Reflections on Image-Based Rendering

Richard Szeliski

The University of Washington

TUM AI Guest Lecture Series

January 28, 2021

Reflections on [25 years of] Image-Based Rendering

Richard Szeliski
The University of Washington

TUM AI Guest Lecture Series

January 28, 2021

CVPR 2020 Tutorial on

Novel View Synthesis: From Depth-Based Warping to Multi-Plane Images and Beyond

Novel view synthesis is a long-standing problem at the intersection of computer graphics and computer vision. Seminal work in this field dates back to the 1990s, with early methods proposing to interpolate either between corresponding pixels from the input images, or between rays in space. Recent deep learning methods enabled tremendous improvements to the quality of the results, and brought renewed popularity to the field. The teaser above shows novel view synthesis from different recent methods. *From left to right: Yoon et al. [1], Mildenhall et al. [2], Wiles et al. [3], and Choi et al. [4]. Images and videos courtesy of the respective authors.*

New edition of my book – almost done

Computer Vision: Algorithms and Applications, 2nd ed.

© 2021 Richard Szeliski, Facebook

https://szeliski.org/Book

New edition of my book

Outline

- Multi-view stereo
- Image-Based Rendering
  - Lumigraphs, Light Fields, Sprites with Depth, and Layers
- Virtual Viewpoint Video
- 360° and 3D Video
- 3D Photos
- Reflections and transparency
- Neural rendering

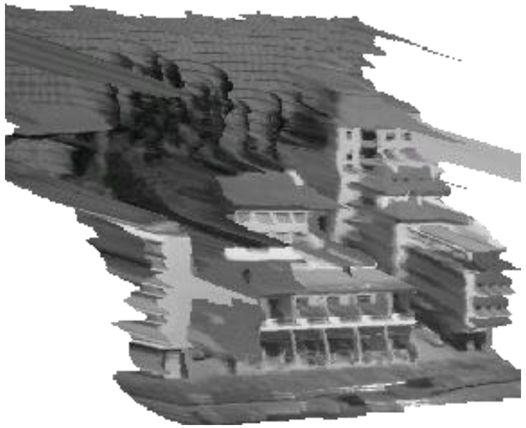
Multi-view Stereo

View Interpolation

- Given two images with correspondences, *morph* (warp and cross-dissolve) between them [Chen & Williams, SIGGRAPH'93]

input

depth image

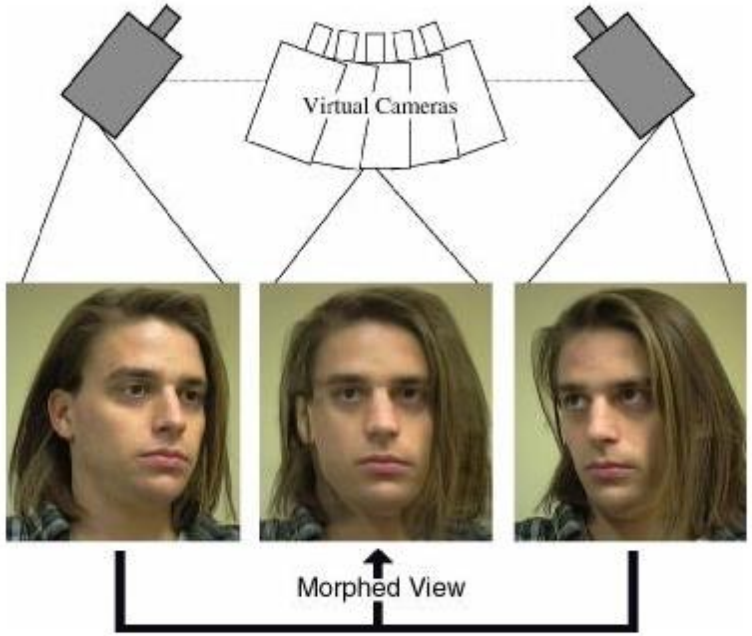


novel view

[Matthies,Szeliski,Kanade'88]

View Morphing

- Morph between pair of images using epipolar geometry [Seitz & Dyer, SIGGRAPH'96]



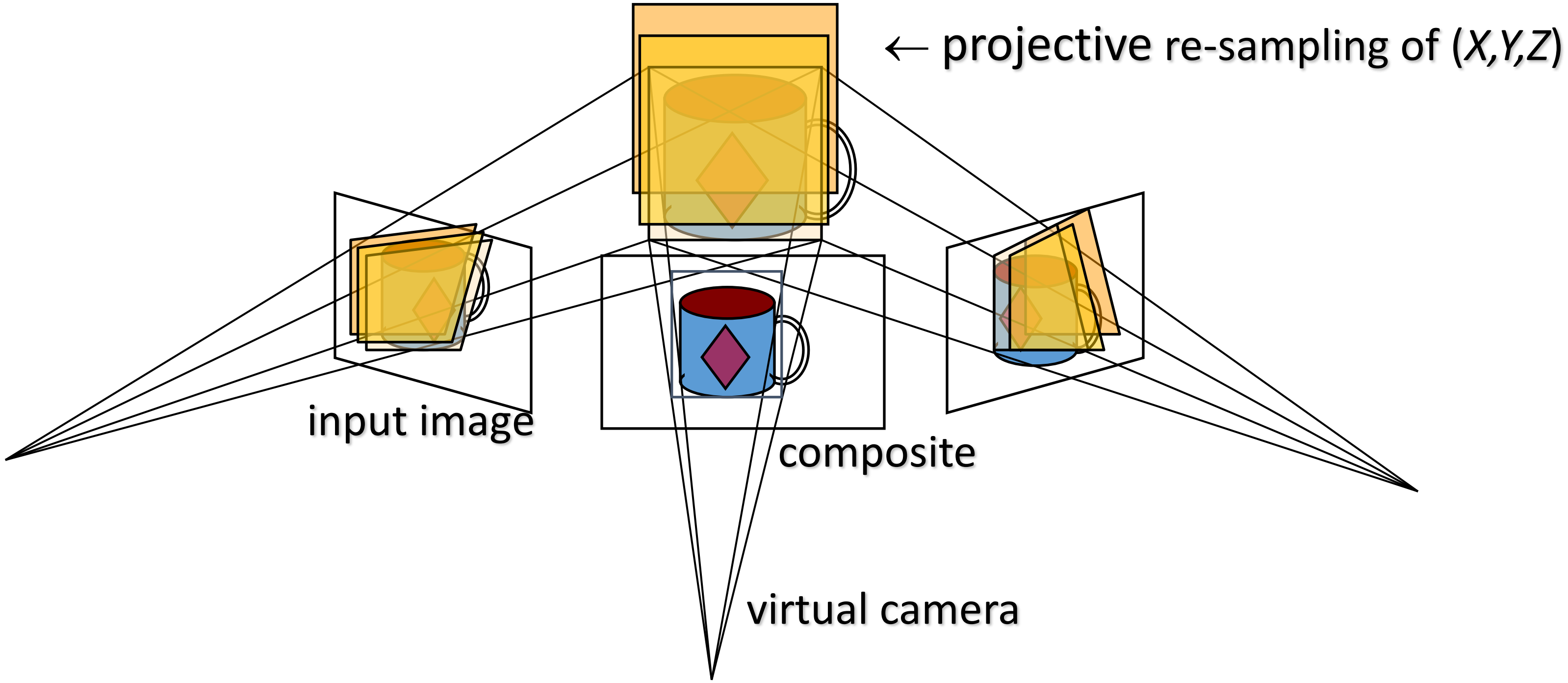
Video view interpolation

Interactive 3D video scenarios

- Sports events, e.g., CMU's 30-camera "EyeVision" system at SuperBowl XXXV) and 2016
- Concert performances, plays, circus acts
- Games
- Instructional video, e.g., golf, skating, martial arts
- Interactive (Internet) video

Plane Sweep Stereo

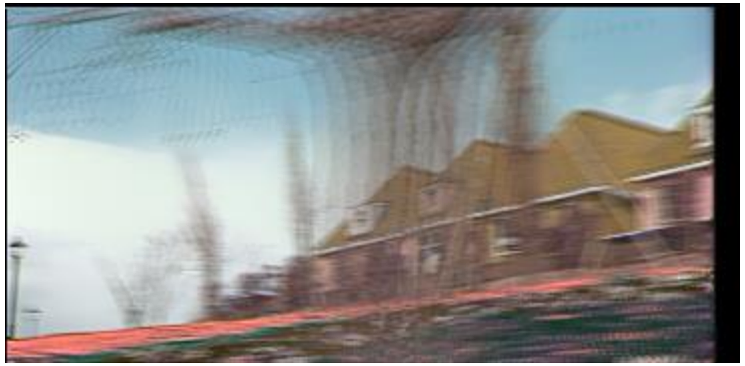
- Sweep family of planes through volume



- each plane defines an image ⇒ composite homography

Plane Sweep Stereo

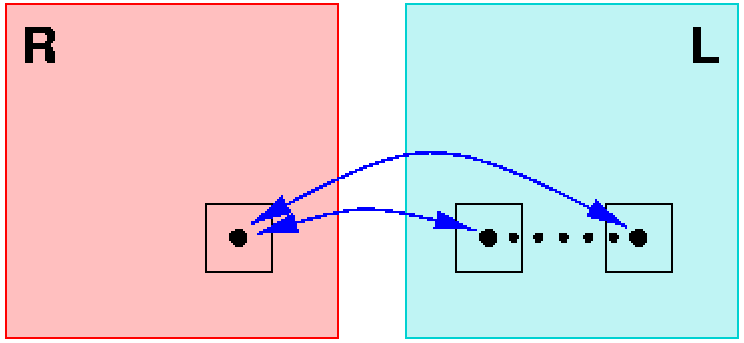
- For each depth plane
  - compute composite (mosaic) image — *mean*

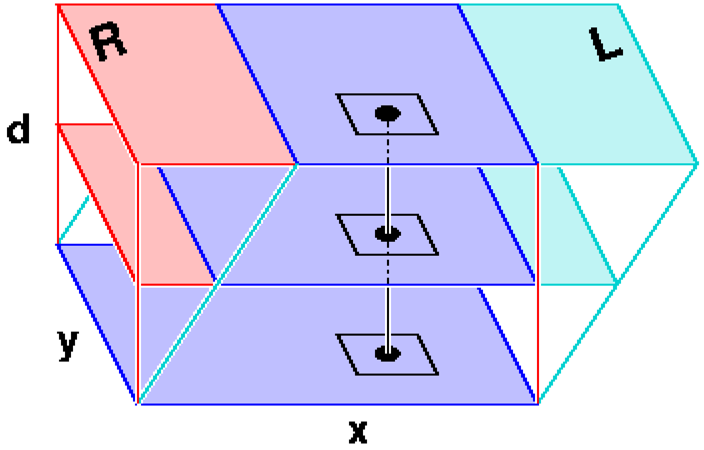


  - compute error image — *variance*
  - convert to confidence and aggregate spatially
- Select winning depth at each pixel

Plane sweep stereo

- Re-order (pixel / disparity) evaluation loops



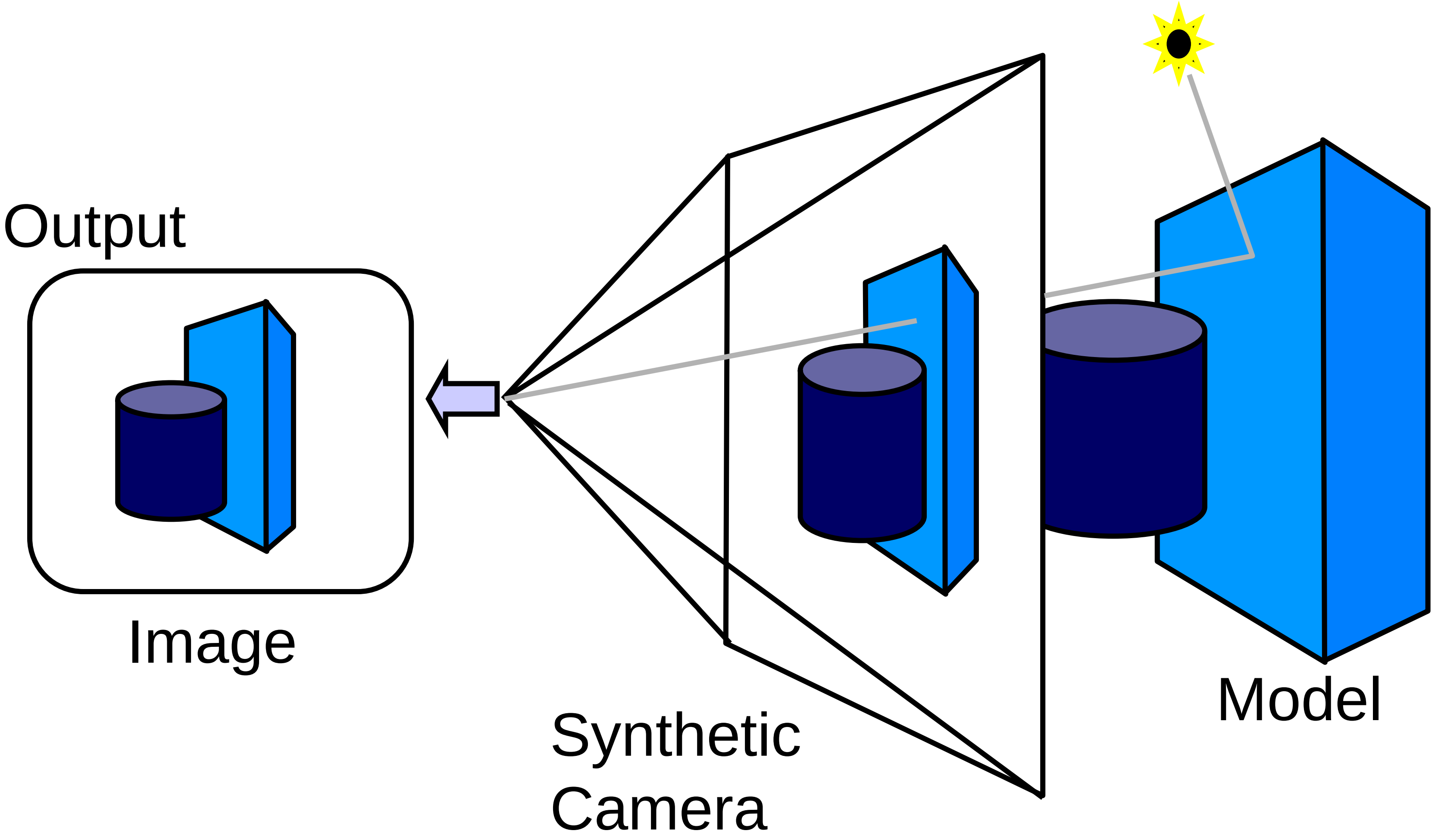


for every pixel,
  for every disparity
    compute cost

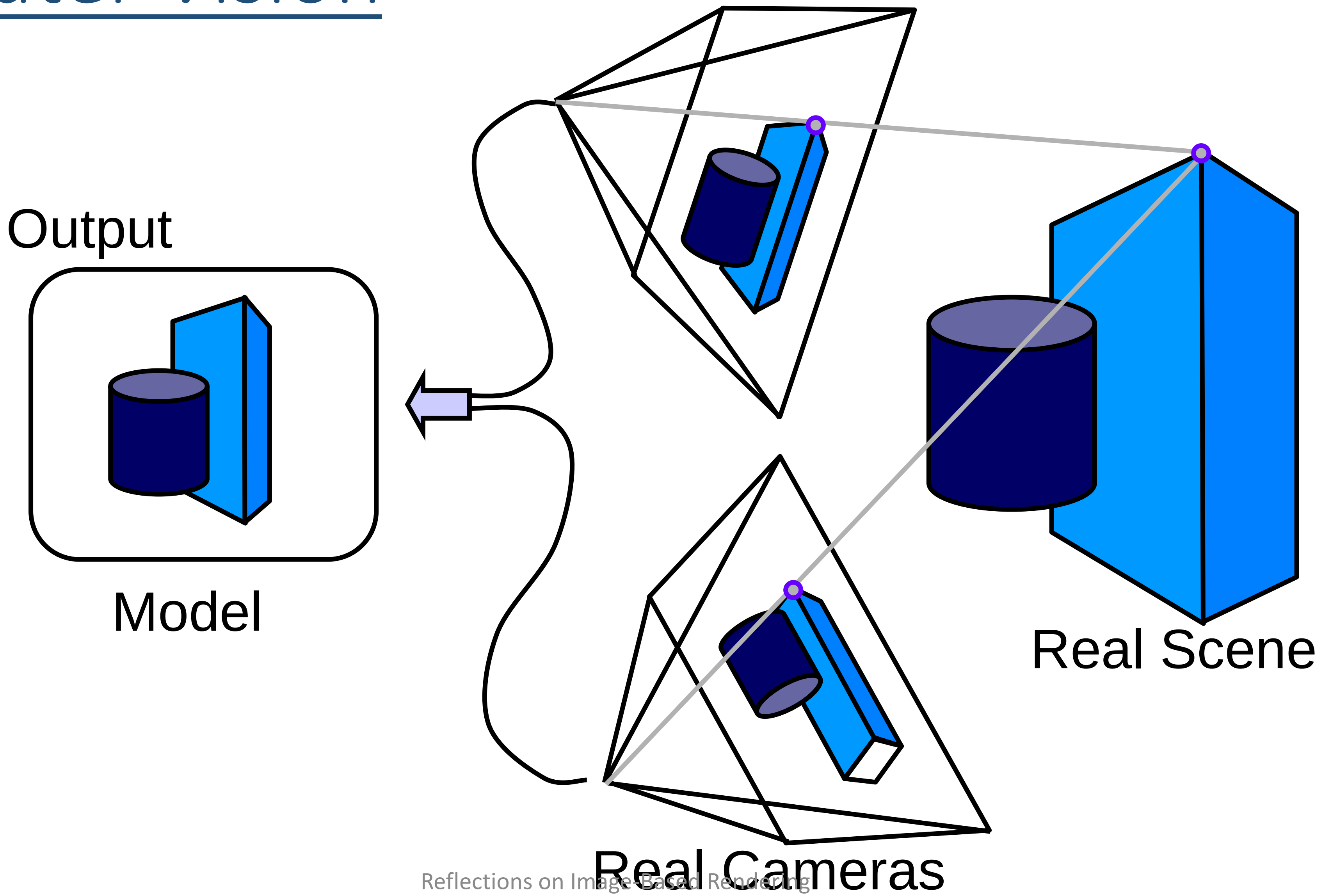
for every disparity
  for every pixel
    compute cost

