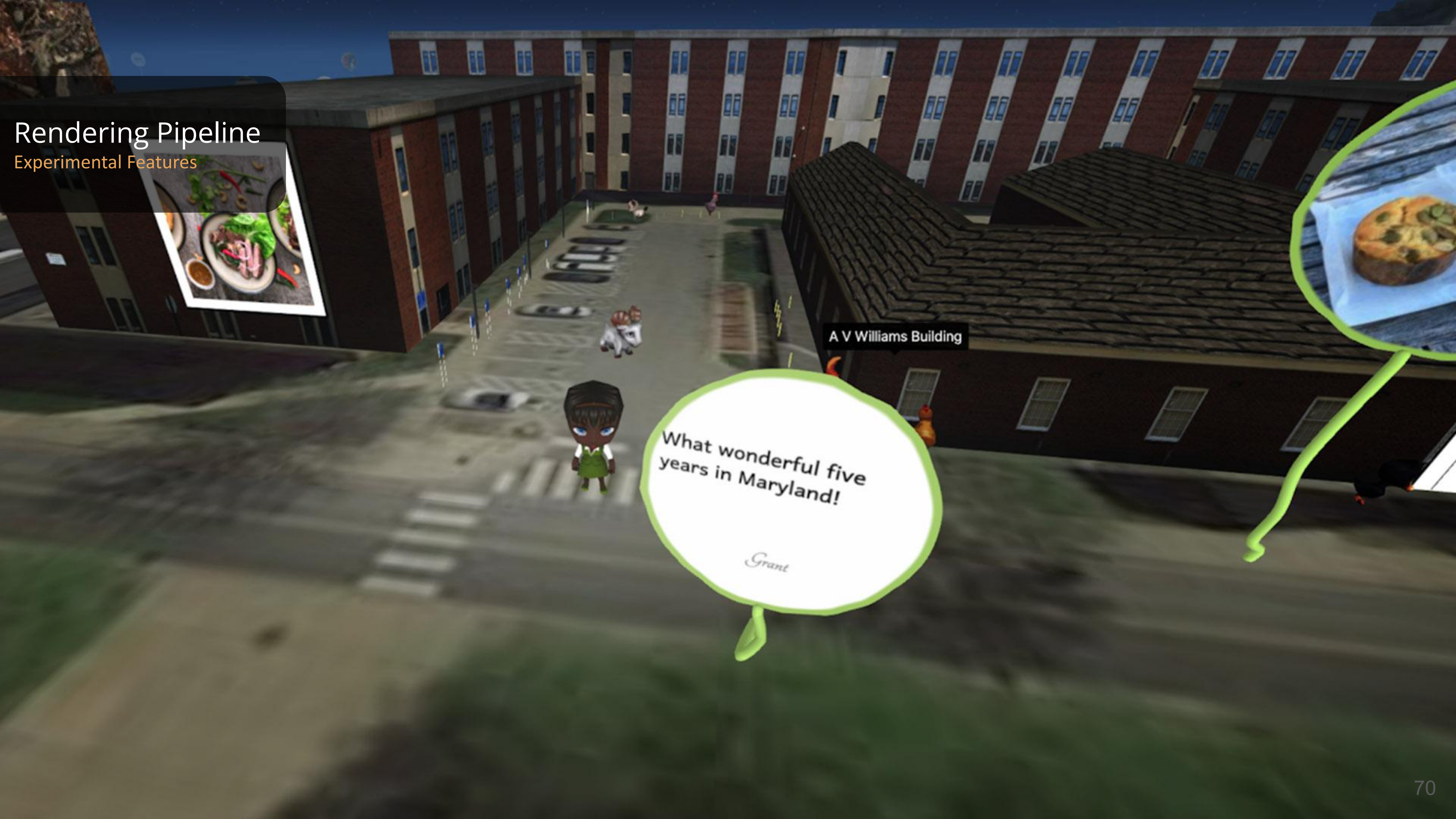
P2 / F

"if there is a way to unify the interaction between them, there will be more realistic buildings [and] you could have more roof structures. Terrains will be interesting to add on."

P18 / M

Rendering Pipeline

Experimental Features



Instant Panoramic Texture Mapping with Semantic Object Matching for Large-Scale Urban Scene Reproduction

TVCG 2021, Jinwoo Park, Ik-beom Jeon, Student Members, Sung-eui Yoon, and Woontack Woo

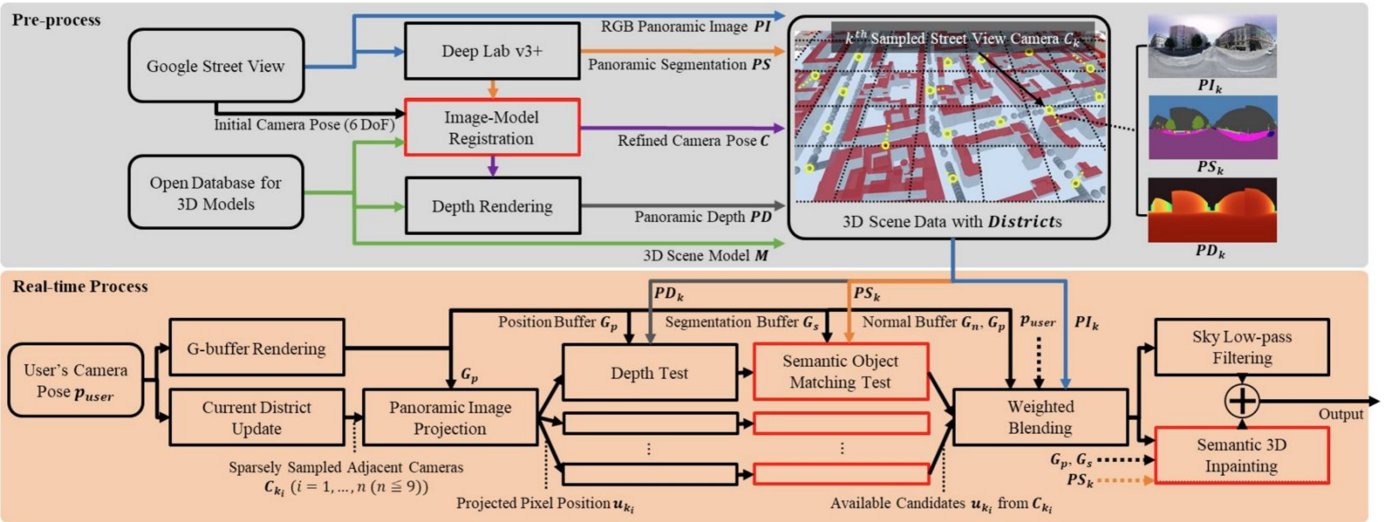


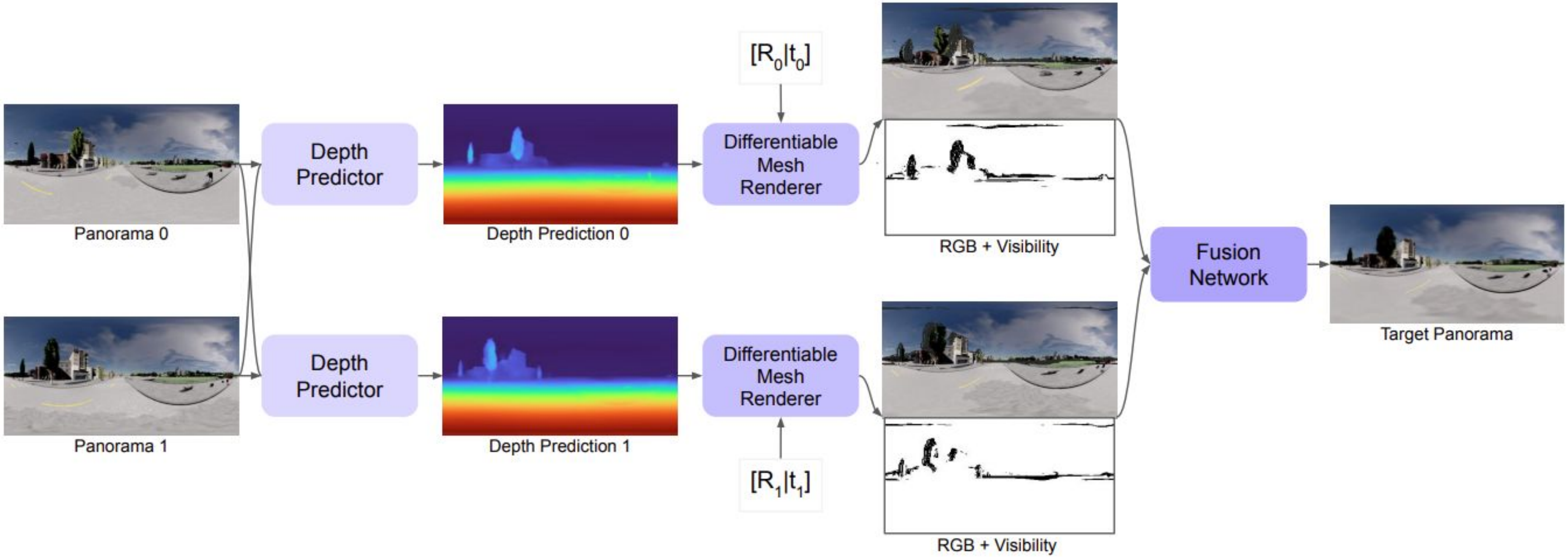
Fig. 2: Overview. In a pre-process, our system constructs *3D Scene Data*, which contains five different input resources: street-view images, 3D models, estimated panoramic-segmentation images, synthetic panoramic-depth images, and refined extrinsic camera parameters. For sparse sampling of street-view images according to a user's current position, *3D Scene Data* is divided into smaller *Districts*. In a real-time process, pixels in a position buffer at a current view are projected onto sampled street-view images to get texture colors. After passing a depth test and a semantic object matching test, filtered candidate colors are properly blended as a final color. As a final step of complete scene rendering without visual holes, low-pass filtering and semantic 3D inpainting are applied to sky and non-sky areas respectively. Note that our main contributions lie on utilizing semantic information in the proper intermediate steps (marked in red boxes) to enhance both rendering quality and performance time.

OmniSyn: Intermediate View Synthesis Between Wide-Baseline Panoramas

David Li, Yinda Zhang, Christian Häne, Danhang Tang, Amitabh Varshney, and Ruofei Du, VR 2022

OmniSyn: Intermediate View Synthesis Between Wide-Baseline Panoramas

David Li, Yinda Zhang, Christian Häne, Danhang Tang, Amitabh Varshney, and Ruofei Du, VR 2022



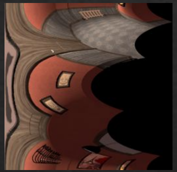
Kernel Foveated Rendering

Xiaoxu Meng, Ruofei Du, Matthias Zwicker and Amitabh Varshney
Augmentarium | UMIACS
University of Maryland, College Park

ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games 2018

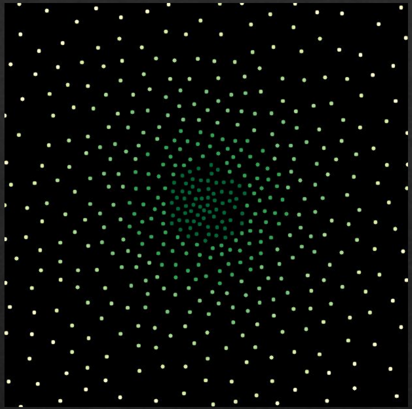
Original Frame

Buffer

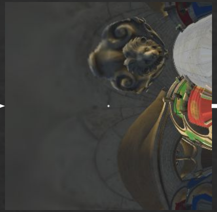


Screen

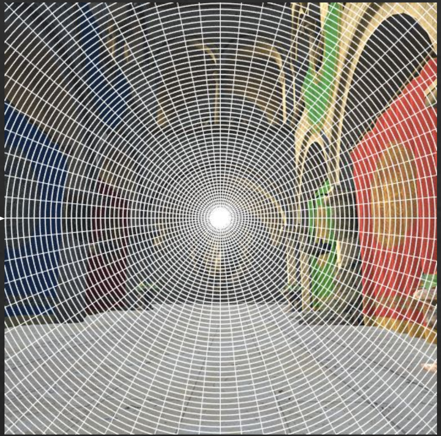
Sample Map



Cartesian coordinates
(x, y)



Kernel log-polar coordinates
(u, v)



Cartesian coordinates
(x, y)

Kernel Log-polar Mapping

$$u = K^{-1}\left(\frac{\log\sqrt{x^2+y^2}}{L}\right)\cdot w$$

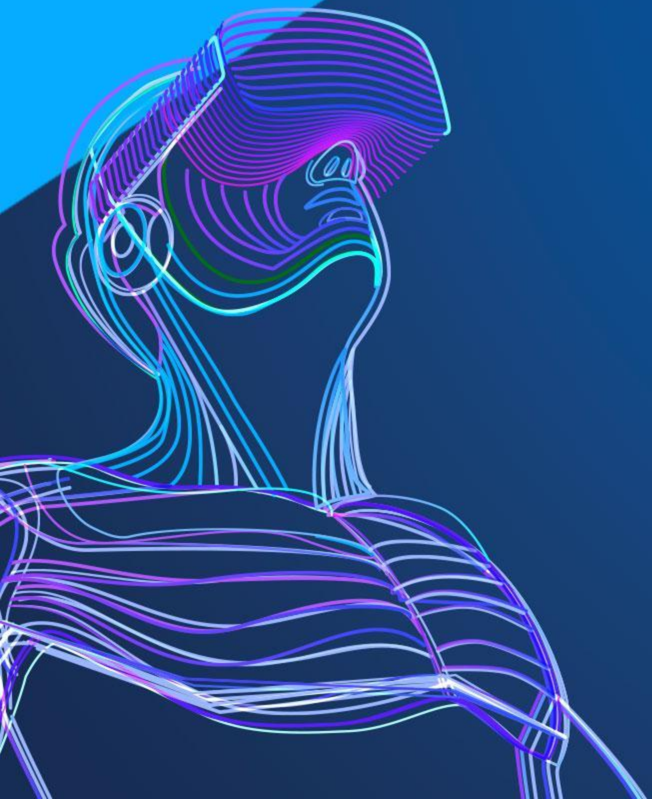
$$v = \frac{\left(\arctan\frac{y}{x} + \mathbf{1}[y-0]\cdot 2\pi\right)}{2\pi}\cdot h$$

$$x = e^{L\cdot K\left(\frac{u}{w}\right)}\cos\left(v\cdot\frac{2\pi}{h}\right)$$

$$y = e^{L\cdot K\left(\frac{u}{w}\right)}\sin\left(v\cdot\frac{2\pi}{h}\right)$$

- W: screen width H: screen height w: buffer width h: bu
- $\mathbf{1}[y<0] = \begin{cases} 1 & y<0 \\ 0 & y>0 \end{cases}$
- $L = \log\sqrt{W^2+H^2}$
- $K(x) = \sum_{i=0}^{\infty}\beta_i x^i$, where $\sum_{i=0}^{\infty}\beta_i = 1$

2020 IEEE VR ATLANTA



Eye-dominance-guided Foveated Rendering

Xiaoxu Meng, Ruofei Du, and Amitabh Varshney

IEEE Transactions on Visualization and Computer Graphics (TVCG)

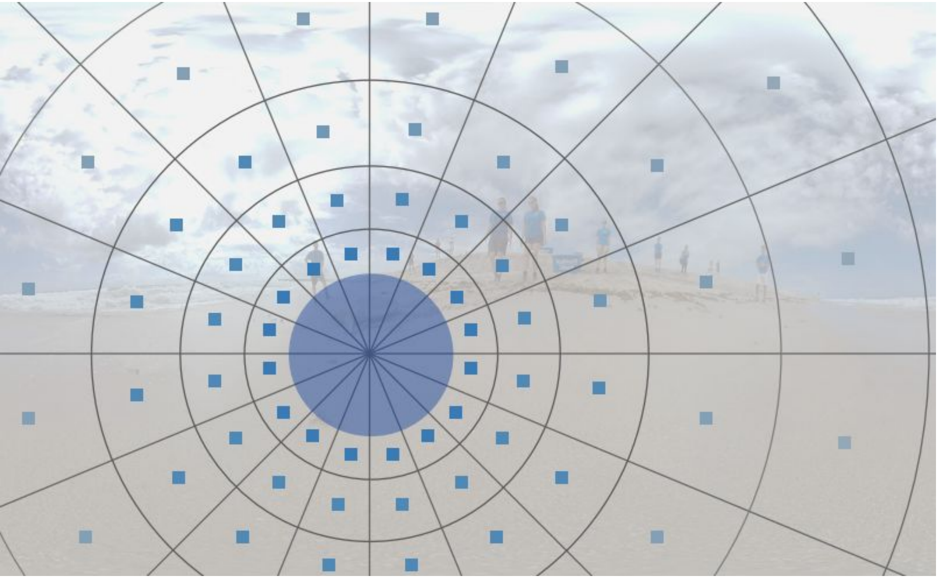
TVCG Honorable Mentions Award

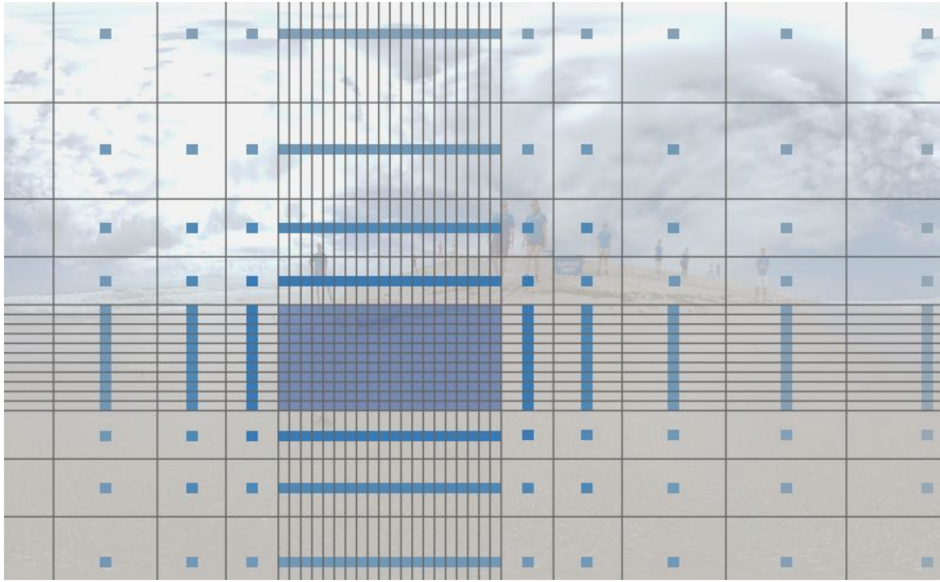
A Log-Rectilinear Transformation for Foveated 360-Degree Video Streaming

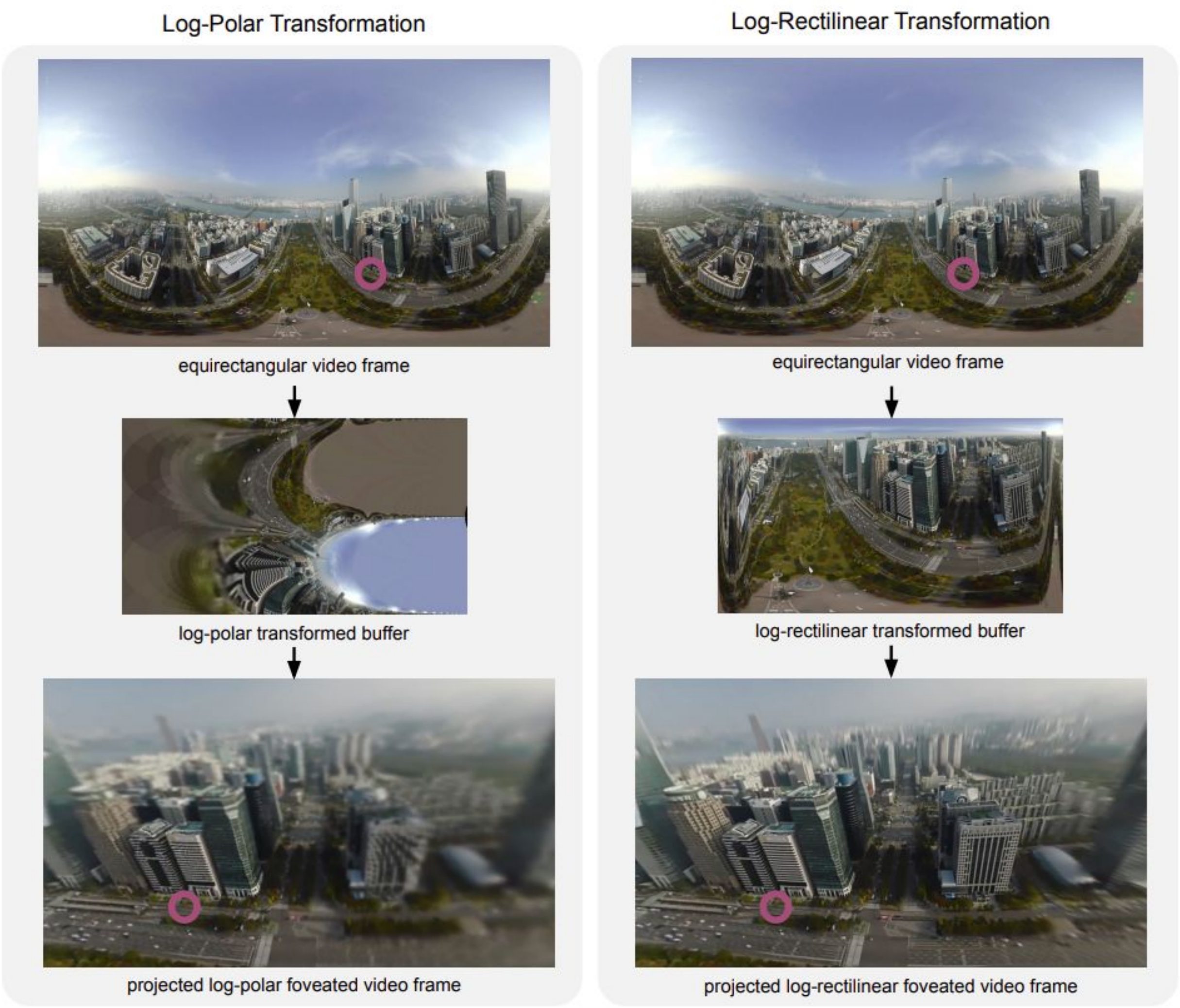
David Li[†], Ruofei Du[‡], Adharsh Babu[†], Camelia Brumar[†], Amitabh Varshney[†]

[†] University of Maryland, College Park [‡] Google

UNIVERSITY OF MARYLAND UMIACS Google






Log-Polar Transformation
equirectangular video frame
log-polar transformed buffer
projected log-polar foveated video frame
Log-Rectilinear Transformation
equirectangular video frame
log-rectilinear transformed buffer
projected log-rectilinear foveated video frame

元宇宙中的交互计算与包容普惠

Computational Interaction for a Universally Accessible Metaverse

第二章 · 交互计算算法与系统

Ruofei Du (杜若飞)
Senior Research Scientist
Google Labs, San Francisco
www.ruofeidu.com
me@duruofei.com

DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality

Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte,
Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces,
Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim

Google | ACM UIST 2020

Introduction
Mobile Augmented Reality

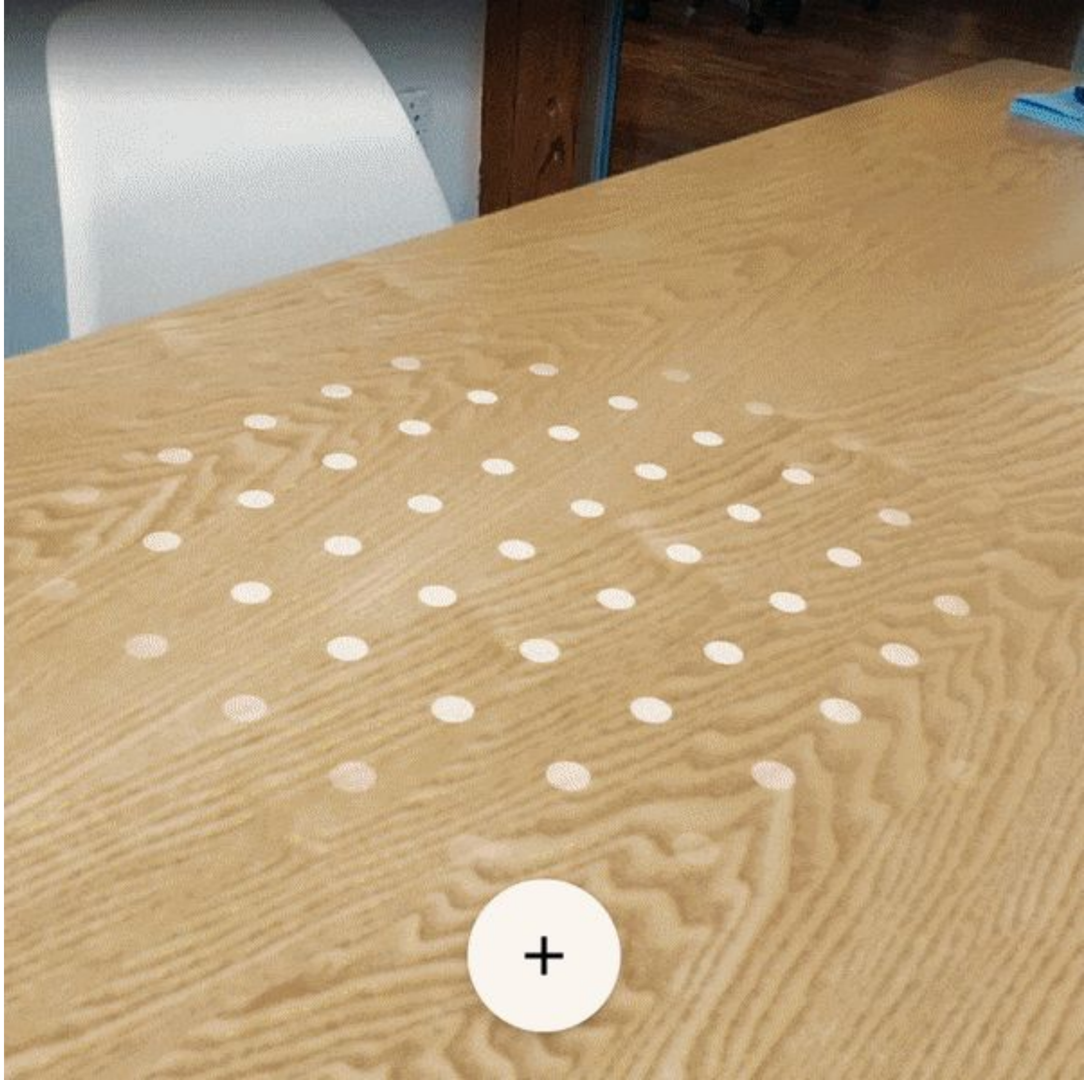
Introduction

Google's ARCore

Introduction

Google's ARCore

ARCore



Introduction

Mobile Augmented Reality

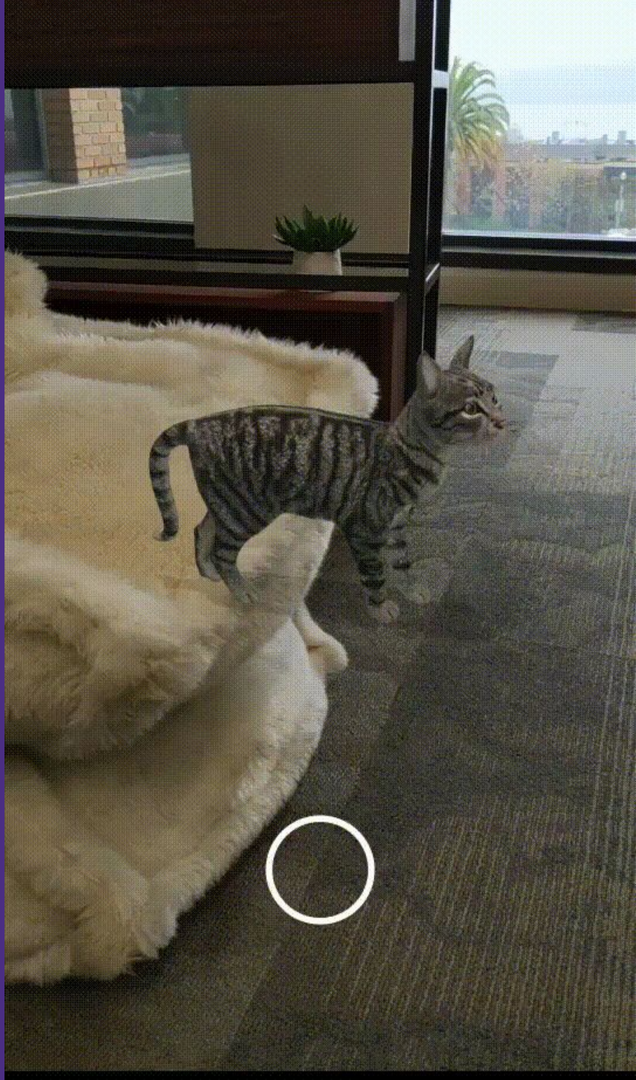
Introduction

Motivation

Is direct placement and rendering of 3D objects sufficient for realistic AR experiences?

