Vid 0

BRINGING CINEMATIC RENDERING INTO CASUALLY-TAKEN PHOTOS AND VIDEOS

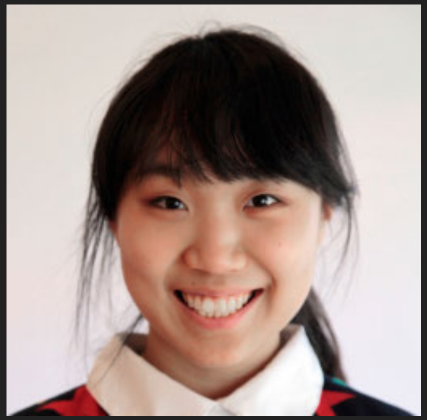
To Make Casual Imaging Context-Aware

Cecilia Zhang

Advisor: Ren Ng

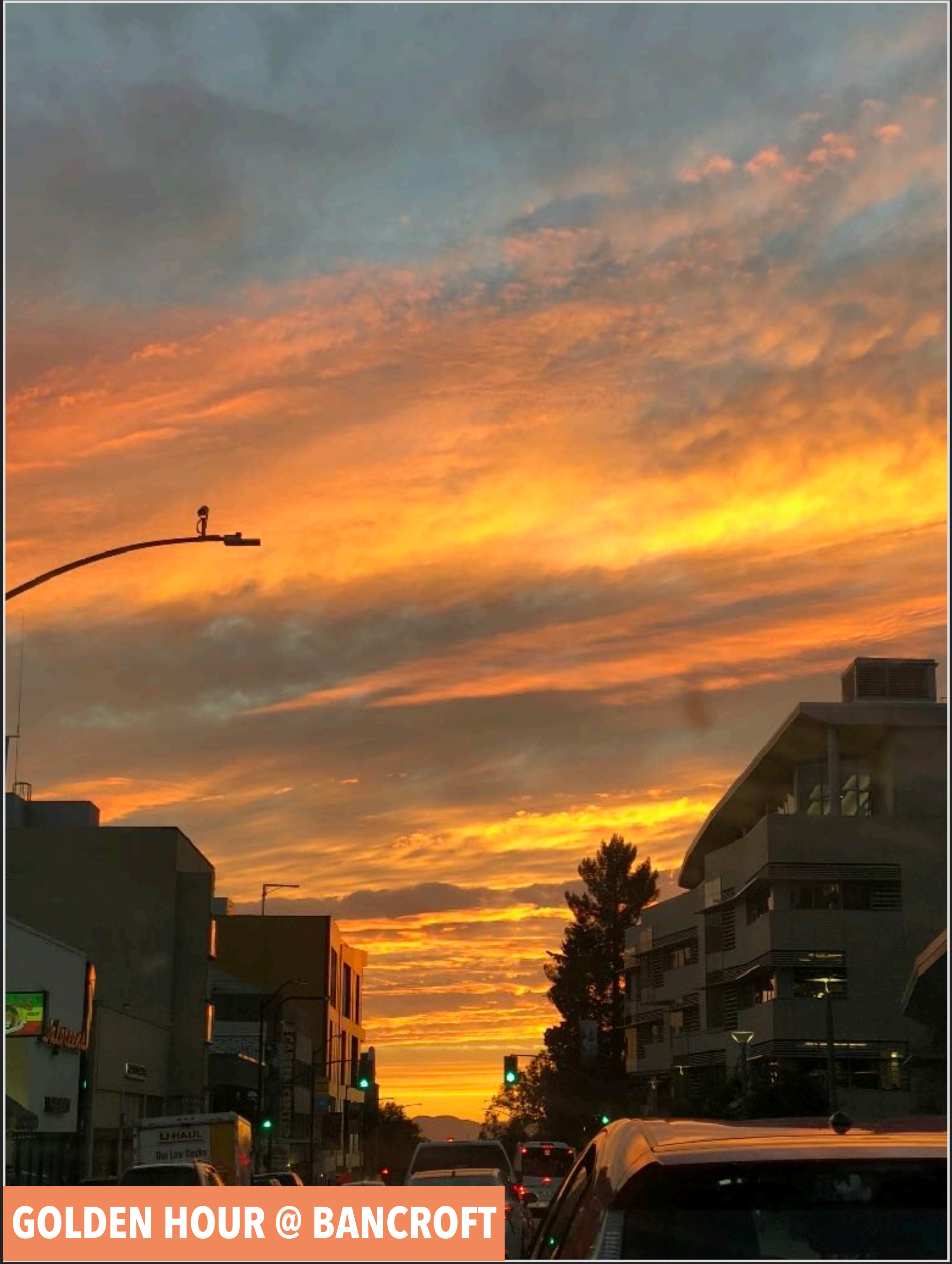
EECS, UC Berkeley

Committee: Alexei Efros, Martin Banks, Ren Ng




Profoto
Umbrella Deep Silver M
Profoto
OCF Softbox 2' Octa
Profoto
Profoto
Profoto
Canon
Canon
SONY
SONY