Image-Based Rendering

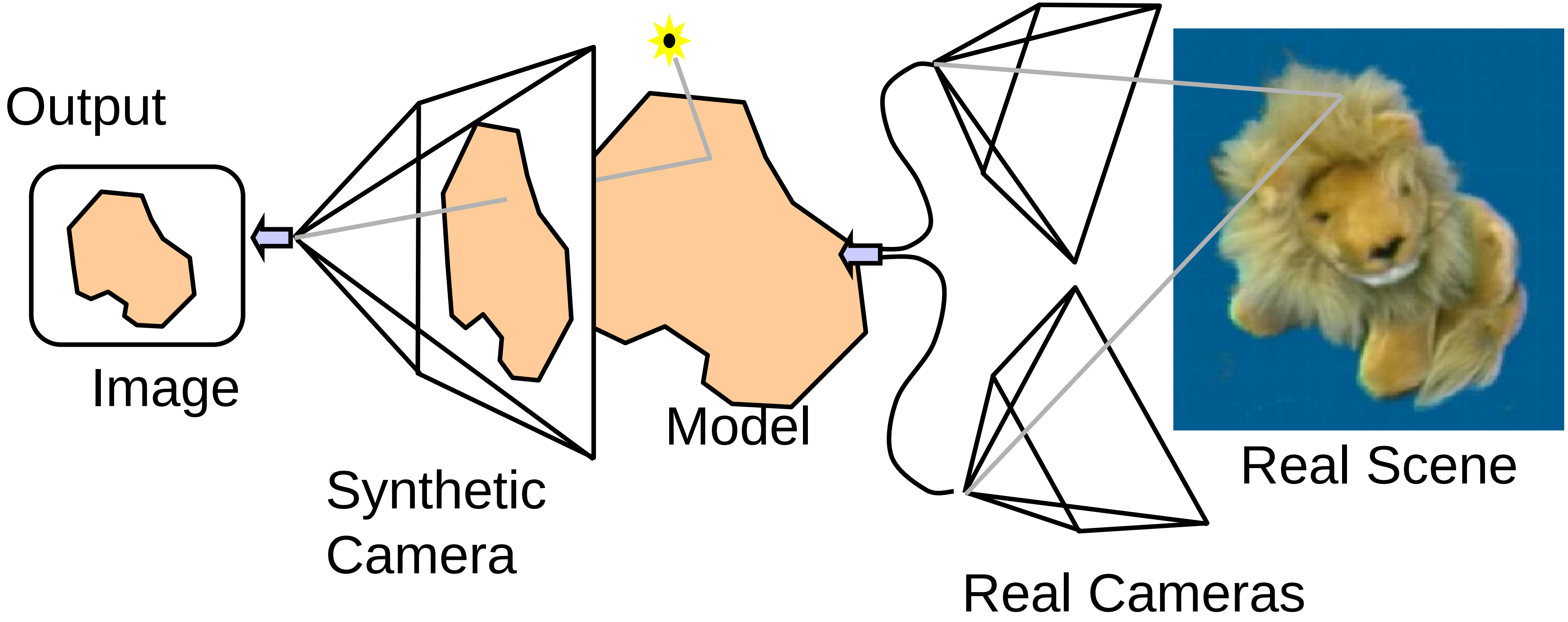
Computer Graphics



Computer Vision



But, vision technology fails



…and so does graphics

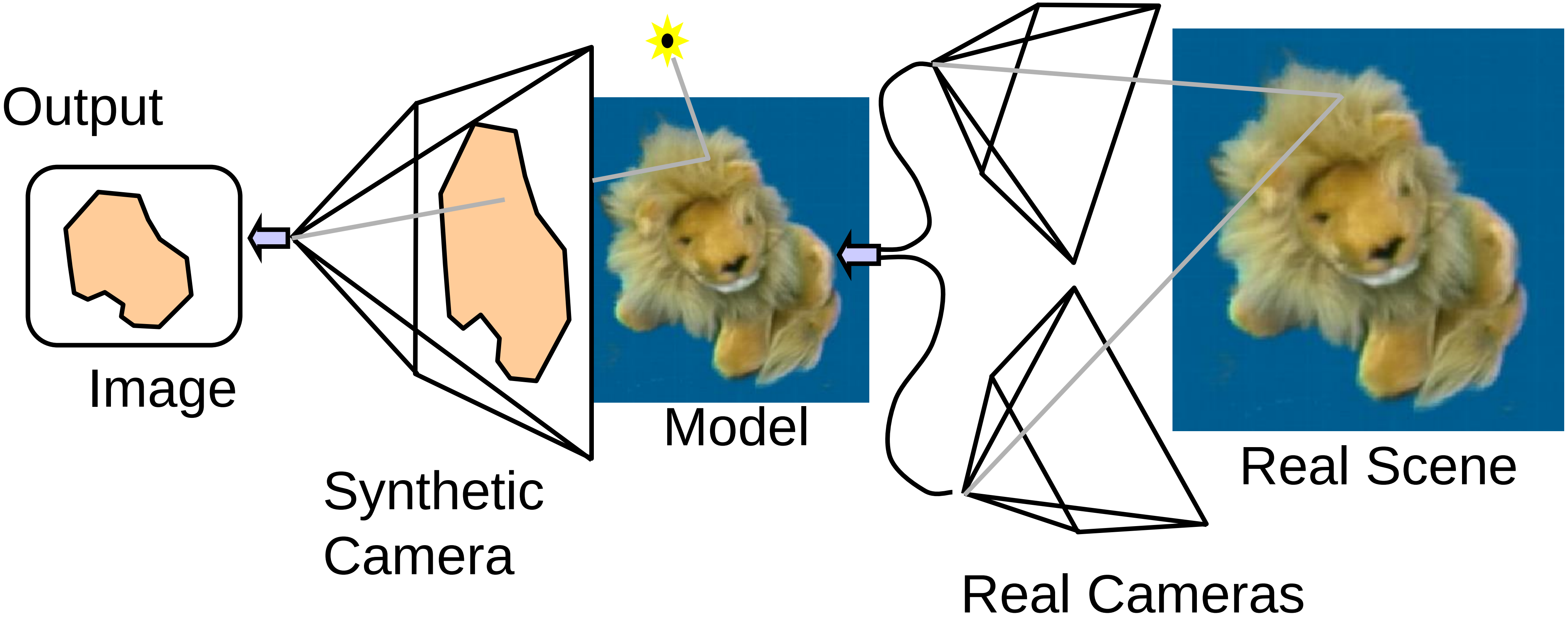
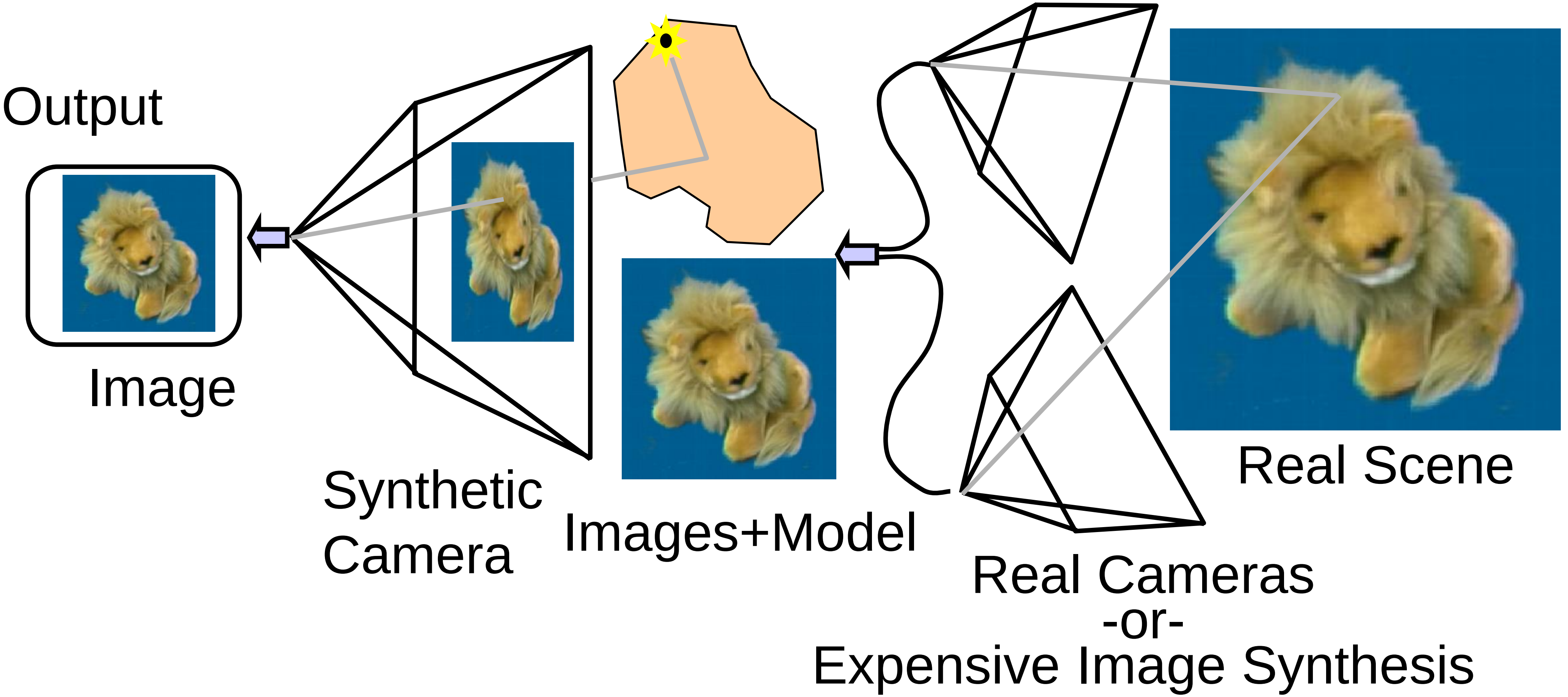
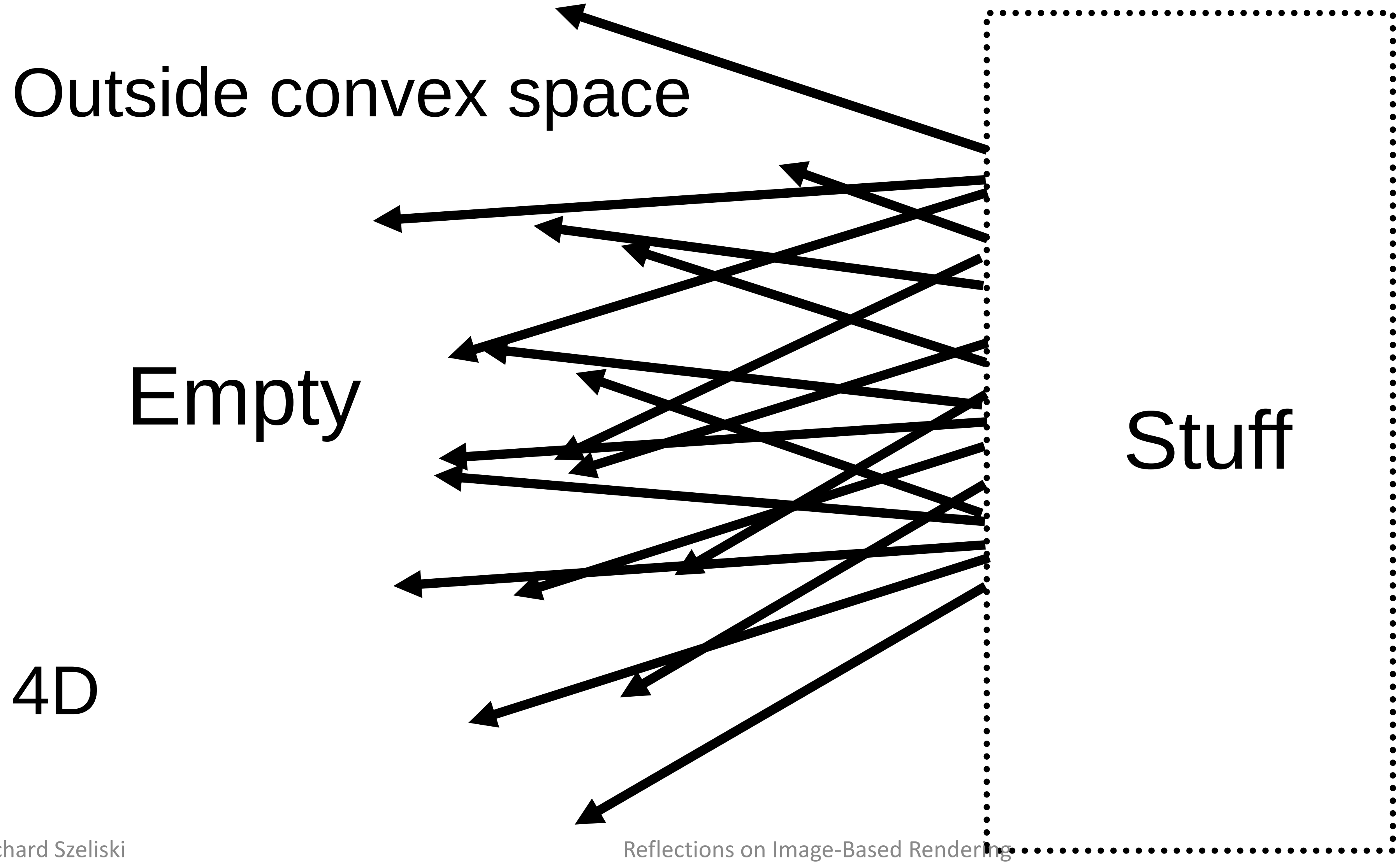


Image-Based Rendering

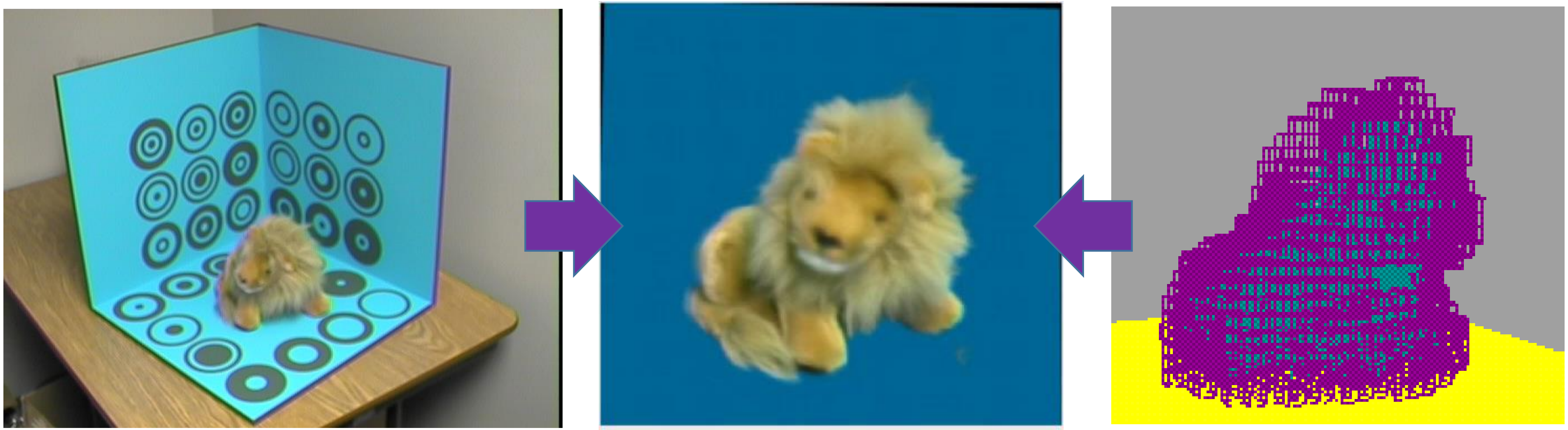


Lumigraph / Light Field [1996]



Lumigraph – Capture

- Convert images into a solid 3D model

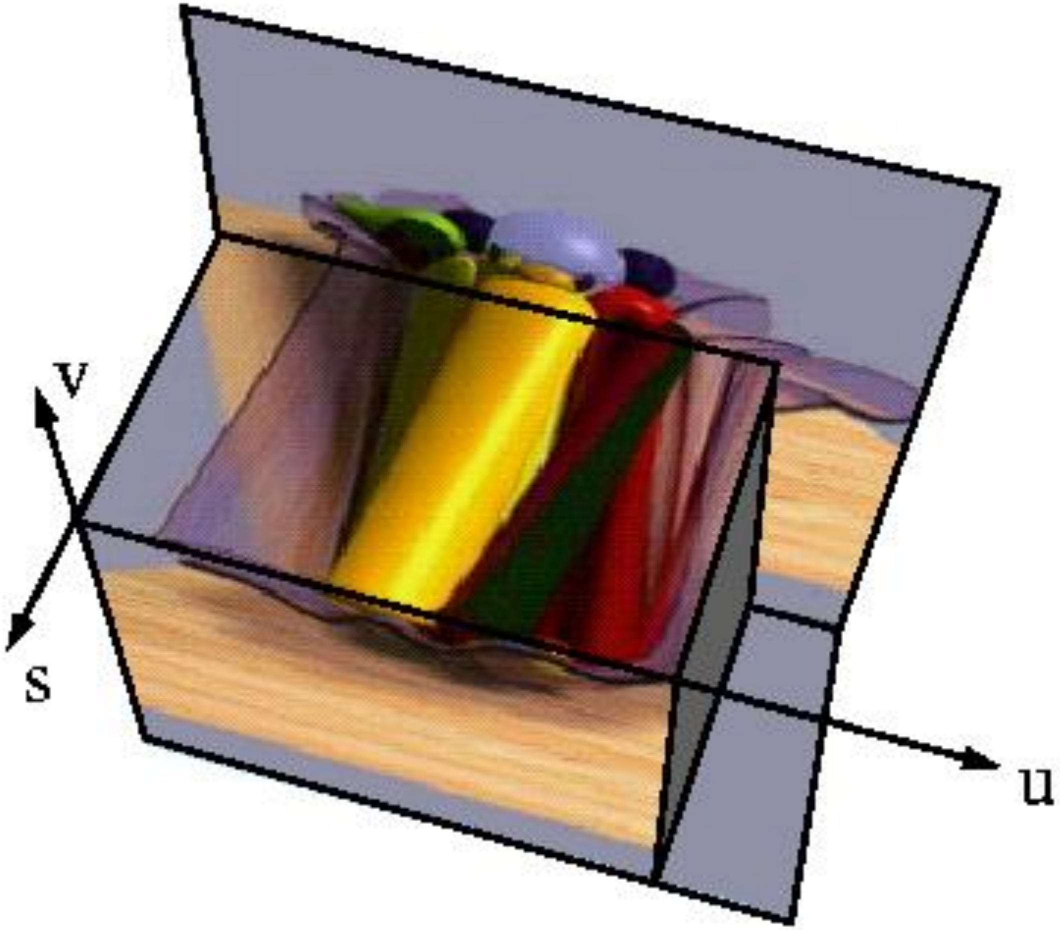


- Render from images and model

Lumigraph – Image Effects

Can model effects such as:

- parallax
- occlusion
- translucency
- refraction
- highlights
- reflections



Unstructured Lumigraph

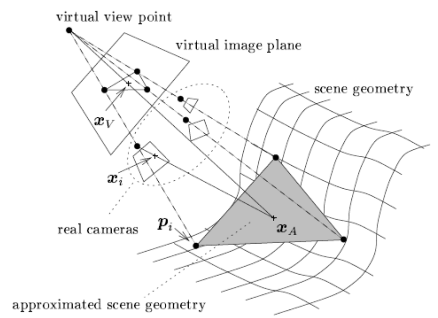
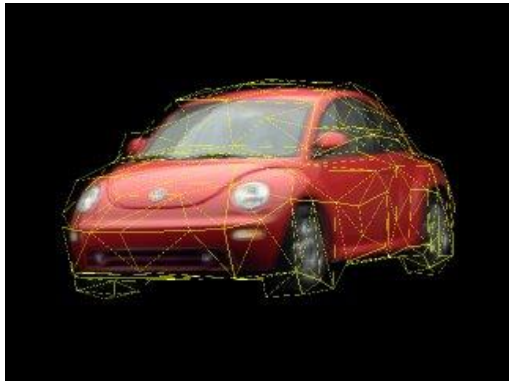
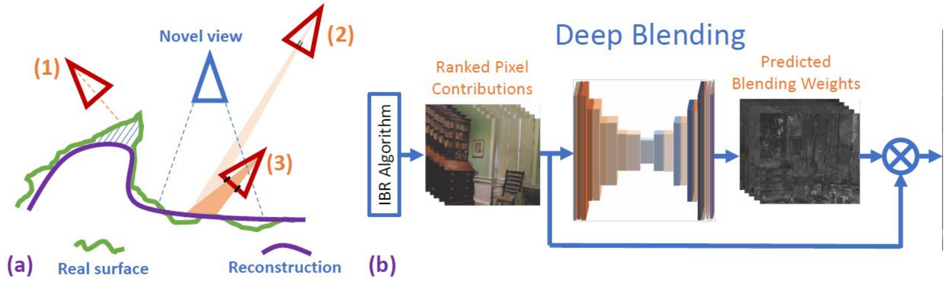


Figure3. Drawing triangles of neighboring projected camera centers and approximating scene geometry by one plane for the whole scene, for one camera triple or by several planes for one camera triple.

- What if the images aren't sampled on a regular 2D grid?
- Can still re-sample rays
- Ray weighting becomes more complex [Heigl *et al.*,DAGM'99]

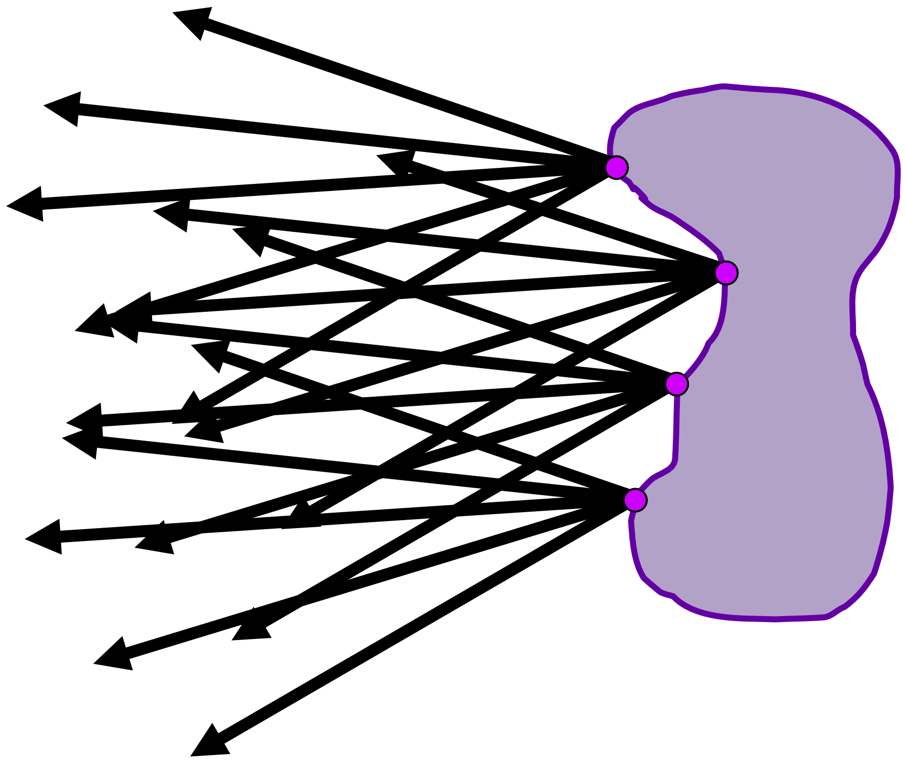


- Unstructured Lumigraph [Buehler *et al.*, SIGGRAPH'2000]
- Deep blending [Hedman *et al.*, SG Asia 2018]
- FVS [Riegler & Koltun, ECCV'2020]



Surface Light Fields

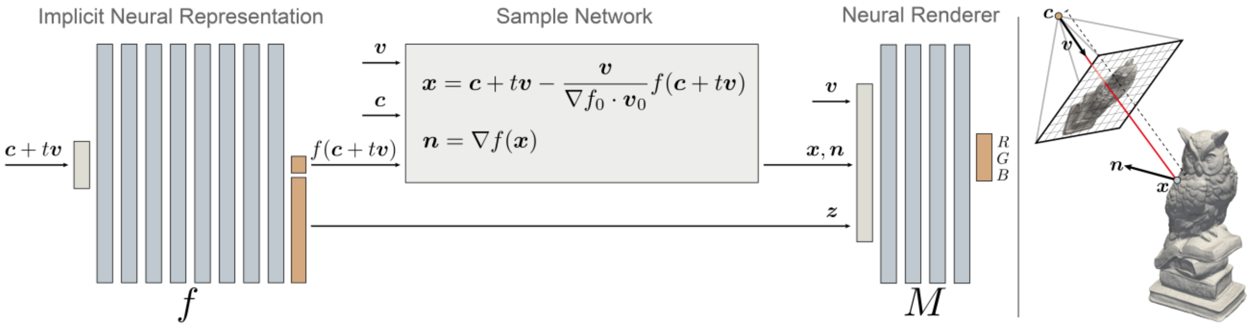
- [Wood et al, SIGGRAPH 2000]
- Turn 4D parameterization around:
  - image @ every surface pt.
- Leverage coherence:
  - compress radiance fn (BRDF * illumination) after rotation by ***n***



Surface Light Fields

- [Wood et al, SIGGRAPH 2000]

- …
- Implicit Differentiable Renderer [Yariv *et al.*, NeurIPS 2020]



Environment Matting [2000]

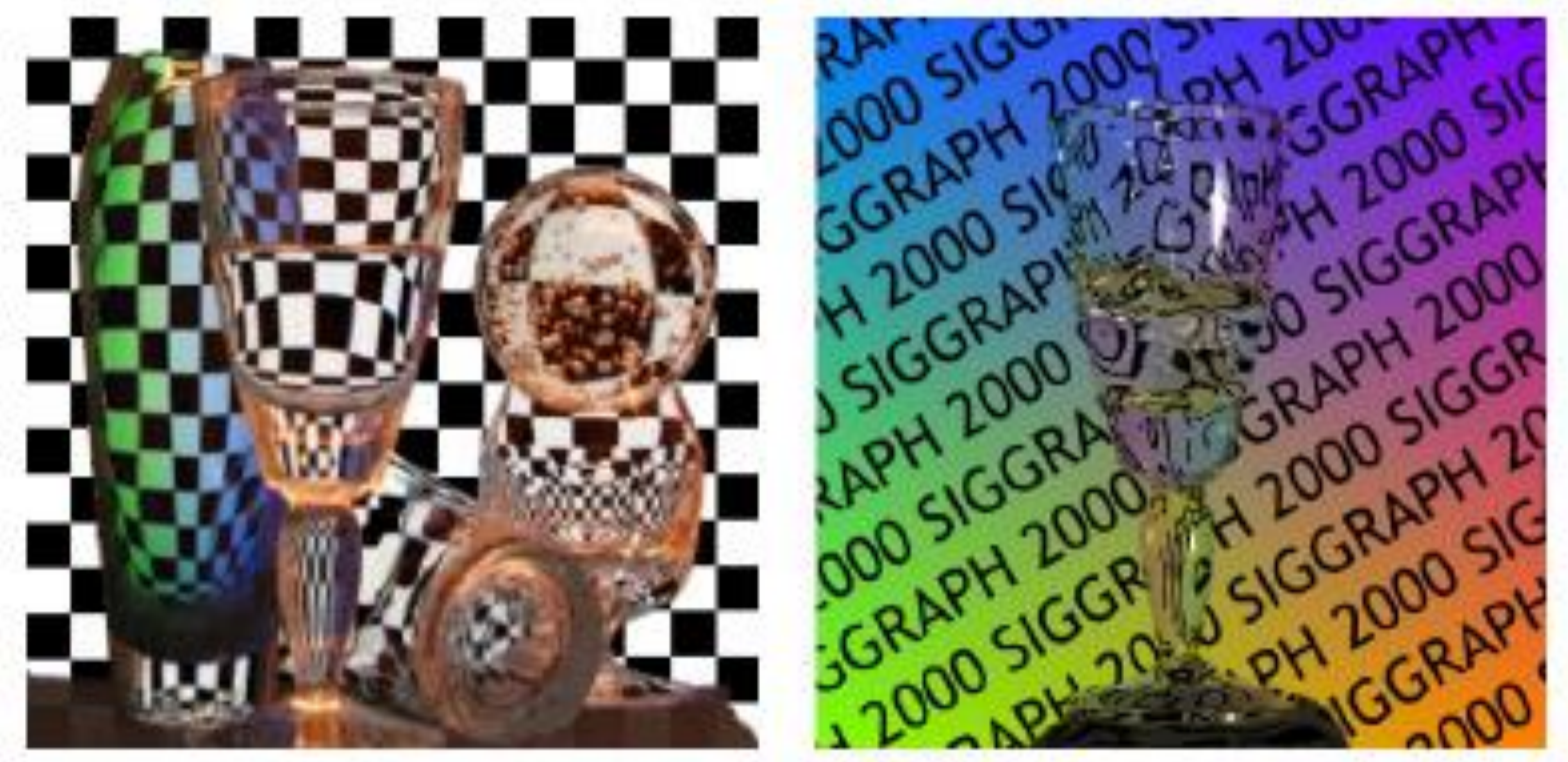
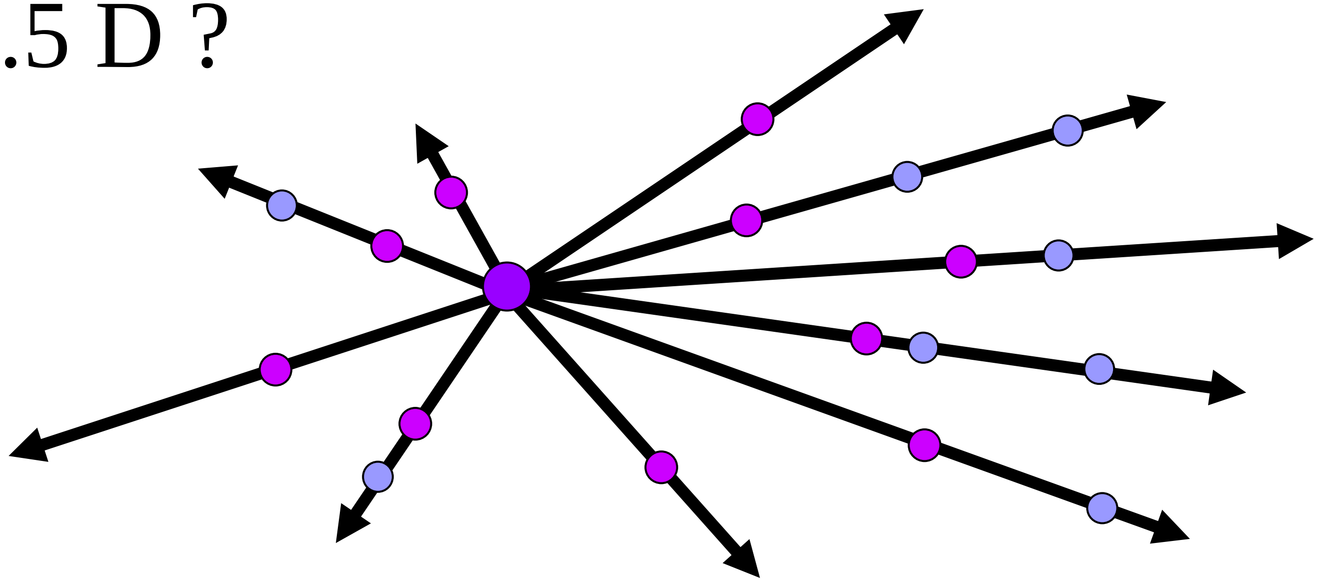


Figure 1 Sample composite images constructed with the techniques of this paper: slow but accurate on the left, and a more restricted example acquired at video rates on the right.