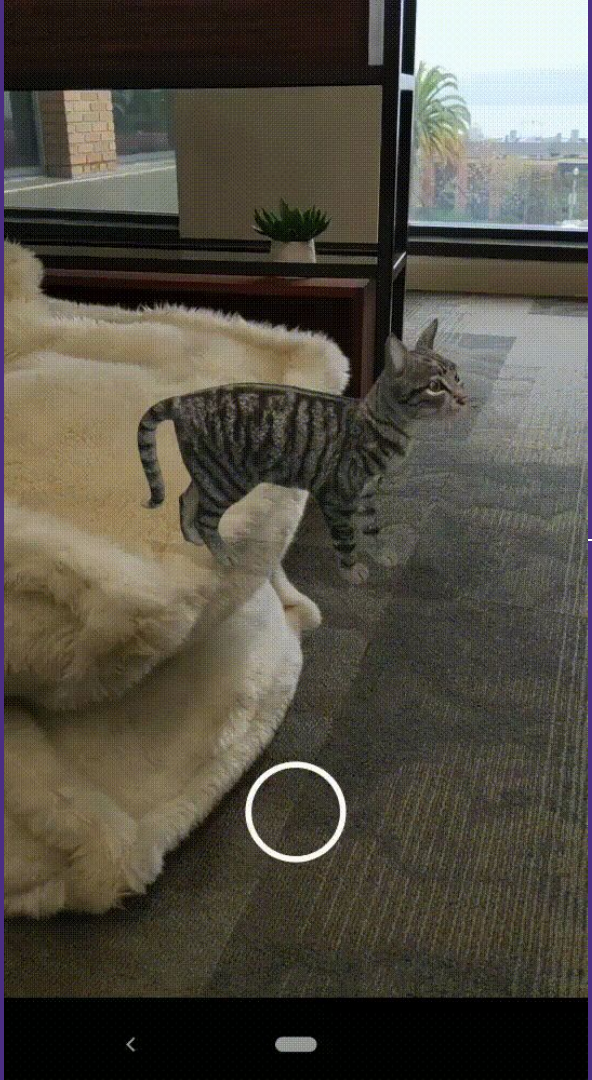
Not always!

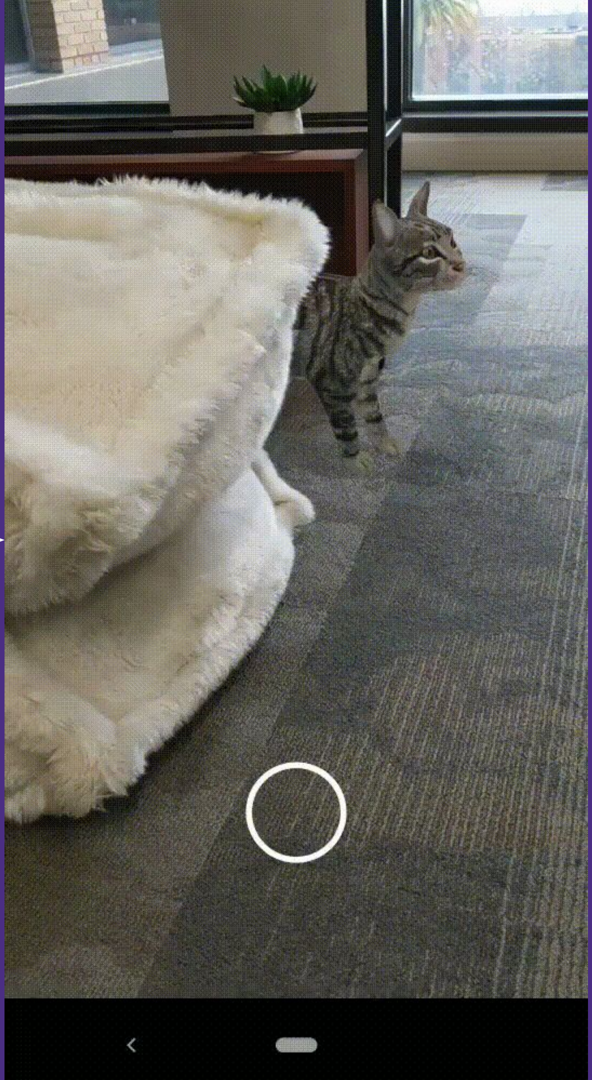
Virtual content looks like it's *"pasted on the screen"* rather than *"in the world"*!



Introduction

Motivation





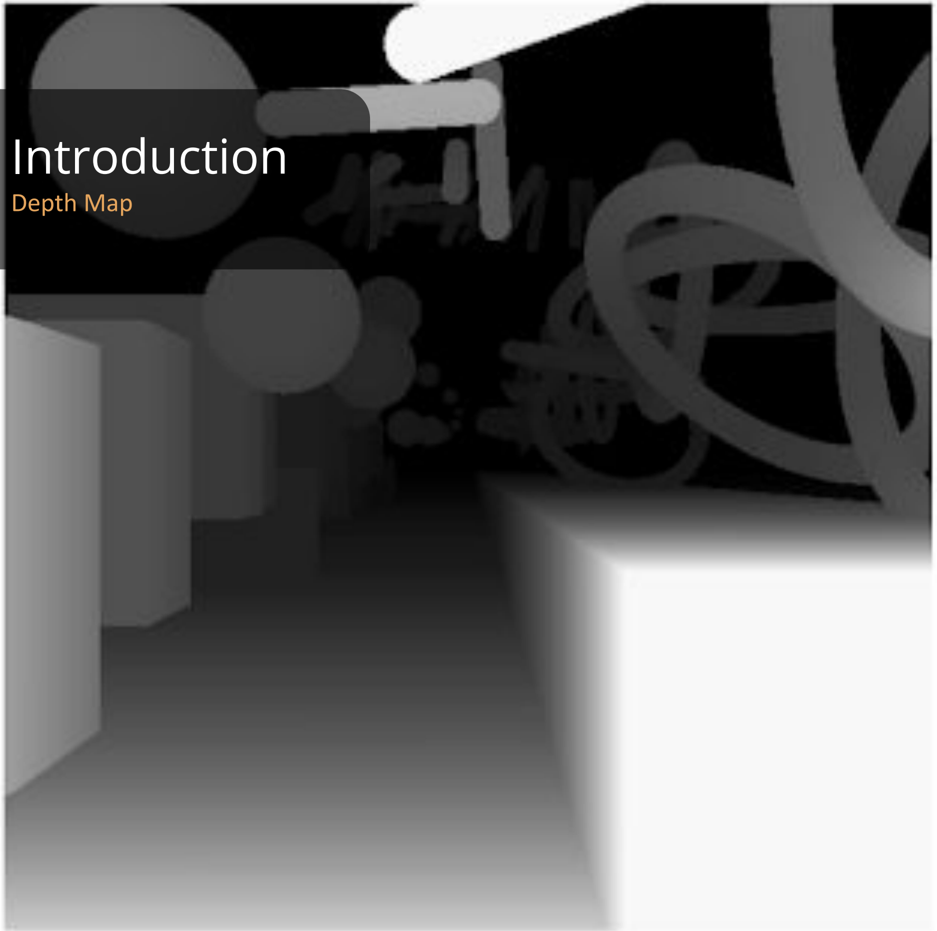
Introduction

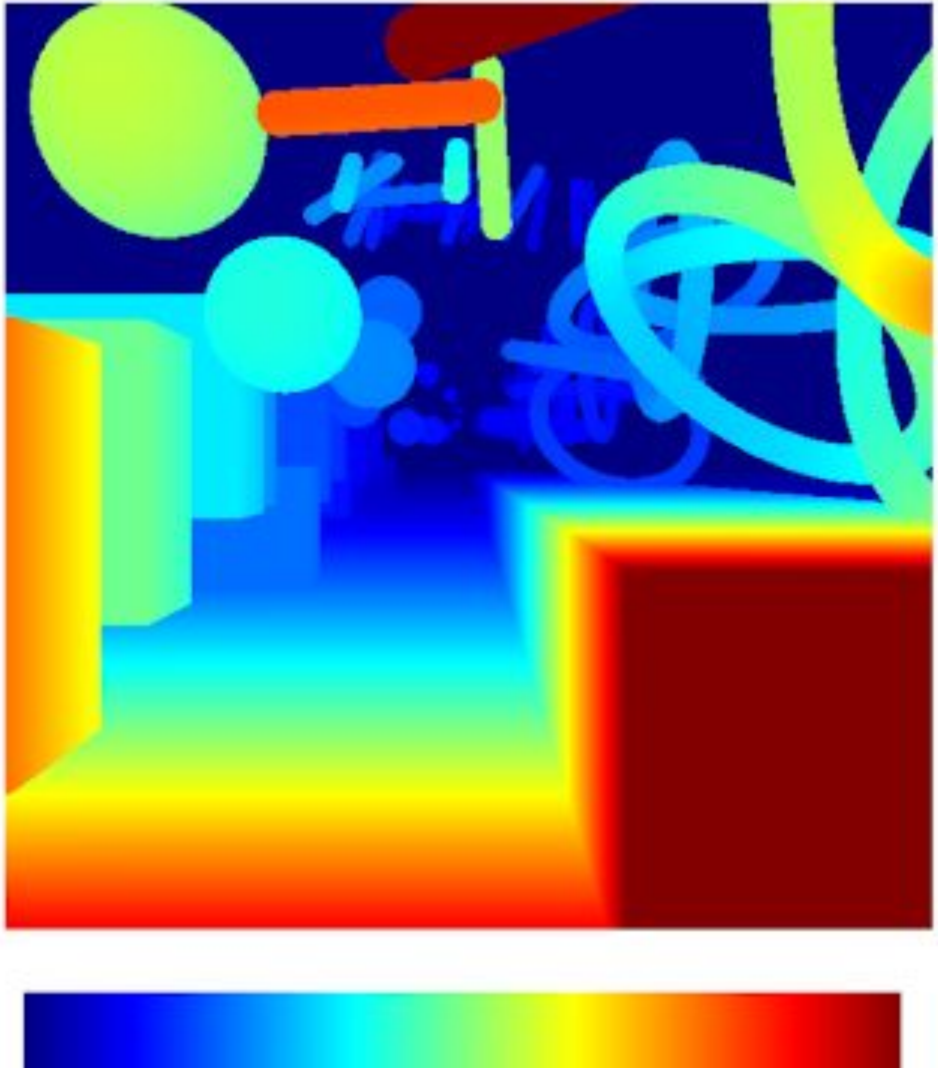
Motivation

How can we bring these advanced features to mobile AR experiences without relying on dedicated sensors or the need for computationally expensive surface reconstruction?

Introduction

Depth Map





Introduction

Depth Lab

Google	• *Pixel 2, Pixel 2 XL, Pixel 3, Pixel 3 XL, Pixel 3a, Pixel 3a XL, Pixel 4, Pixel 4 XL*
Huawei	• *Honor 10, Honor V20, Mate 20 Lite, Mate 20, Mate 20 X, Nova 3, Nova 4, P20, P30, P30 Pro*
LG	• *G8X ThinQ, V35 ThinQ, V50S ThinQ, V60 ThinQ 5G*
OnePlus	• *OnePlus 6, OnePlus 6T, OnePlus 7, OnePlus 7 Pro, OnePlus 7 Pro 5G, OnePlus 7T, OnePlus 7T Pro*
Oppo	• *Reno Ace*
Samsung	• *Galaxy A80, Galaxy Note8, Galaxy Note9, Galaxy Note10, Galaxy Note10 5G, Galaxy Note10+, Galaxy Note10+ 5G, Galaxy S8, Galaxy S8+, Galaxy S9, Galaxy S9+, Galaxy S10e, Galaxy S10, Galaxy S10+, Galaxy S10 5G, Galaxy S20, Galaxy S20+ 5G, Galaxy S20 Ultra 5G*
Sony	• *Xperia XZ2, Xperia XZ2 Compact, Xperia XZ2 Premium, Xperia XZ3*
Xiaomi	• *Pocophone F1*

And growing…

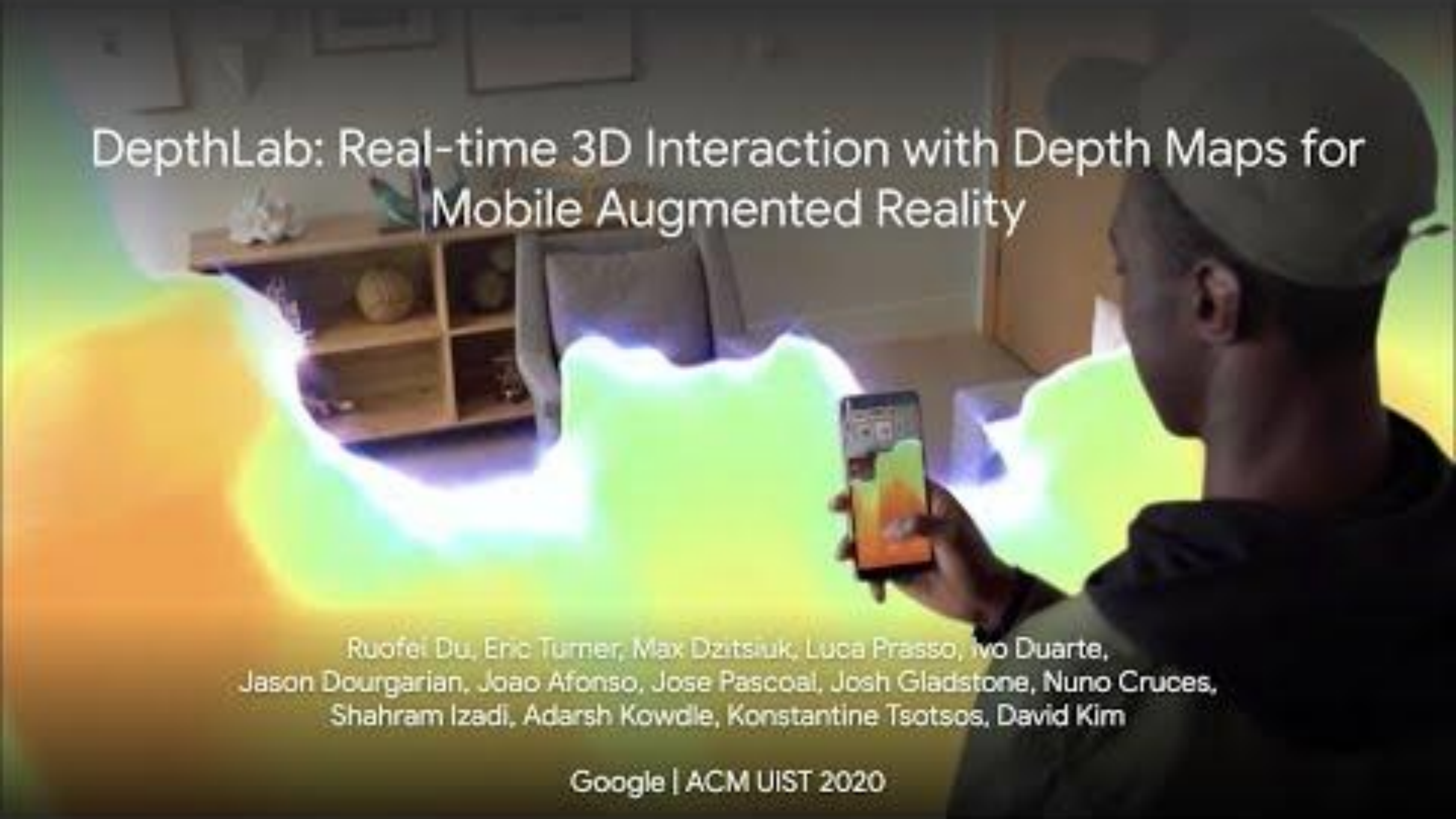
https://developers.google.com/ar/discover/supported-devices

Is there *more* to realism than occlusion?

Surface interaction?

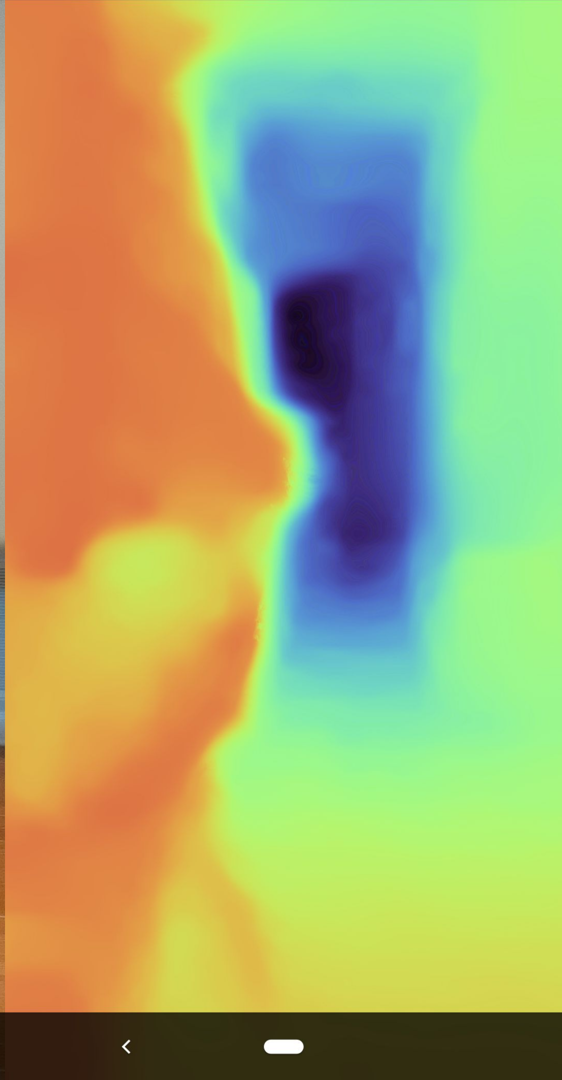
Realistic Physics?

Path Planning?


DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality
Ruofei Du, Eric Turner, Max Dzitsiuk, Luca Prasso, Ivo Duarte,
Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces,
Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim
Google | ACM UIST 2020

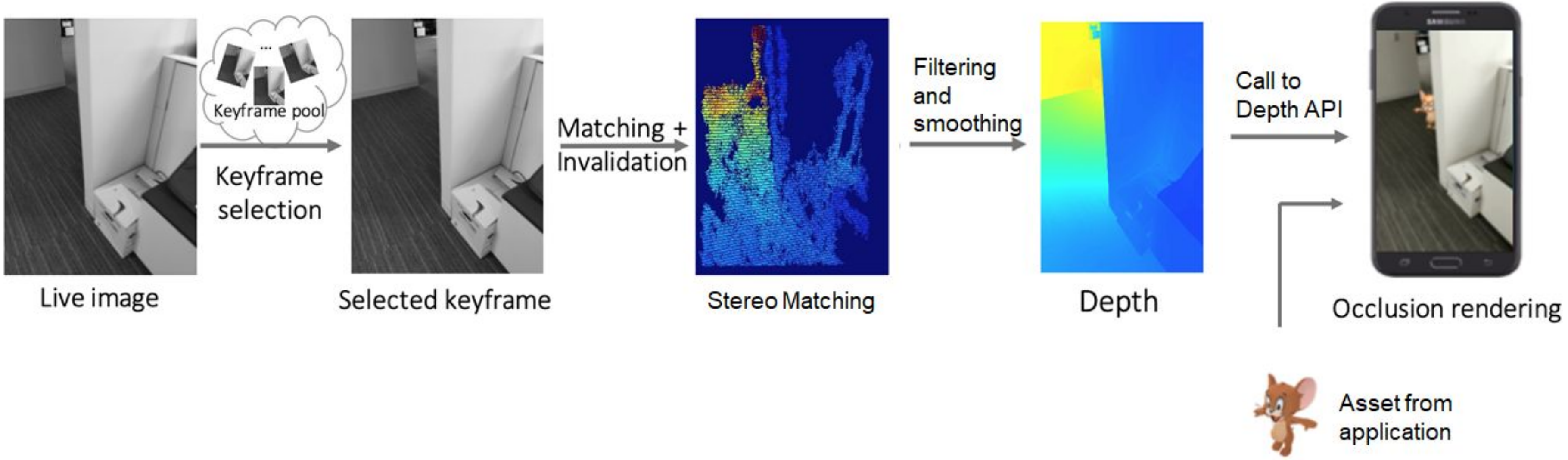
Introduction

Depth Lab



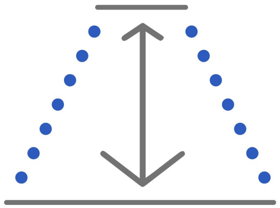
Related Work

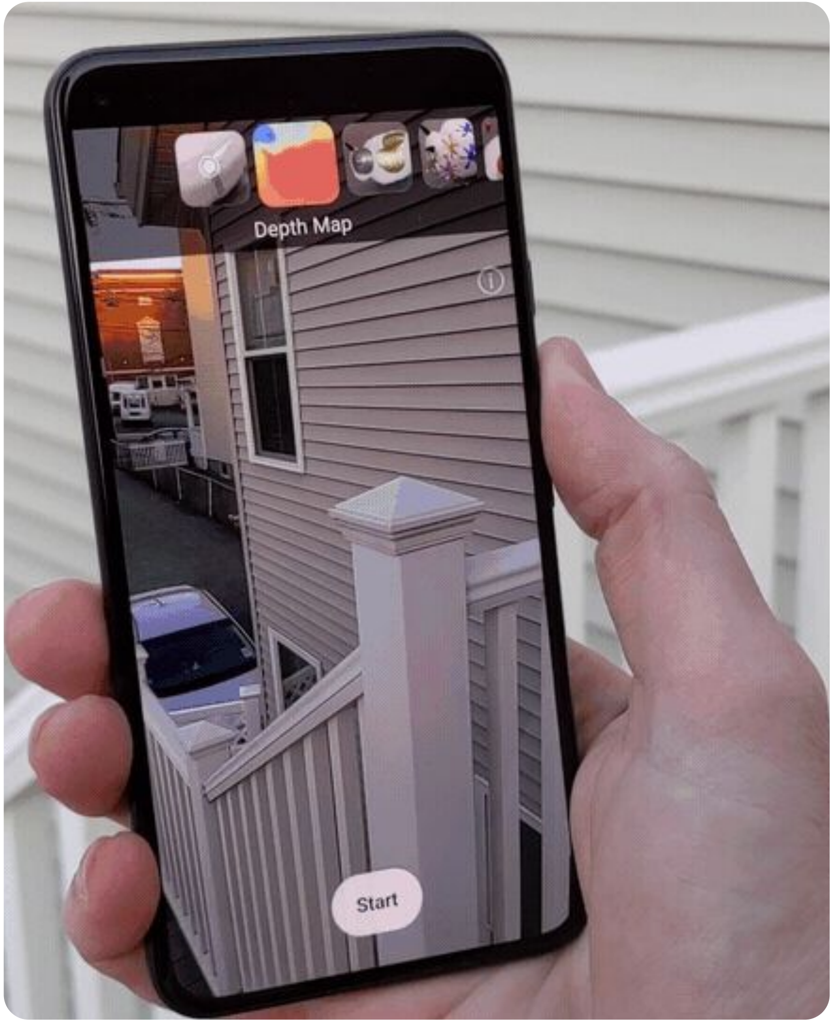
Valentin et al.



Depth from Motion for Smartphone AR (SIGGRAPH Asia 2018)