Photography Becomes a Casual Activity



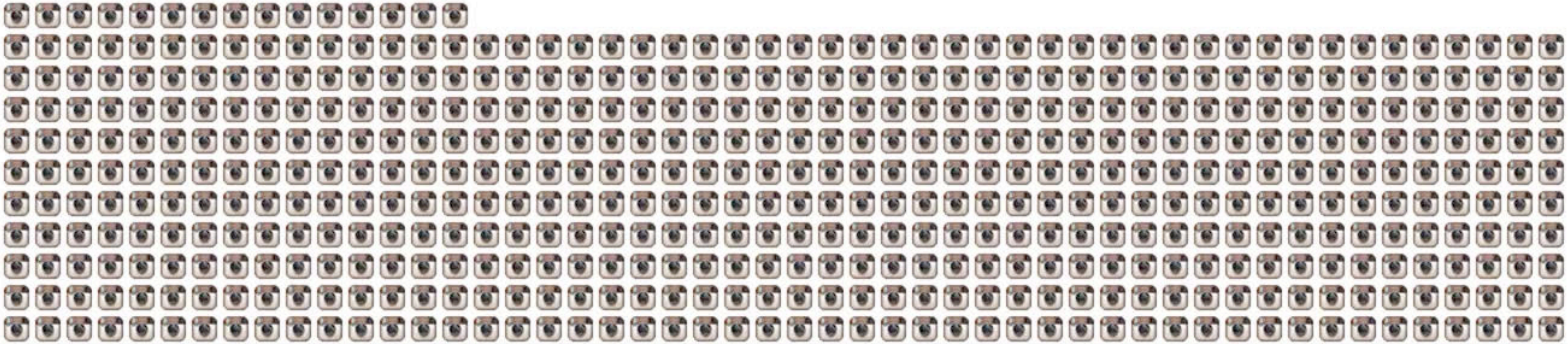
GOLDEN HOUR @ BANCROFT

VINTAGE CAR @ BANCROFT

WALL FLOWER @ SIBLEY

1,015 Instagram photos uploaded in 1 second

4,587 New photos since opening this page
0:00:03 seconds ago

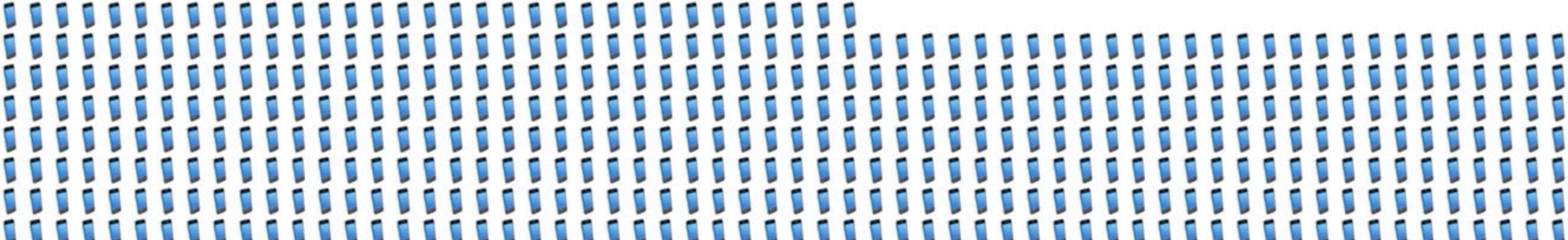


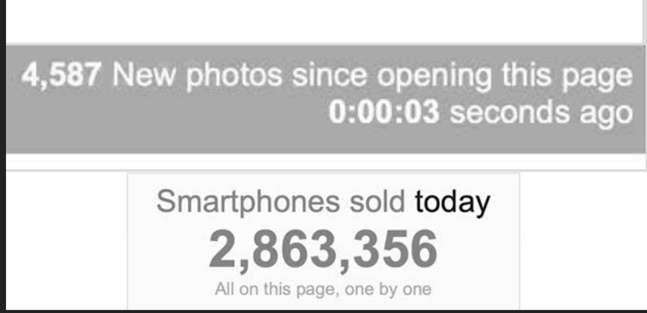
Smartphones sold **today**

2,863,356

All on this page, one by one

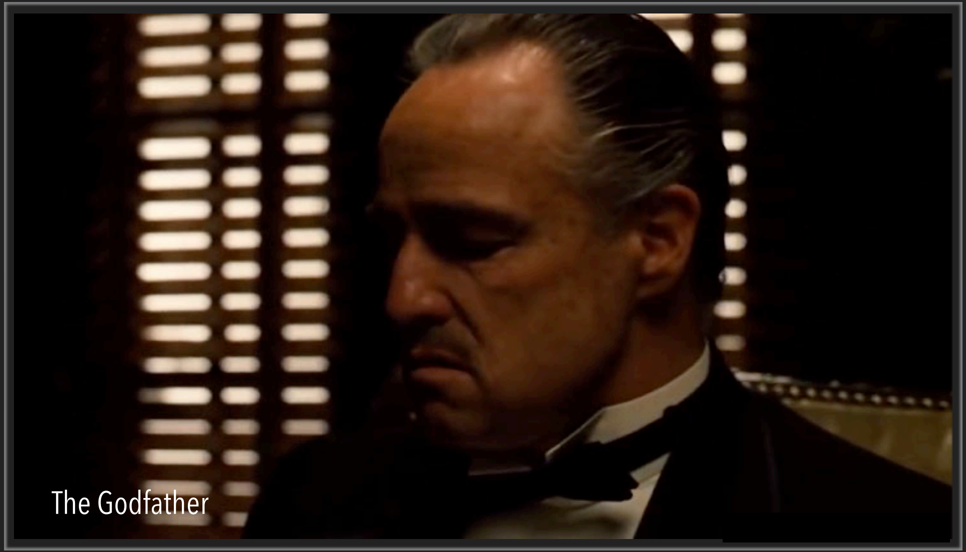
scroll ↓





- Visual Human Beings
- Cheap Cloud Storage
- High Speed Telecommunication

Film — the Best Visual Storytelling Formats



What Makes the Cinematic Look

Image Quality

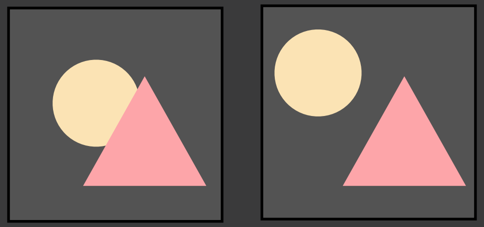
Focus

Lighting

Other Cinematic-Look

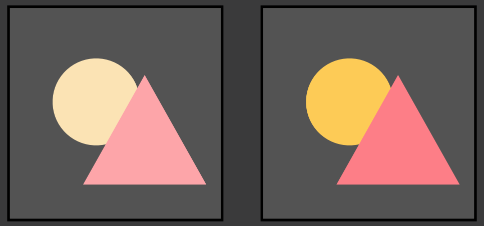