Layered Depth Image

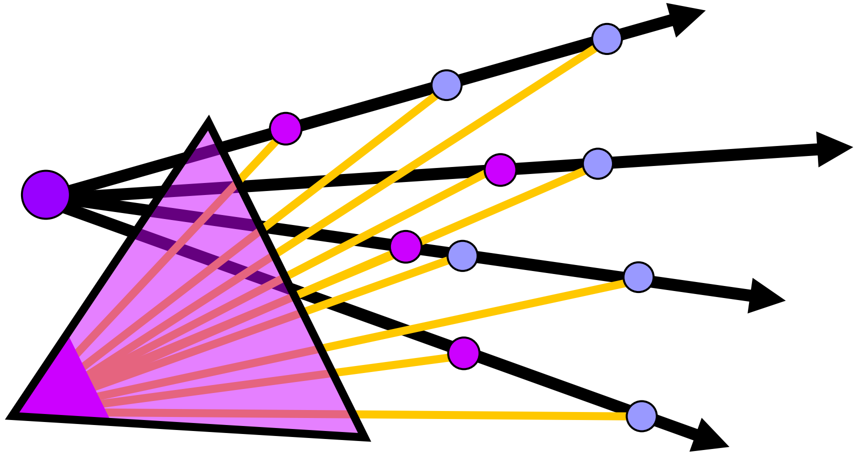
2.5 D ?



Layered Depth Image

Layered Depth Image

- Rendering from LDI
  [Shade et al., SIGGRAPH'98]



- Incremental in LDI X and Y
- Guaranteed to be in back-to-front order

Sprites with Depth

- Represent scene as collection of cutouts with depth (planes + parallax)
- Render back to front with fwd/inverse warping [Shade *et al.*, SIGGRAPH’98]
- Basis of Virtual Viewpoint Video [Zitnick *et al.* 2004]

Multiplane images

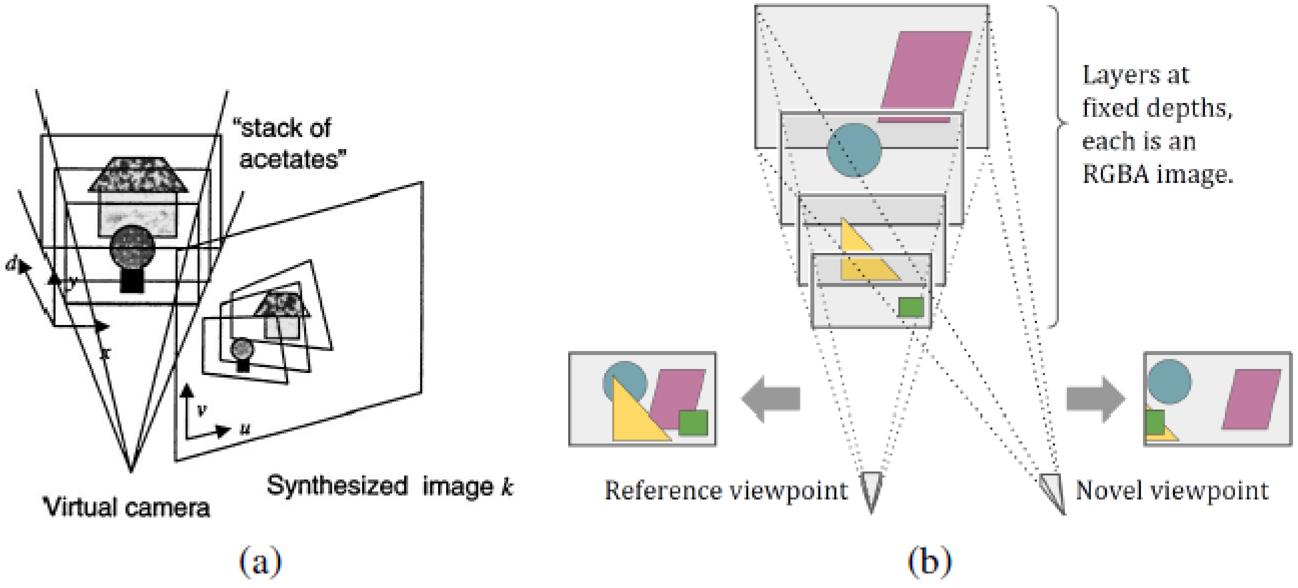


Figure 14.7 *Finely sliced fronto-parallel layers: (a) stack of acetates (Szeliski and Golland 1999) © 1999 Springer and (b) multiplane images (Zhou, Tucker, Flynn* et al. *2018) © 2018 ACM.*

Multiplane images

Multi-sphere and layered meshes

Immersive Light Field Video with a Layered Mesh Representation

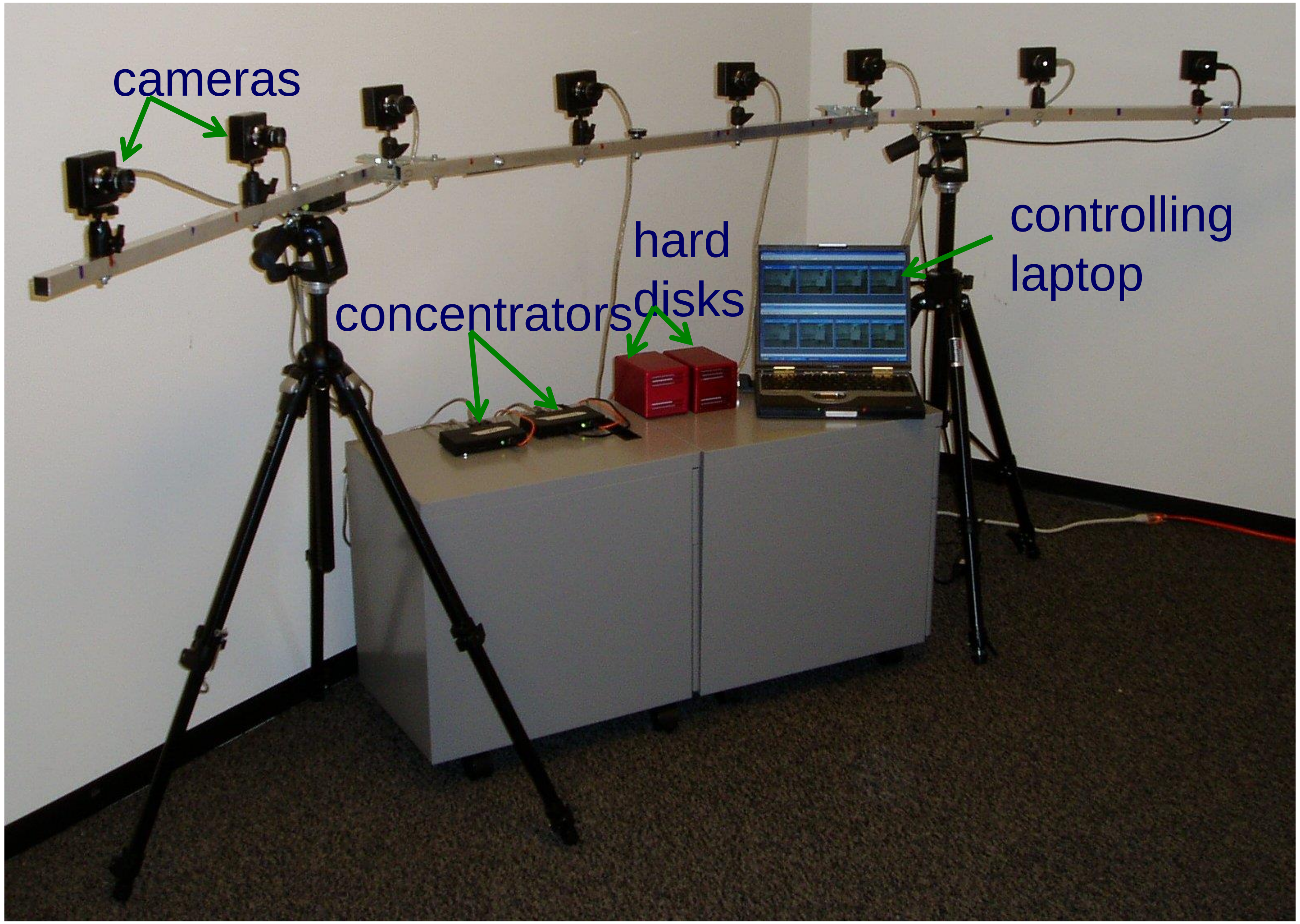
MICHAEL BROXTON*, JOHN FLYNN*, RYAN OVERBECK*, DANIEL ERICKSON*, PETER HEDMAN, MATTHEW DUVALL, JASON DOURGARIAN, JAY BUSCH, MATT WHALEN, and PAUL DEBEVEC, Google

(a) Capture Rig (b) Multi-Sphere Image (c) Layered Mesh Representation

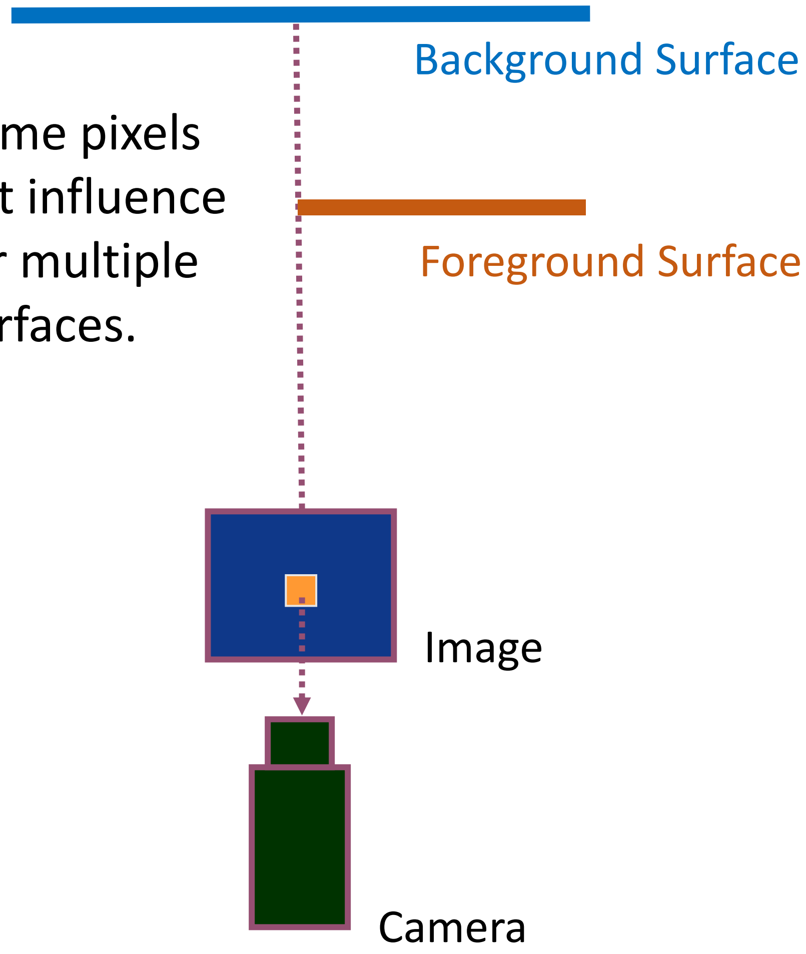
[SIGGRAPH'2020]

Virtual Viewpoint Video

Virtual Viewpoint Video [SIGGRAPH 2004]



Matting



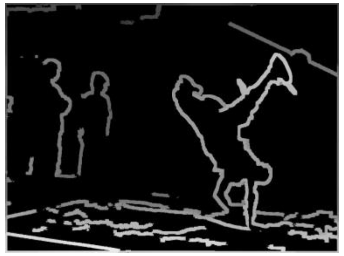
Some pixels get influence for multiple surfaces.

Close up of real image:

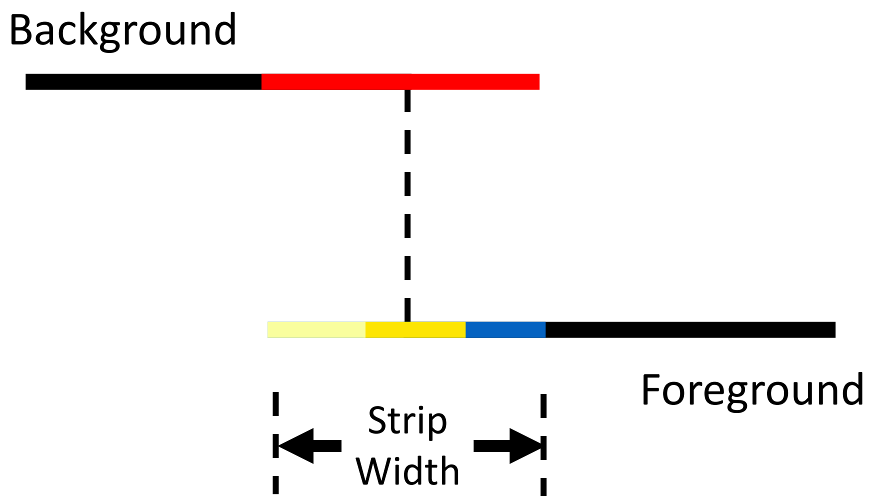
Multiple colors and depths at boundary pixels…

Find matting information:

1. Find boundary strips using depth.



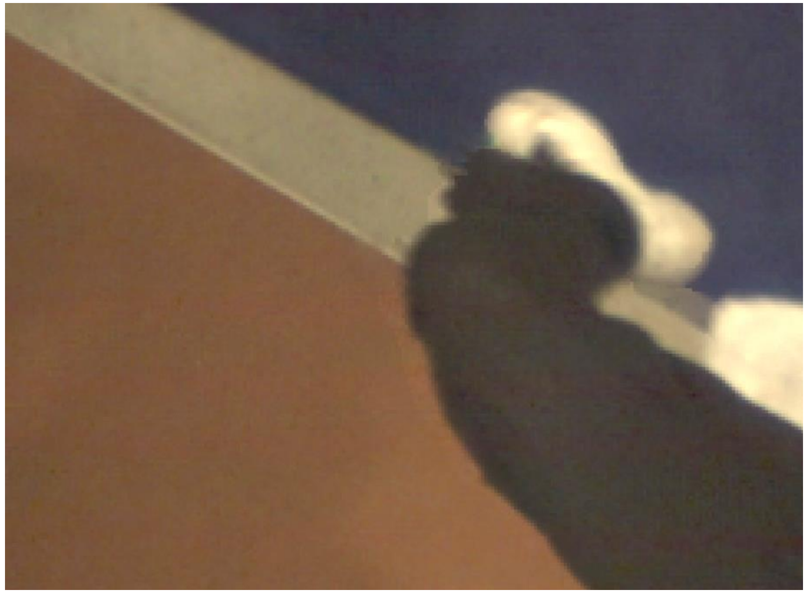
2. Within boundary strips compute the colors and depths of the foreground and background object.



Why matting is important

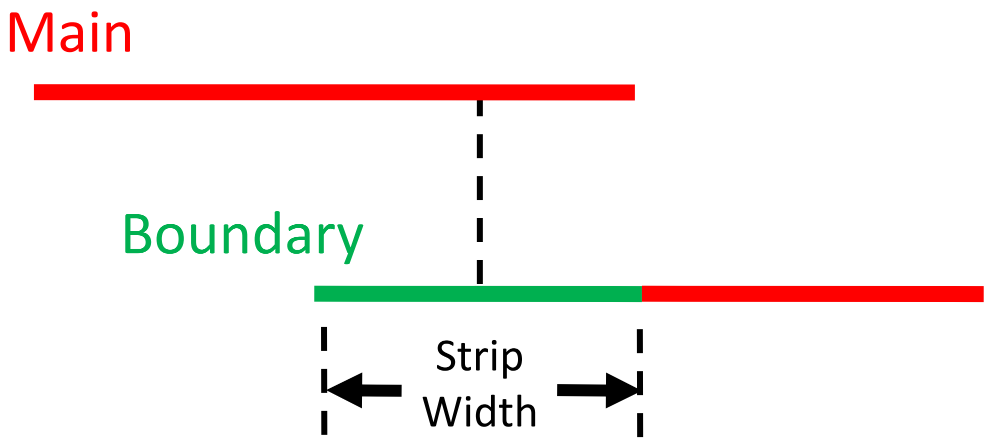
No Matting

Matting

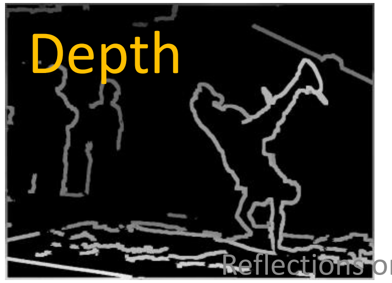


Virtual Viewpoint Video

Two-layer model with thin boundary strips [Zitnick *et al.*, SIGGRAPH'04]



Main Layer: Boundary Layer:

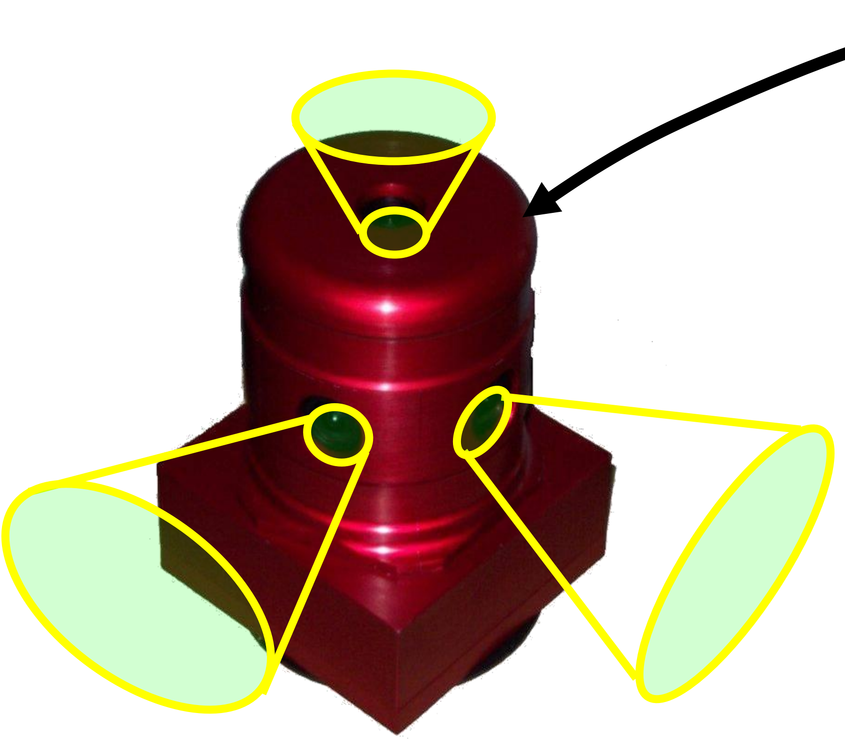



Massive Arabesque

360° Video

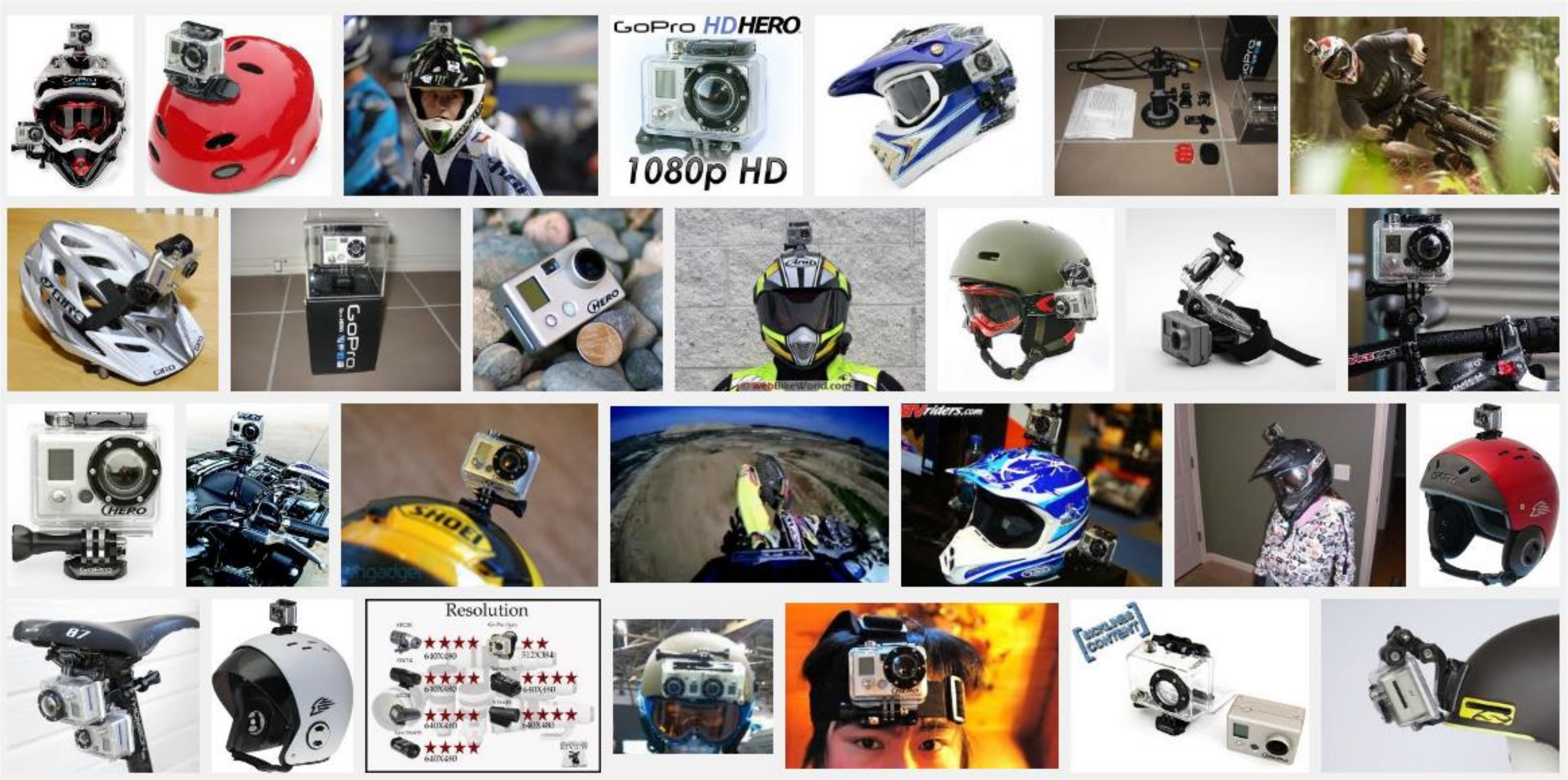
360 Video

[Uyttendaele et al. 2004]