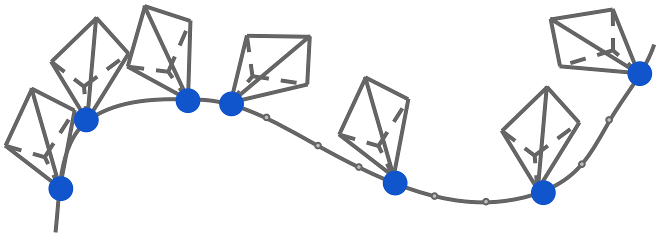
Depth Maps

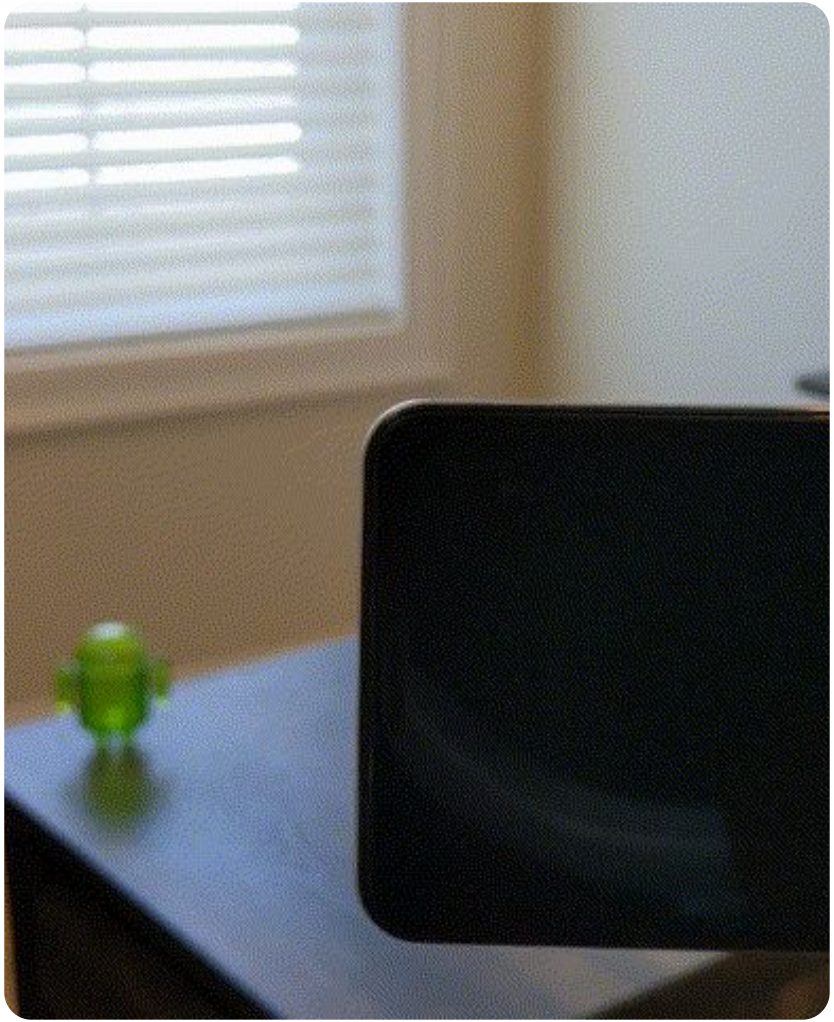




Depth From a Single Camera

Depth from Motion





Depth From a Single Camera

Best Practices

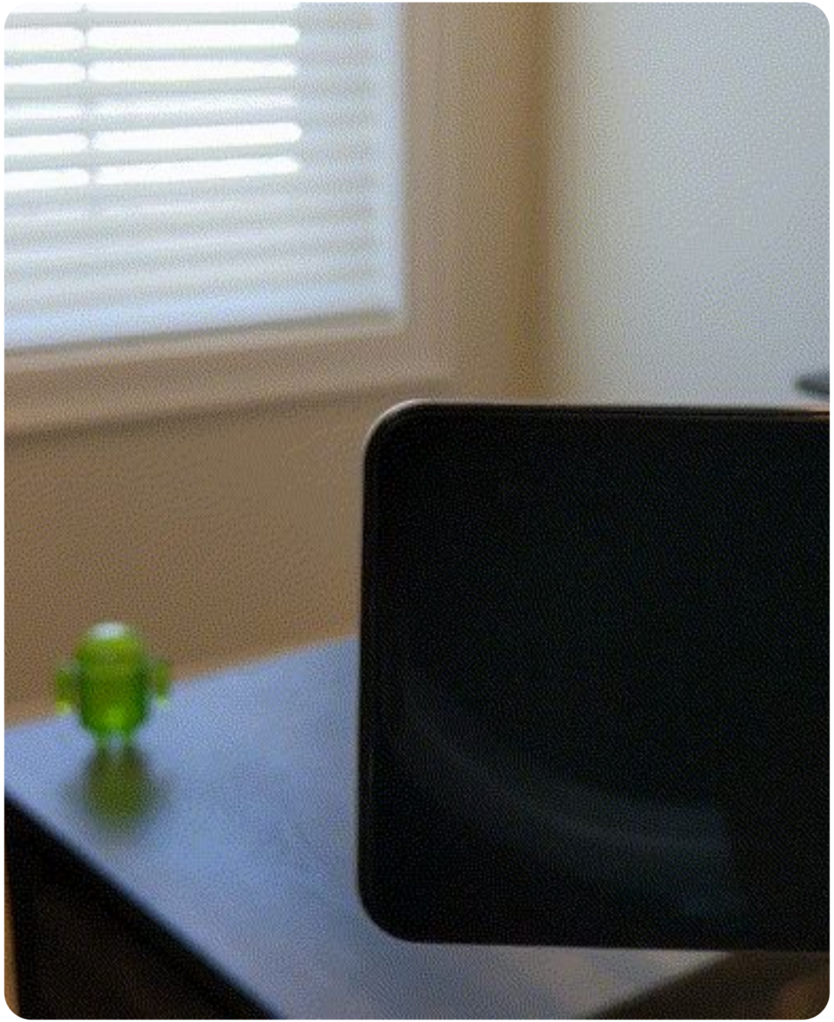
Use depth-certified ARCore devices

Minimal movement in the scene

Encourage users to move the device

Depth from 0 to 8 meters

Best accuracy 0.5 to 5 meters

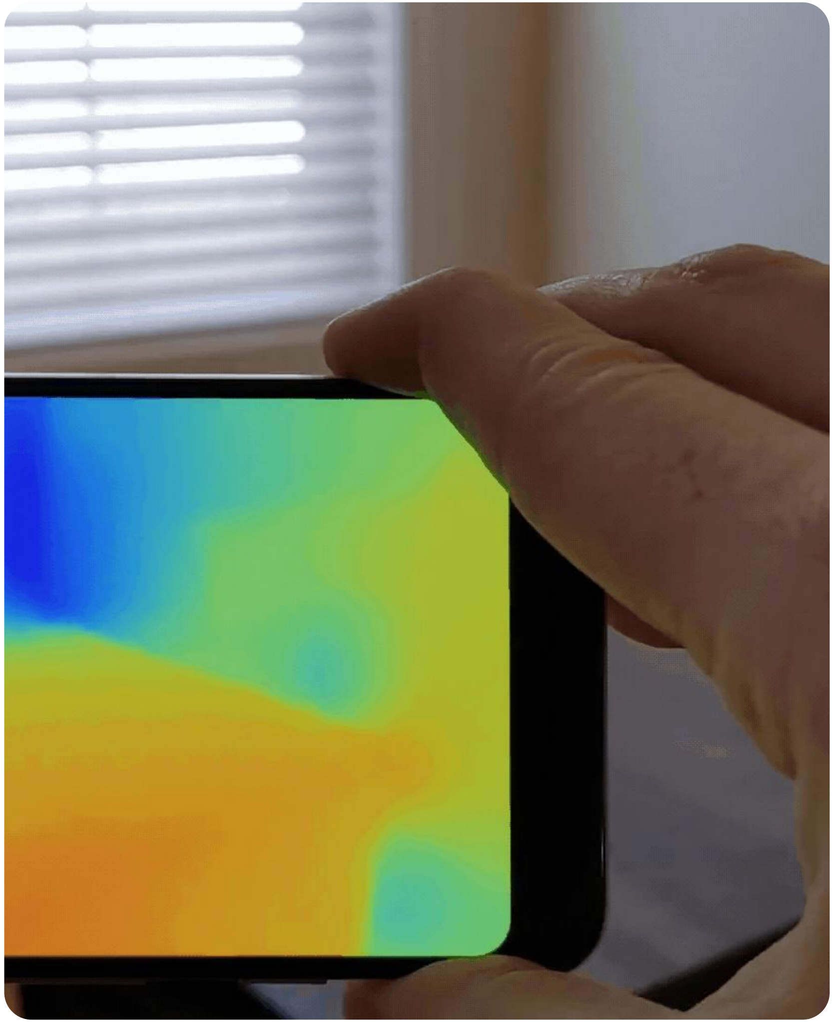


Enhancing Depth

Optimized to give you the best depth

Depth from Motion is fused with state-of-the-art Machine Learning

Depth leverages specialized hardware like a Time-of-Flight sensor when available



Introduction

Depth Lab

Introduction

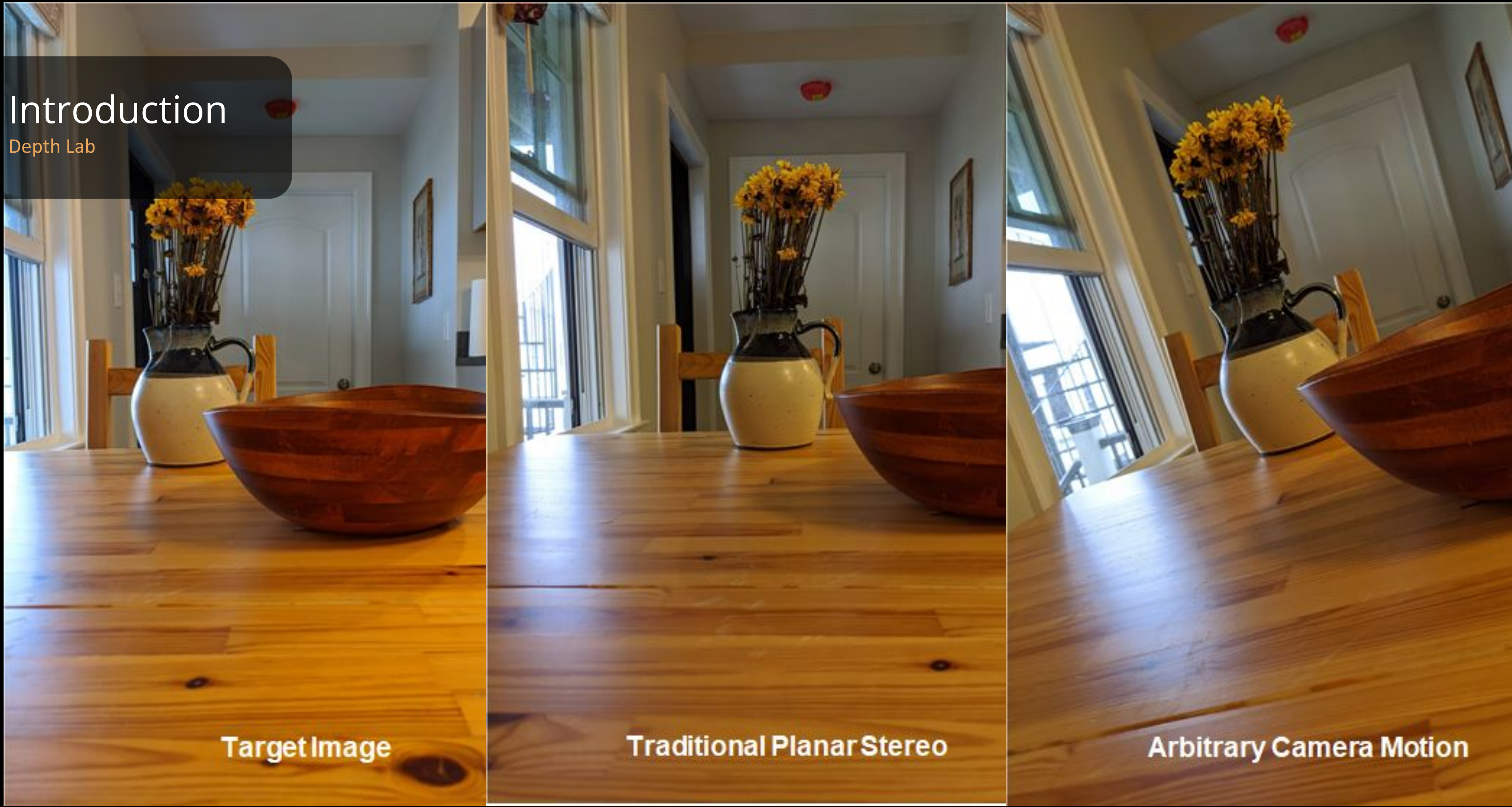
Depth Lab

Introduction

Depth Generation

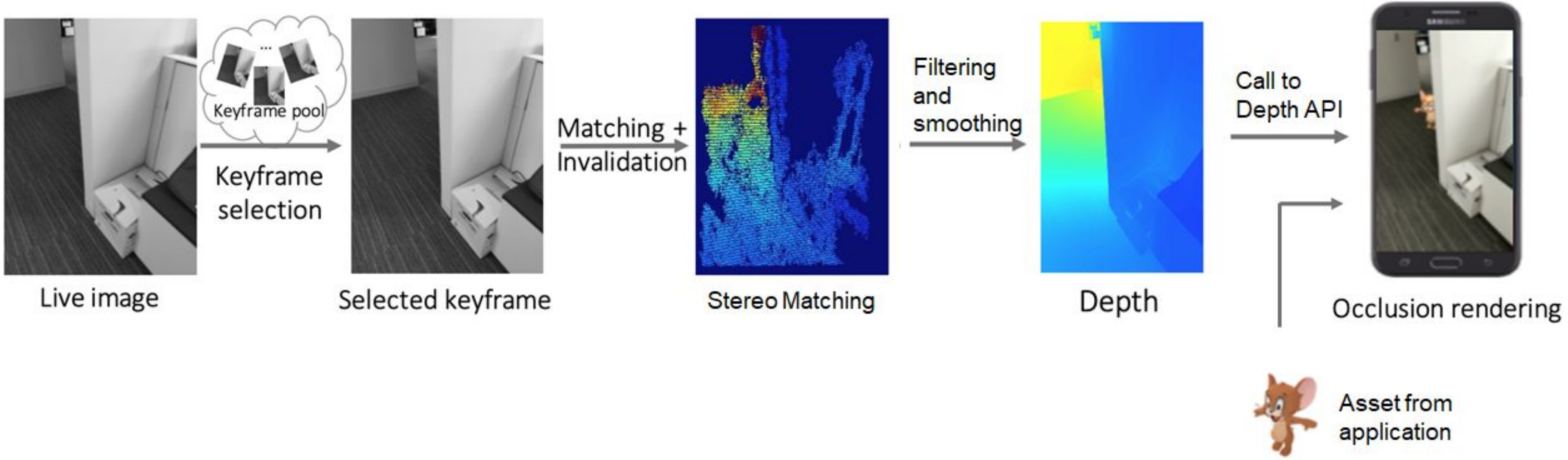
Introduction

Depth Lab



Related Work

Valentin et al.



Depth from Motion for Smartphone AR (SIGGRAPH Asia 2018)

Introduction

Depth Lab

Introduction

Depth Lab

Up to 8 meters, with
the best within 0.5m to 5m

Motivation

Gap from raw depth to applications

Introduction

Depth Lab

Design Process

3 Brainstorming Sessions

3 brainstorming sessions
18 participants
39 aggregated ideas

Design Process

3 Brainstorming Sessions

Supplementary Material for DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality

Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte, Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces, Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim†
Google LLC

GEOMETRY-AWARE AR FEATURES

In this section, we list all ideas from our brainstorming sessions and discuss their depth representation requirements, use cases, and whether each is implemented in DepthLab [5]. Note that ideas 9, 21, 24, 25 are not available as open source code yet, but can be easily reproduced with the provided algorithms.

Depth Representation Requirement: Localized Depth

Figure 1. Implementation examples of geometry-aware AR features 1–5 with localized depth use cases. Please refer to the supplementary video for live demonstration.

1. **3D oriented cursor:** Render a 3D cursor centered in the screen center. The 3D cursor should change its orientation and scale according to the surface normal and distance when moving along physical surfaces. Implemented in DepthLab: Yes.

4. **Avatar path planning:** Navigate a virtual object to move naturally between two points in physical environments. Implemented in DepthLab: Yes.
5. **Collision-aware placement:** Test if a virtual object's volume collides with observed environment surfaces. Implemented in DepthLab: Yes.

Depth Representation Requirement: Surface Depth

Figure 2. Implementation examples of geometry-aware AR features 6–9.

6. **Virtual shadows:** Render geometry-aware shadows [11] that are cast onto physical surfaces. The shadow may be

8. **Physical simulation:** Simulate physical phenomena for augmented reality objects, e.g. collision. Implemented in DepthLab: Yes.
9. **AR graffiti:** Allow the user to touch on the screen and sketch/spray/paint virtual drawings onto physical objects. Implemented in DepthLab: Yes.
10. **AR paintball:**

Figure 3. Implementation examples of geometry-aware AR features 15–25 with dense depth.

19. **Rain effects:** Similiar in behavior to the snow effect, the rain particles should also splat on the surface using the estimated normal vector from the localized depth. Implemented in DepthLab: Yes.
20. **Fog effects:** Render screen-space post-processing effects, where far objects are overlaid with thicker fog. The user may interactively adjust the fog intensity in real time. Implemented in DepthLab: Yes.
21. **Edge highlighting:** Highlight the edges of the observed environment according to the depth map. Unlike edge detection in a color image, highlighting depth edges may offer a clean segmentation of physical objects regardless of their texture. Implemented in DepthLab: Yes.
22. **Depth-based segmentation:** Segment the foreground, background, or objects between a certain range of depth values from the color image. It may be useful for teleconferencing tasks. Implemented in DepthLab: Yes.
23. **False-color visualization and animated transition effects:** Visualize the depth map based on a specific transfer function and animate the transition

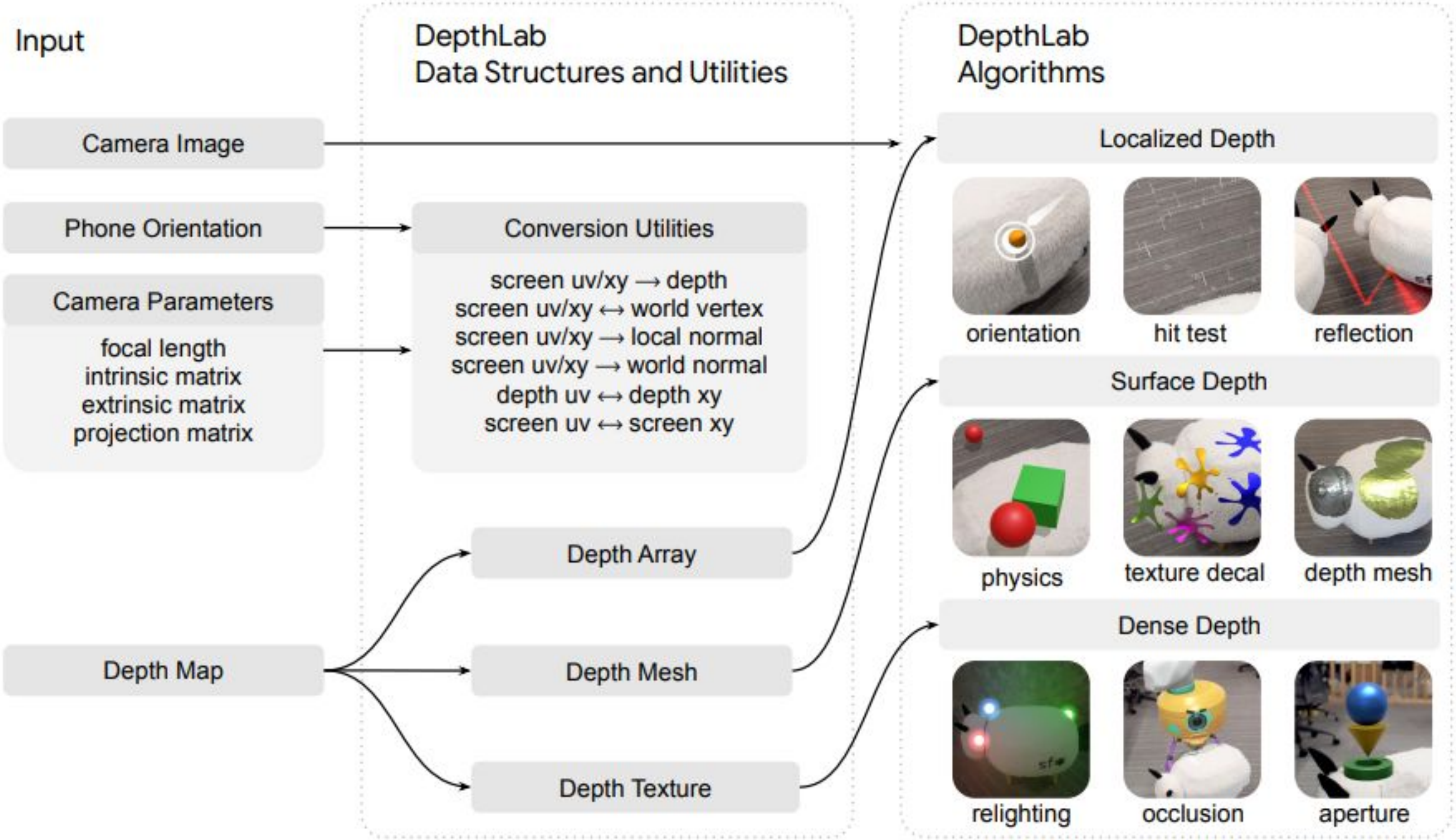
35. **Enable multitouch on surfaces [12]:** User may annotate sticky notes and papers with a pen and "program" them to control smart lights, music, and other digital functions of the environment.
36. **Person capture:** Enable self-scanning with the frontal camera [1] and teleconference with the rear camera.
37. **AR board game:** Design an AR-based board game [2] that overlays digital assets upon physical cards with aware of users' gestures and actions.
38. **Interactive surface editing:** Apply simple 3D distortion (pinch, twist, taper, bend) to captured colored voxels of the physical environment [4].
39. **Interactive music experience:** Design in-air instruments (guitar, piano) with dynamic gesture recognition [8]. Virtual targets are placed in 3D space, such as a drum set, big piano keys, etc. Upon contact detection, the app plays a sound.

REFERENCES

[1] Karan Ahuja, Chris Harrison, Mayank Goel, and Robert Xiao. 2019. MeCap: Whole-Body Digitization for Low-Cost VR/AR Headsets. In *Proceedings of the 32Nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. ACM, ACM, New York, NY, USA, 453–462. DOI: http://dx.doi.org/10.1145/3332165.3347889
[2] Troels L Andersen, Sune Kristensen, Bjørn W Nielsen, and Kaj Grønbæk. 2004. Designing an augmented reality board game with children: the battleboard 3D experience. In *Proceedings of the 2004 conference on Interaction design and children: building a community*. 137–138.
[3] Sujal Bista, Icaro Lins Leitao da Cunha, and Amitabh Varshney. 2017. Kinetic Depth Images: Flexible Generation of Depth Perception. *The Visual Computer* 33, 10 (01 October 2017), 1357–1369.
[4] Ming Chuang and Michael Kazhdan. 2011. Interactive and anisotropic geometry processing using the screened Poisson equation. In *ACM SIGGRAPH 2011 papers*. 1–10.
[5] Ruofei Du, Eric Turner, Max Dzitsiuk, Luca Prasso, Ivo Duarte, Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces, Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim

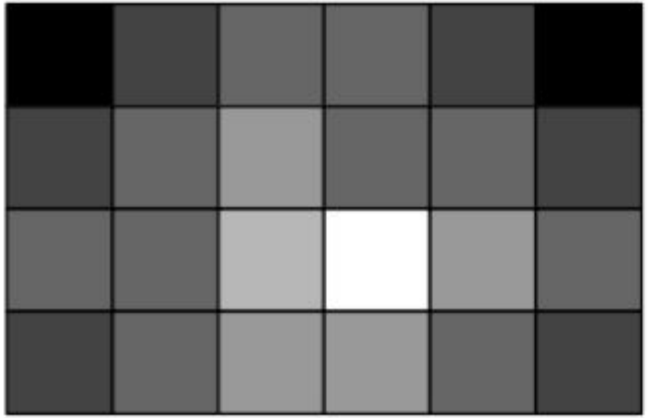
System

Architecture overview



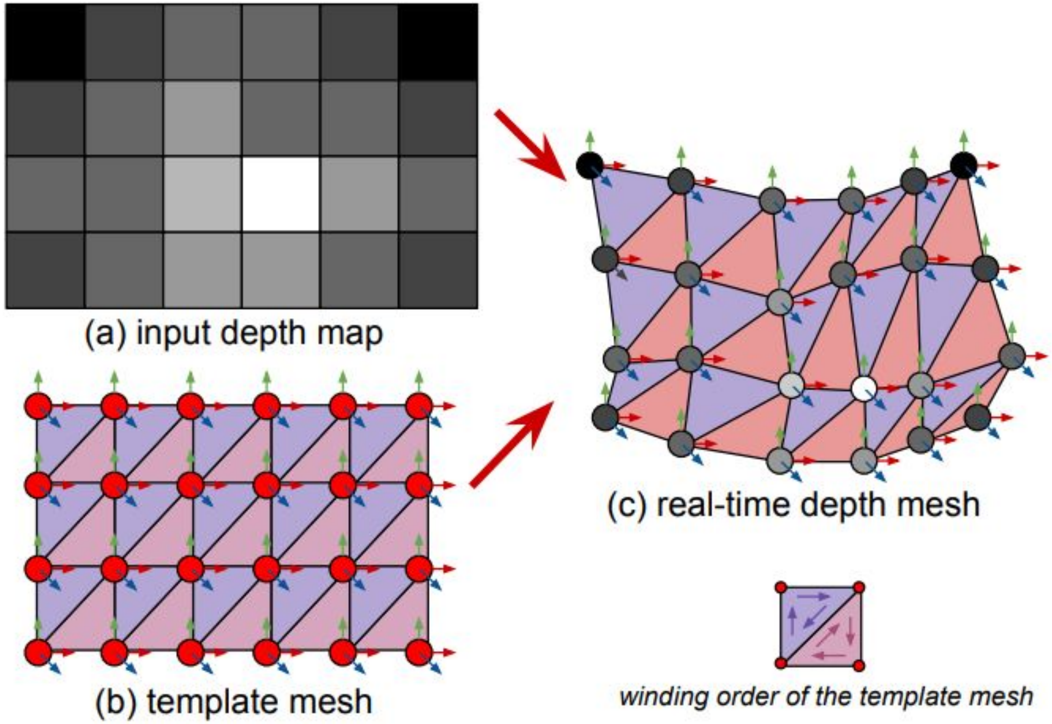
Data Structure

Depth Array



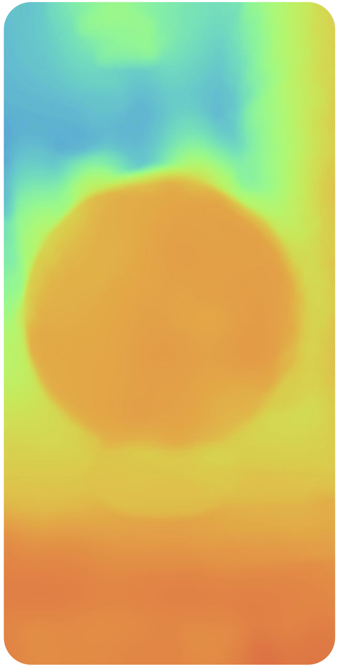
2D array (160x120 and above) of 16-bit integers

Data Structure
Depth Mesh



Data Structure

Depth Texture



System
Architecture

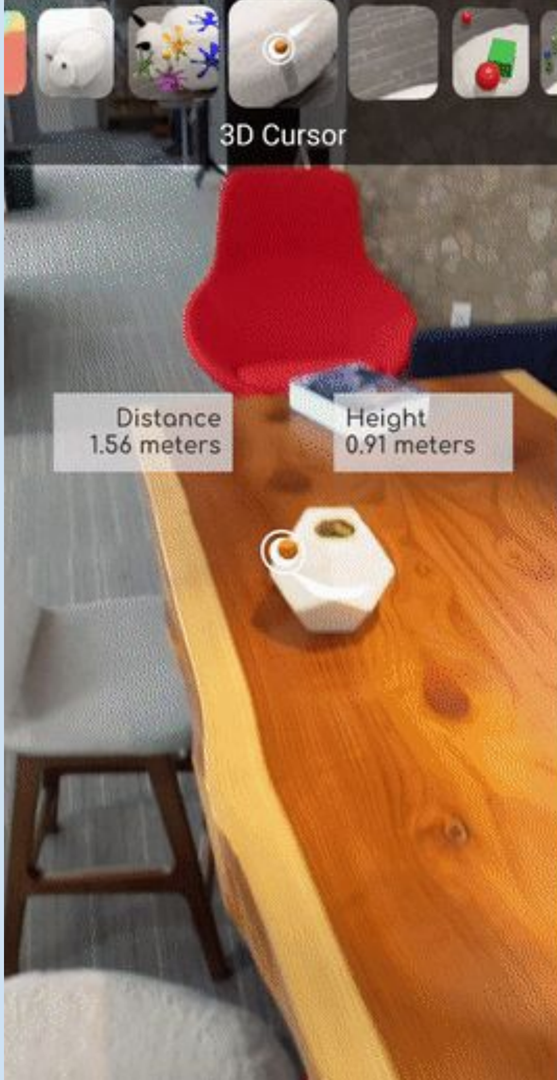
	Localized Depth	Surface Depth	Dense Depth
CPU	✓	✓	✗ (non-real-time)
GPU	N/A	✓ (compute shader)	✓ (fragment shader)
Prerequisite	point projection normal estimation	depth mesh triplanar mapping	anti-aliasing multi-pass rendering
Data Structure	depth array	depth mesh	depth texture
Example Use Cases	physical measure oriented 3D cursor path planning	collision & physics virtual shadows texture decals	scene relighting aperture effects occluded objects

Localized Depth

Coordinate System Conversion

Conversion Utilities

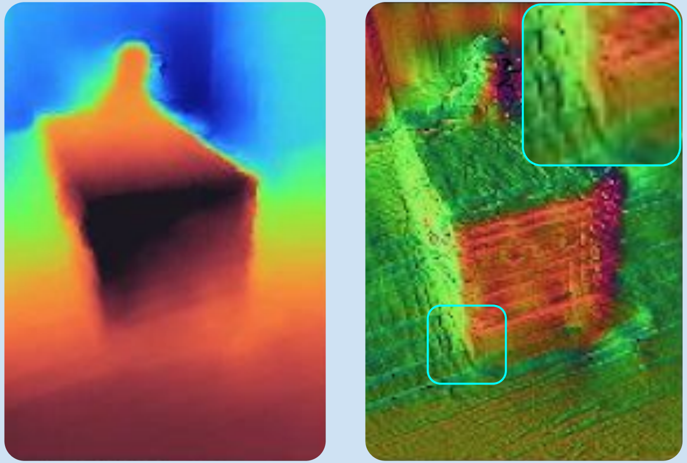
screen uv/xy ⟶ depth
screen uv/xy ⟷ world vertex
screen uv/xy ⟶ local normal
screen uv/xy ⟶ world normal
depth uv ⟷ depth xy
screen uv ⟷ screen xy



Localized Depth

Normal Estimation

$$\mathbf{n_p} = \left(\mathbf{v_p} - \mathbf{v_{p+(1,0)}}\right) \times \left(\mathbf{v_p} - \mathbf{v_{p+(0,1)}}\right)$$



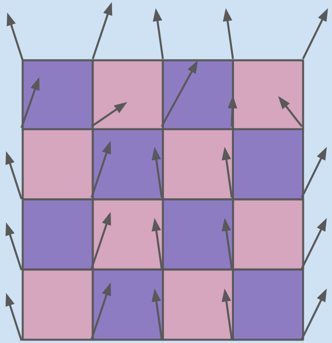
Localized Depth

Normal Estimation

Algorithm 1: Estimation of the Normal Vector of a Screen Point in DepthLab.

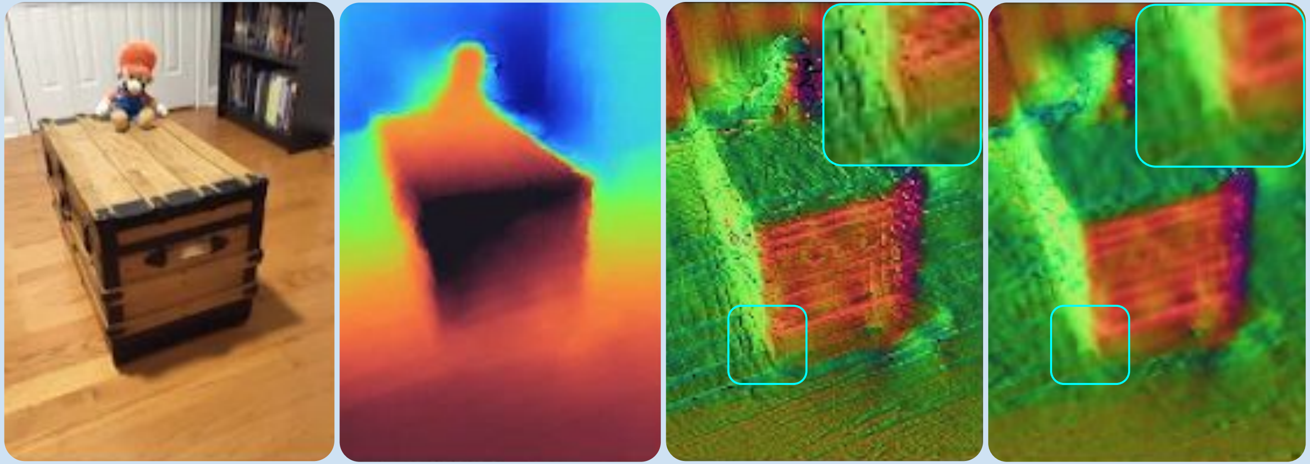
Input : A screen point $\mathbf{p} \leftarrow (x, y)$ and focal length f.
Output : The estimated normal vector $\mathbf{n}$.

Set the sample radius: $r \leftarrow 2$ pixels.
Initialize the counts along two axes: $c_X \leftarrow 0, c_Y \leftarrow 0$.
Initialize the correlation along two axes: $\rho_X \leftarrow 0, \rho_Y \leftarrow 0$.
for $\Delta x \in [-r, r]$ **do**
    for $\Delta y \in [-r, r]$ **do**
        Continue if $\Delta x = 0$ **and** $\Delta y = 0$.
        Set neighbor's coordinates: $\mathbf{q} \leftarrow [x + \Delta x, y + \Delta y]$.
        Set $\mathbf{q}$'s distance in depth: $d_{\mathbf{pq}} \leftarrow \|\mathbf{D}(\mathbf{p}), \mathbf{D}(\mathbf{q})\|$.
        Continue if $d_{\mathbf{pq}} = 0$.
        if $\Delta x \neq 0$ **then**
            $c_X \leftarrow c_X + 1$.
            $\rho_X \leftarrow \rho_X + d_{\mathbf{pq}} / \Delta x$.
        end
        if $\Delta y \neq 0$ **then**
            $c_Y \leftarrow c_Y + 1$.
            $\rho_Y \leftarrow \rho_Y + d_{\mathbf{pq}} / \Delta y$.
        end
    end
end
Set pixel size: $\lambda \leftarrow \frac{\mathbf{D}(\mathbf{p})}{f}$.
return the normal vector $\mathbf{n}$: $\left(-\frac{\rho_Y}{\lambda c_Y}, -\frac{\rho_X}{\lambda c_X}, -1\right)$.