Ladybug (six-camera head)

Acquisition platforms (today)



360 Video


ULTRACAM
mobile
K 980 CW

ULTRACAM
mobile

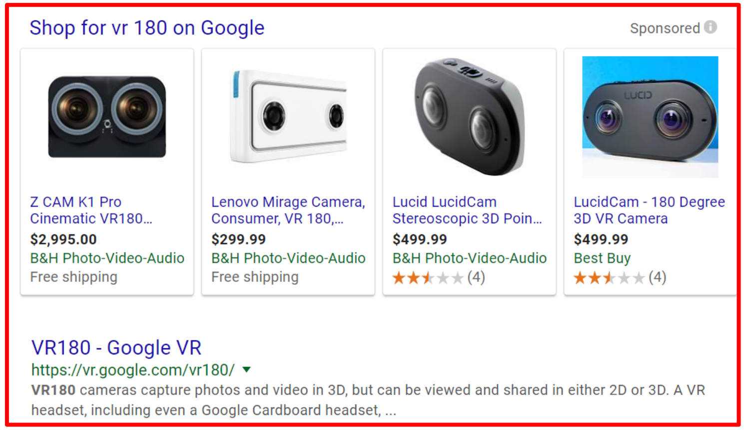
K 980 CW
www.europcar.at

360 Video


RICOH


kolor | GoPro
360°


ulus
Oculus
Development Kit


Shop for vr 180 on Google
Sponsored
Z CAM K1 Pro Cinematic VR180...
$2,995.00
B&H Photo-Video-Audio
Free shipping
Lenovo Mirage Camera, Consumer, VR 180,...
$299.99
B&H Photo-Video-Audio
Free shipping
Lucid LucidCam Stereoscopic 3D Poin...
$499.99
B&H Photo-Video-Audio
(4)
LucidCam - 180 Degree 3D VR Camera
$499.99
Best Buy
(4)
VR180 - Google VR
https://vr.google.com/vr180/
VR180 cameras capture photos and video in 3D, but can be viewed and shared in either 2D or 3D. A VR headset, including even a Google Cardboard headset, ...


LG
RICOH
VIDEO
HERO4
$200
$1,000
Insta360 Nano
SAMSUNG
bubl

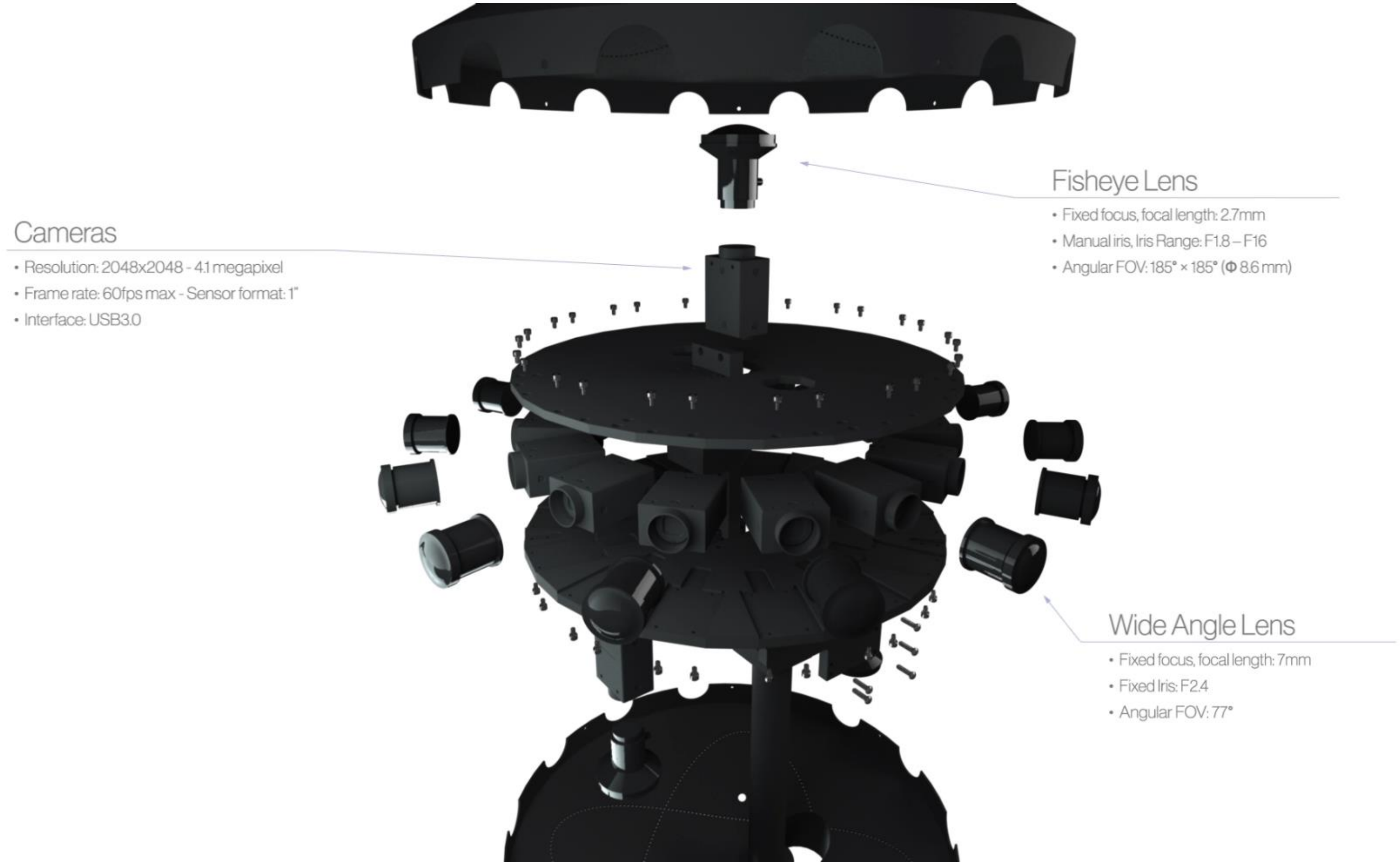
Google Jump [2015]

ODYSSEY

+

JUMP

Facebook Surround 360 [2016]



Facebook Surround 360 [2017]

Facebook’s new Surround 360 video cameras let you move around inside live-action scenes

The freedom of VR with the fidelity of real life

By Nick Statt | @nickstatt | Apr 19, 2017, 1:15pm EDT

Facebook today announced the second generation of its Surround 360 video camera design, and this time the company is serious about helping potential customers purchase it as an actual product. The Surround 360, which Facebook unveiled last year as an open-source spec guide for others to build off of, has been upgraded as both a larger, more capable unit and a smaller, more portable version.

An Integrated 6DoF Video Camera and System Design

ALBERT PARRA POZO, MICHAEL TOKSVIG, TERRY FILIBA SCHRAGER, and JOYCE HSU, Facebook Inc.
UDAY MATHUR, RED Digital Cinema
ALEXANDER SORKINE-HORNUNG, RICK SZELISKI, and BRIAN CABRAL, Facebook Inc.

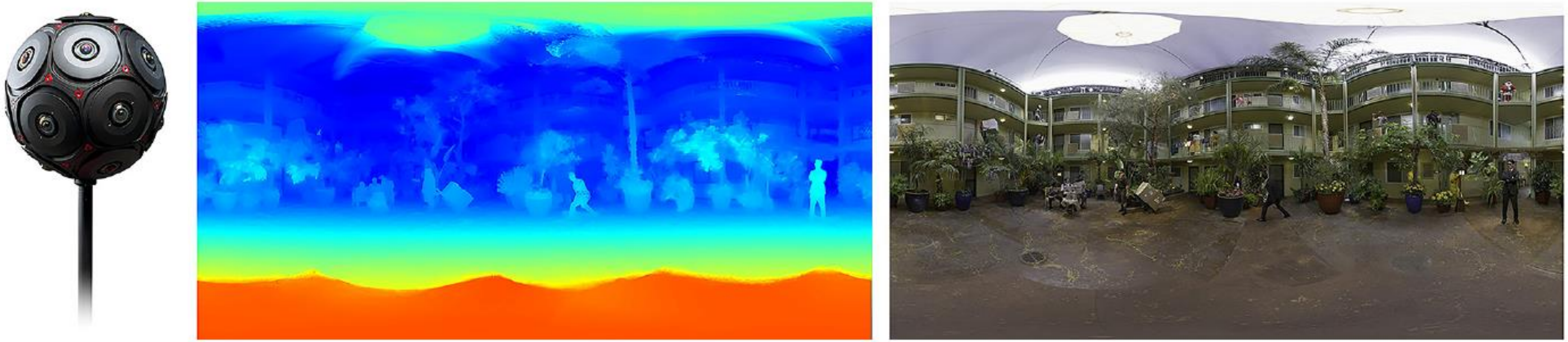


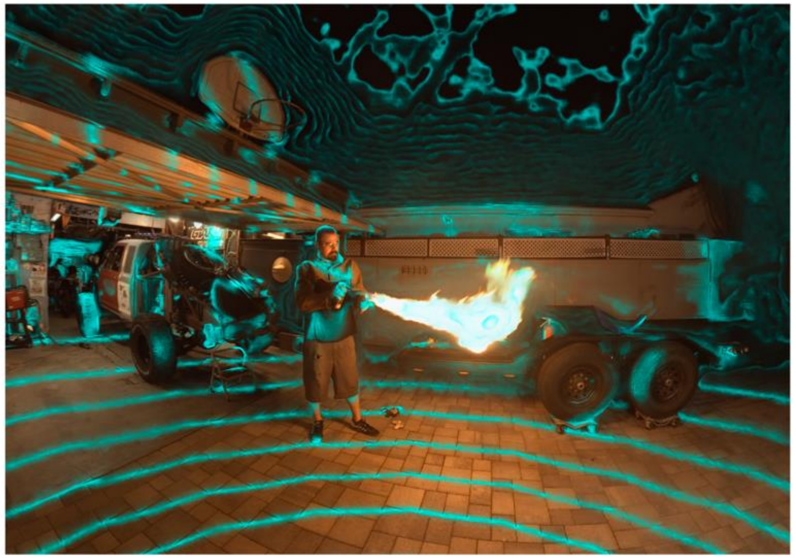
Fig. 1. The commercial 16 camera system, an equirectangular depth map, and final color rendering produced from our system.

[Video](#)

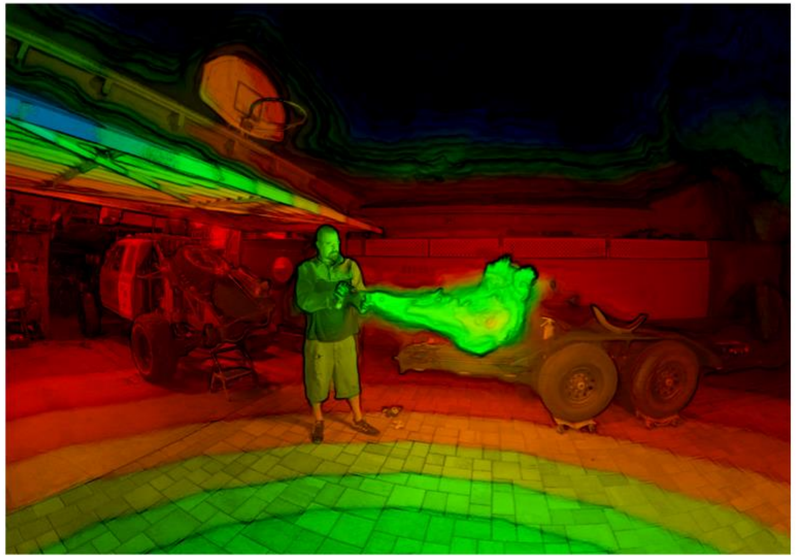
[SIGGRAPH Asia 2019]

Hemispherical light field capture & playback

(a) Capture Rig



(b) Multi-Sphere Image



(c) Layered Mesh Representation

IMMERSIVE LIGHT FIELD VIDEO WITH A LAYERED MESH REPRESENTATION

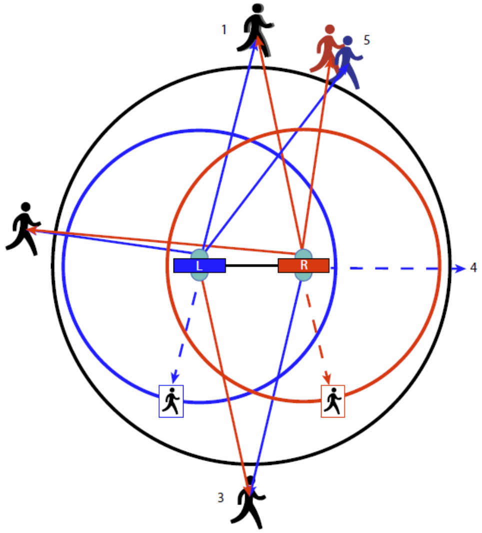
SIGGRAPH 2020 Technical Paper

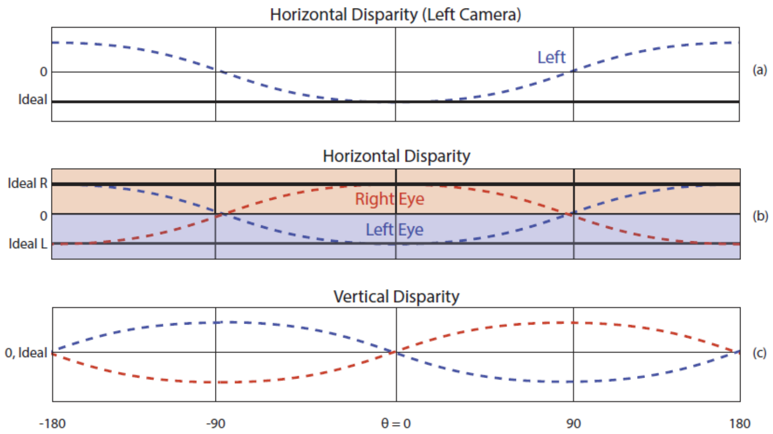
Download PDF

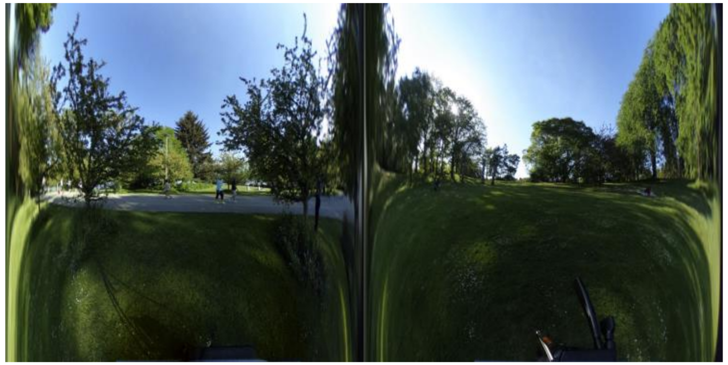
Michael Broxton*, John Flynn*, Ryan Overbeck*, Daniel Erickson*,
Peter Hedman, Matthew DuVall, Jason Dourgarian, Jay Busch, Matt Whalen,
Paul Debevec

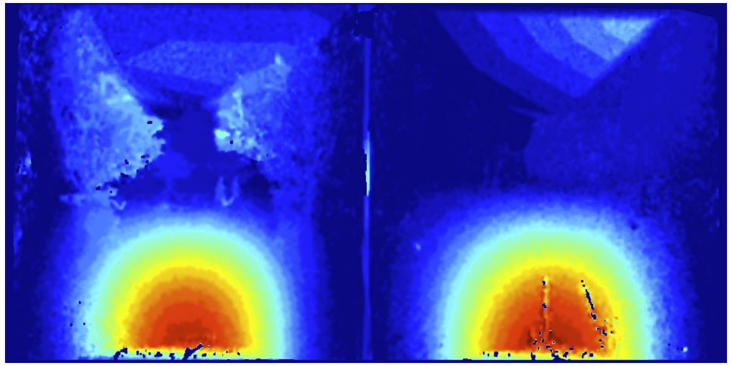
Stereo from *two* 360 cameras

Low-Cost 360 Stereo Photography and Video Capture, *Matzen, Cohen, Evans, Kopf, Szeliski*, SIGGRAPH 2017.









Immersive Video Stabilization

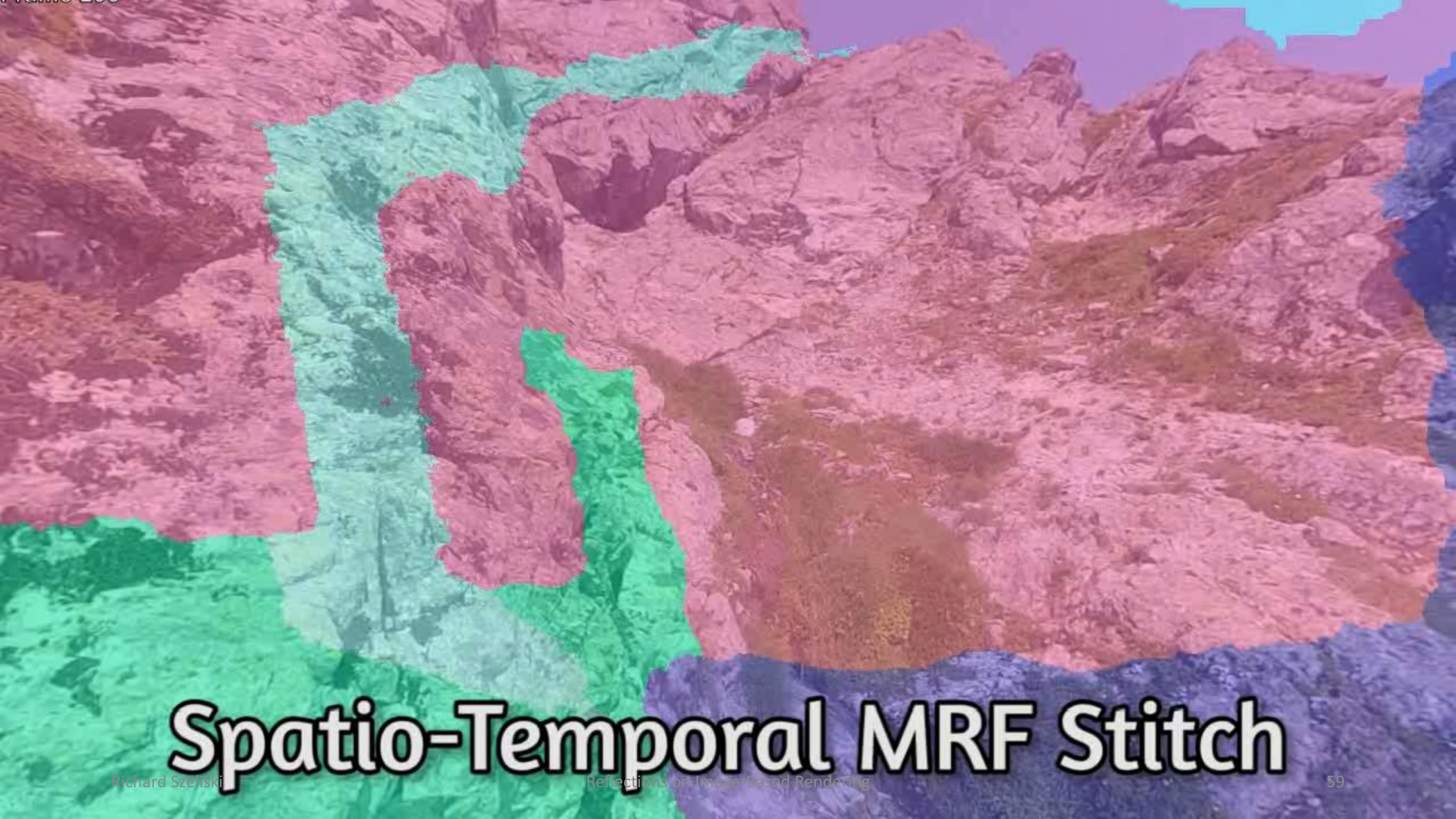
First-person Hyperlapse

Create buttery-smooth "fast forwards" from action videos

(a) Scene reconstruction (b) Proxy geometry (c) Stitched & blended

[Kopf, Cohen, Szeliski, SIGGRAPH 2014]

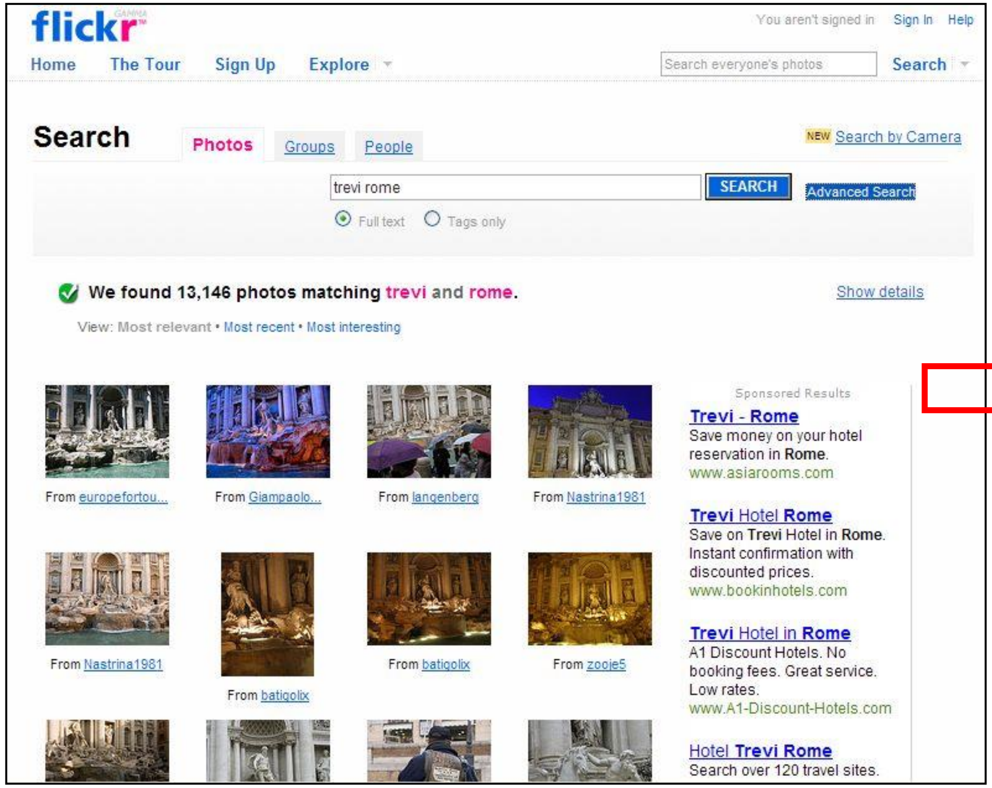

Proxy Geometry
(for a single video frame)

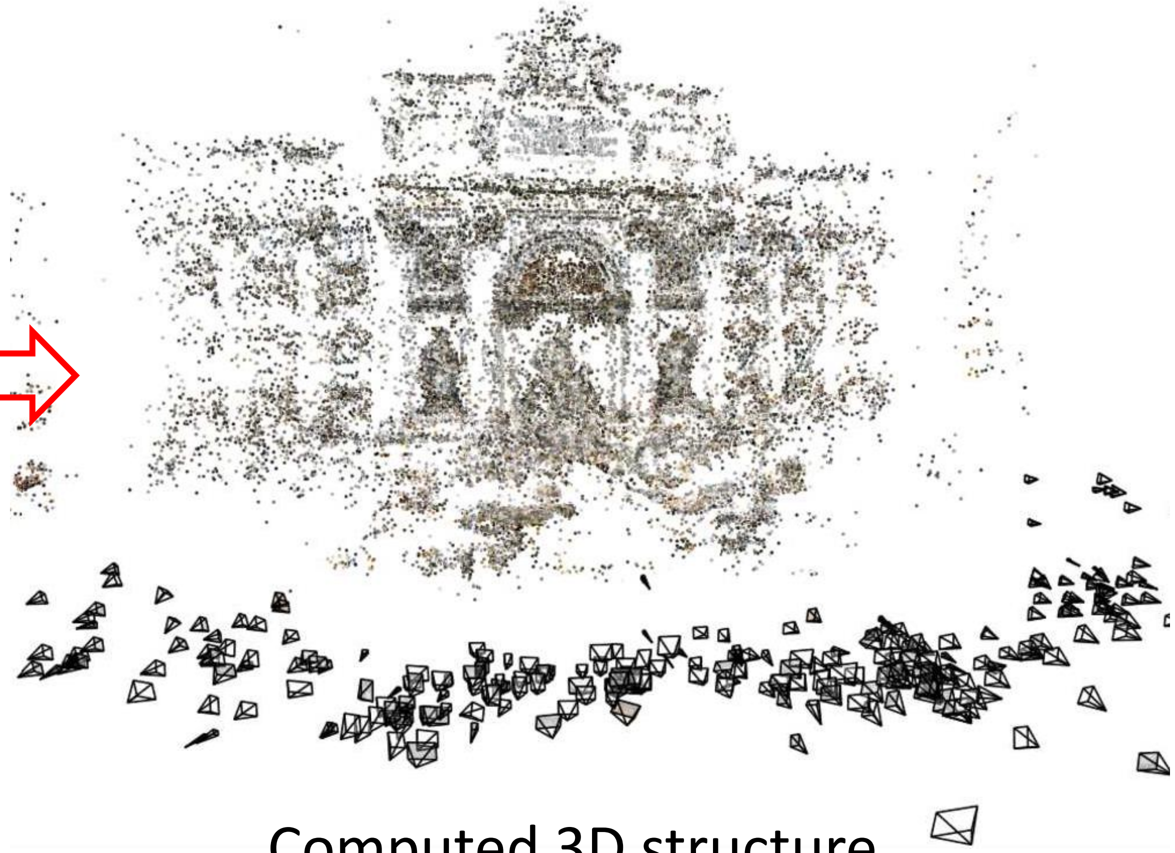

Spatio-Temporal MRF Stitch


Input Video

Large-Scale Reconstruction

Photo Tourism



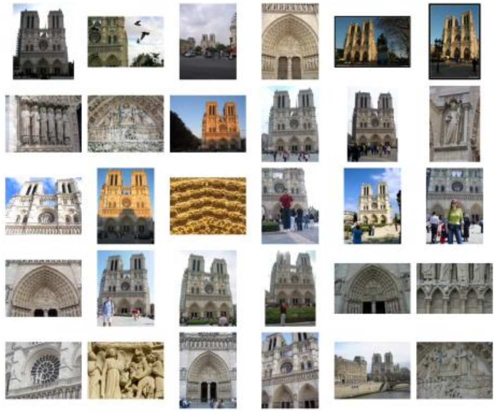


Internet images

Computed 3D structure

[Snavely, Seitz, Szeliski, SIGGRAPH 2006]

System overview



Input photographs

Scene reconstruction

Relative camera positions and orientations

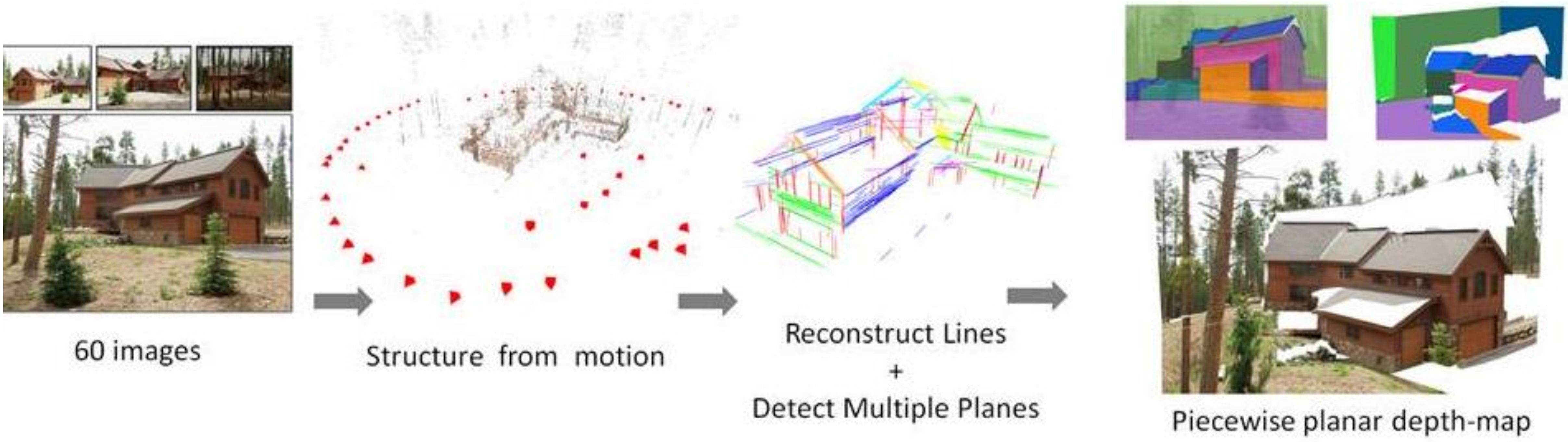
Point cloud

Sparse correspondence

Photo Explorer

Navigation: Prague Old Town Square

Piecewise planar proxies

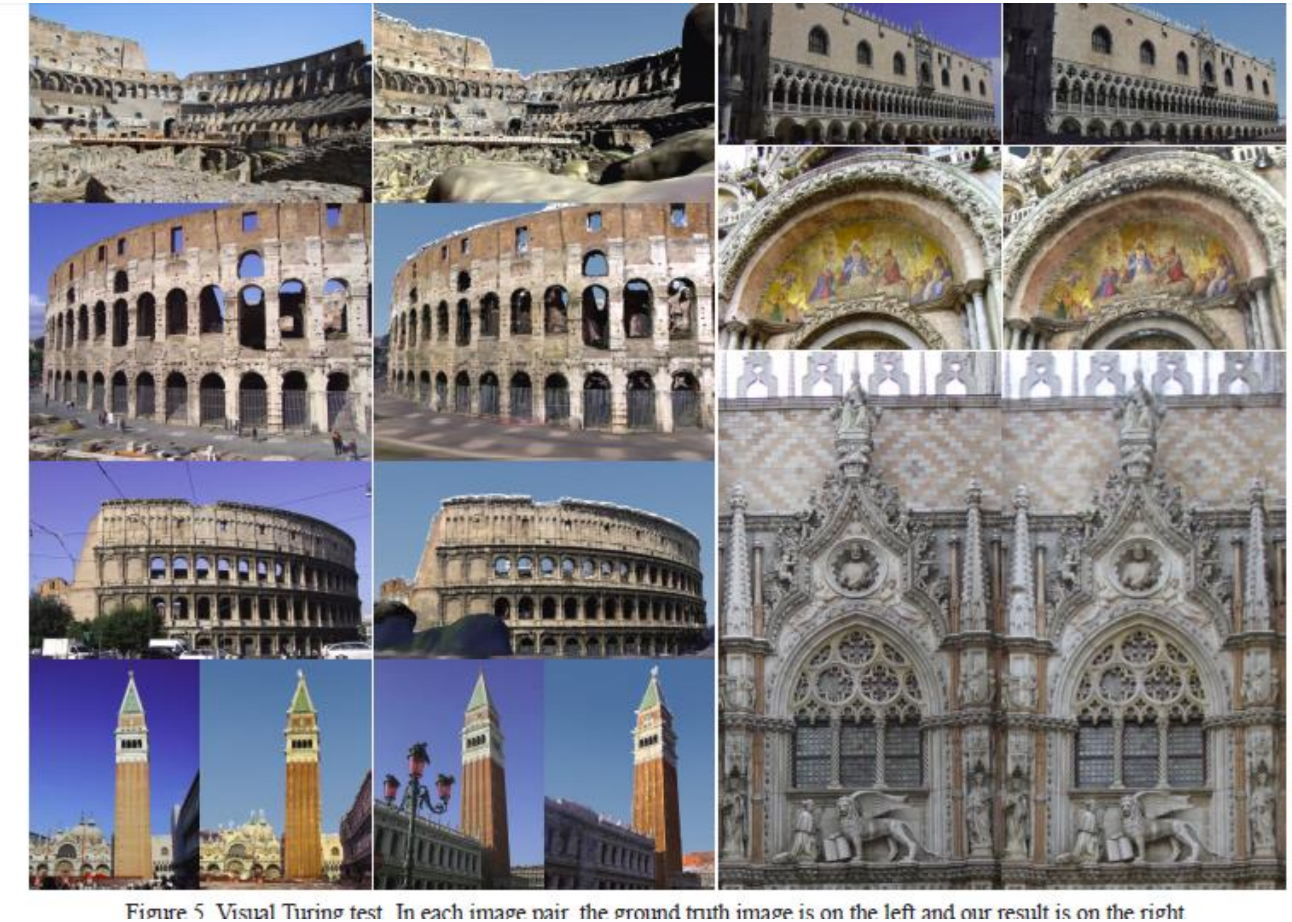


[Sinha, Steedly, Szeliski ICCV'09]

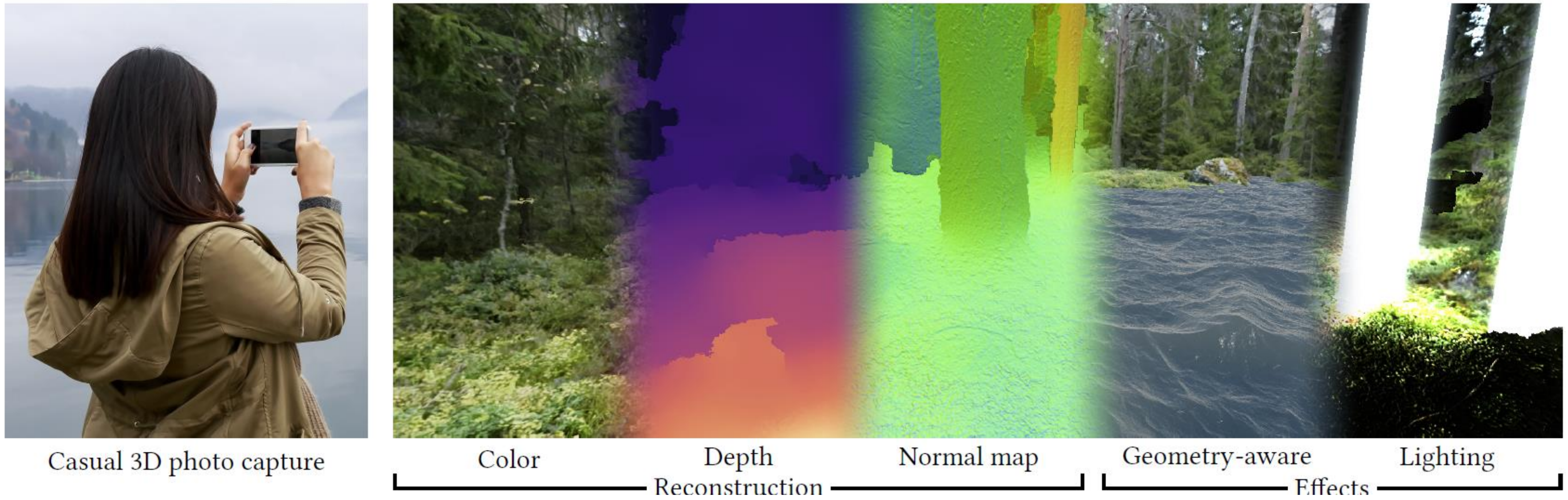
Photo Tours - 2012

[Kushal *et al.*, 3DIMPVT 2012]

The Visual Turing Test - 2013



[Shan *et al.*, 3DV 2013]



Casual 3D Photography

Peter Hedman, Suhib Alsisan, Richard Szeliski, Johannes Kopf

SIGGRAPH Asia 2017

Casual 3D Photography

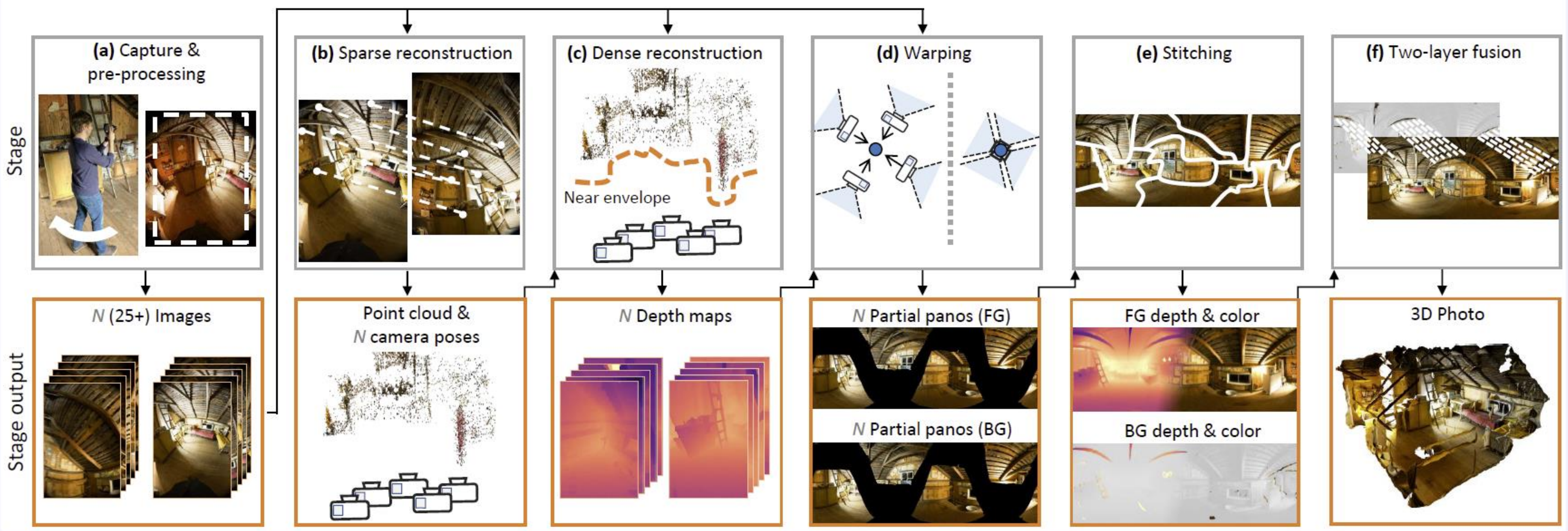
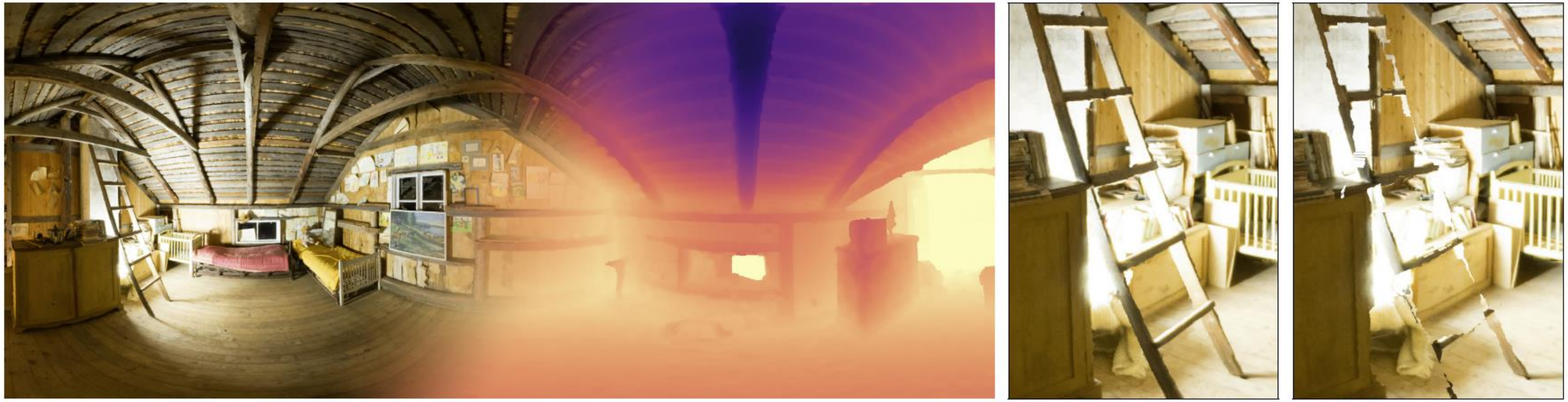


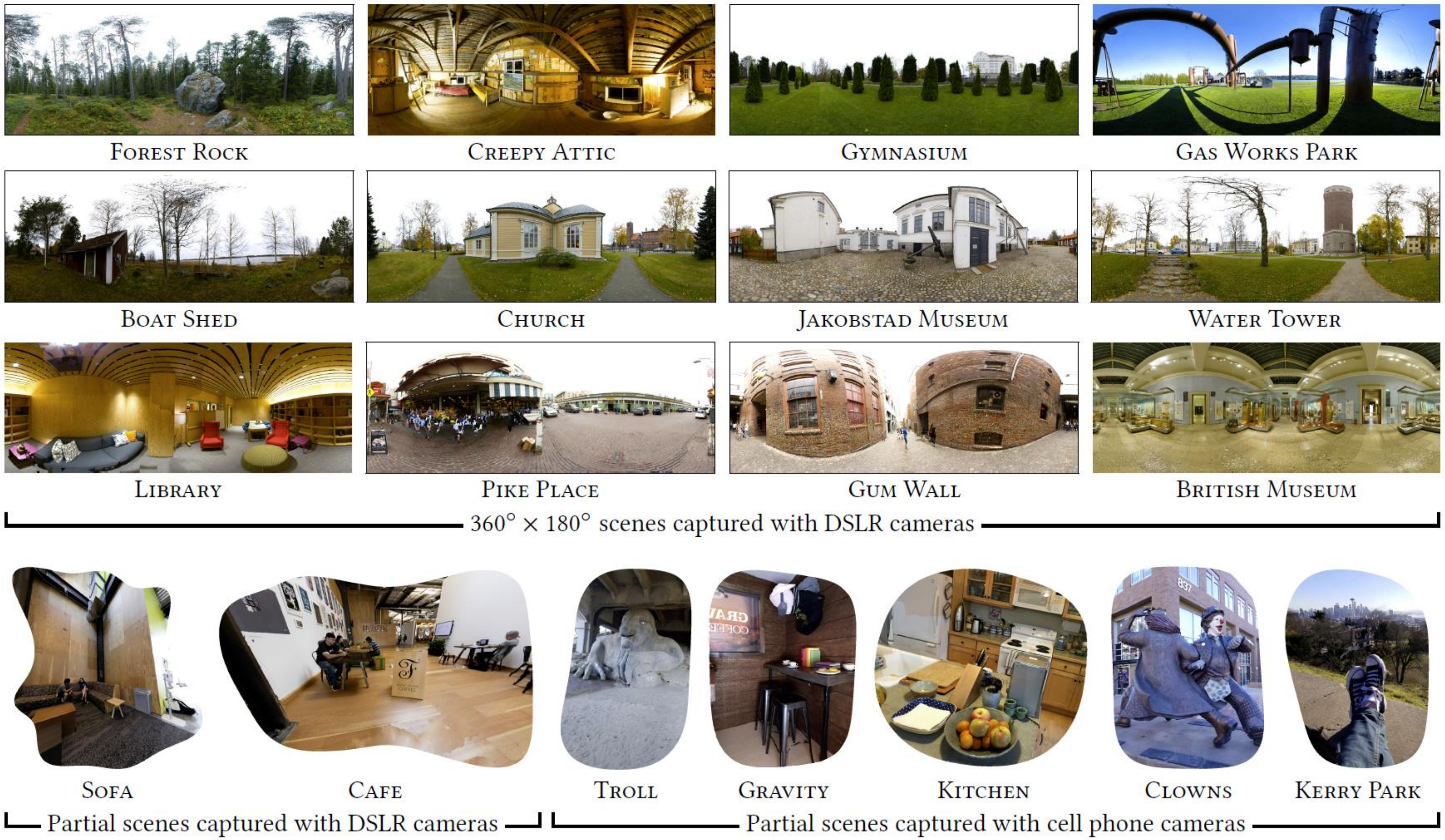
Figure 2: A breakdown of the 3D photo reconstruction algorithm into its six stages, with corresponding inputs and outputs: (a) Capture and pre-processing, Sec. 4.1; (b) Sparse reconstruction, Sec. 4.2; (c) Dense reconstruction, Sec. 4.3; (d) Warping into a central panorama, Sec. 4.4.1; (e) Parallax-tolerant Stitching, Sec. 4.4.2; (f) Two-layer fusion, Sec. 4.4.3.