Localized Depth

Normal Estimation

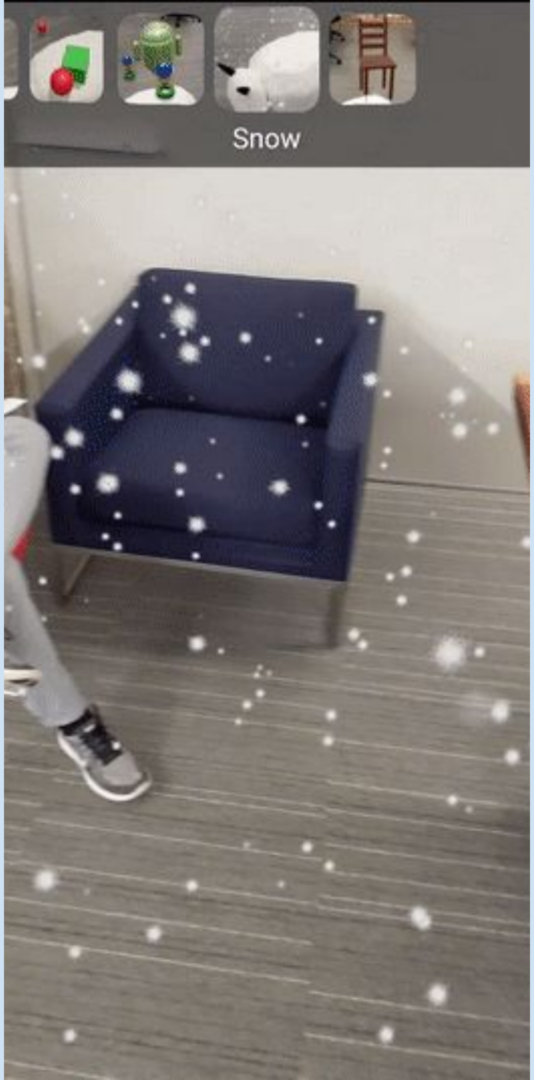


Localized Depth

Avatar Path Planning

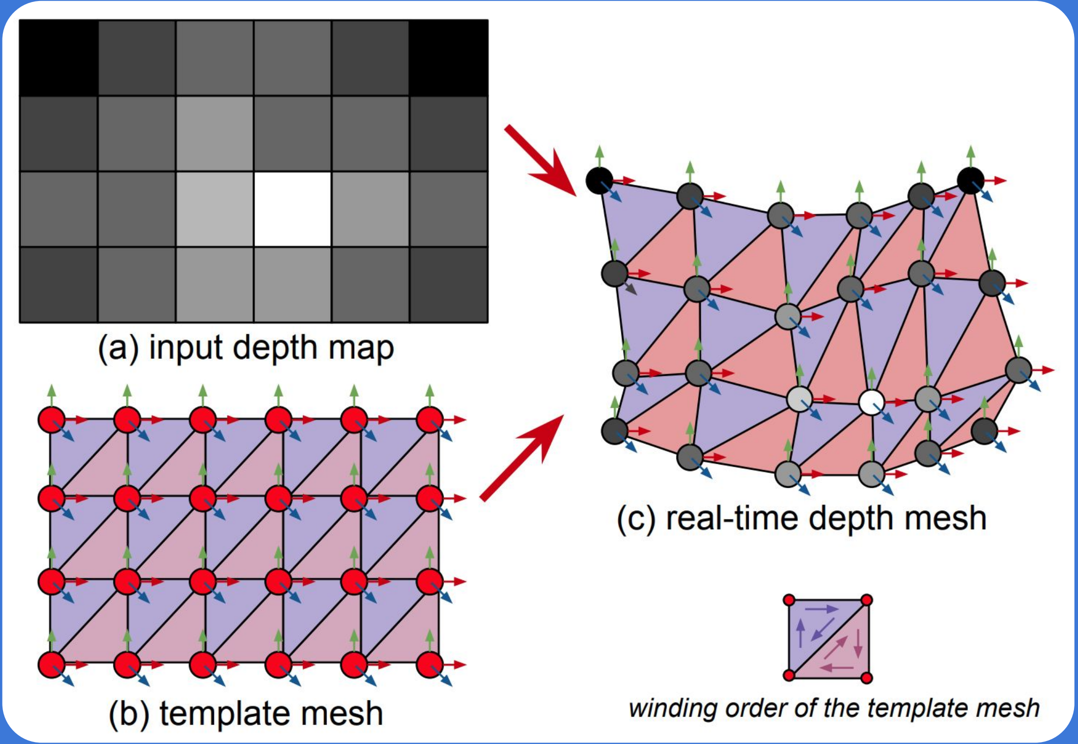
Localized Depth

Rain and Snow



Surface Depth

Use Cases



Surface Depth

Physics collider

Physics with depth mesh.

Surface Depth
Texture decals

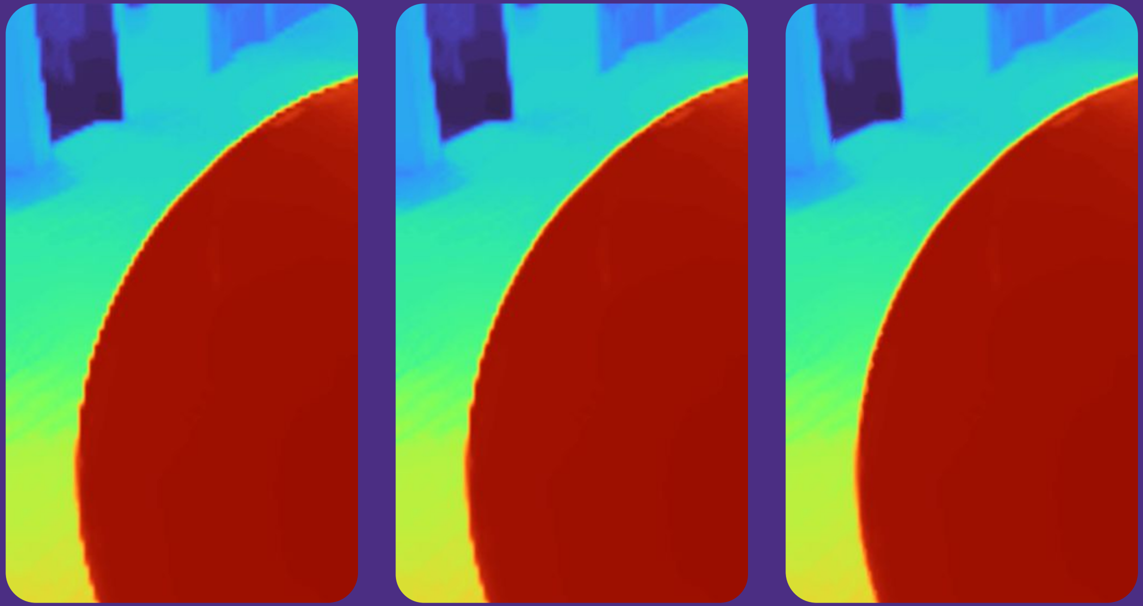
Texture decals with depth mesh.

Surface Depth
3D Photo

Projection mapping with depth mesh.

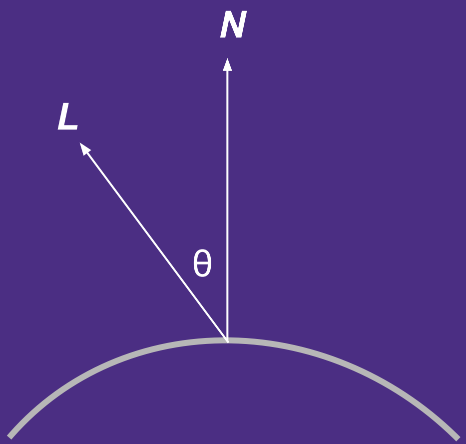
Dense Depth

Depth Texture - Antialiasing



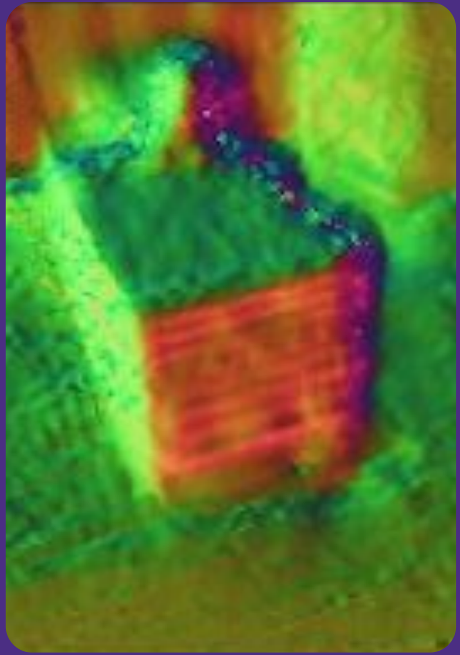
Dense Depth

Real-time relighting



Dense Depth

Why normal map does not work?



Dense Depth

Real-time relighting

Algorithm 3: Ray-marching-based Real-time Relighting.

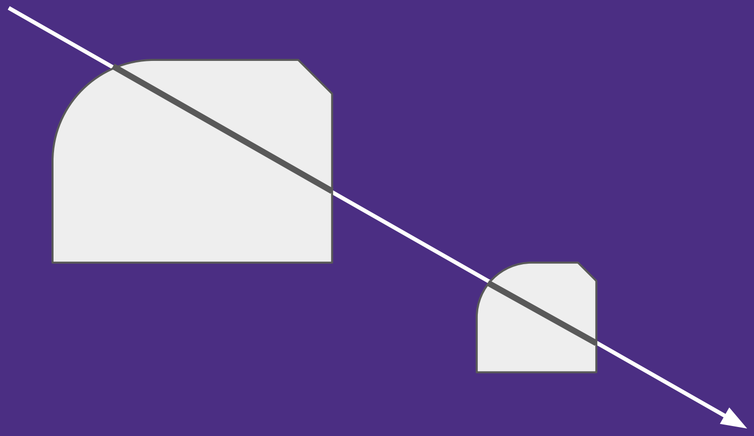
Input : Depth map $\mathbf{D}$, the camera image $\mathbf{I}$, camera intrinsic matrix $\mathbf{K}$, L light sources $\mathbb{L} = \{\mathscr{L}^i, i \in L\}$ with each light's location $\mathbf{v}_{\mathscr{L}}$ and intensity in RGB channels $\phi_{\mathscr{L}}$.

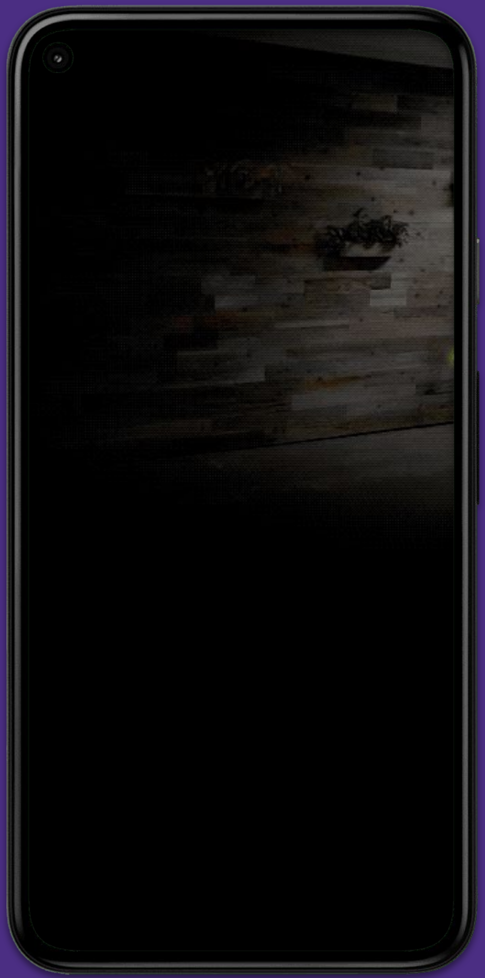
Output : Relighted image $\mathbf{O}$.

for *each image pixel* $\mathbf{p} \in$ *depth map* $\mathbf{D}$ *in parallel* **do**
    Sample $\mathbf{p}$'s depth value $d \leftarrow \mathbf{D}(\mathbf{p})$.
    Compute the corresponding 3D vertex $\mathbf{v_p}$ of the screen point $\mathbf{p}$ using the camera intrinsic matrix $\mathbf{v_p}$ with $\mathbf{K}$: $\mathbf{v_p} = \mathbf{D}(\mathbf{p}) \cdot \mathbf{K}^{-1}[\mathbf{p}, 1]$
    Initialize relighting coefficients of $\mathbf{v_p}$ in RGB: $\phi_\mathbf{p} \leftarrow \mathbf{0}$.
    for *each light* $\mathscr{L} \in$ *light sources* $\mathbb{L}$ **do**
        Set the current photon coordinates $\mathbf{v}_o \leftarrow \mathbf{v_p}$.
        Set the current photon energy $E_o \leftarrow 1$.
        while $\mathbf{v}_o \neq \mathbf{v}_{\mathscr{L}}$ **do**
            Compute the weighted distance between the photon to the physical environment $\Delta d \leftarrow \alpha|\mathbf{v}_o^{xy} - \mathbf{v}_{\mathscr{L}}^{xy}| + (1-\alpha)|\mathbf{v}_o^z - \mathbf{v}_{\mathscr{L}}^z|, \alpha = 0.5$.
            Decay the photon energy: $E_o \leftarrow 95\% E_o$
            Accumulate the relighting coefficients: $\phi_\mathbf{p} \leftarrow \phi_\mathbf{p} + \Delta d E_o \phi_{\mathscr{L}}$.
            March the photon towards the light source: $\mathbf{v}_o \leftarrow \mathbf{v}_o + (\mathbf{v}_{\mathscr{L}} - \mathbf{v}_o)/S$, here $S = 10$, depending on the mobile computing budget.
        end
    end
    Sample pixel's original color: $\Phi_\mathbf{p} \leftarrow \mathbf{I}(\mathbf{p})$.
    Apply relighting effect: $\mathbf{O}(\mathbf{p}) \leftarrow \gamma \cdot |\mathbf{0.5} - \phi_\mathbf{p}| \cdot \Phi_\mathbf{p}^{1.5-\phi_\mathbf{p}} - \Phi_\mathbf{p}$, here $\gamma \leftarrow 3$.
end

Dense Depth

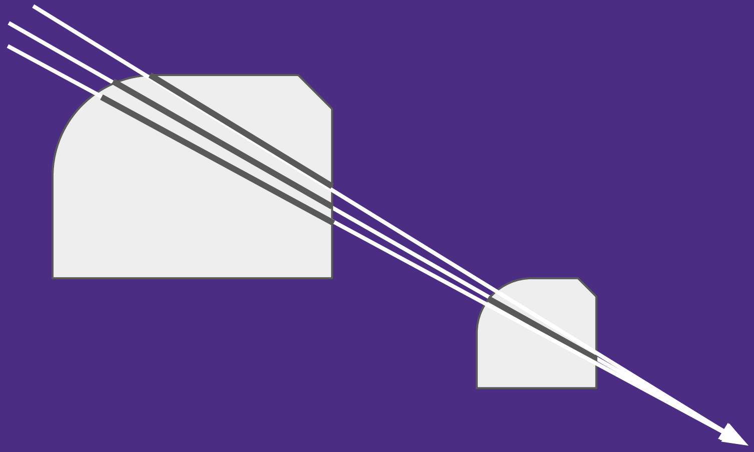
Real-time relighting

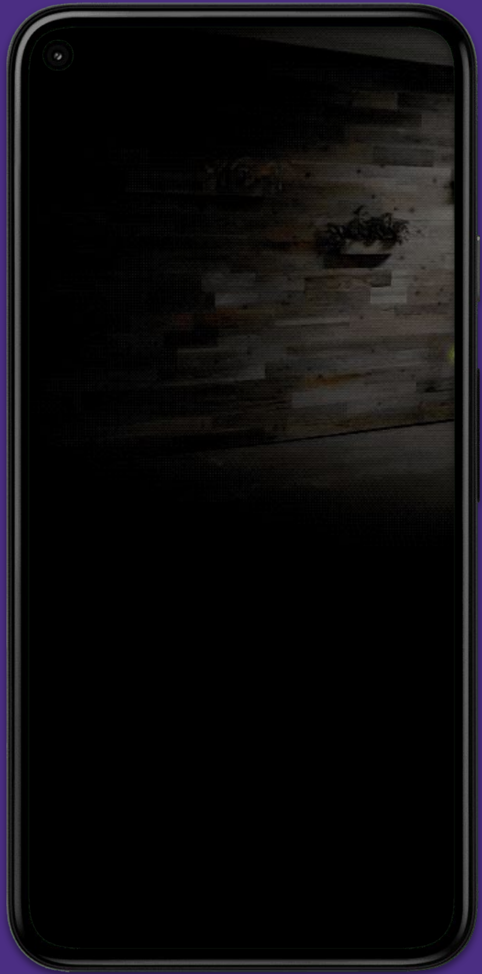




Dense Depth

Real-time relighting

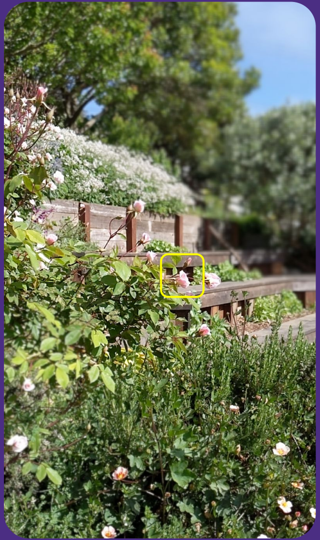


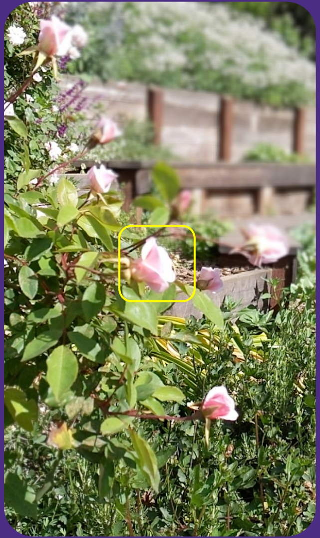


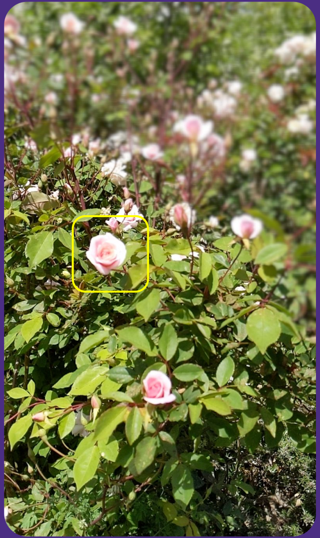
go/realtime-relighting, go/relit

Dense Depth

Wide-aperture effect

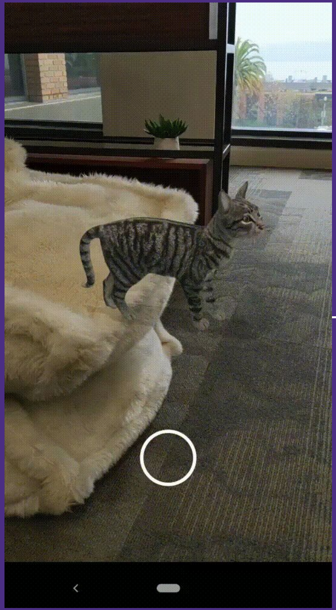


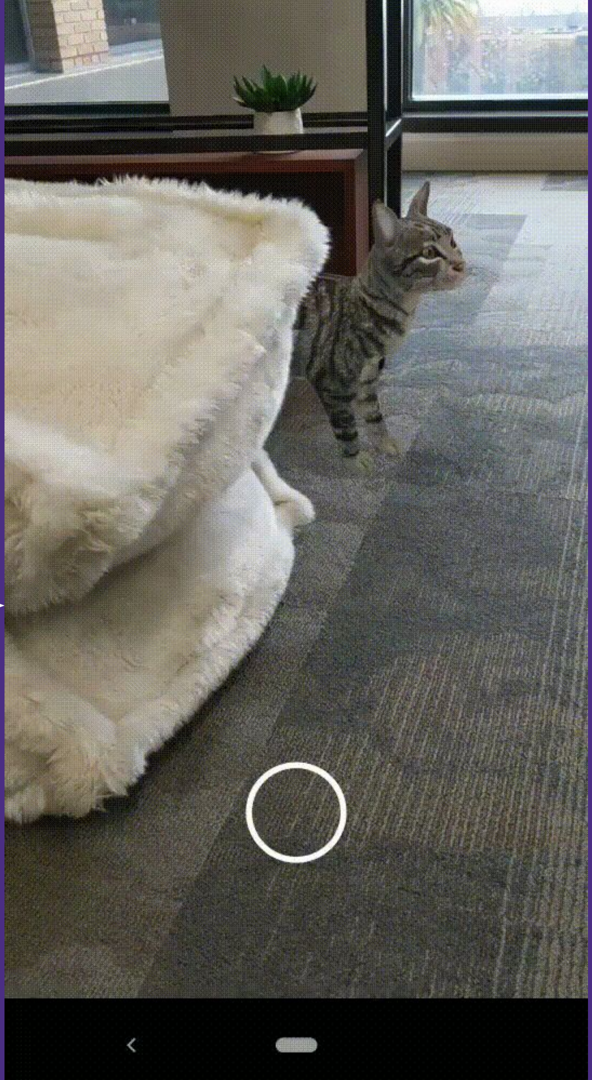


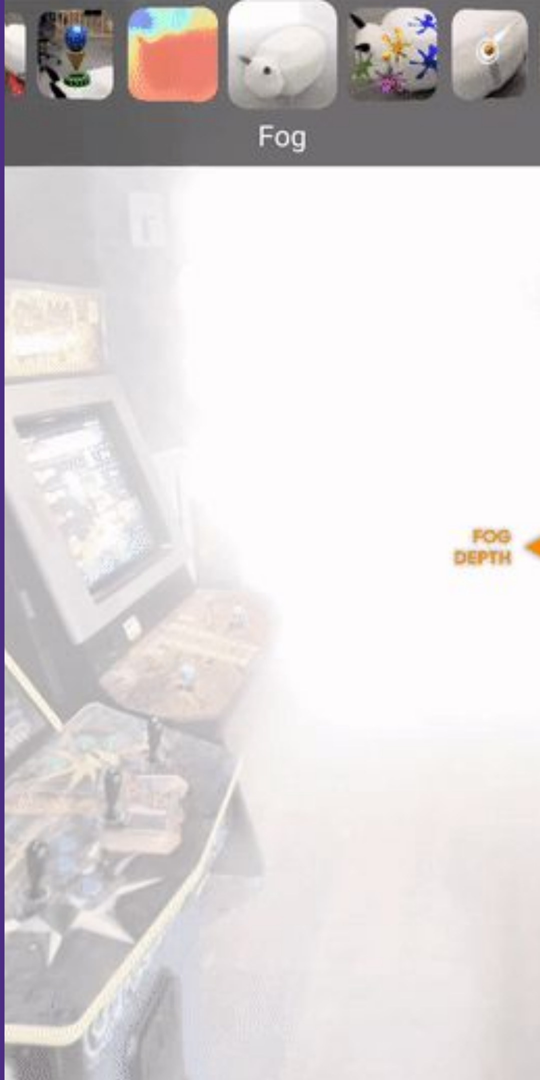


Dense Depth

Occlusion-based rendering

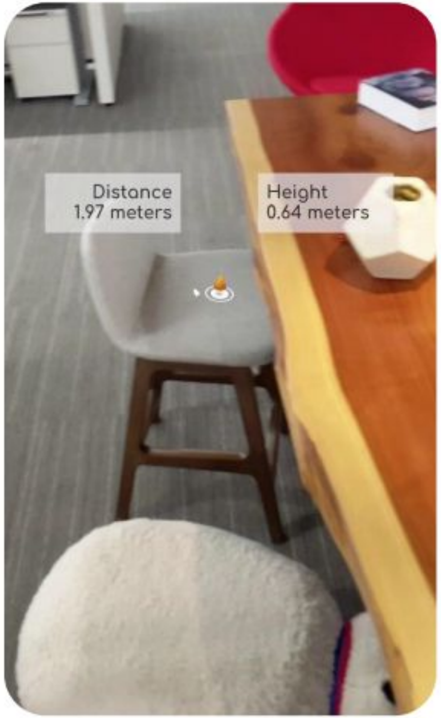


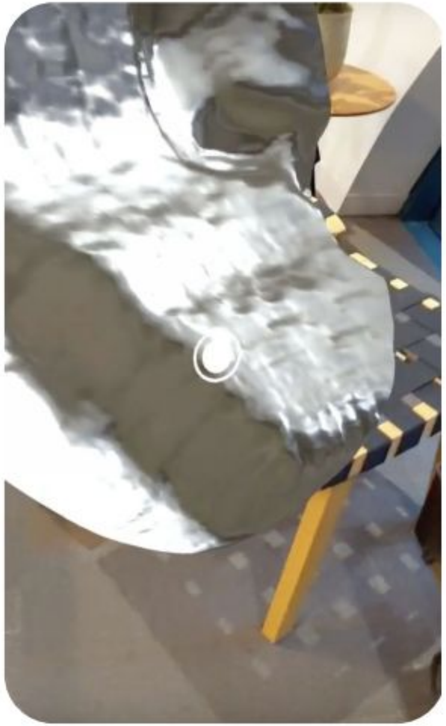


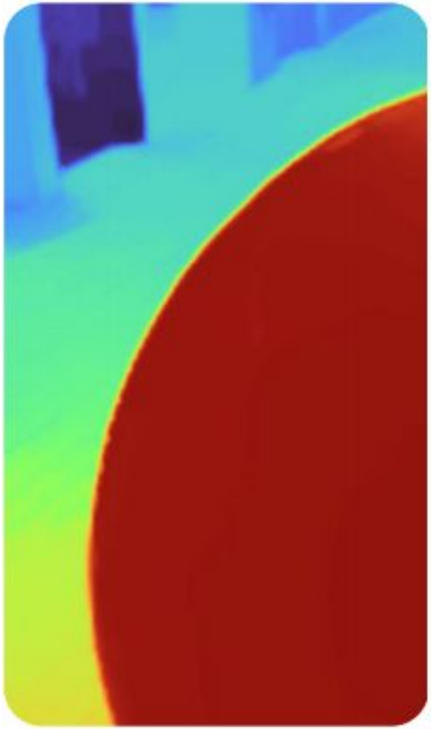


Experiments

DepthLab minimum viable application







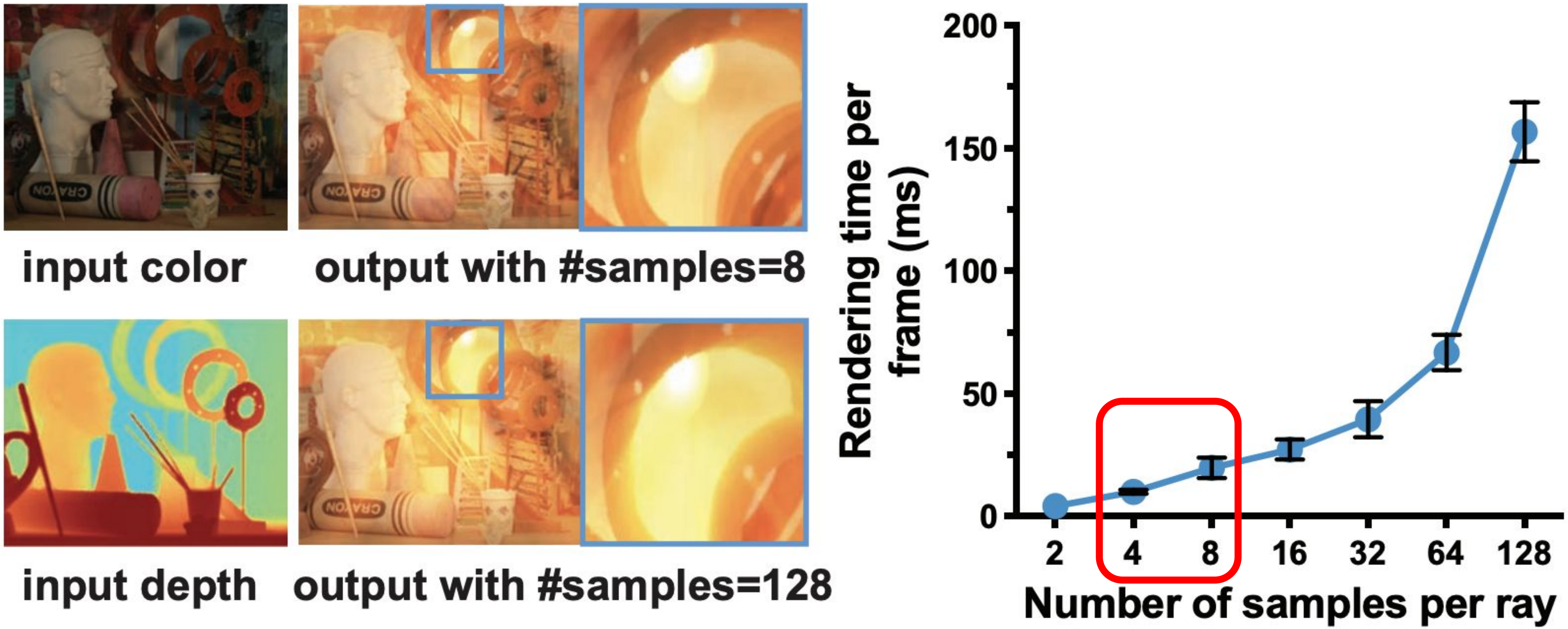
Experiments

General Profiling of MVP

Procedure	Timings (ms)
DepthLab's overall processing and rendering in Unity	8.32
DepthLab's data structure update and GPU uploading	1.63
Point Depth: normal estimation algorithm	< 0.01
Surface Depth: depth mesh update algorithm	2.41
Per-pixel Depth: visualization with single texture fetch	0.32