Casual 3D Photography



(a) Front color-and-depth panorama (b) Front detail (c) Back detail

Casual 3D Photography

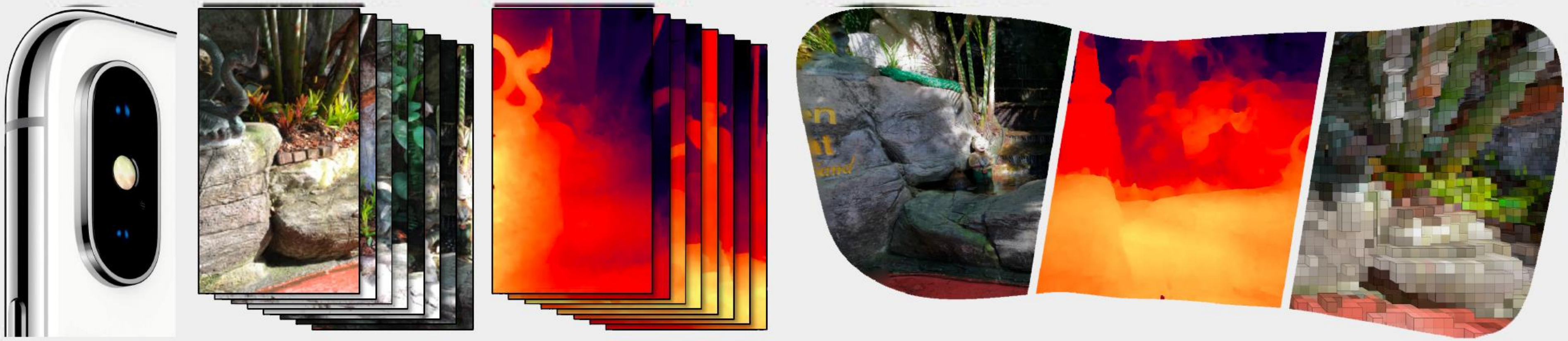


Instant 3D Photography

Peter Hedman
University College London *

Johannes Kopf
Facebook

* This work was done while Peter was working as a contractor for Facebook.



Dual camera phone

Input burst of 34 color-and-depth photos, captured in 34.0 seconds

Our 3D panorama (showing color, depth, and a 3D effect), generated in 34.7 seconds.

Our work enables practical and casual 3D capture with regular dual camera cell phones. Left: A burst of input color-and-depth image pairs that we captured with a dual camera cell phone at a rate of one image per second. Right: 3D panorama generated with our algorithm in about the same time it took to capture. The geometry is highly detailed and enables viewing with binocular and motion parallax in VR, as well as applying 3D effects that interact with the scene, e.g., through occlusions (right).

Practical 3D Photography

Johannes Kopf, Suhib Alsisan, Francis Ge, Yangming Chong, Kevin Matzen
Ocean Quigley, Josh Patterson, Jossie Tirado, Shu Wu, Michael F. Cohen
Facebook

(a) Input (setup) (100 ms)
(b) LDI (inpainted color / depth) (1100 ms)
(d) Triangle Mesh (100 ms)
(e) Novel view (30fps)

Figure 1. 3D Photo Creation. Runtime measured on iPhone X.

Practical 3D Photography

Johannes Kopf, Suhib Alsisan, Francis Ge, Yangming Chong, Kevin Matzen, Ocean Quigley, Josh Patterson, Jossie Tirado, Shu Wu, Michael F. Cohen

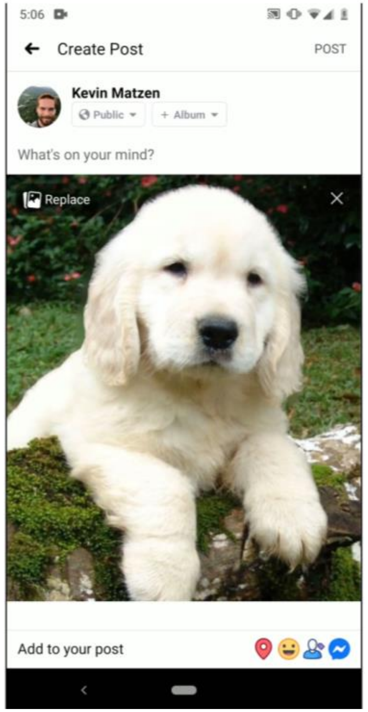
CVPR Workshop on Computer Vision for Augmented and Virtual Reality, Long Beach, CA, 2019.

PDF

#spotlight, #demo

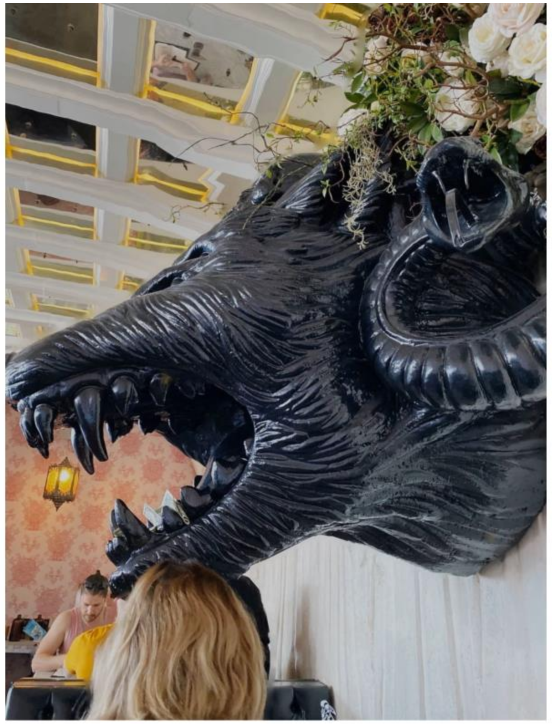
3D Photos on Facebook

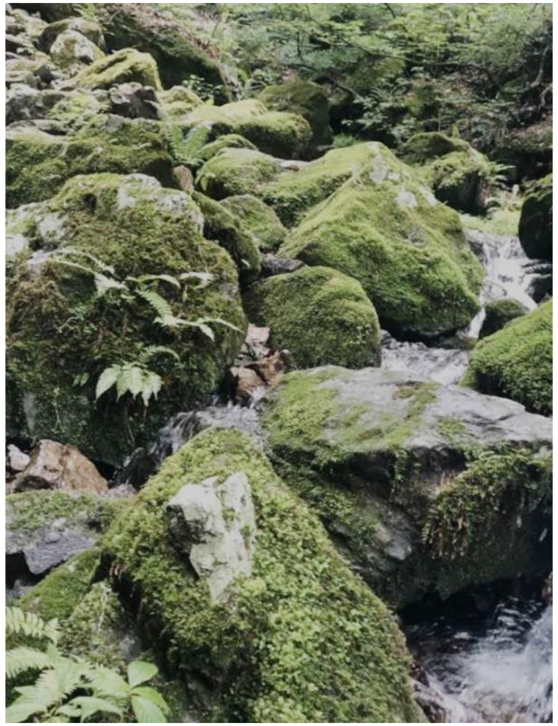
Estimate depth map from photo to create an interactive animation



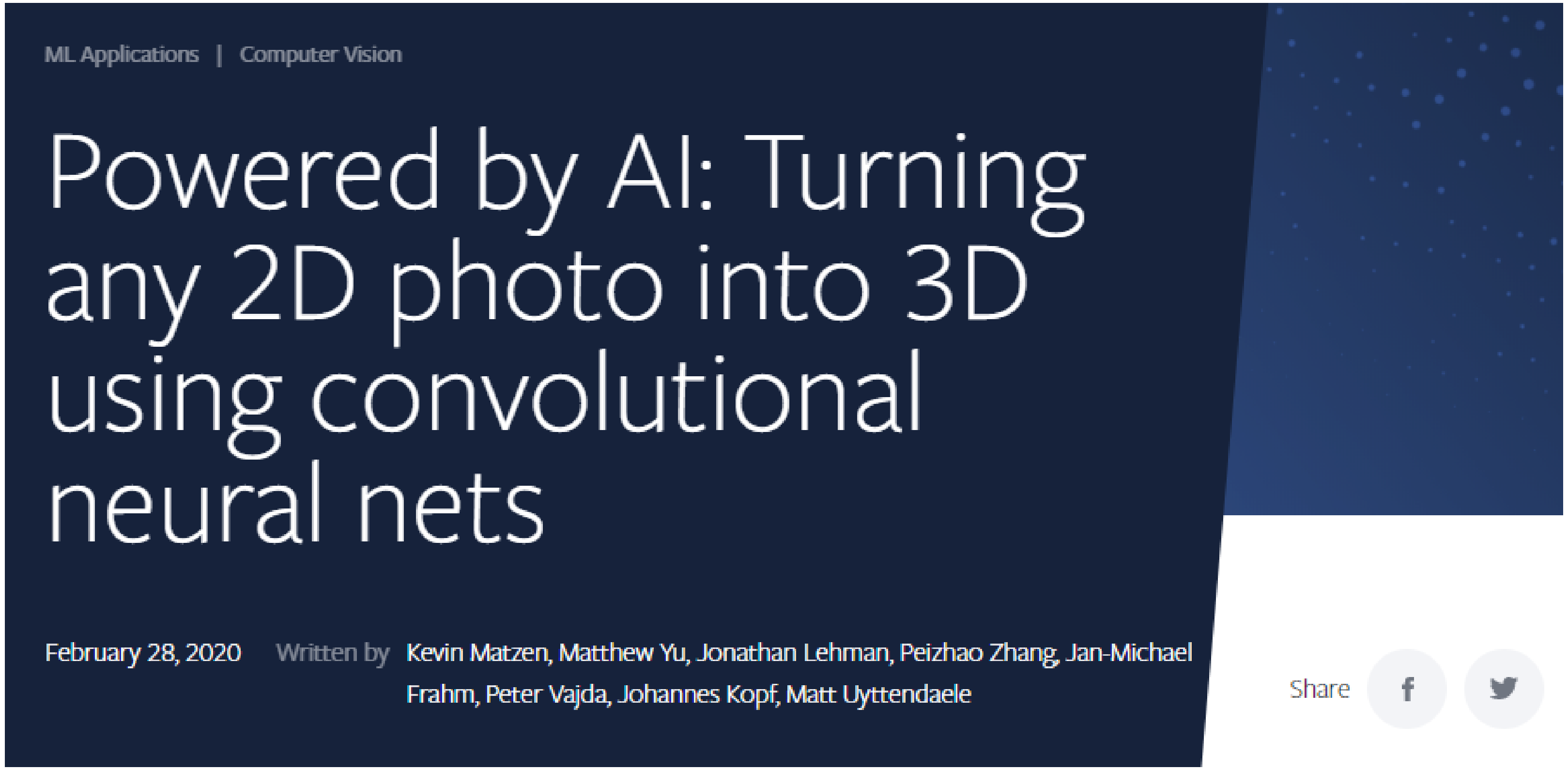
3D Photos on Facebook

Estimate depth map from photo to create an interactive animation





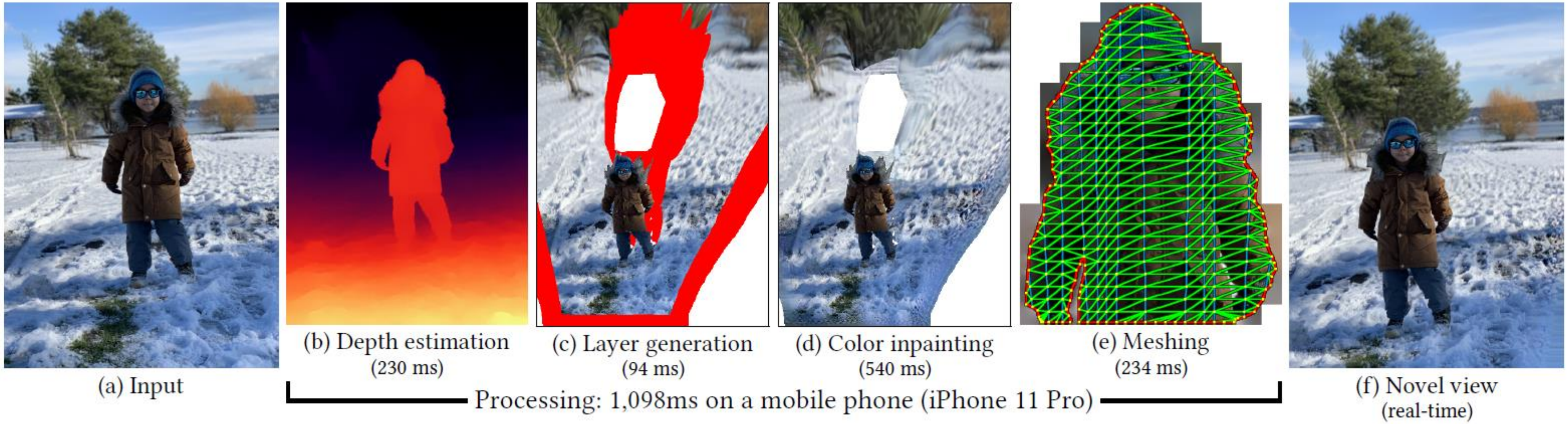
3D Photos blog post



https://ai.facebook.com/blog/-powered-by-ai-turning-any-2d-photo-into-3d-using-convolutional-neural-nets/

One Shot 3D Photography

JOHANNES KOPF, KEVIN MATZEN, SUHIB ALSISAN, OCEAN QUIGLEY, FRANCIS GE, YANGMING CHONG, JOSH PATTERSON, JAN-MICHAEL FRAHM, SHU WU, MATTHEW YU, PEIZHAO ZHANG, ZIJIAN HE, PETER VAJDA, AYUSH SARAF, and MICHAEL COHEN, Facebook



[SIGGRAPH 2020]

3D Photography using Context-aware Layered Depth Inpainting CVPR'2020

Google Photos cinematic effect

Jamie Aspinall

Product Manager, Google Photos

Published Dec 15, 2020

Relive the moment with Cinematic photos

Cinematic photos help you relive your memories in a way that feels more vivid and realistic—so you feel like you're transported back to that moment. To do this, we use machine learning to predict an image's depth and produce a 3D representation of the scene—even if the original image doesn't include depth information from the camera. Then we animate a virtual camera for a smooth panning effect—just like out of the movies.

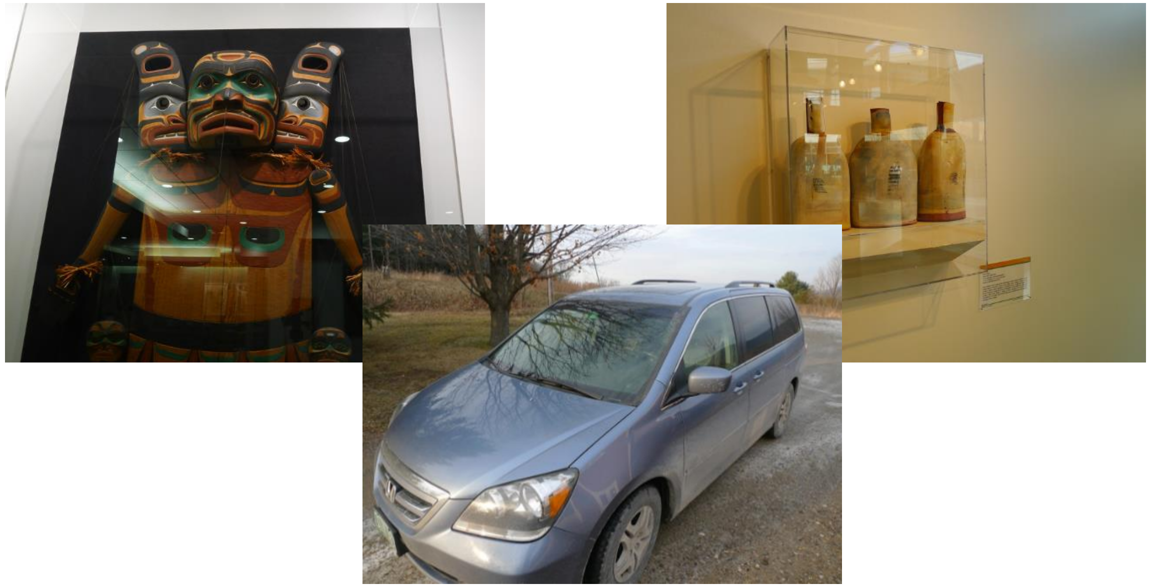
https://blog.google/products/photos/new-cinematic-photos-and-more-ways-relive-your-memories/

What's missing?

Reflections and Transparency

Image-Based Rendering with Reflections

- Reflections, gloss, and highlights are everywhere



- How do these affect image-based modeling / rendering?

  [Sinha *et al.*, SIGGRAPH 2012]

Standard IBR with Reflections

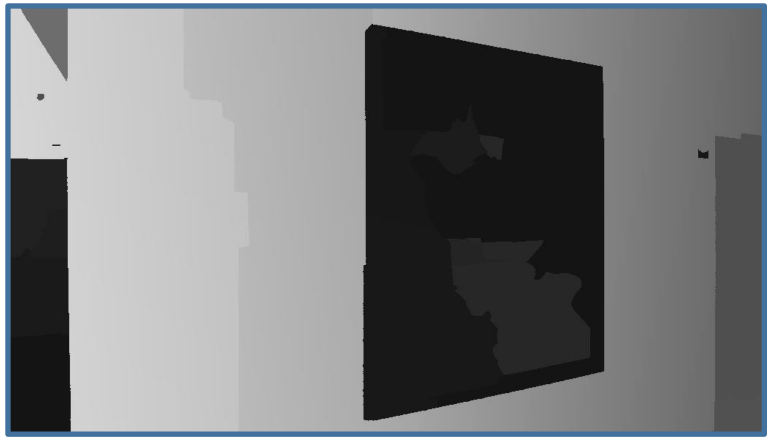
Our New Rendering System


Two Layers
One Layer


Our Result
Two Layers
Standard IBR
One Layer

Input

Front Depth



Rear Depth

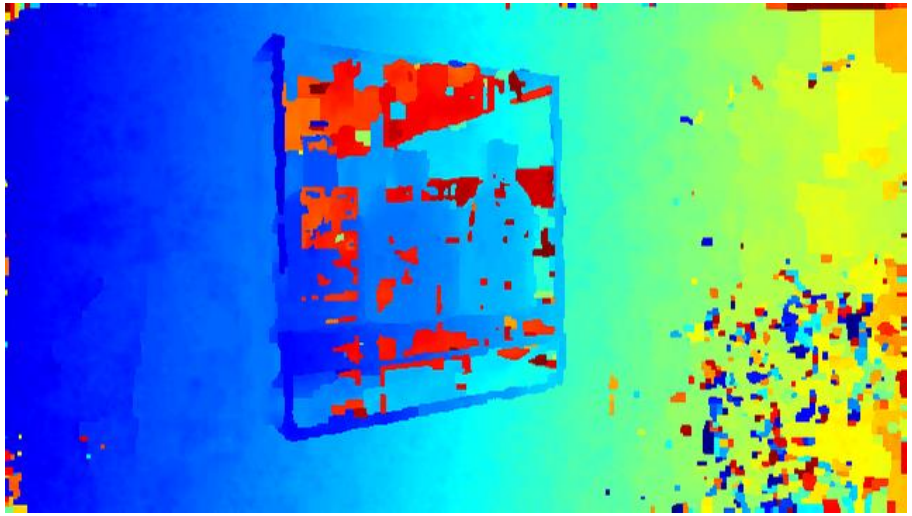
Input

Front Layer

Rear Layer

Image-Based Rendering in the Gradient Domain

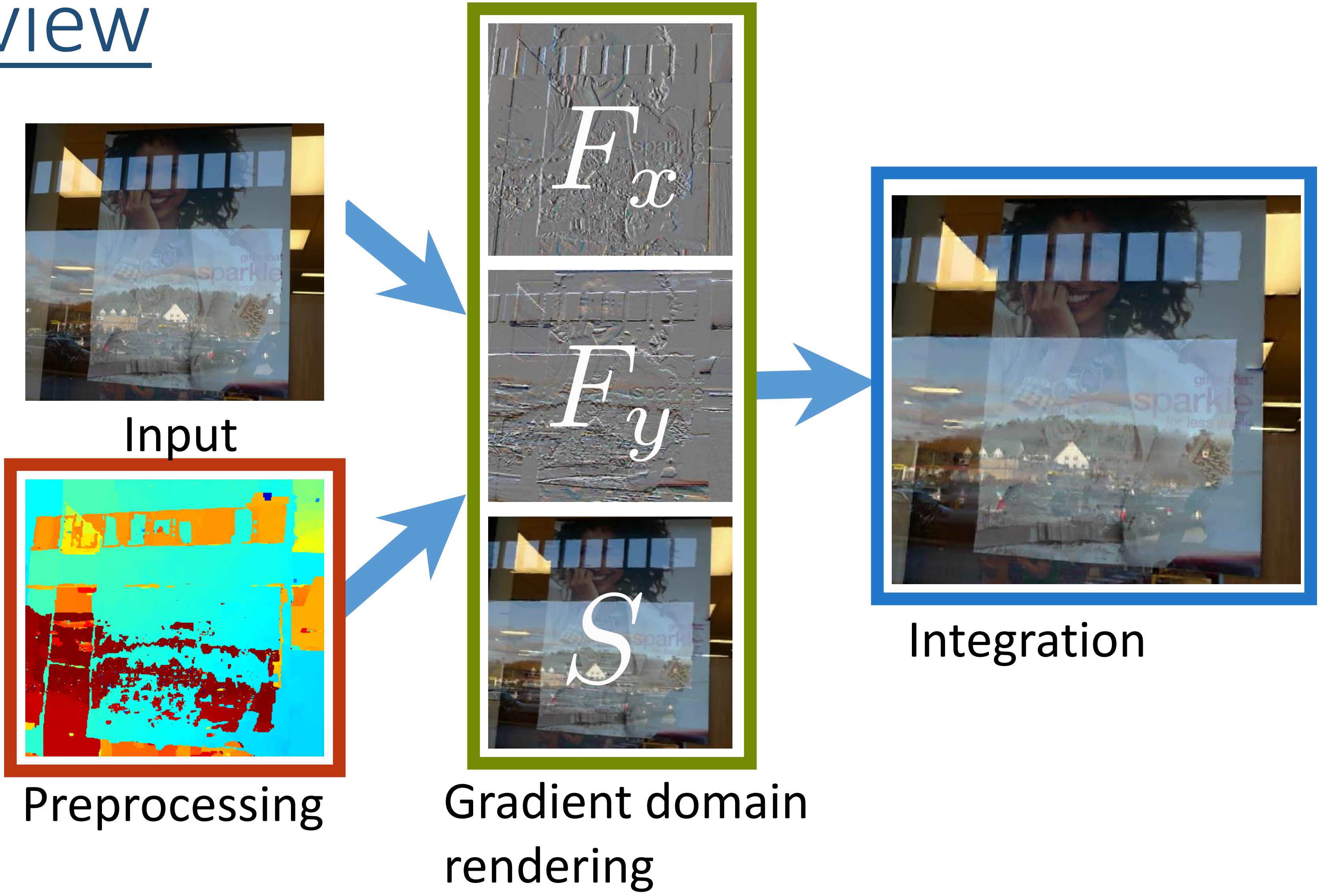
- Wrong depth for textureless or transparent areas