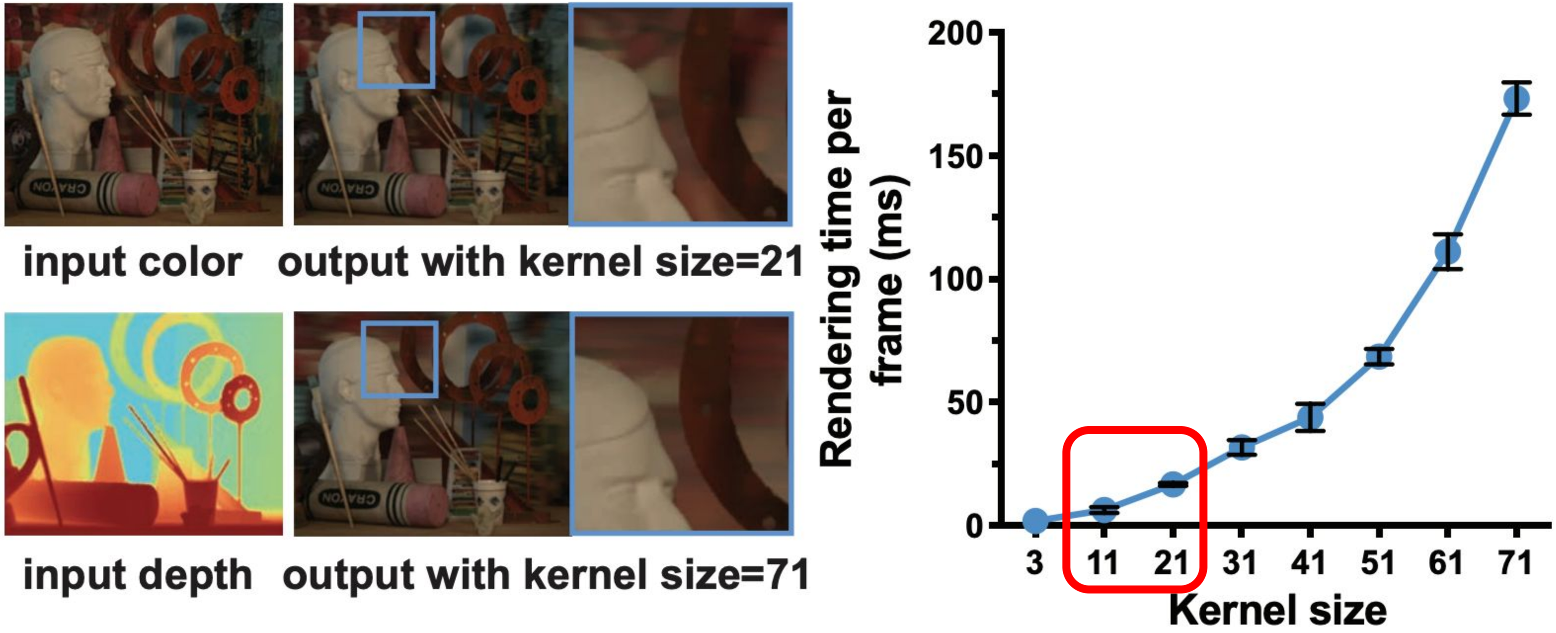
Experiments

Relighting



Experiments

Aperture effects



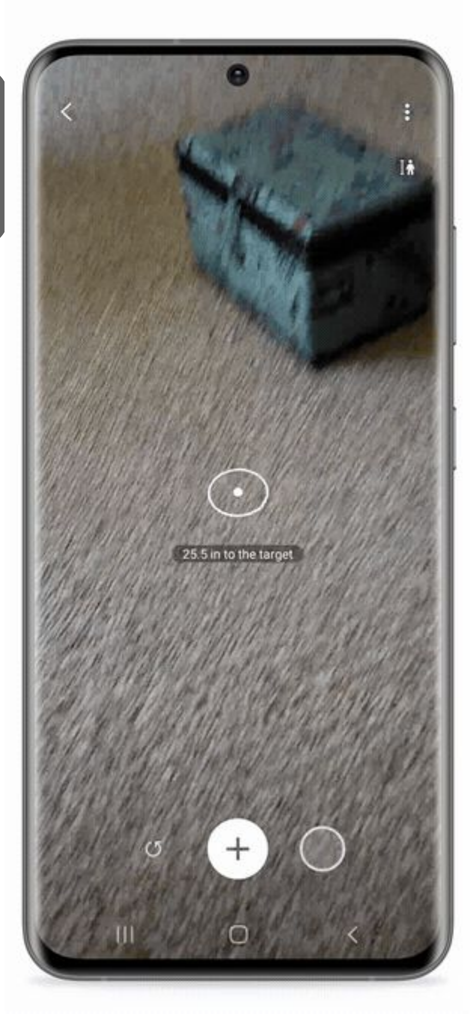
(a) examples of aperture effects (b) performance benchmark

Discussion

Deployment with partners

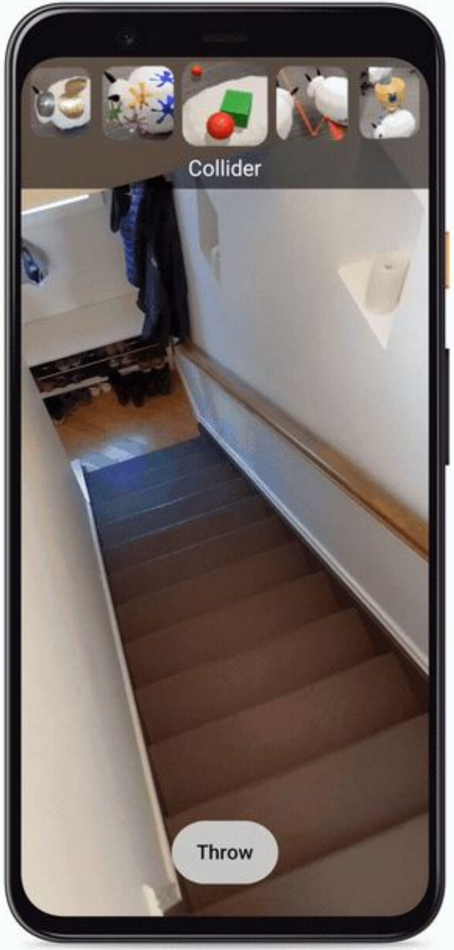
Discussion

Deployment with partners



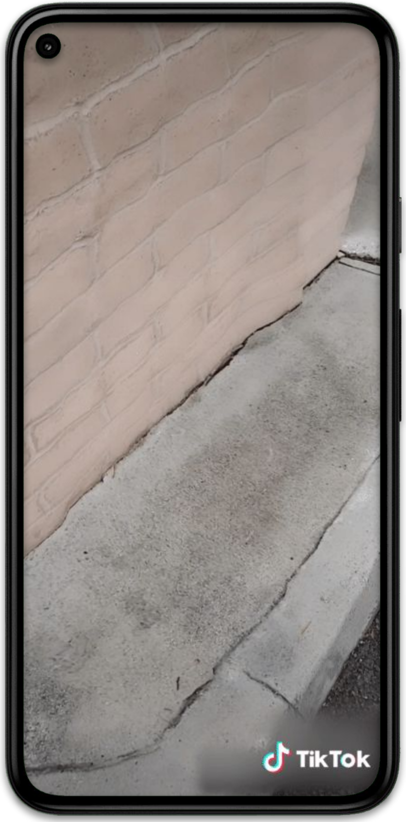
Discussion

Deployment with partners



AR Realism

In TikTok



AR Realism

Built into Lens Studio for Snapchat Lenses

Snap
Dancing Hotdog

Kevaid
Saving Chelon

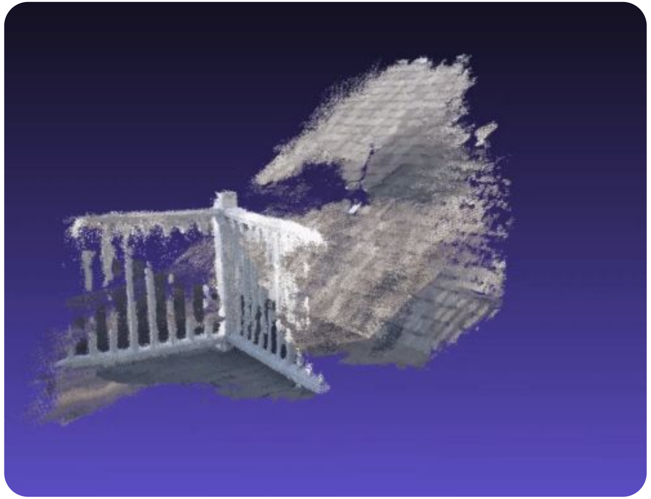
Quixotical
The Seed: World of Anthrotopia

New depth capabilities

Raw Depth API

Provides a more **detailed** representation of the geometry of the objects in the scene.

Camera Image

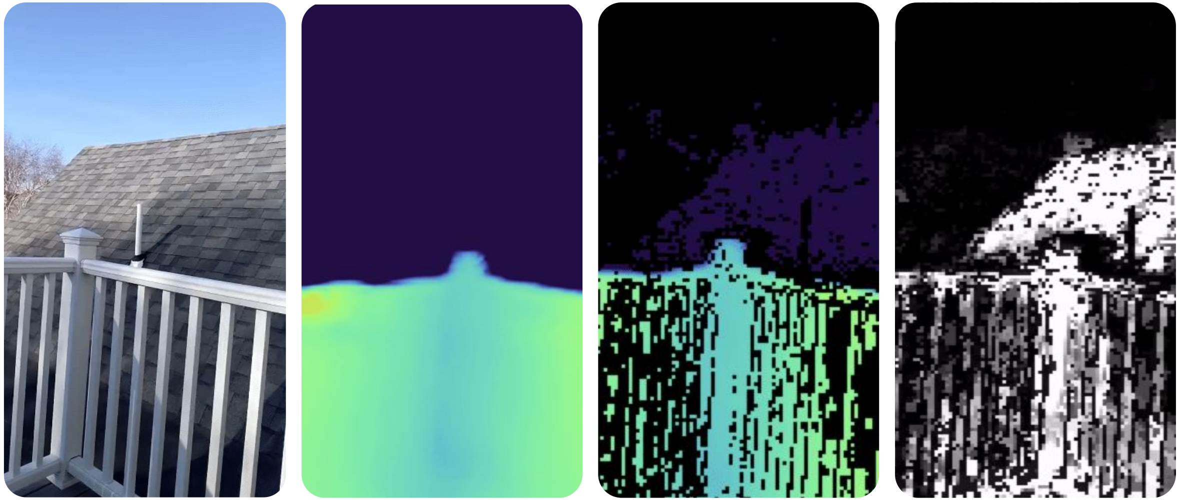


3D Point Cloud

New depth capabilities

Raw Depth API

Provides a more **detailed** representation of the geometry of the objects in the scene.

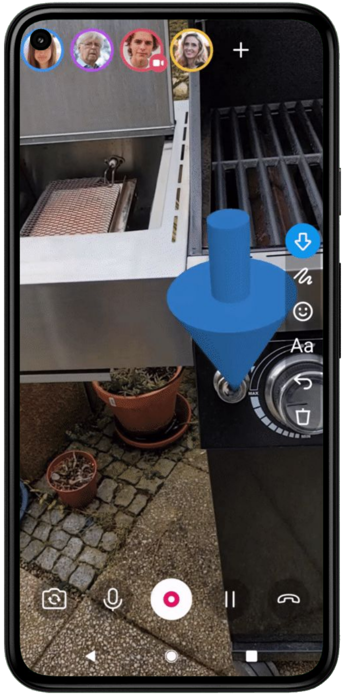


Camera Image | Depth Image | Raw Depth Image | Confidence Image

New depth capabilities

Depth Hit Test

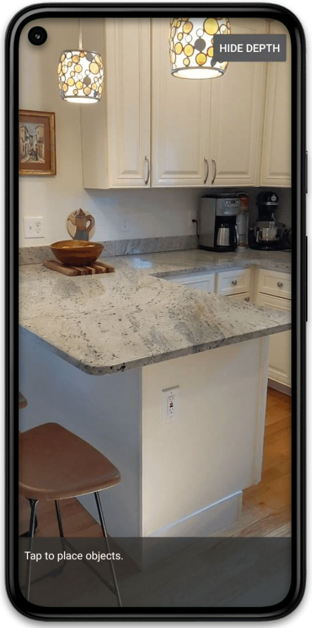
Try it yourself!



TeamViewer
LifeAR App

ARCore
Depth Lab App

ARCore
Depth Lab App



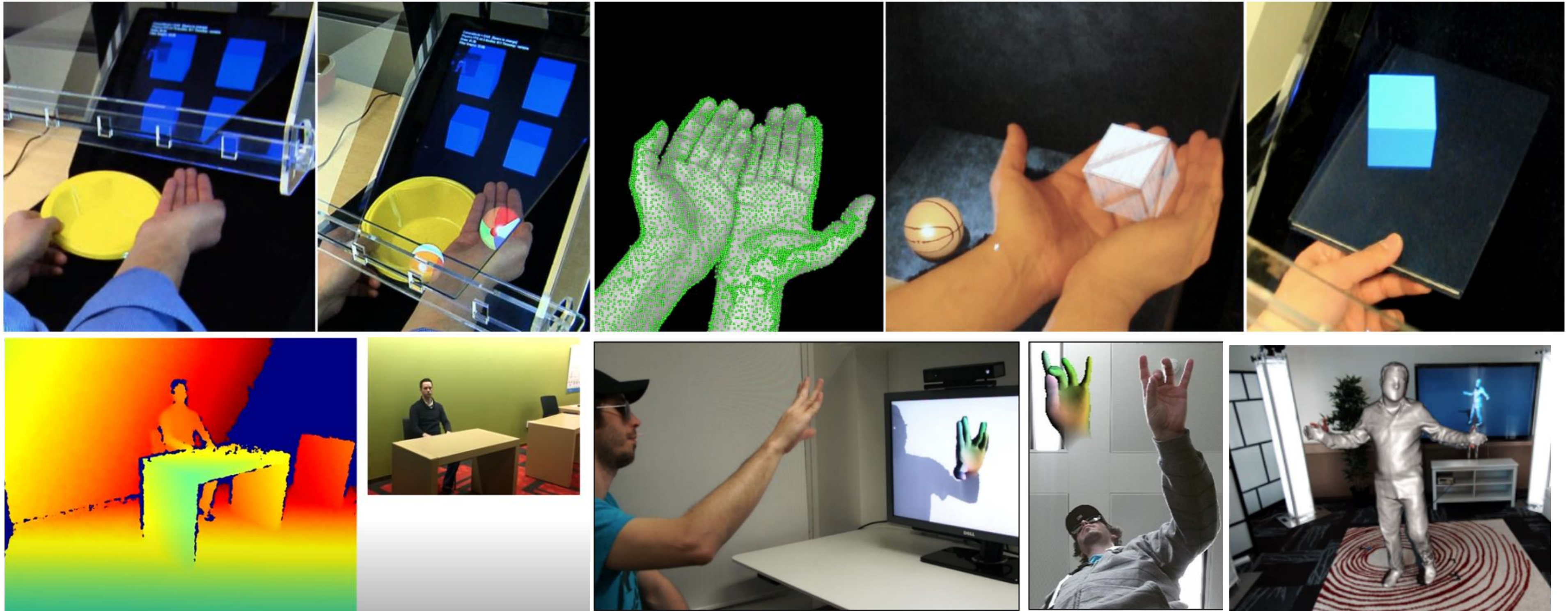
Depth API
Codelab

Raw Depth API
Codelab

Limitations

Design space of dynamic depth

Dynamic Depth? HoloDesk, HyperDepth, Digits, Holoportation for mobile AR?

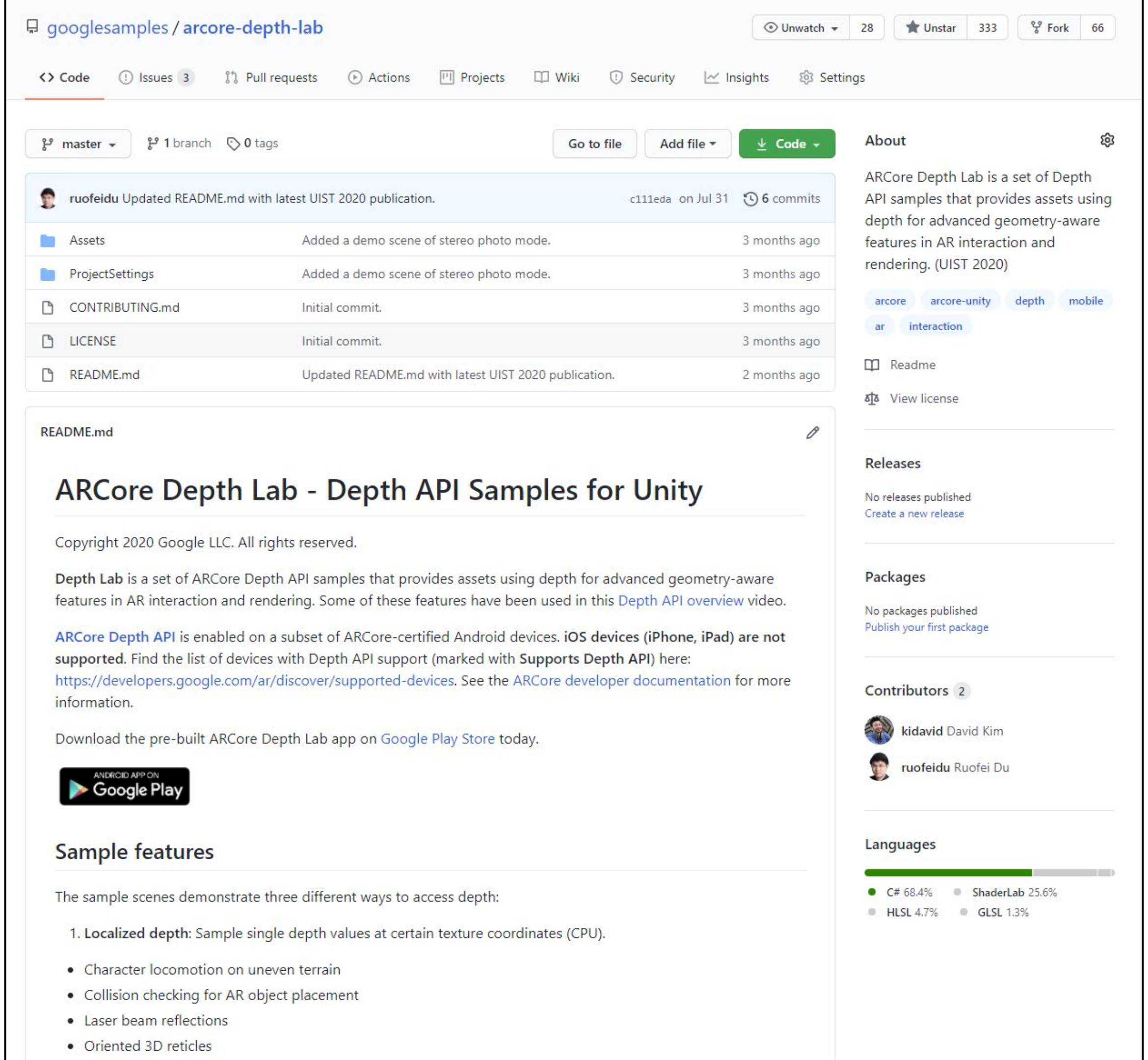


Envision

Design space of dynamic depth

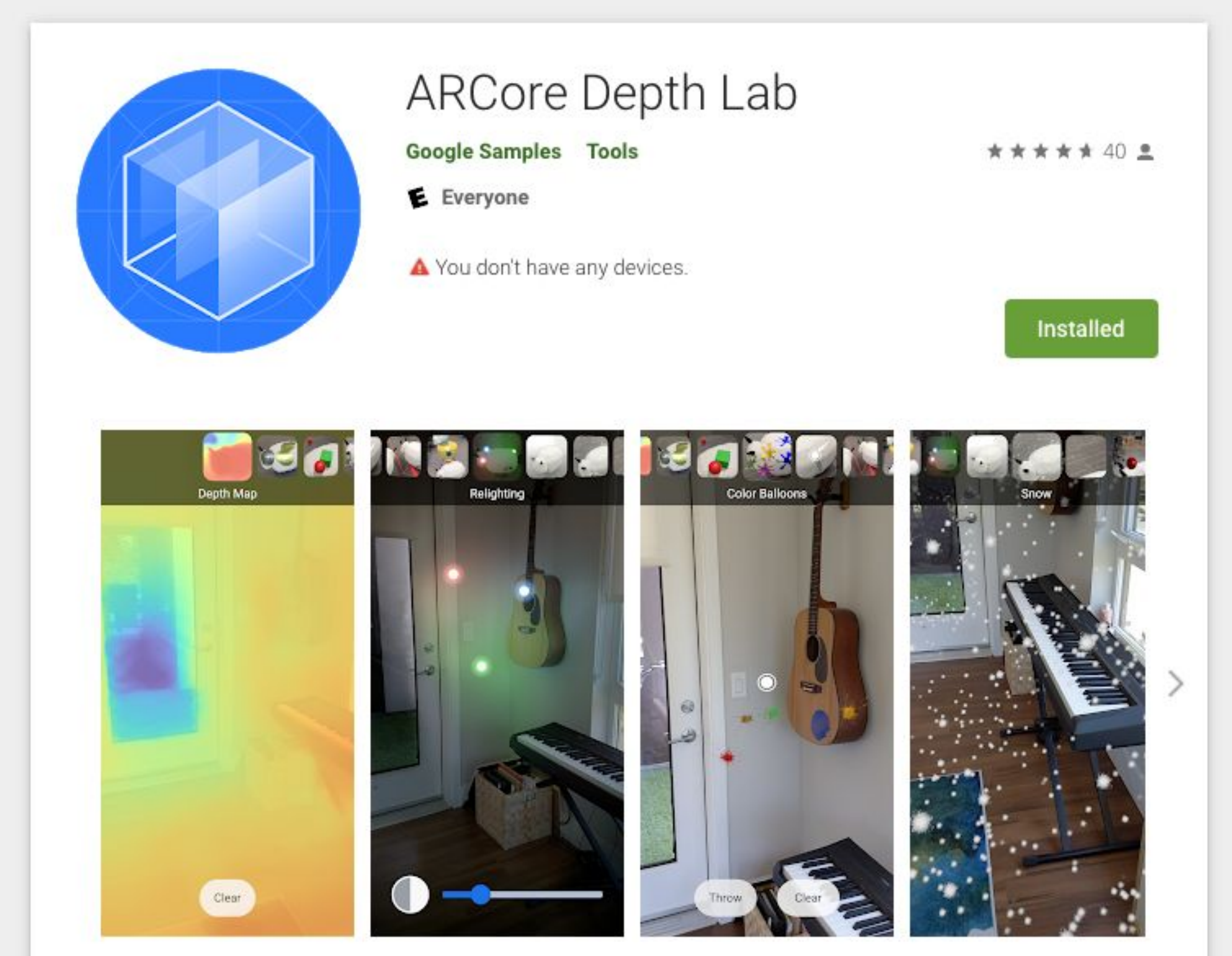
GitHub

Please feel free to fork!



Play Store

Try it yourself!



Impact

Significant Media Coverage

KEY QUOTES

"The result is a more believable scene, because the depth detection going on under the hood means your smartphone better understands every object in a scene and how far apart each object is from one another. Google says it's able to do this through optimizing existing software, so you won't need a phone with a specific sensor or type of processor. It's also all happening on the device itself, and not relying on any help from the cloud." - The Verge

"Occlusion is arguably as important to AR as positional tracking is to VR. Without it, the AR view will often "break the illusion" through depth conflicts." - UploadVR

"Alone, that feature (creating a depth map with one camera) would be impressive, but Google's intended use of the API is even better: occlusion, a trick by which digital objects can appear to be overlapped by real-world objects, blending the augmented and real worlds more seamlessly than with mere AR overlays." - VentureBeat

"Along with the Environmental HDR feature that blends natural light into AR scenes, ARCore now rivals ARKit with its own exclusive feature. While ARKit 3 offers People Occlusion and Body Tracking on compatible iPhones, the Depth API gives ARCore apps a level of environmental understanding that ARKit can't touch as of yet." - Next Reality

"More sophisticated implementations make use of multiple cameras...That's what makes this new Depth API almost magical. With just one camera, ARCore is able to create 3D depth maps ... in real-time as you move your phone around." - Slash Gear

Impact

Significant Media Coverage

COVERAGE LINKS

- A New Wave of AR Realism with the ARCore Depth API. Google Developers. June 25, 2020.
- Google Makes Its AR-Centric Depth API Available to All Developers. Engadget. June 25, 2020.
- AR Realism with the ARCore Depth API (Video). Google Developers. June 25, 2020.
- Introducing the ARCore Depth API for Android and Unity. Google AR & VR. June 25, 2020.
- ARCore's new Depth API is out of beta, bringing the next generation of hide-and-seek to phones. Android Police. June 25, 2020.
- Google is improving its augmented reality tool so virtual cats can hide behind your sofa. ZDNet. December 10, 2019.
- ARCore's Depth API helps create depth maps using a single camera. XDA Developers. December 10, 2019.
- Google's New Phone AR Update Can Hide Virtual Things in the Real World. CNET. December 9, 2019.
- Google Shows off Stunning New AR Features Coming to Web and Mobile Apps Soon. The Verge. December 9, 2019.
- Google's ARCore Depth API Enables AR Depth Maps and Occlusion with One Camera. VentureBeat. December 9, 2019.
- Google's ARCore Is Getting Full Occlusion For More Real AR. UploadVR. December 9, 2019.
- Google ARCore Depth API Now Available, Letting Devs Make AR More Realistic. RoadToVR. December 9, 2019.
- ARCore Depth API Takes Android AR Experiences To A Whole New Level. VRScout. December 9, 2019.
- Google Update Adds Real-World Occlusion to ARCore with Depth API. Next Reality. December 9, 2019.
- ARCore phones can now detect depth with a single camera. 9To5Google. December 9, 2019.
- ARCore Depth API: How it will fundamentally transform your AR experiences. Android Authority. December 9, 2019.
- ARCore Depth API lets you hide cats behind sofas even with one camera. SlashGear. December 9, 2019.
- Google's Latest ARCore API Needs Just One Camera For Depth Detection. HotHardware. December 9, 2019.
- Get Ready for the ARCore Depth API (Video). Google AR & VR. December 9, 2019.
- Blending Realities with the ARCore Depth API. Google Developers. December 9, 2019.

More Links

Significant Media Coverage

WebXR + ARCore Depth:
https://storage.googleapis.com/chromium-webxr-test/r991081/proposals/index.html

Hugging Face Depth:
https://huggingface.co/spaces/Detomo/Depth-Estimation

ARCore Depth Lab Play Store App:
https://play.google.com/store/apps/details?id=com.google.ar.unity.arcore_depth_lab

DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality

Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte,
Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces,
Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim

Google | ACM UIST 2020

Thank you!

DepthLab | UIST 2020

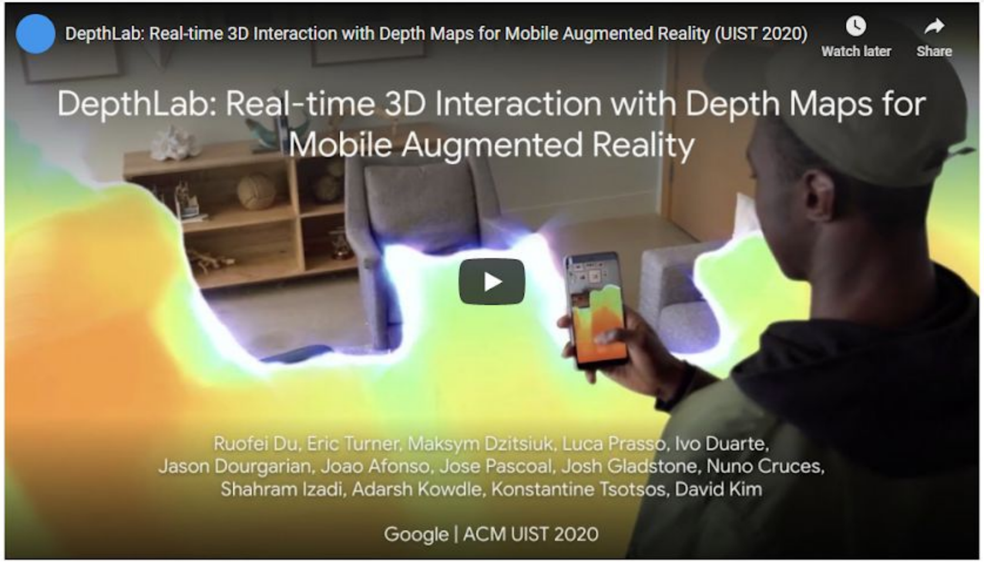
DEPTHLAB: REAL-TIME 3D INTERACTION WITH DEPTH MAPS FOR MOBILE AUGMENTED REALITY

ACM UIST 2020

Download PDF (6 MB) or Low-Res PDF (4 MB)

Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte,
Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces,
Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim

Google LLC



Demo

DepthLab | UIST 2020

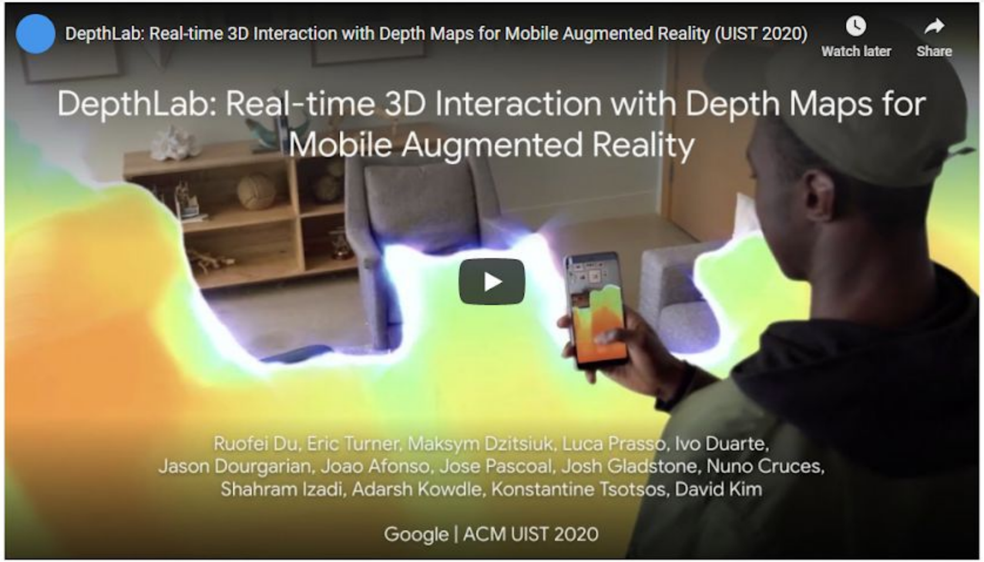
DEPTHLAB: REAL-TIME 3D INTERACTION WITH DEPTH MAPS FOR MOBILE AUGMENTED REALITY

ACM UIST 2020

Download PDF (6 MB) or Low-Res PDF (4 MB)

Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte, Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces, Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim

Google LLC



DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality

Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte,
Jason Dourgarian, Joao Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces,
Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, David Kim

Google | ACM UIST 2020

Ad hoc UI: On-the-fly Transformation of Everyday Objects into Tangible 6DOF Interfaces for AR

Ruofei Du, Alex Olwal, Mathieu Le Goc, Shengzhi Wu, Danhang Tang, Yinda Zhang, Jun Zhang, David Joseph Tan, Federico Tombari, David Kim

Google | CHI 2022 Interactivity


Opportunistic Interfaces for Augmented Reality: Transforming Everyday Objects into Tangible 6DoF Interfaces Using Ad hoc UI
Ruofei Du, Alex Olwal, Mathieu Le Goc, Shengzhi Wu, Danhang Tang, Yinda Zhang, Jun Zhang, David Joseph Tan, Federico Tombari, David Kim
Google | ACM CHI 2022

Introduction

Motivation

How can we allow users to instantly transform arbitrary everyday objects into Tangible User Interfaces?

Representations

Physical objects - shape - 2D

Representations

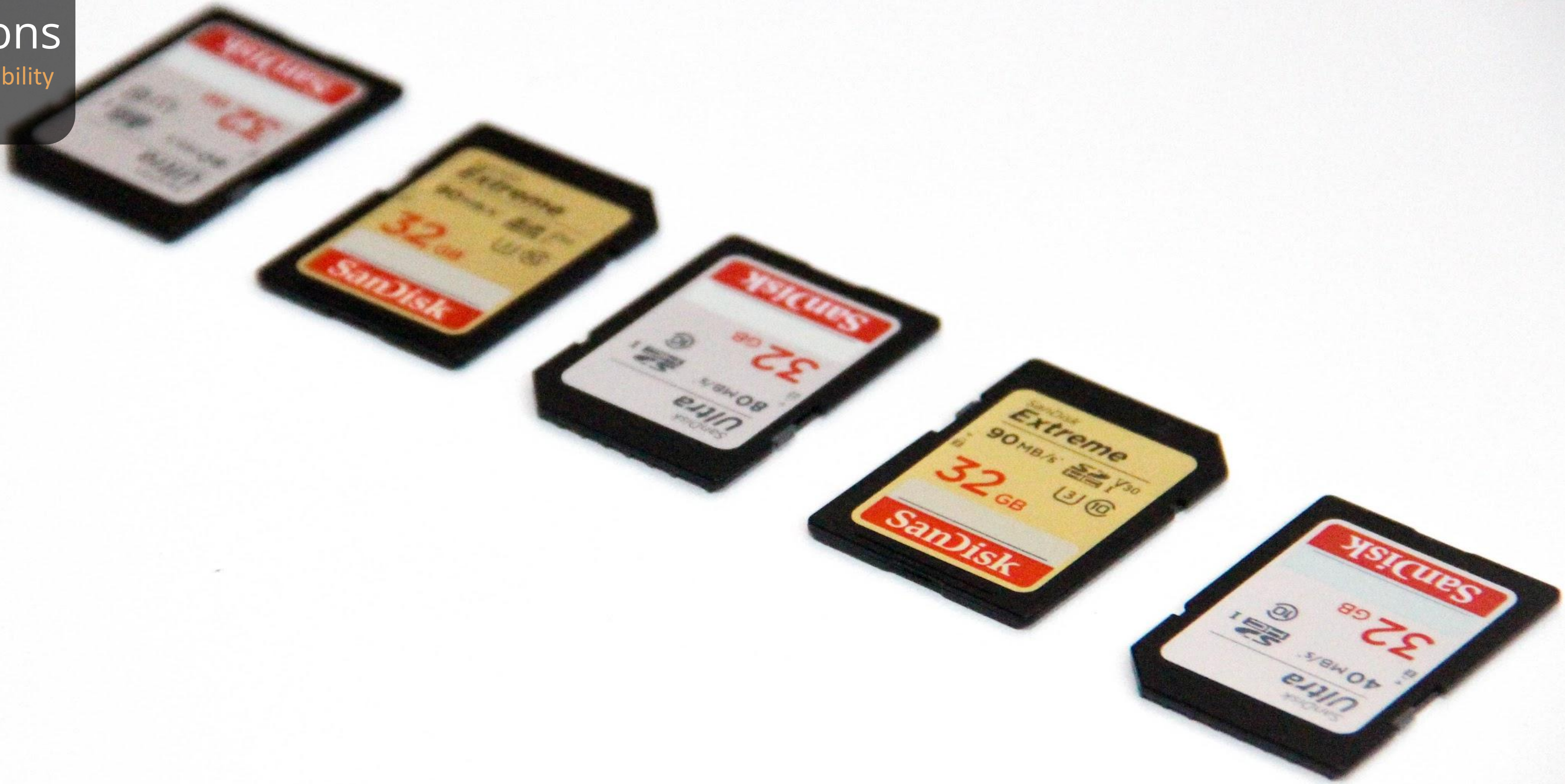
Physical objects - shape - 3D

Representations

Physical objects - portability

Representations

Physical objects - Deformability



Representations

Physical objects

Representations

Physical objects

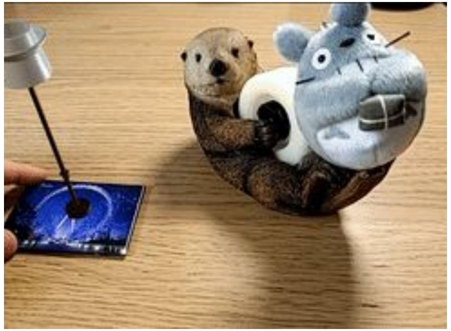
Representations

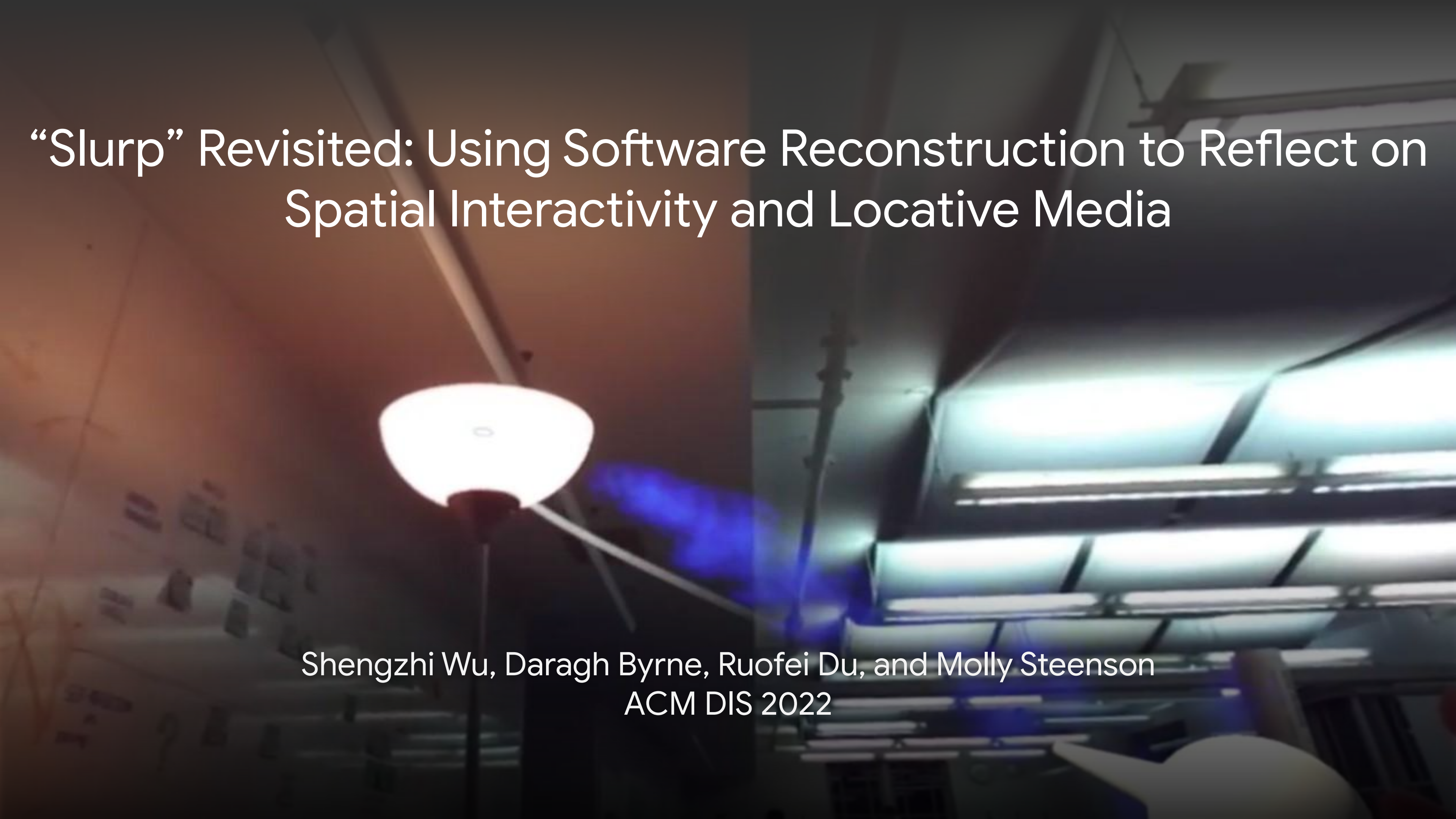
Physical objects

Representations

Physical objects

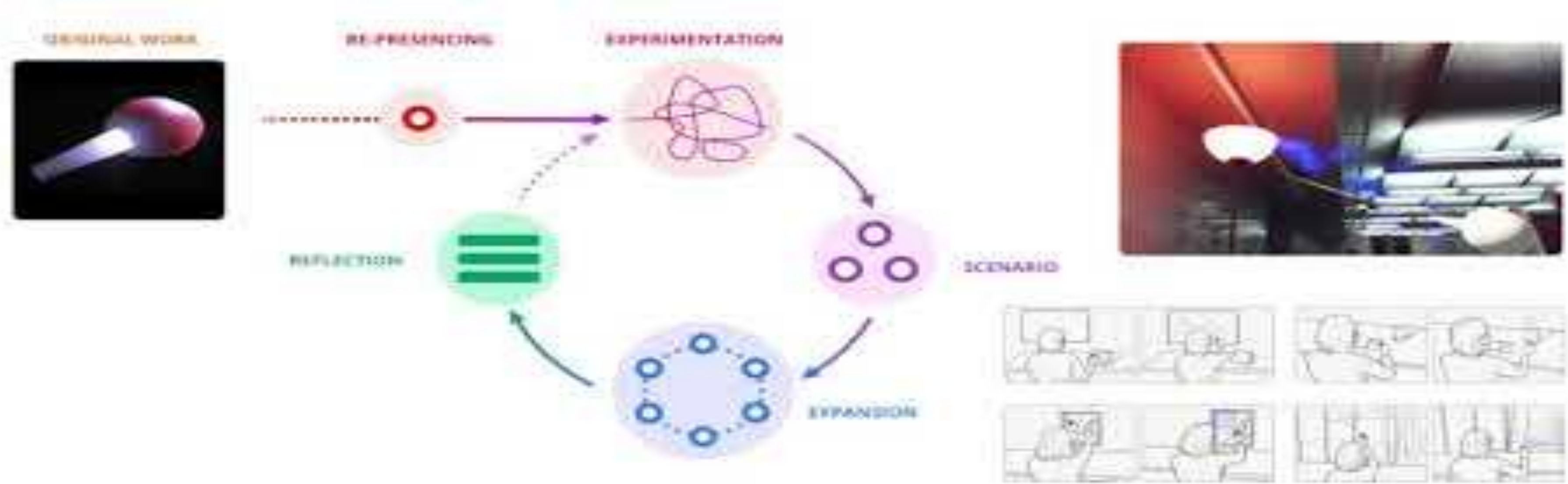
Applications





“Slurp” Revisited: Using Software Reconstruction to Reflect on Spatial Interactivity and Locative Media

Shengzhi Wu, Daragh Byrne, Ruofei Du, and Molly Steenson
ACM DIS 2022



Overview of our design process and artifacts

RetroSphere: Self-Contained Passive 3D Controller Tracking for Augmented Reality

Ananta Narayanan Balaji, Clayton Kimber, David Li, Shengzhi Wu, Ruofei Du, David Kim

ACM Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT) 2022

元宇宙中的交互计算与包容普惠

Computational Interaction for a Universally Accessible Metaverse

第三章 · 数字人与增强交互

Ruofei Du (杜若飞)
Senior Research Scientist
Google Labs, San Francisco
www.ruofeidu.com
me@duruofei.com

Montage4D: Interactive Seamless Fusion of Multiview Video Textures

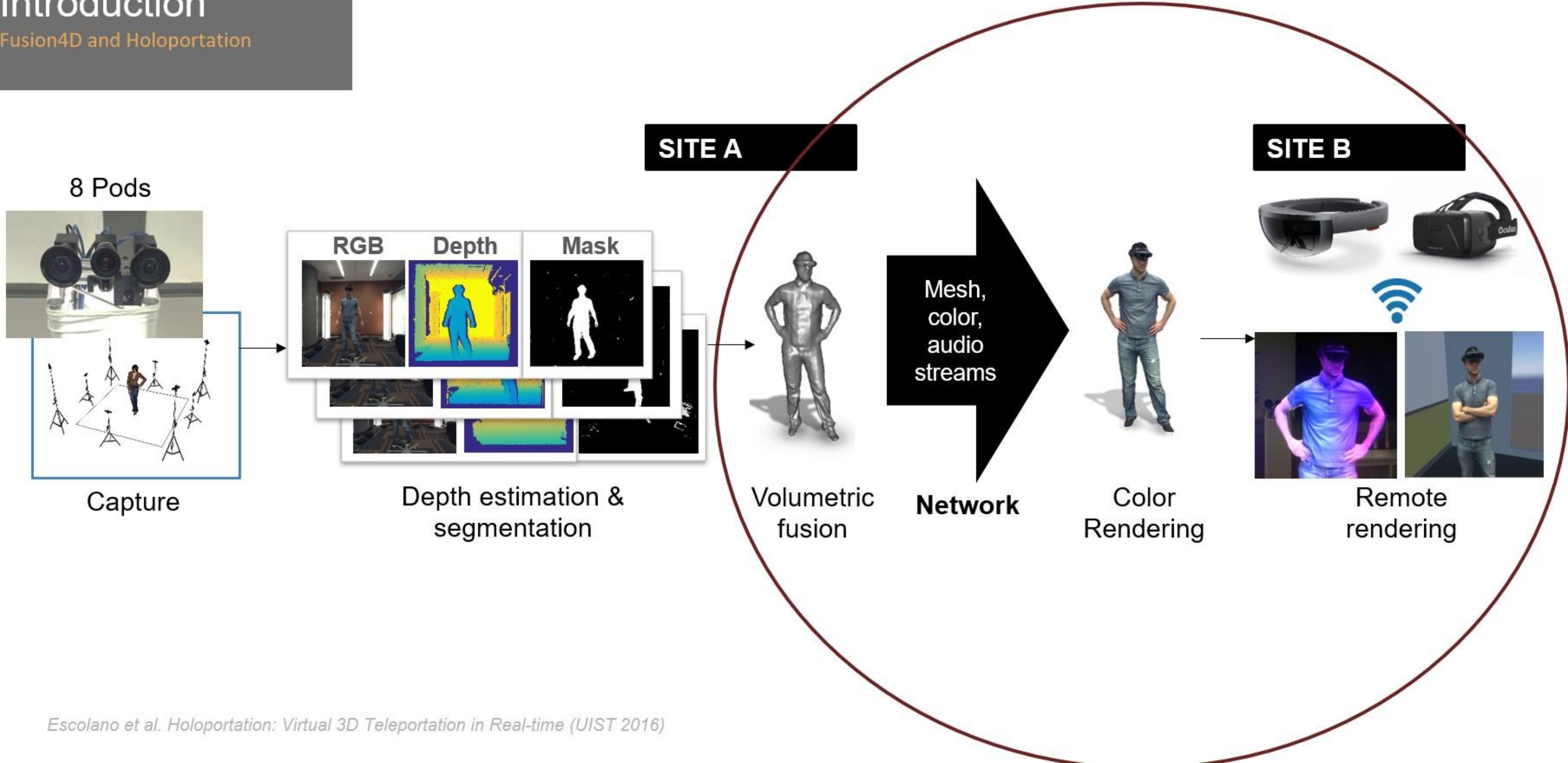
Ruofei Du[†‡], Ming Chuang[‡¶], Wayne Chang[‡], Hugues Hoppe[‡§], and Amitabh Varshney[†]

[†]Augmentarium | UMIACS | University of Maryland, College Park

[‡]Microsoft Research, Redmond [¶]PerceptIn Inc. [§]Google Inc.

Introduction

Fusion4D and Holoportation



Escolano et al. Holoportation: Virtual 3D Teleportation in Real-time (UIST 2016)

Fusing multiview video textures onto dynamic task with real-time constraint is **a challenging task**

30%

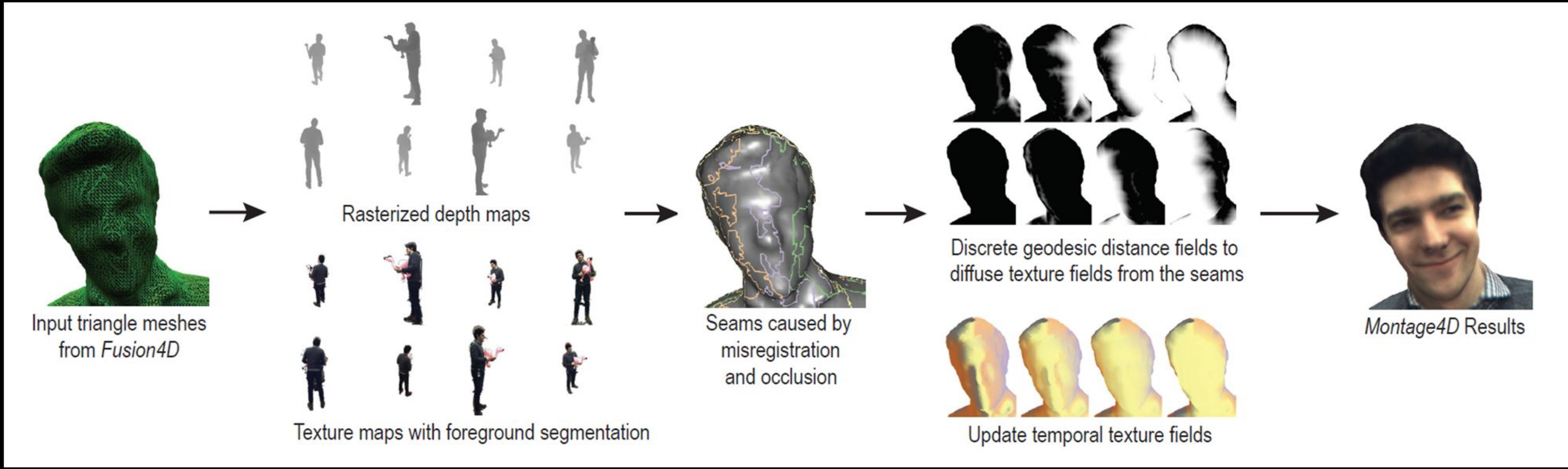
of the users does not believe the 3D reconstructed person looks real

Motivation

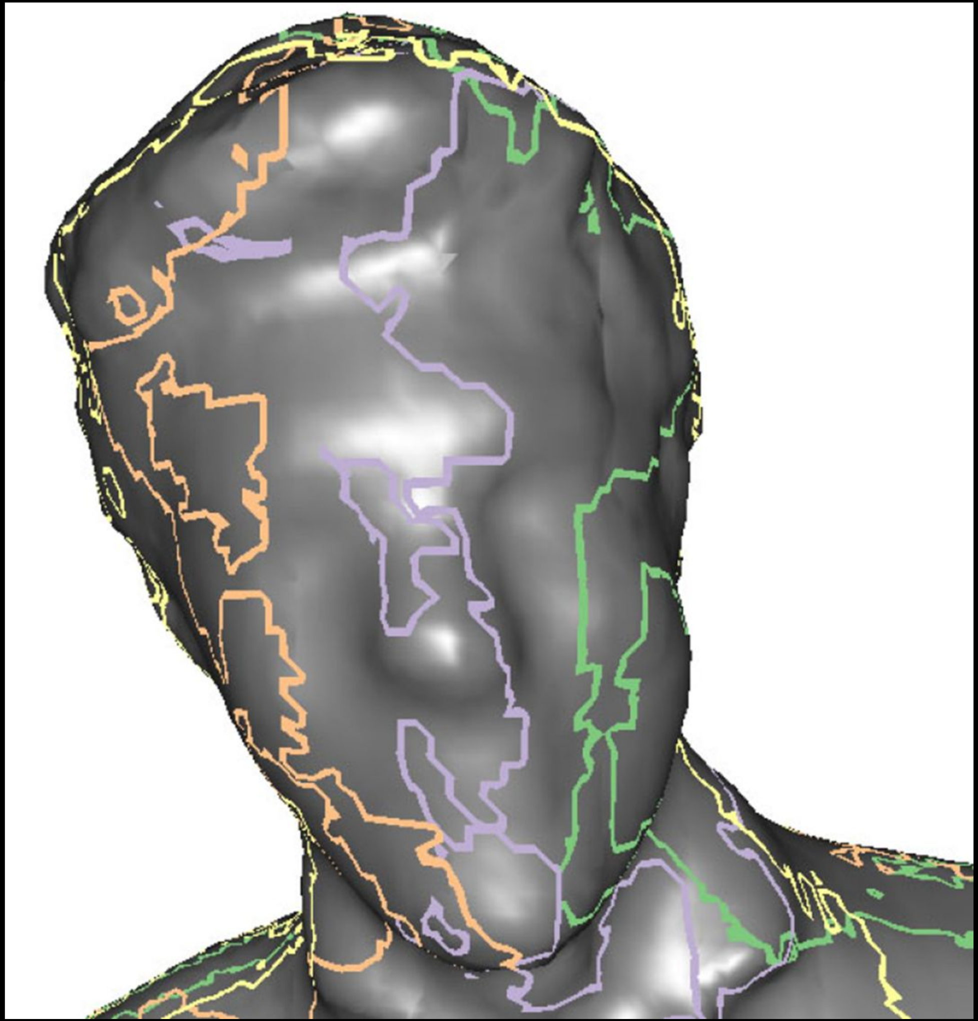
Visual Quality Matters

Workflow

Identify and diffuse the seams



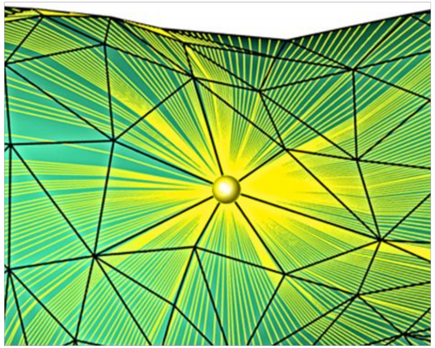
Seams
Causes



Geodesics

For diffusing the seams

Geodesic is the **shortest** route between two points on the surface.



On triangle meshes, this is challenging because of the computation of **tangent directions**.
And shortest paths are defined on **edges** instead of the vertices.

Approximate Geodesics

For diffusing the seams

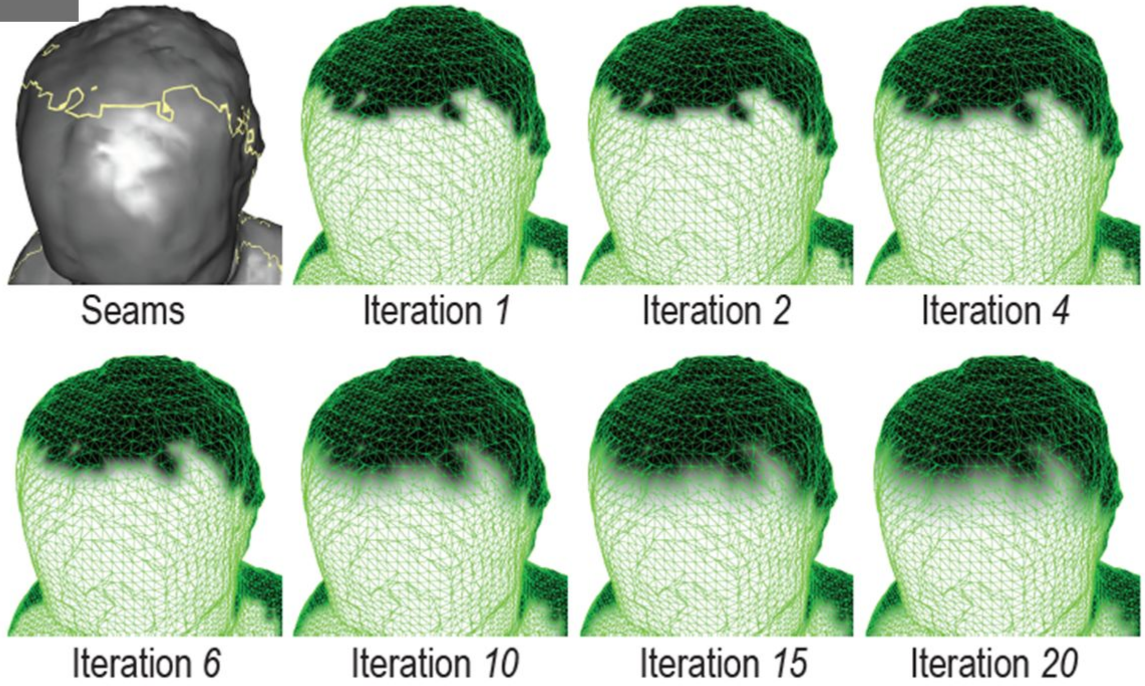
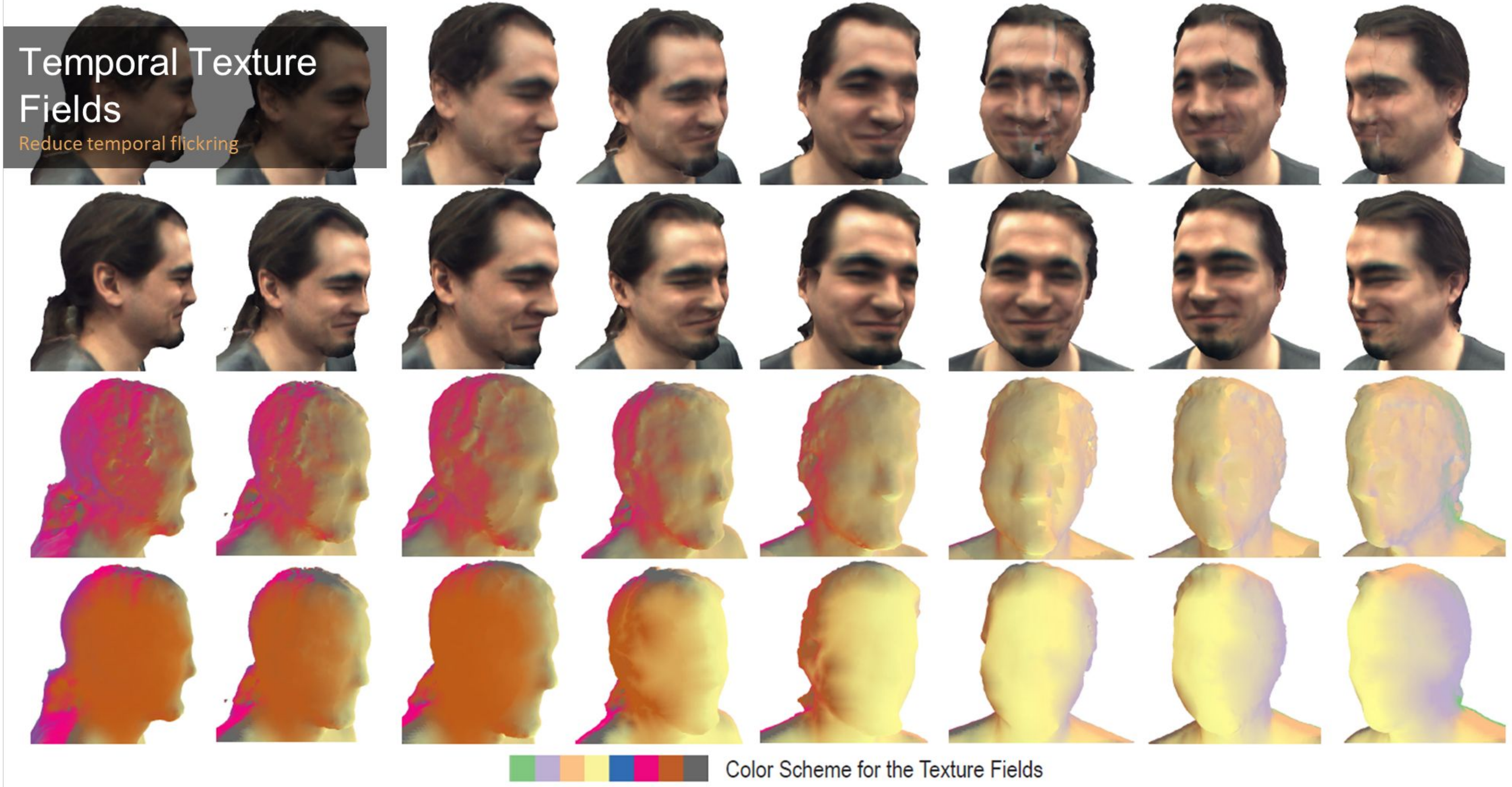


Figure 6: *Examples of the initial seam triangles and the propagation process for updating the geodesics.*