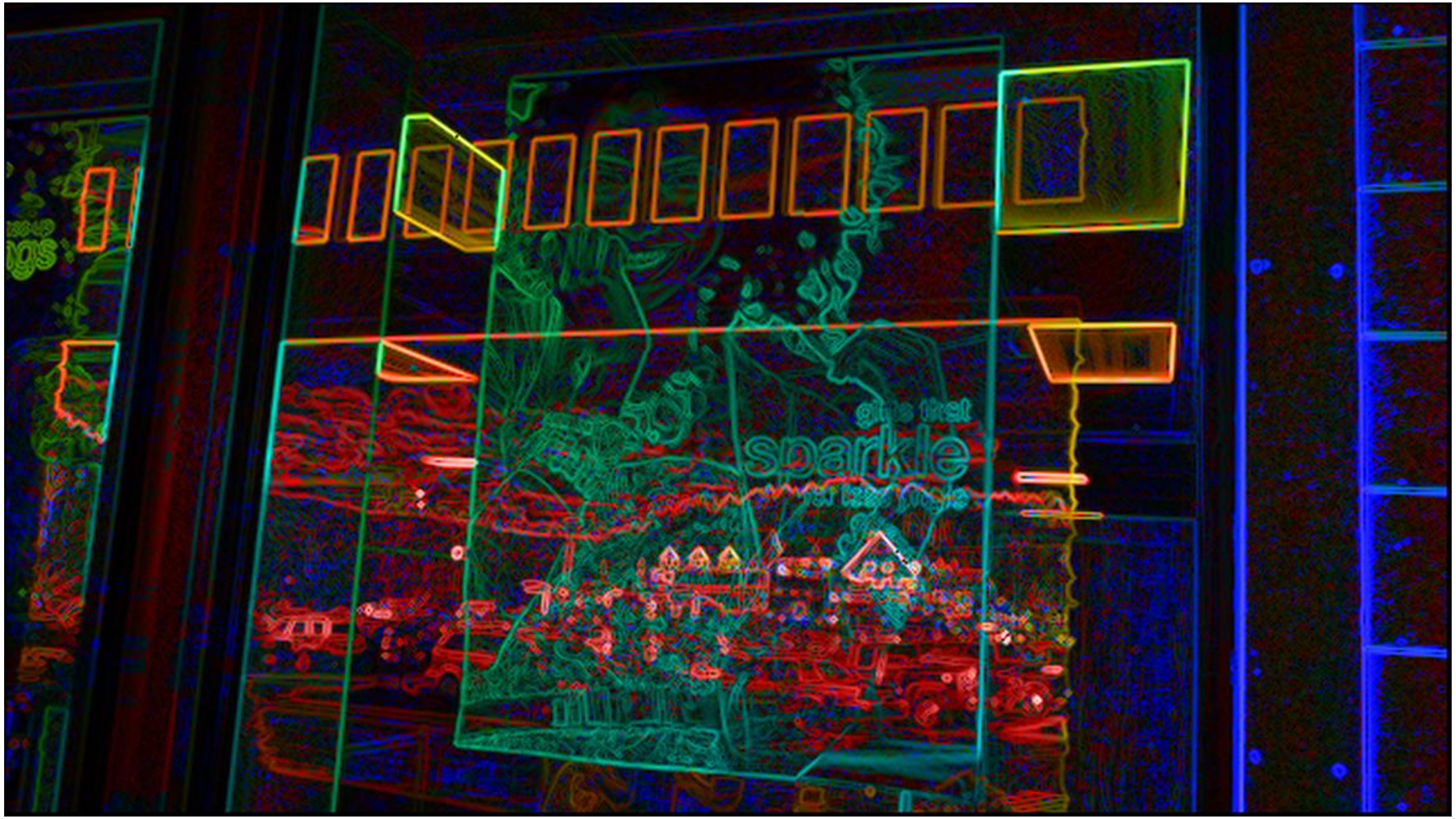
- Solve by reconstructing depth at gradients and re-integrating

    [Kopf *et al.* SIGGRAPH Asia 2013]

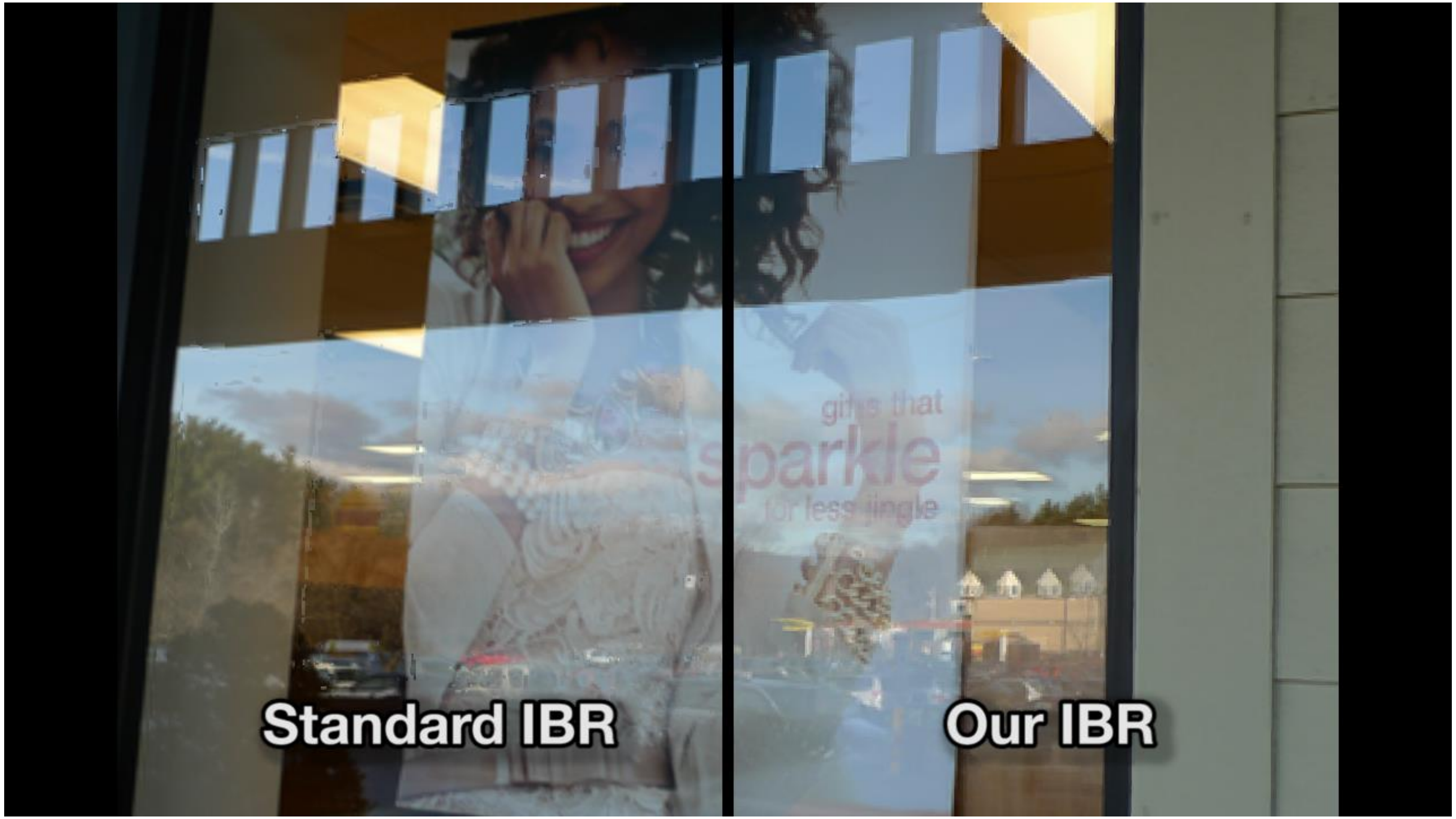
Overview


Input
Preprocessing
F_x
F_y
S
Gradient domain rendering
Integration

Gradient Domain



Our Method



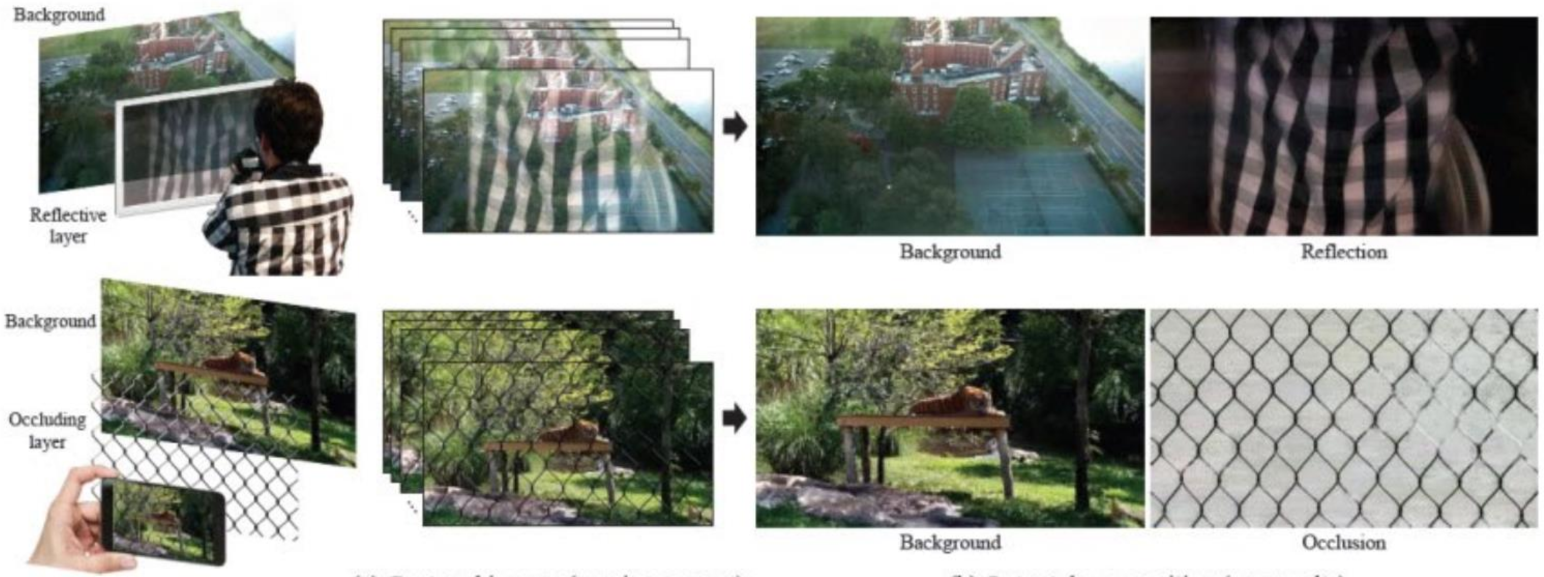
SIGGRAPH 2015

A Computational Approach for Obstruction-Free Photography

Tianfan Xue[1*] Michael Rubinstein[2*] Ce Liu[2*] William T. Freeman[1,2]

[1]MIT CSAIL [2]Google Research

* Part of this work was done while Michael Rubinstein and Ce Liu were at Microsoft Research, and when Tianfan Xue was an intern at Microsoft Research New England.

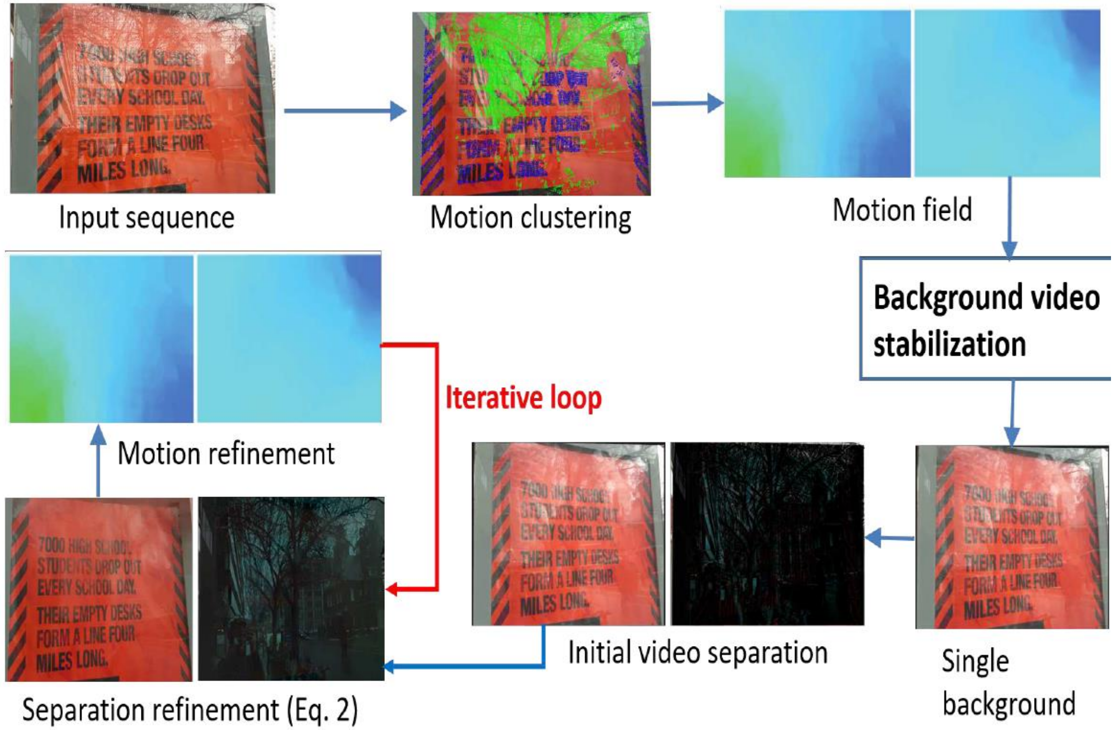


(a) Captured images (moving camera) (b) Output decomposition (our results)

Video Reflection Removal Through Spatio-Temporal Optimization

Ajay Nandoriya*[1], Mohamed Elgharib*[1], Changil Kim[2], Mohamed Hefeeda[3], and Wojciech Matusik[2]

[1]Qatar Computing Research Institute, HBKU [2]MIT CSAIL [3]Simon Fraser University



[ICCV 2017]

Reflection Removal Using a Dual-Pixel Sensor

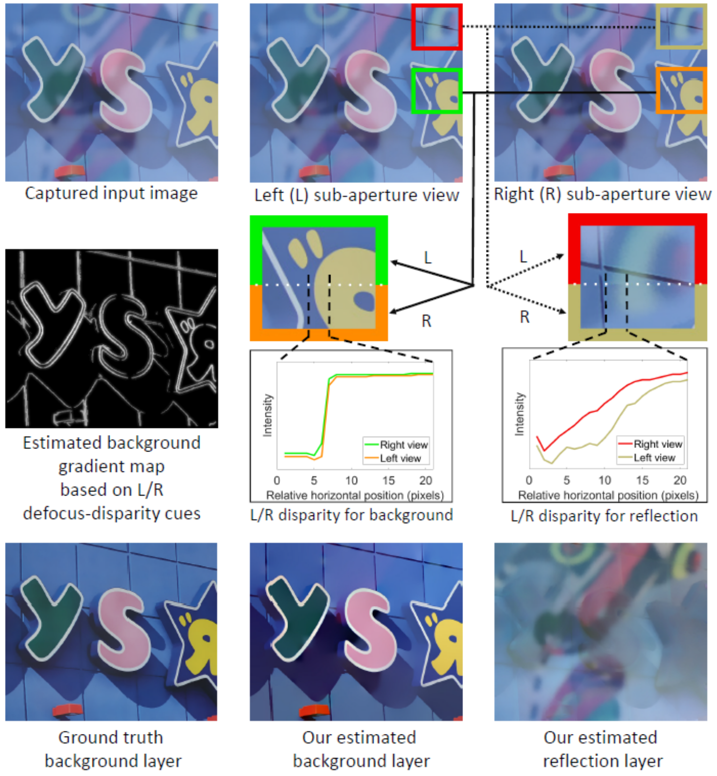
Abhijith Punnappurath
York University
pabhijith@eecs.yorku.ca

Michael S. Brown
York University
mbrown@eecs.yorku.ca

mailto:pabhijith@eecs.yorku.ca

Abstract

Reflection removal is the challenging problem of removing unwanted reflections that occur when imaging a scene that is behind a pane of glass. In this paper, we show that most cameras have an overlooked mechanism that can greatly simplify this task. Specifically, modern DLSR and smartphone cameras use dual pixel (DP) sensors that have two photodiodes per pixel to provide two sub-aperture views of the scene from a single captured image. "Defocus-disparity" cues, which are natural by-products of the DP sensor encoded within these two sub-aperture views, can be used to distinguish between image gradients belonging to the in-focus background and those caused by reflection interference. This gradient information can then be incorporated into an optimization framework to recover the background layer with higher accuracy than currently possible from the single captured image. As part of this work, we provide the first image dataset for reflection removal consisting of the sub-aperture views from the DP sensor.



[CVPR 2019]

Open issues

- Improve stereo matching
  - Plane + parallax representation
- Reflectivity (β) estimation
  - Iterative Refinement
- Handle distorted reflections
  - [See next slide]
- Model real-valued reflectivity
  - Fresnel reflection

Real-World Normal Map Capture for Nearly Flat Reflective Surfaces

Bastien Jacquet[1], Christian Häne[1], Kevin Köser[12*], Marc Pollefeys[1]

ETH Zürich[1]
Zürich, Switzerland

GEOMAR Helmholtz Centre for Ocean Research[2]
Kiel, Germany

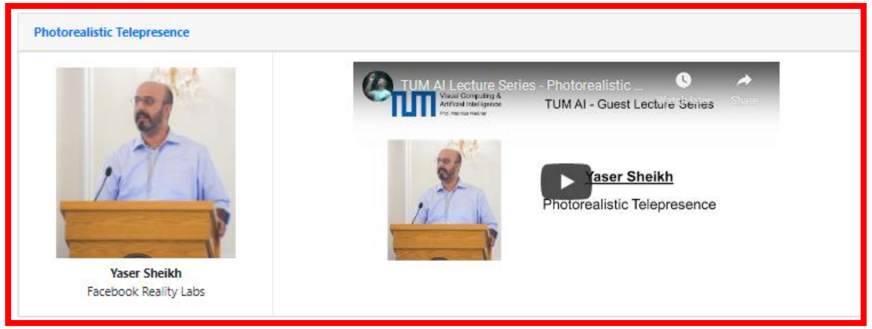
Abstract

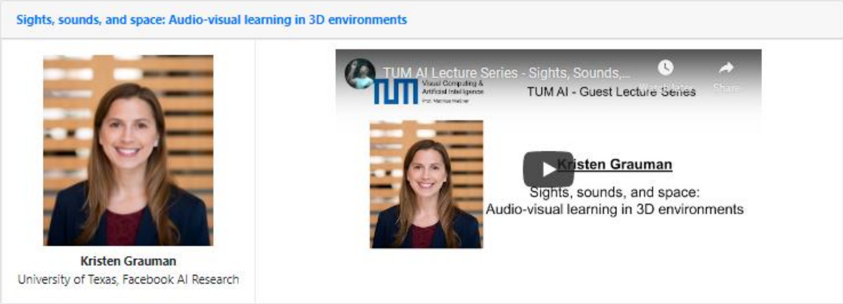
Although specular objects have gained interest in recent years, virtually no approaches exist for markerless reconstruction of reflective scenes in the wild. In this work, we present a practical approach to capturing normal maps in real-world scenes using video only. We focus on nearly planar surfaces such as windows, facades from glass or metal, or frames, screens and other indoor objects and show how normal maps of these can be obtained without the use of an artificial calibration object. Rather, we track the reflections of real world straight lines, while moving with a hand held

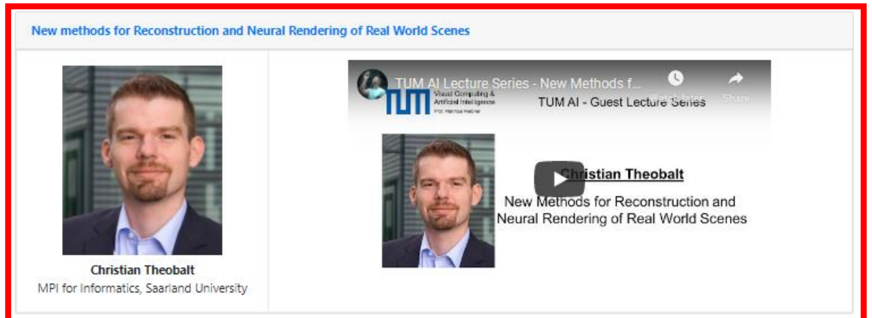
Figure 1. Real-world glass reflection. Notice that reflection in different windows on the same facade can appear very different due to minor deformations and normal variations. Our goal is to capture normal maps of real windows to faithfully reproduce this effect.

Neural Rendering

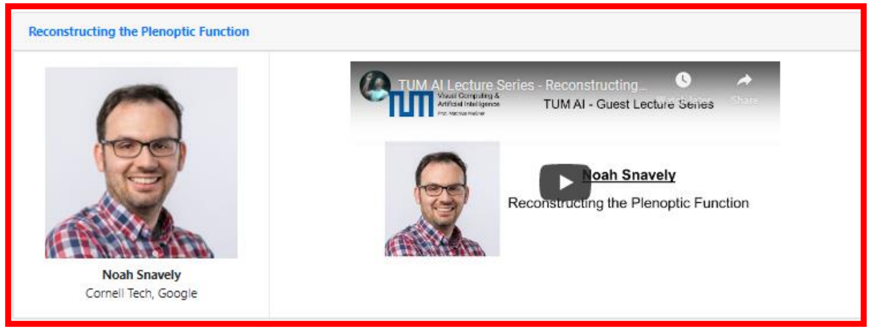
TUM AI Lecture series

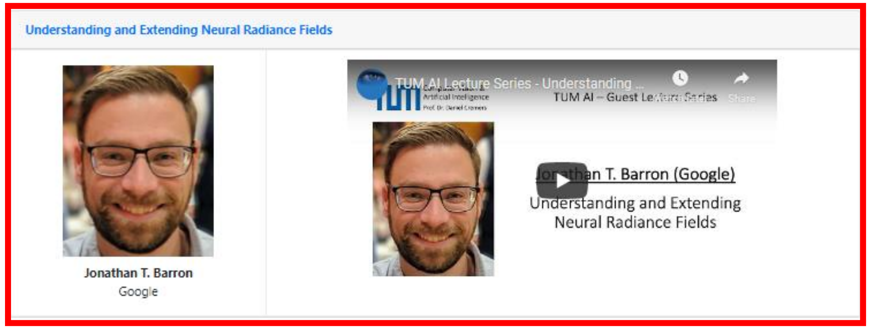


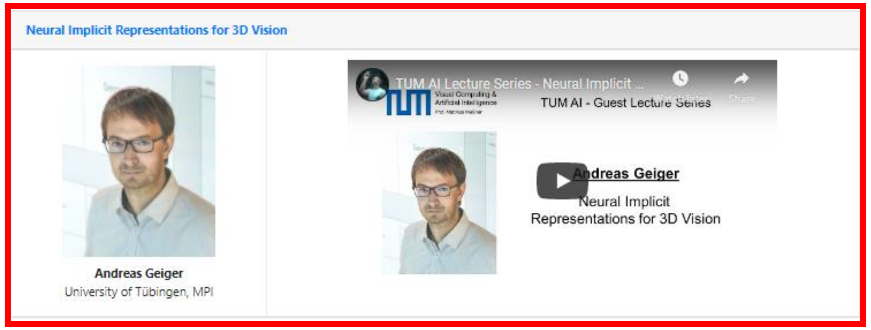




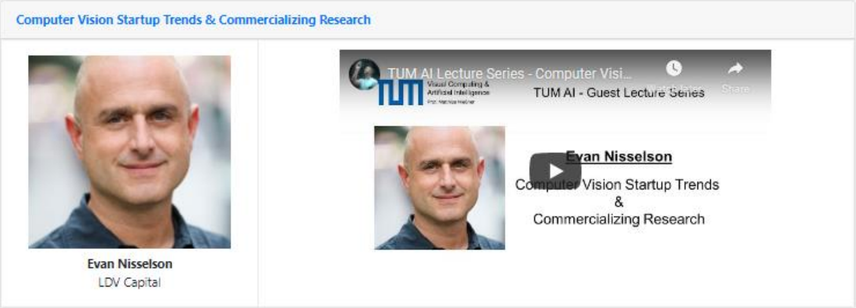
TUM AI Lecture series







TUM AI Lecture series



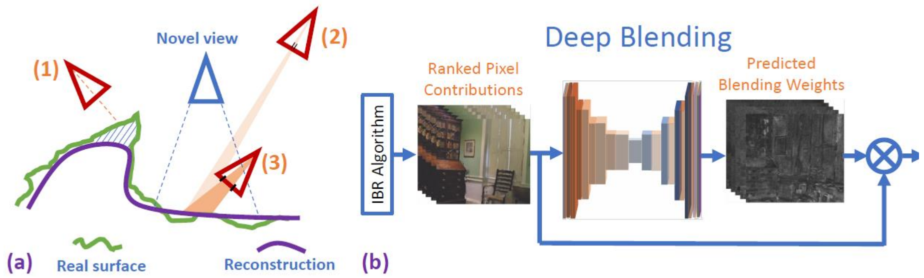
Neural Rendering

CVPR 2020 tutorial.