Temporal Texture Fields

Reduce temporal flickring



Exmperiment

Cross-validation

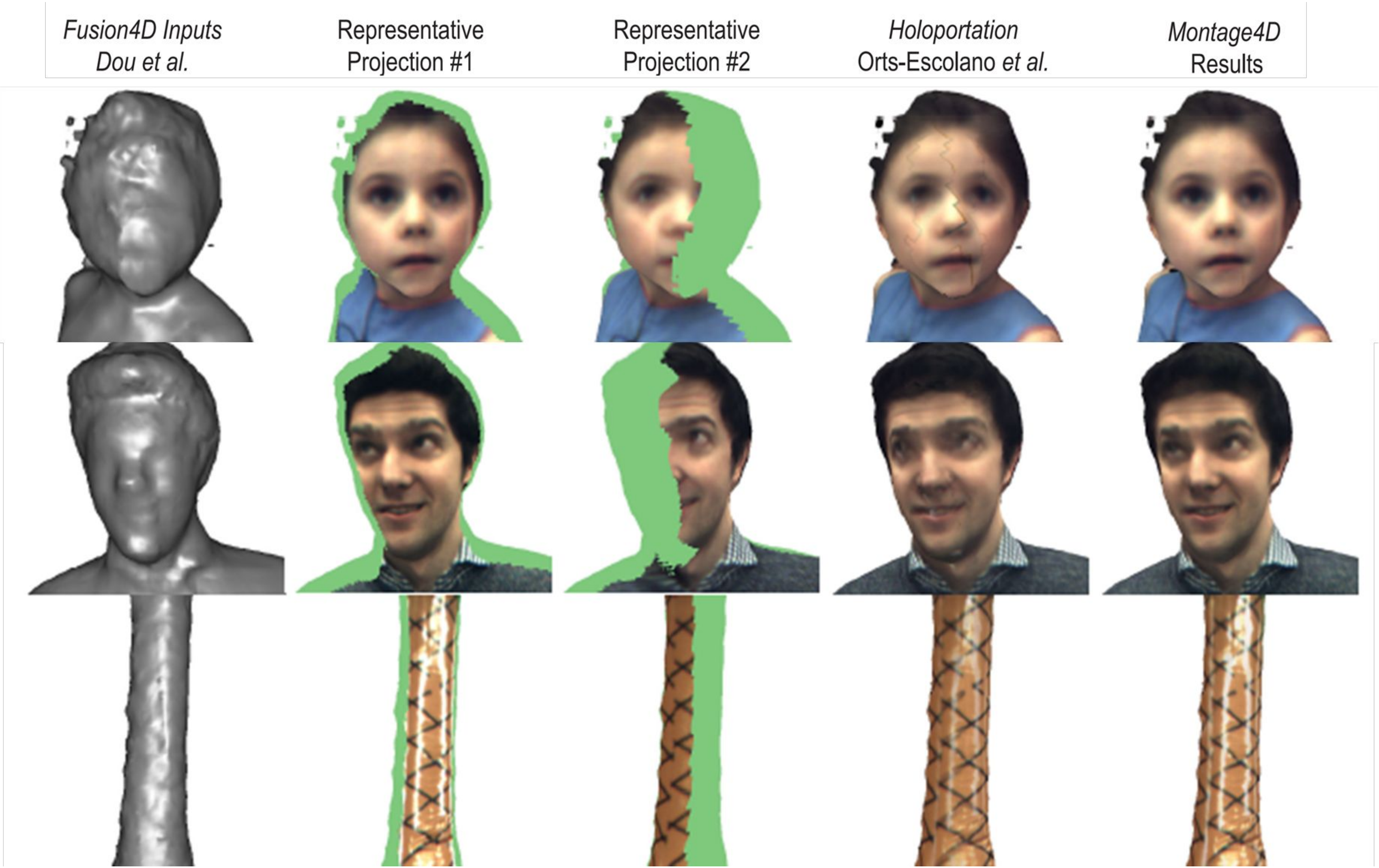
Table 1: Comparison between *Holoportation* and *Montage4D* in cross-validation experiments

Dataset	Frames	#vertices / frame	#triangles / frame	*Holoporation*				*Montage4D*			
				RMSE	SSIM	PSNR	FPS	RMSE	SSIM	PSNR	FPS
Timo	837	131K	251K	5.63%	0.9805	38.60dB	227.2	3.27%	0.9905	40.23dB	135.0
Yury	803	132K	312K	5.44%	0.9695	39.20dB	222.8	3.01%	0.9826	40.52dB	130.5
Sergio	837	215K	404K	7.74%	0.9704	29.84dB	186.8	4.21%	0.9813	30.09dB	114.3
Girl	1192	173K	367K	7.16%	0.9691	36.28dB	212.56	3.73%	0.9864	36.73dB	119.4
Julien	526	157K	339K	12.63%	0.9511	33.94dB	215.18	6.71%	0.9697	35.05dB	120.6

Montage4D achieves **better quality with over 90 FPS**

- Root mean square error (RMSE) ↓
- Structural similarity (SSIM) ↑
- Signal-to-noise ratio (PSNR) ↑

Fusion4D Inputs
Dou et al.
Representative
Projection #1
Representative
Projection #2
Holoportation
Orts-Escolano et al.
Montage4D
Results


Fusion4D Inputs
Dou et al.
Representative
Projection #1
Representative
Projection #2
Holoportation
Orts-Escolano et al.
Montage4D
Results

Before

After

In conclusion, **Montage4D** provides a practical texturing solution for real-time 3D reconstructions. In the future, we envision that **Montage4D** is useful for **fusing the massive multi-view video data** into VR applications like remote business meeting, remote training, and broadcasting industries.

Total Relighting:
Learning to Relight Portraits for Background Replacement

ROHIT PANDEY*, SERGIO ORTS ESCOLANO*, CHLOE LEGENDRE*, CHRISTIAN HÄNE, SOFIEN BOUAZIZ, CHRISTOPH RHEMANN, PAUL DEBEVEC, and SEAN FANELLO, Google Research

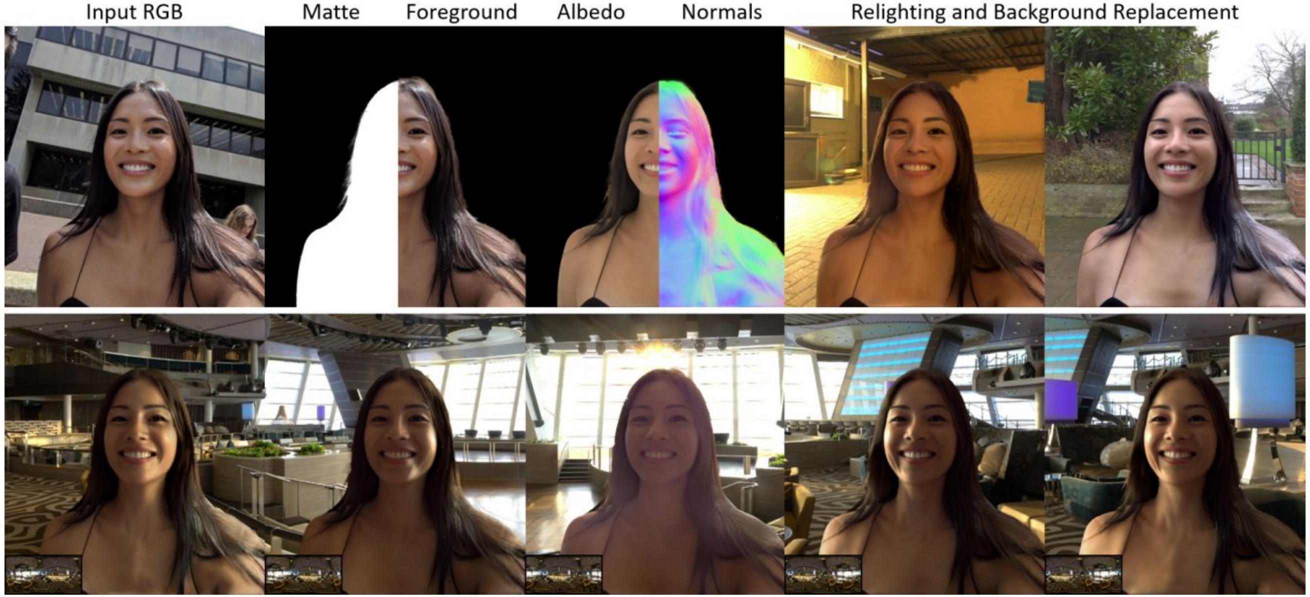
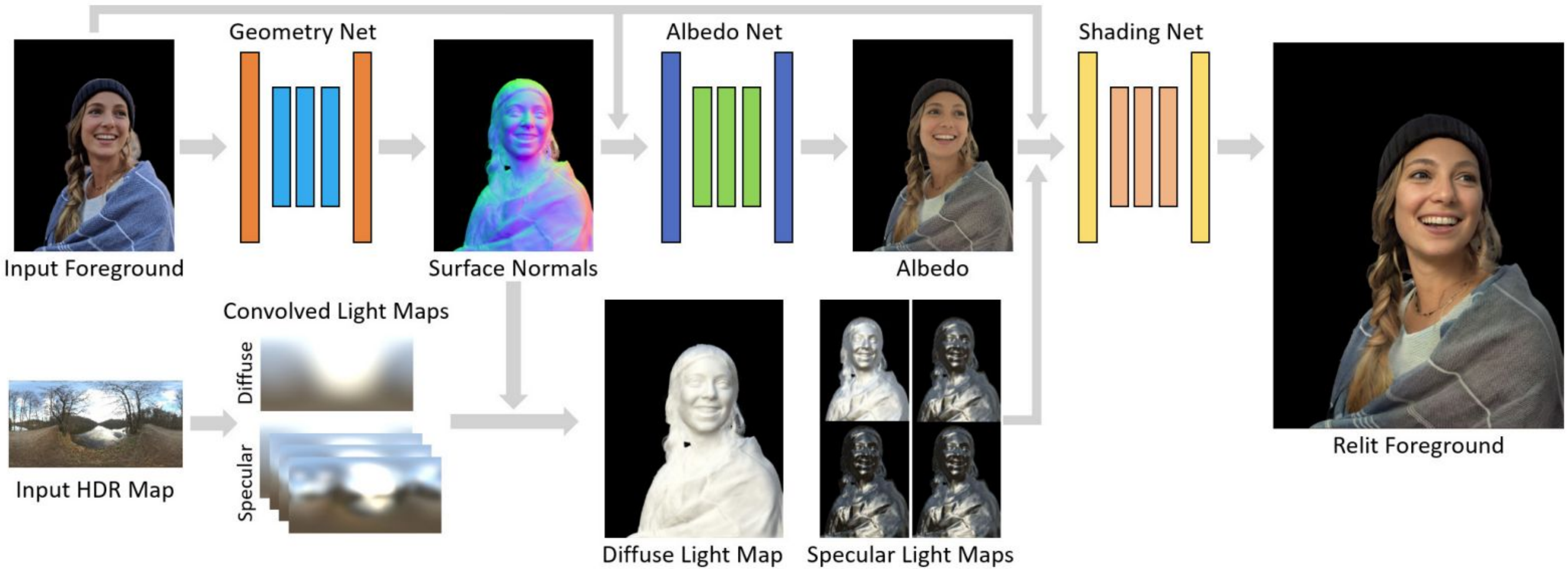


Fig. 1. Given a portrait and an arbitrary high dynamic range lighting environment, our framework uses machine learning to composite the subject into a new scene, while accurately modeling their appearance in the target illumination condition. We estimate a high quality alpha matte, foreground element, albedo map, and surface normals, and we propose a novel, per-pixel lighting representation within a deep learning framework.

ACM Trans. Graph., Vol. 40, No. 4, Article 1. SIGGRAPH 2021


Geometry Net
Albedo Net
Shading Net
Input Foreground
Surface Normals
Albedo
Convolved Light Maps
Diffuse
Specular
Input HDR Map
Diffuse Light Map
Specular Light Maps
Relit Foreground

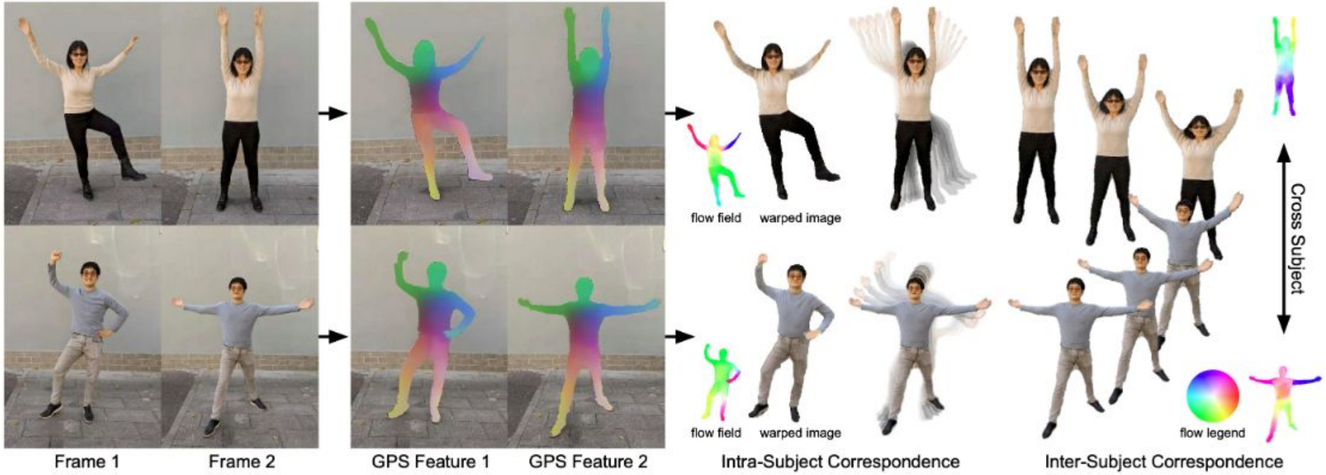
HumanGPS: Geodesic PreServing Feature for Dense Human Correspondences

CVPR 2021

Feitong Tan[1,2] Danhang Tang[1] Mingsong Dou[1] Kaiwen Guo[1] Rohit Pandey[1] Cem Keskin[1]
Ruofei Du[1] Deqing Sun[1] Sofien Bouaziz[1] Sean Fanello[1] Ping Tan[2] Yinda Zhang[1]

[1] Google

[2] Simon Fraser University



Live Demo

1. Click Choose File to upload a human image and a mask image (recommended w : h = 256 : 384) or use the example images: 1 2 3
2. Click 'Process' button to run the model. You may use Pen, Brush, Eraser, and Clear to doodle on the mask image.

Note that the first time of 'Process' takes longer time for initialization. The model may take a few seconds to load in some region.

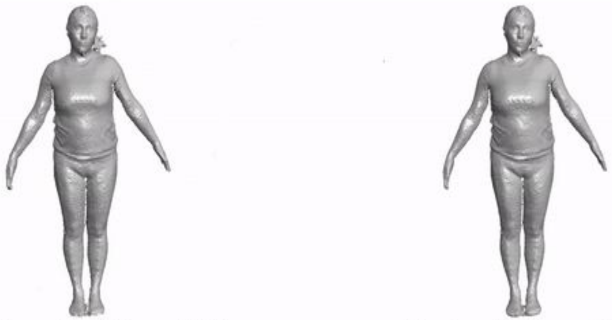
HumanGPS

Process

12061 ms

Choose File No file chosen

Choose File No file chosen



Fusion without HumanGPS Fusion with HumanGPS

Fusion without HumanGPS Fusion with HumanGPS

GazeChat

Enhancing Virtual Conferences With Gaze-Aware 3D Photos

Zhenyi He[†], Keru Wang[†], Brandon Yushan Feng[‡], Ruofei Du[⸸], Ken Perlin[†]

[†] New York University
[‡] University of Maryland, College Park
[⸸] Google

ruofei

keru

zhenyi

brandon

Introduction

VR headset & video streaming

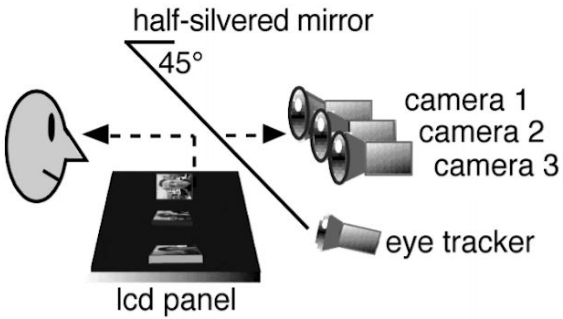
ruofei

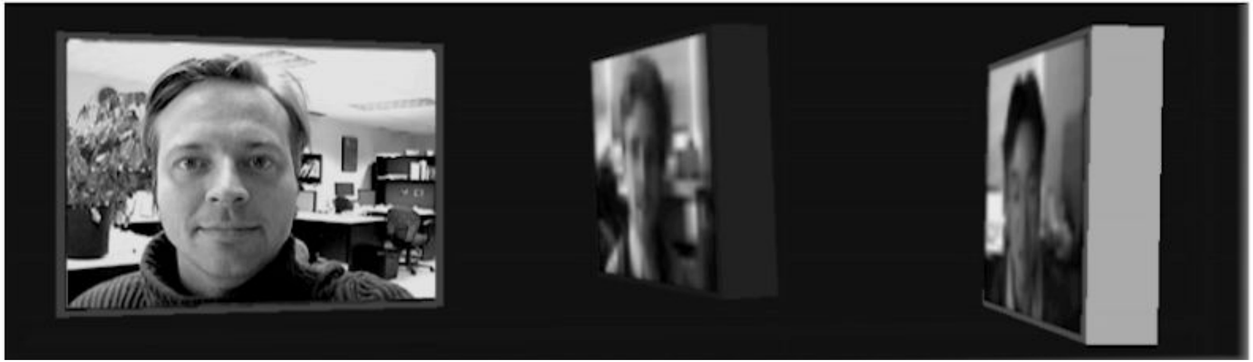
keru

zhenyi

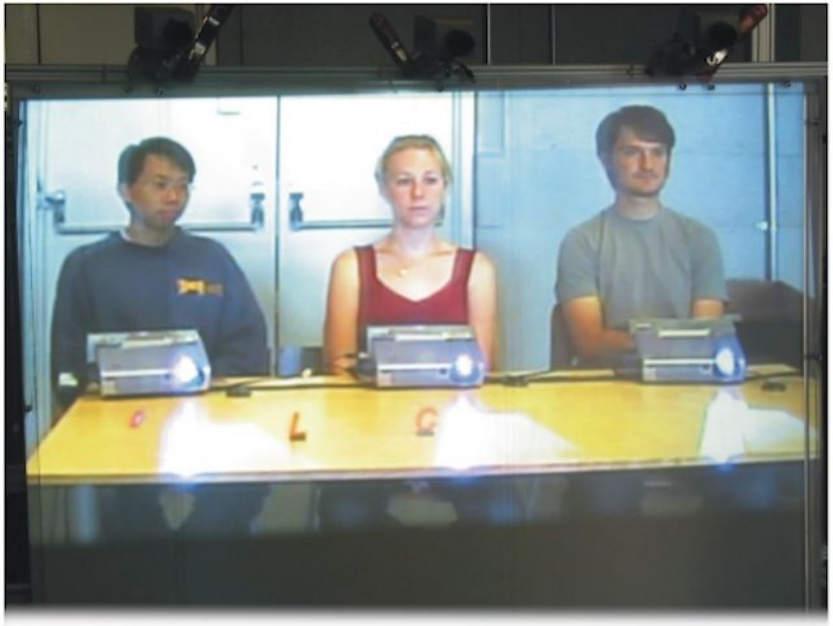
brandon

Related Work
Gaze-2 (2003)





Related Work
MultiView (2005)



Related Work

MMSpace (2016)

Our Work
GazeChat (UIST 2021)

Gaze Awareness

Definition

Gaze awareness, defined here as *knowing what someone is looking at*.

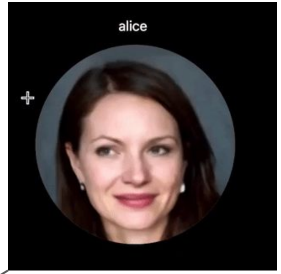
Gaze Awareness
Definition

raw input image

gaze correction

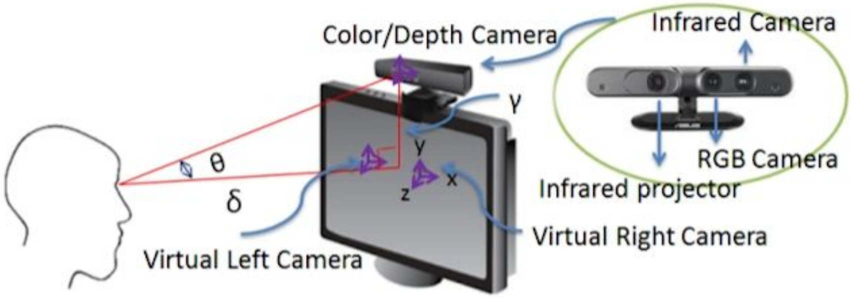
gaze redirection

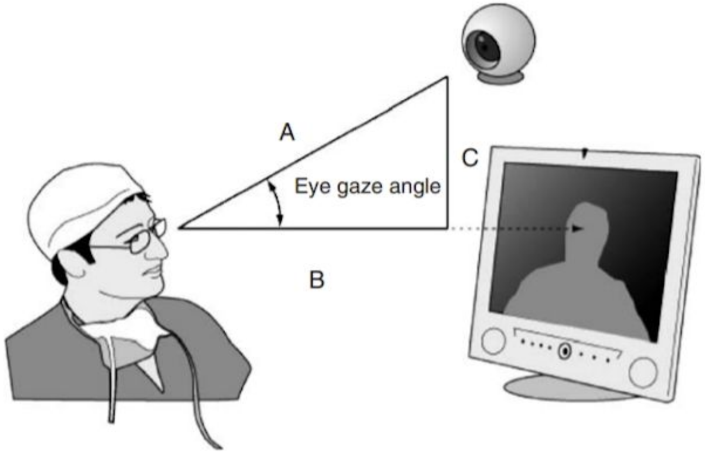
GazeChat



Gaze Correction

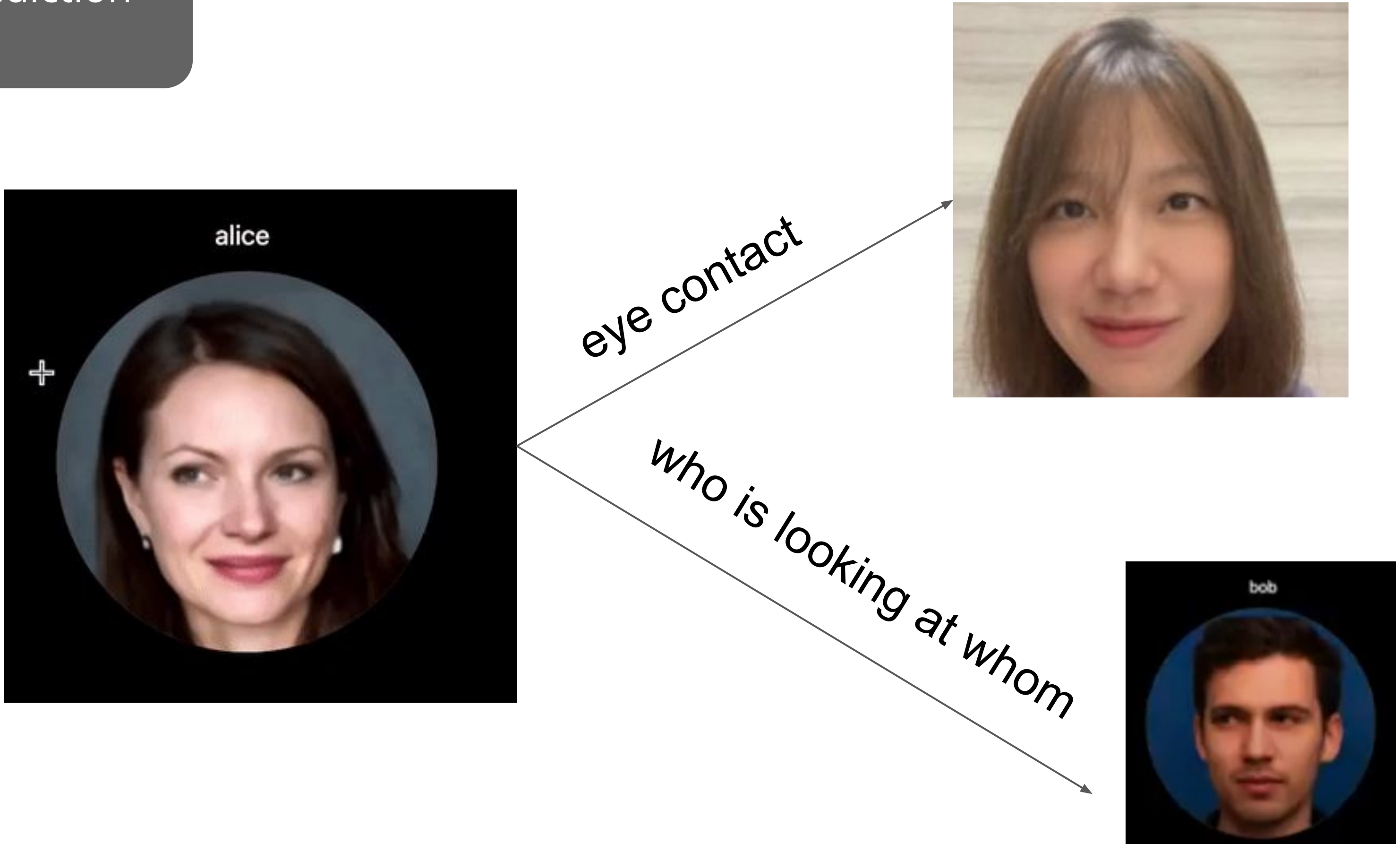
Definition



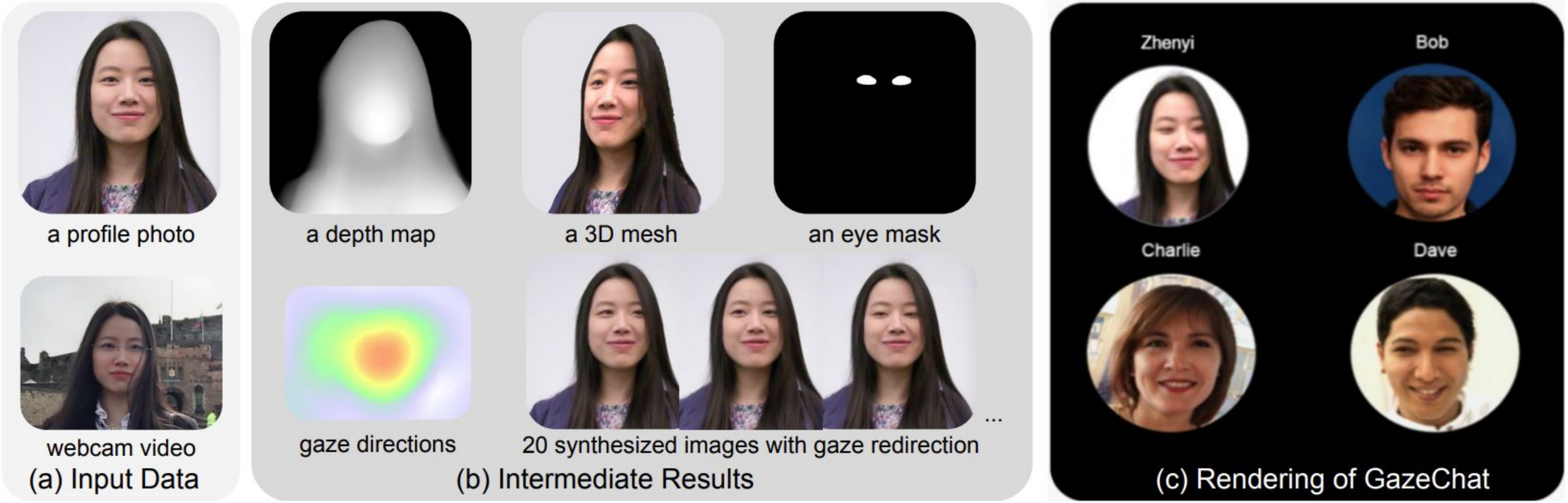


Gaze Rediction

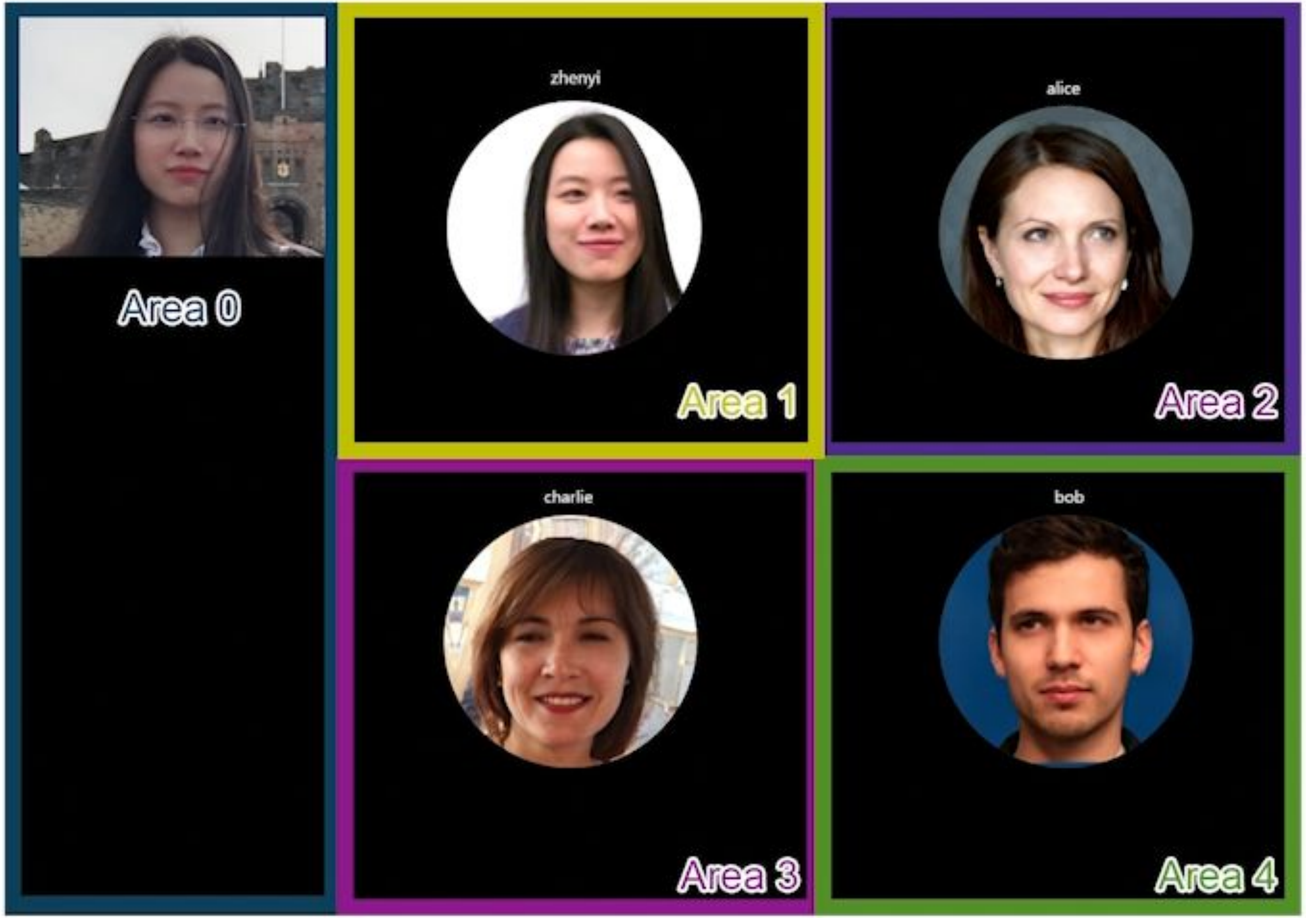
Definition



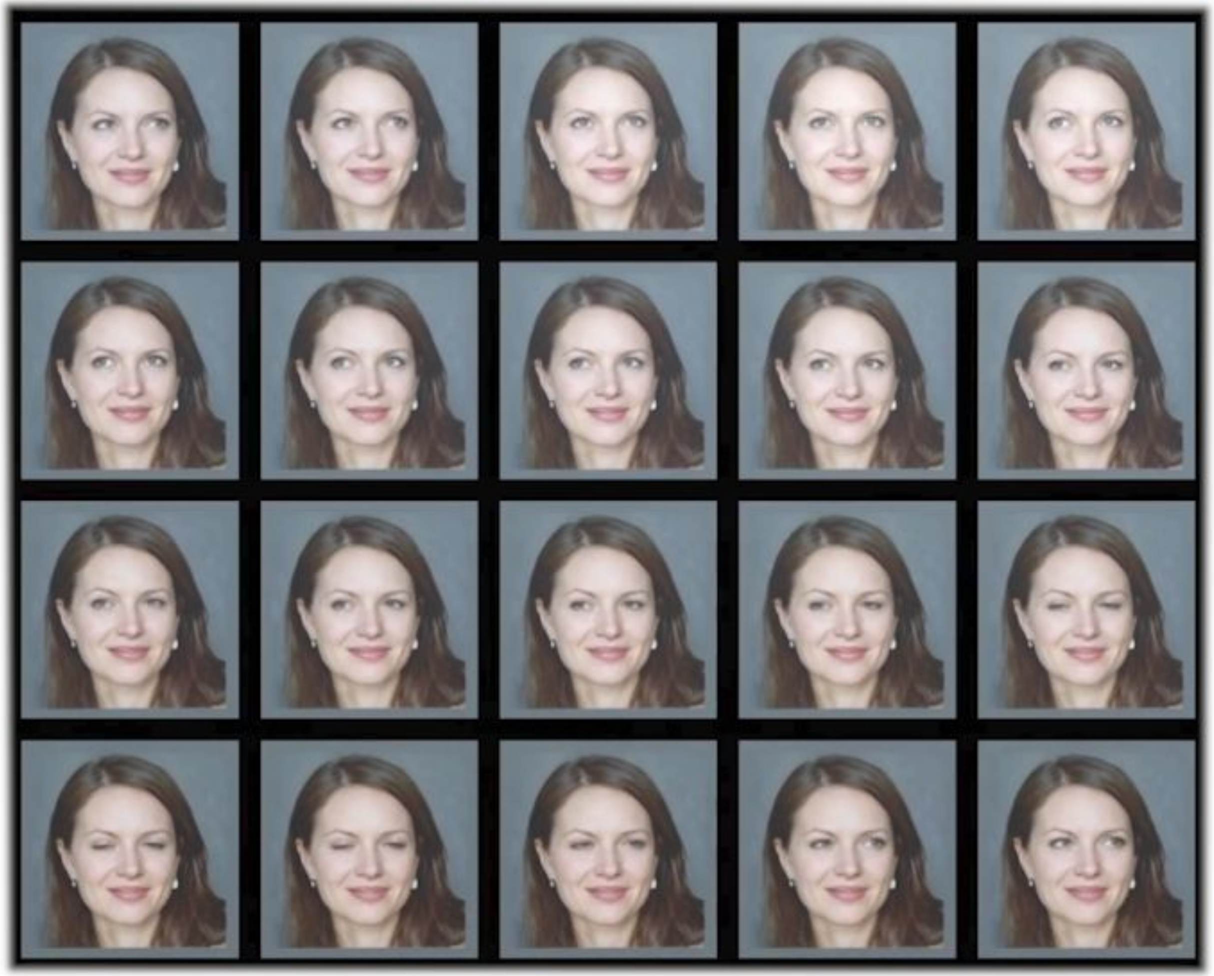
Pipeline
System



Eye Tracking
WebGazer..js

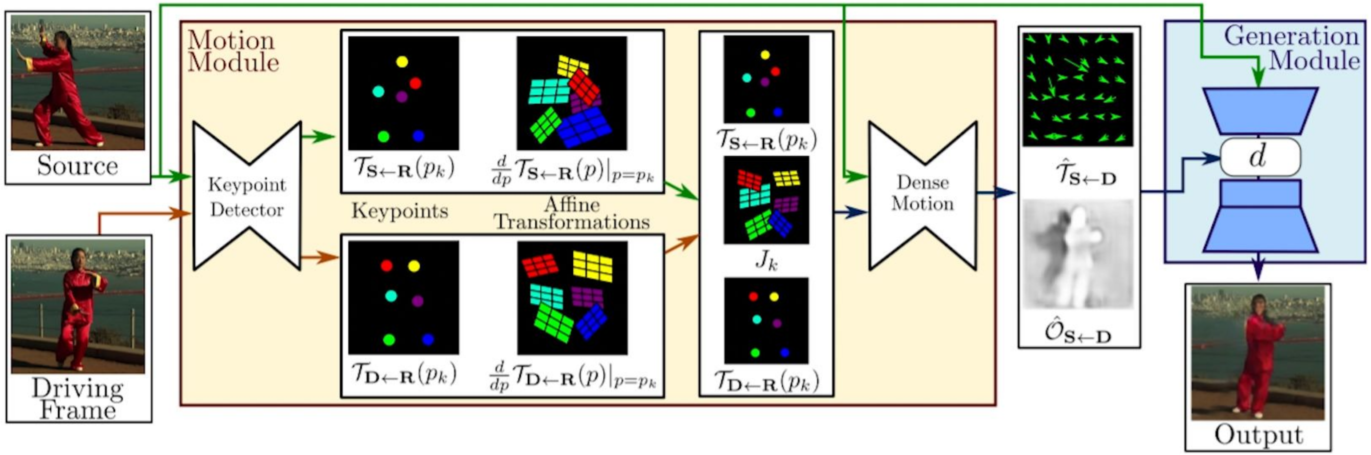


Neural Rendering
Eye movement



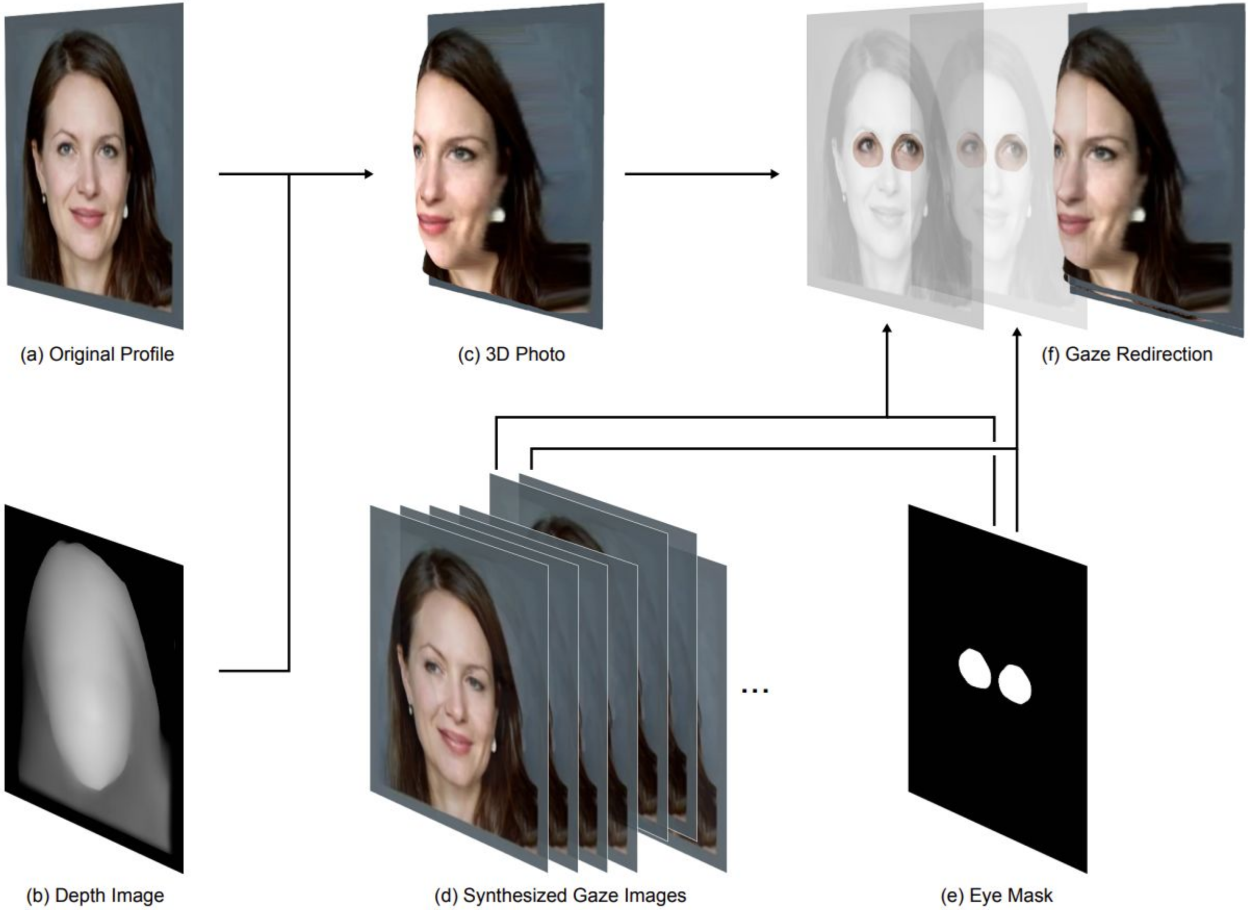
Eye Movement Synthesis

First Order Motion Model



3D Photo Rendering

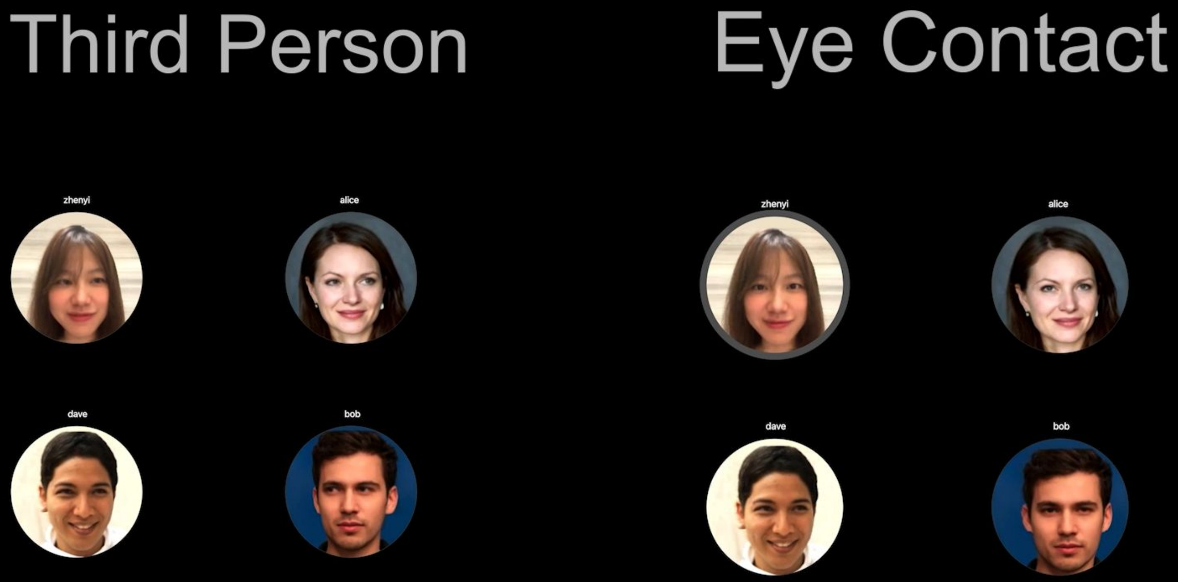
3D photos



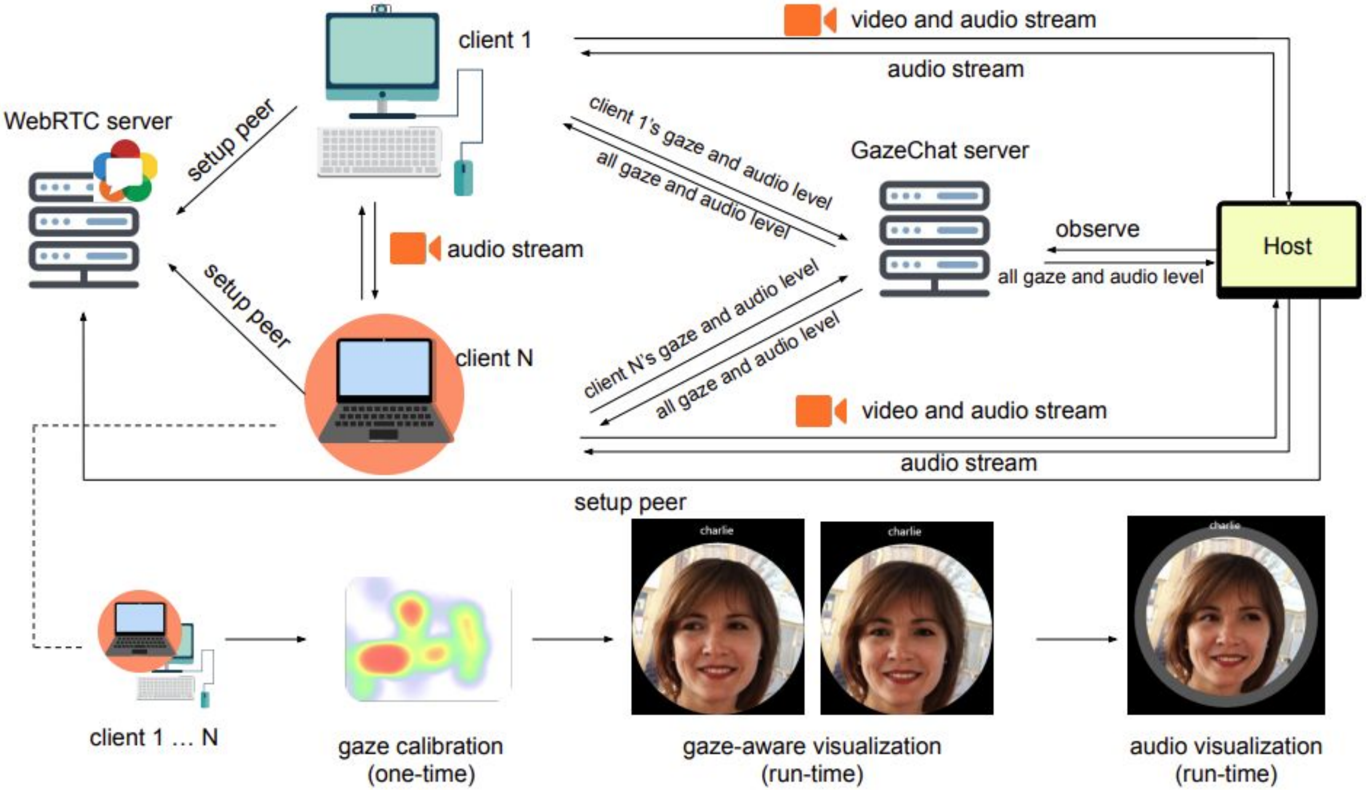
3D Photo Rendering
3D photos

Layouts

UI


Third Person
zhenyi
alice
dave
bob
Eye Contact
zhenyi
alice
dave
bob

Networking
WebRTC



CollaboVR: A Reconfigurable Framework for Creative Collaboration in Virtual Reality

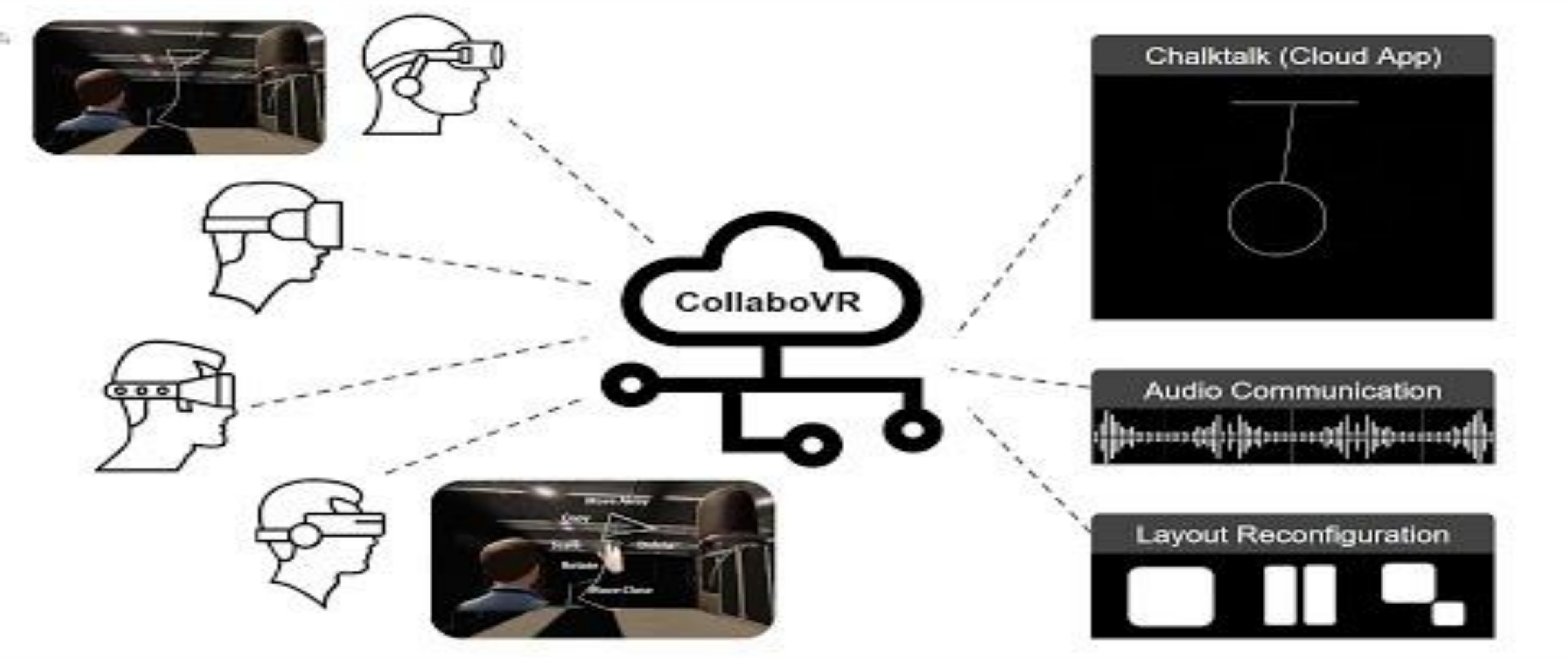
Zhenyi He[*]

Ruofei Du[†]

Ken Perlin[*]

[*]Future Reality Lab, New York University [†]Google LLC

NYU FUTURE REALITY LAB Google ISMAR2020
IEEE International Symposium on Mixed and Augmented Reality


Chalktalk (Cloud App)
CollaboVR
Audio Communication
Layout Reconfiguration

SketchyScene: Richly-Annotated Scene Sketches

Changqing Zou, Qian Yu, Ruofei Du, Haoran Mo, Yi-Zhe Song, Tao Xiang, Chengying Gao, Baoquan Chen, and Hao Zhang (ECCV 2022)

Fig. 1. A scene sketch from our dataset SketchyScene that is user-generated based on the reference image shown, a segmentation result (middle) obtained by a method trained on SketchyScene, and a typical application: sketch captioning.

Language-based Colorization of Scene Sketches

Changqing Zou, Haoran Mo, Chengying Gao, Ruofei Du, and Hongbo Fu (ACM Transaction on Graphics, SIGGRAPH Asia 2019)

Scene sketch

*"the car is **red** with **black** windows"*

*"the road is **black**" / "colorize the road with **black**" / "**black** road"*

*"all the trees are **green**"*
*"the sun in the sky is **yellow**"*
......
*"the grasses are **dark green**"*

*"the sky is **blue** and the ground is **green**"*

ProtoSound: A Personalized and Scalable Sound Recognition System for Deaf and Hard-of-Hearing Users

ACM CHI 2012 · Dhruv Jain, Khoa Nguyen, Steven Goodman, Rachel Grossman-Kahn, Hung Ngo, Aditya Kusupati, Ruofei Du, Alex Olwal, Leah Findlater, and Jon Froehlich

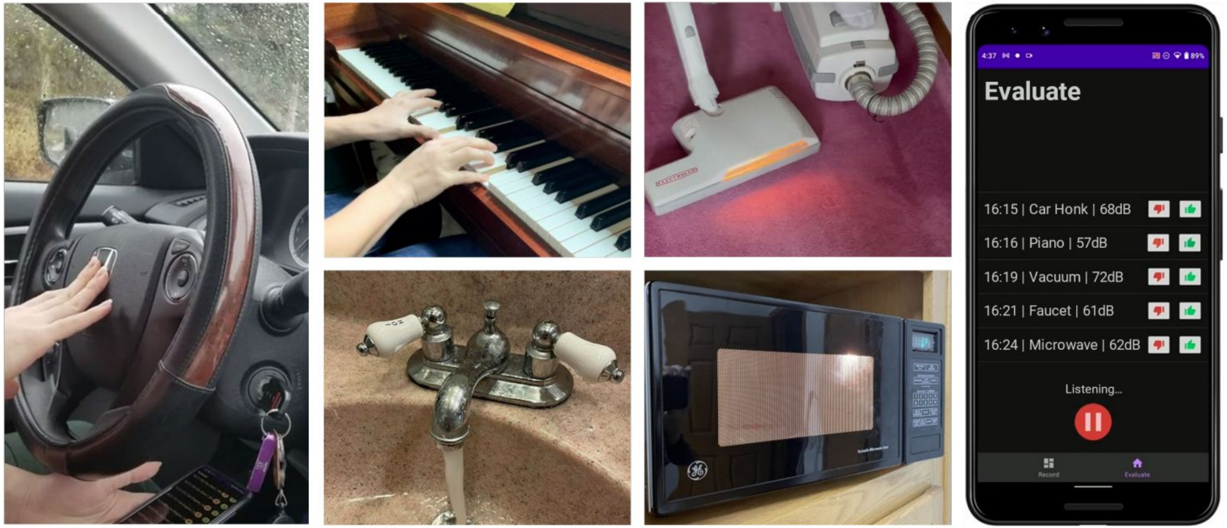


Figure 1: ProtoSound is a technique to customize a sound recognition model using very few recordings, enabling the model to scale across contextual variations of sound (*e.g.,* water flowing on a stainless steel *vs.* a porcelain sink) and support new user-specific sound classes (*e.g.,* a piano). Images show some example sound categories that were trained and recognized during our field evaluation using an experimental mobile app built off ProtoSound. See our supplementary video for details.


15:27:33| Car Door | 60 dB
15:27:48| Car Door | 69 dB
15:27:56| Car Honk | 69 dB
Listening...
Record


Google I/O
2022 Keynote

Wearable Subtitles

Augmenting Spoken Communication with Lightweight Eyewear for All-day Captioning

Figure 1. Our Wearable Subtitles proof-of-concept shows how eyewear could benefit people who are deaf or hard of hearing. We explore hands-free access to spoken communication, situational and speaker awareness, and improved understanding while engaged in a primary task. Our lightweight (54 g) 3D-printed eyewear prototype augments the user's perception of speech and sounds in a socially acceptable form factor with an architecture that could enable up to 15 hours of continuous transcription.

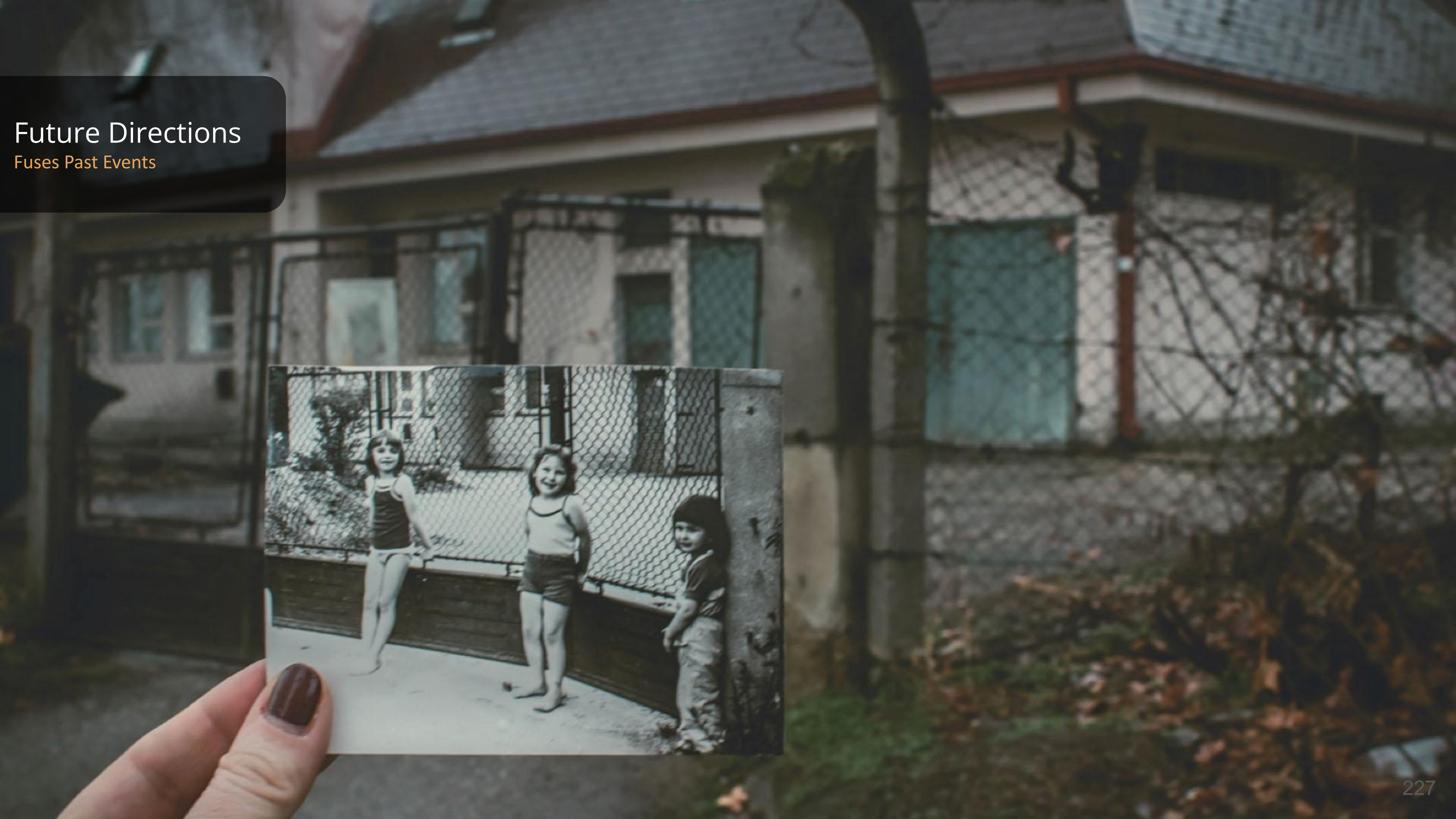
Google

Future Directions

The Ultimate XR Platform

Future Directions

Fuses Past Events



Future Directions

With the present

Future Directions

And look into the future

Future Directions

Change the way we communicate in 3D and consume the information

Future Directions

Consume the information throughout the world

元宇宙中的交互计算与包容普惠

Computational Interaction for a Universally Accessible Metaverse

Ruofei Du (杜若飞)
Senior Research Scientist
Google Labs, San Francisco
www.ruofeidu.com
me@duruofei.com