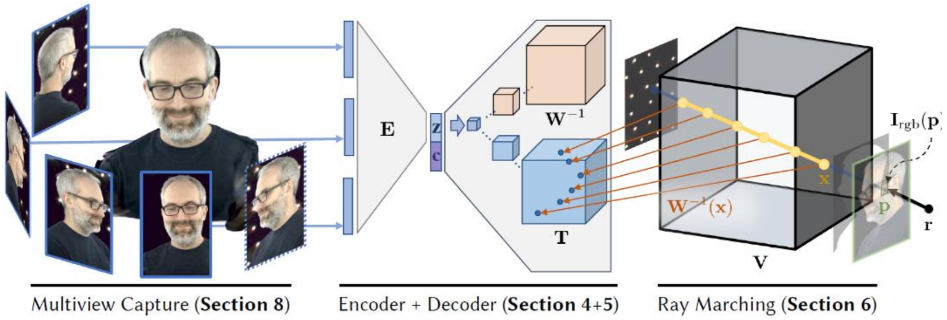
09:00–09:15	Welcome and Introduction	Michael Zollhöfer
09:15–09:30	Fundamentals, Taxonomy, Neural Rendering	Ayush Tewari
Semantic Photo Synthesis and Manipulation		
09:30–09:40	Overview	Jun-Yan Zhu
09:40–10:00	Semantic Image Synthesis with Spatially-Adaptive Normalization	Taesung Park
10:00–10:30	Coffee Break	
Facial Reenactment & Body Reenactment		
10:25–10:35	Overview	Justus Thies
10:35–11:00	Neural Rendering for High-Quality Synthesis of Human Portrait Video and Images	Christian Theobalt
11:00–11:20	Neural Rendering for Virtual Avatars	Aliaksandra Shysheya
Novel View Synthesis		
11:20–11:35	Overview	Vincent Sitzmann
11:30–11:50	Neural Rerendering in the Wild	Moustafa Meshry
11:50–12:10	NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis	Ben Mildenhall
12:10–13:20	Lunch Break	
Learning to Relight		
13:20–13:30	Overview	Zexiang Xu
13:30–13:50	Multi-view Relighting Using a Geometry-Aware Network	Julien Philip
13:50–14:10	Neural Inverse Rendering	Abhimitra Meka
Free Viewpoint Videos		
14:10–14:20	Overview	Sean Fanello
14:20–14:40	Neural Rendering for Performance Capture	Rohit K. Pandey
14:40–15:00	Neural Volumes: Learning Dynamic Renderable Volumes from Images	Stephen Lombardi
15:00–15:30	Coffee Break	
15:30–15:45	Social Implications, Open Challenges, Conclusion	Ohad Fried
15:45–16:15	Followup Discussion	

3D representations for neural rendering

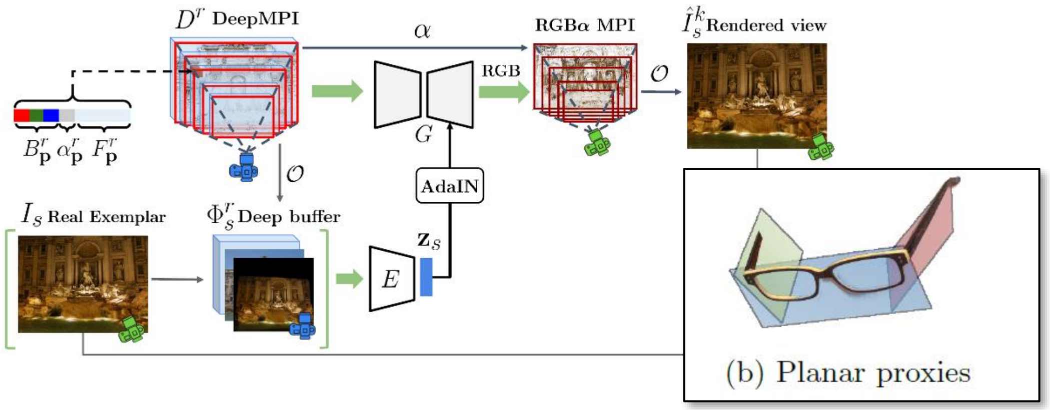
- 3D models & textures



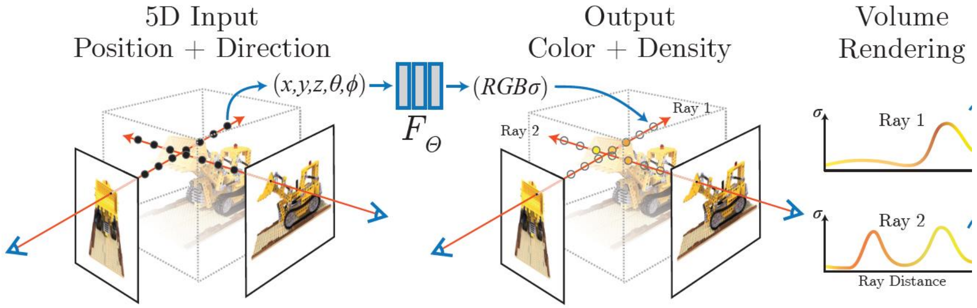
- Voxels



- Depth images and layers



- Implicit functions (MLPs)



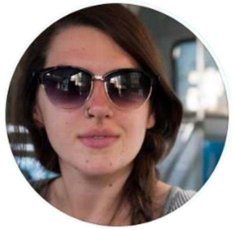
SynSin: view synthesis from a single image

SynSin: End-to-end View Synthesis from a Single Image

Olivia Wiles[1*] Georgia Gkioxari[2] Richard Szeliski[3] Justin Johnson[2,4]

[1]University of Oxford [2]Facebook AI Research [3]Facebook [4]University of Michigan

Figure 1: **End-to-end view synthesis.** Given a *single* RGB image (red), SynSin generates images of the scene at new viewpoints (blue). SynSin predicts a 3D point cloud, which is projected onto new views using our differentiable renderer; the rendered point cloud is passed to a GAN to synthesise the output image. SynSin is trained end-to-end, without 3D supervision.



SynSin: view synthesis from a single image

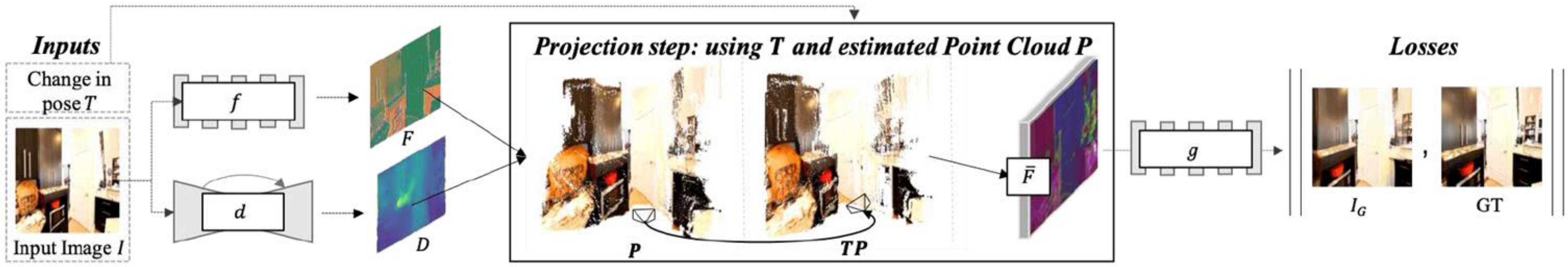
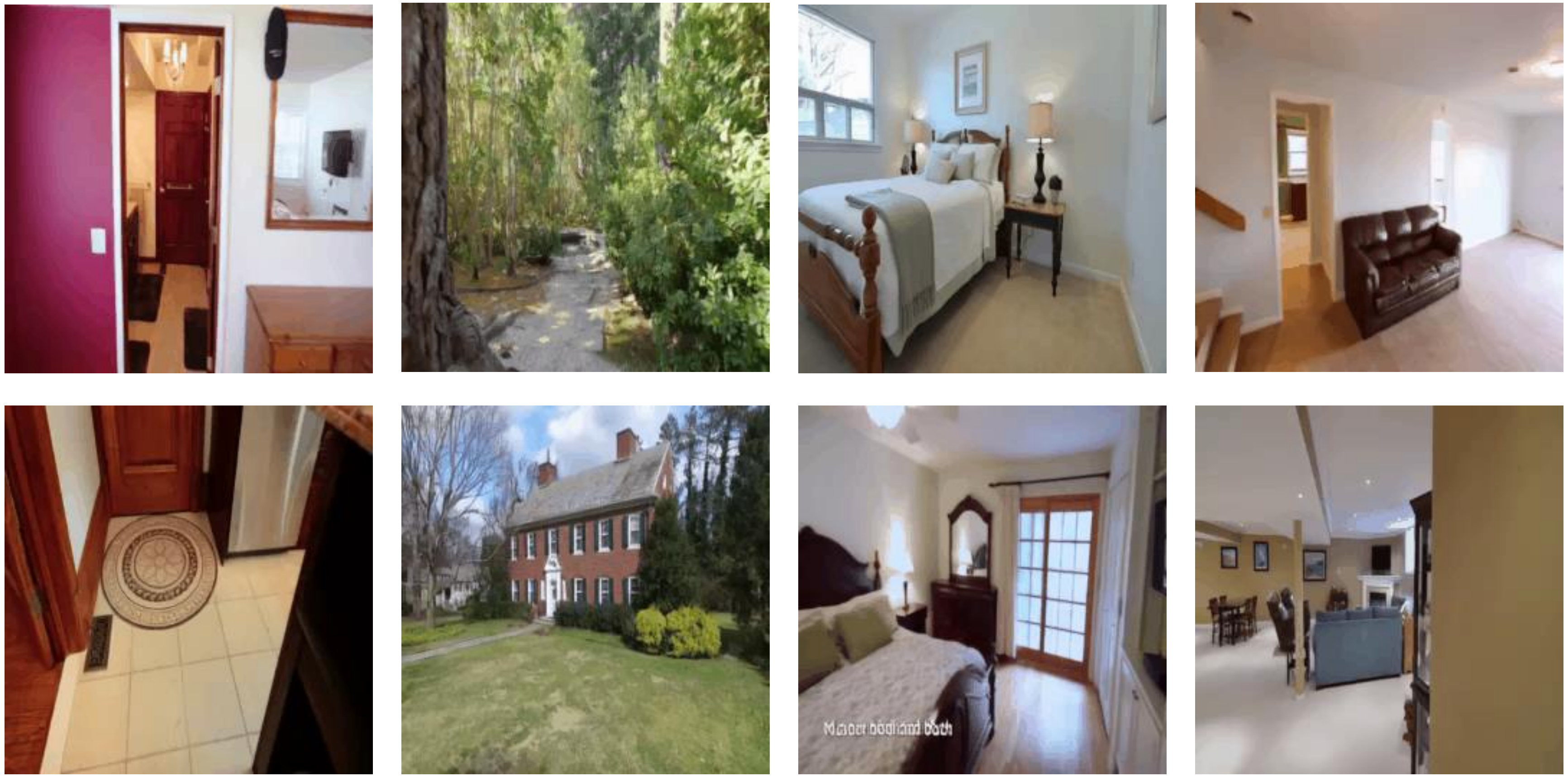


Figure 2: **Our end-to-end system.** The system takes as input an image I of a scene and change in pose $\mathtt{T}$. The *spatial feature predictor* (f) learns a set of features F (visualised by projecting features using PCA to RGB) and the *depth regressor* (d) a depth map D. F are projected into 3D (the diagram shows RGB for clarity) to give a point cloud $\mathcal{P}$ of features. $\mathcal{P}$ is transformed according to $\mathtt{T}$ and rendered. The rendered features $\bar{F}$ are passed through the *refinement network* (g) to generate the final image I_G. I_G should match the target image, which we enforce using a set of discriminators and photometric losses.

SynSin: view synthesis from a single image

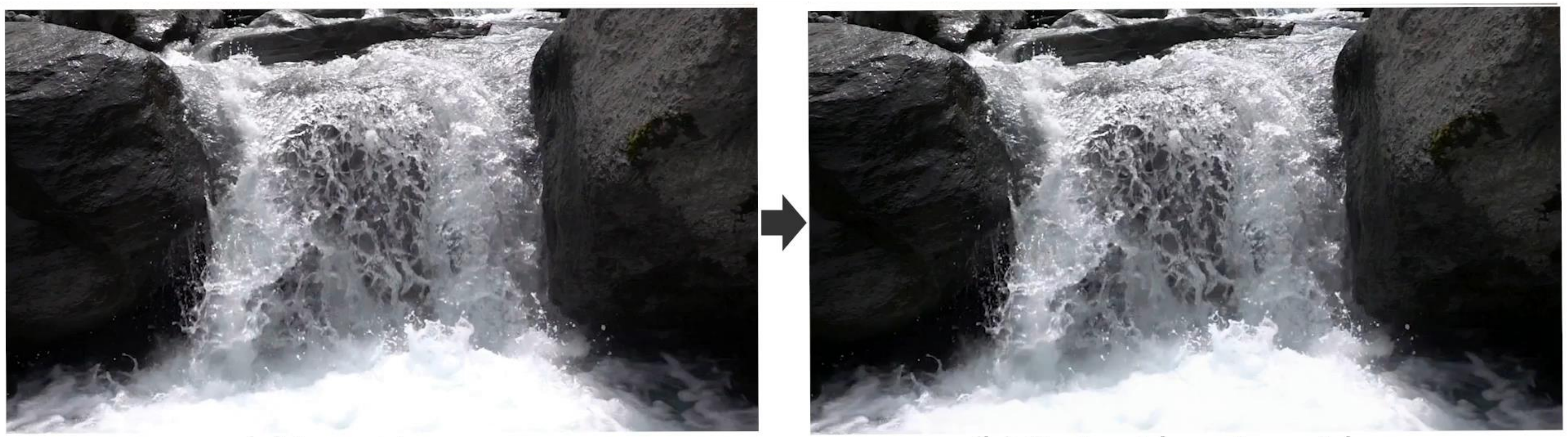


Animating Pictures

Animating Pictures with Eulerian Motion Fields

Aleksander Holynski[1], Brian Curless[1], Steven M. Seitz[1], Richard Szeliski[2]
[1]University of Washington, [2]Facebook

Paper | arXiv | Video | Code (coming soon)



(a) Input image (b) Output looping video

https://eulerian.cs.washington.edu/

Animating Pictures

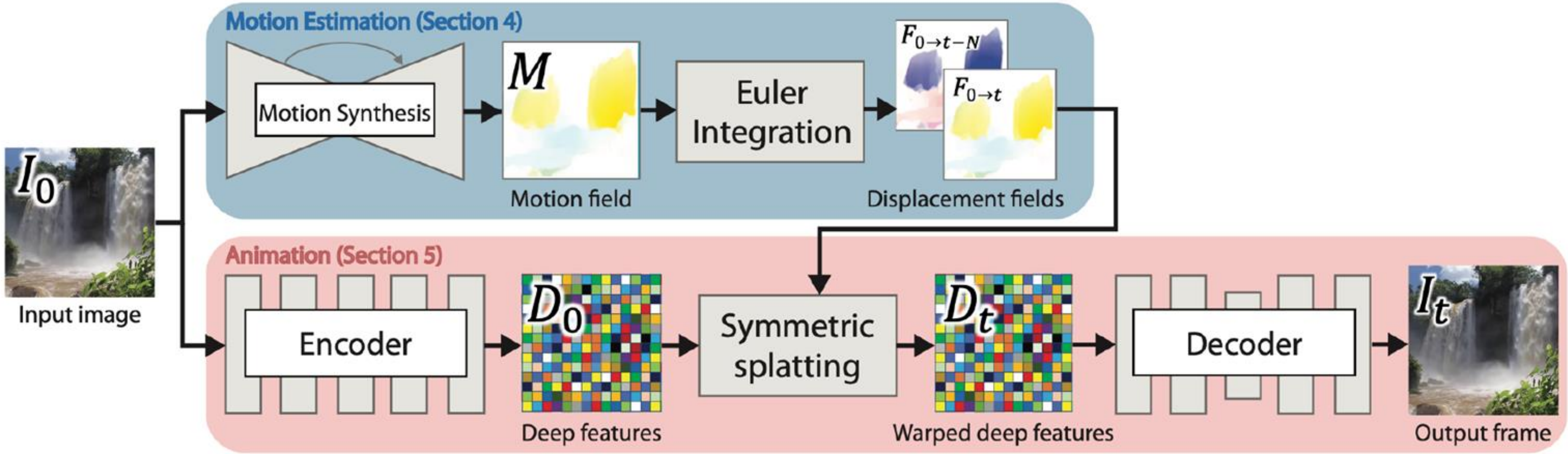


Figure 2: **Overview:** Given an input image I_0, our motion estimation network predicts a motion field M. Through Euler integration, M is used to generate future and past displacement fields $F_{0\to t}$ and $F_{0\to t-N}$, which define the source pixel locations in all other frames t. To animate the input image using our estimated motion, we first use a feature encoder network to encode the image as a feature map D_0. This feature map is warped by the displacement fields (using a novel symmetric splatting technique) to produce the corresponding warped feature map D_t. The warped features are provided to the decoder network to create the output video frame I_t.

Animating Pictures

Animating Pictures

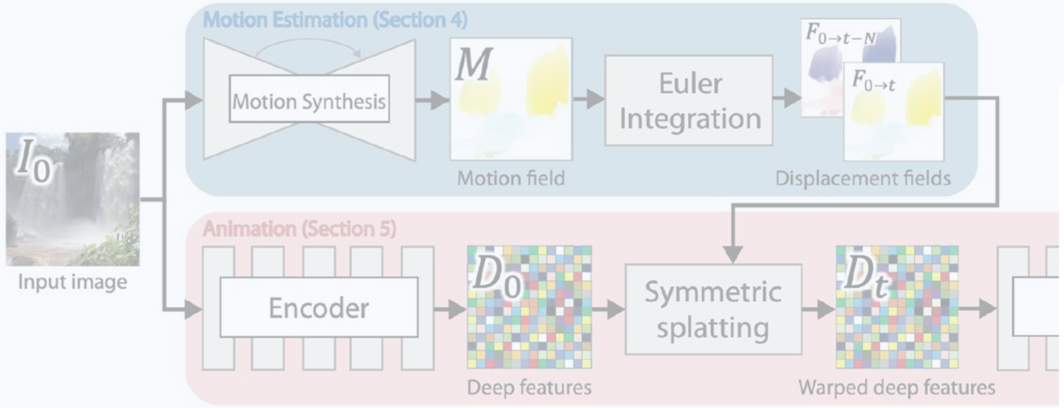


Figure 2: **Overview:** Given an input image I_0, our motion estimation network predicts a motion field
is used to generate future and past displacement fields $F_{0\to t}$ and $F_{0\to t-N}$, which define the source pi
To animate the input image using our estimated motion, we first use a feature encoder network to encod
This feature map is warped by the displacement fields (using a novel symmetric splatting technique) to p
feature map D_t. The warped features are provided to the decoder network to create the output video fra

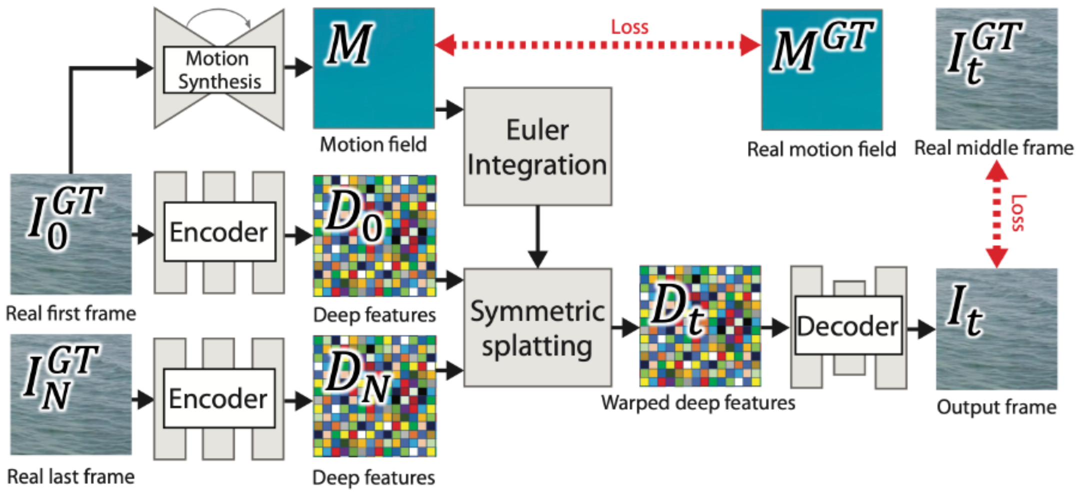
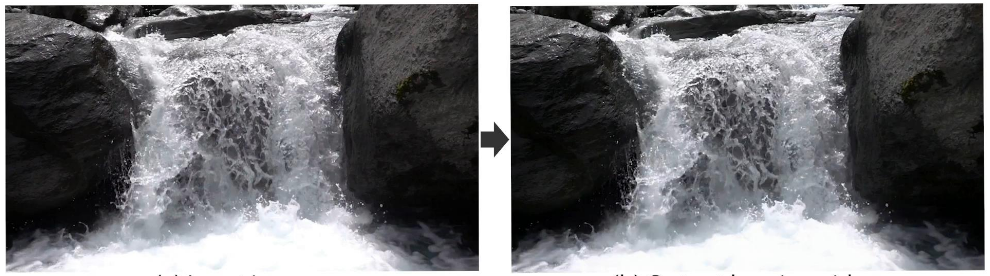


Figure 5: **Training:** As described in Section 5.1, each frame in our generated looping video is composed of textures from two warped frames. To supervise this process during training, i.e., to have a real frame to compare against, we perform our symmetric splatting using the features from two different frames, I_0 and I_N (instead of I_0 twice, as in inference). We enforce the motion field M to match

… wrapping up …

Outline

- Multi-view stereo
- Image-Based Rendering
  - Lumigraphs, Light Fields, Sprites with Depth, and Layers
- Virtual Viewpoint Video
- 360° and 3D Video
- 3D Photos
- Reflections and transparency
- Neural rendering



(a) Input image (b) Output looping video


TEXTS IN COMPUTER SCIENCE
Computer Vision
Algorithms and Applications
Second Edition
Richard Szeliski
Springer
1
2
3
4
5
6

Thank you