Composition

– how the elements on screen appear in respect to each other and within the frame itself.



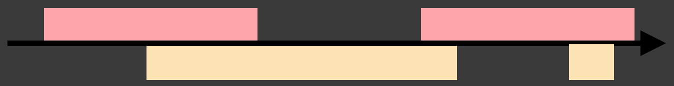
Color Grading

– the process of creatively altering the appearance of your footage to give it the desired emotional impact.



Timeline Editing

– cuts from one shot to another to create a seamless finished product, with no trace of their tampering.



Sound

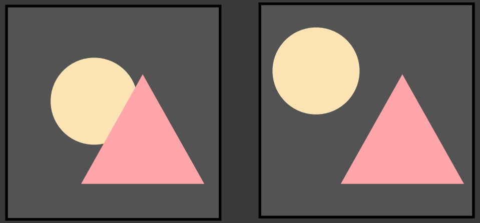
– how to reinforce the drama and lead the audience's emotions.

Other Cinematic Aspects

Composition

– Badki et al, “Computational Zoom: A Framework for Post-Capture Image Composition”, SIGGRAPH **2017**

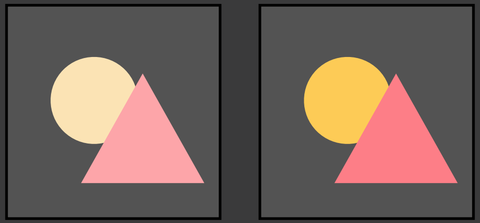
– Niklaus et al, “3D Ken Burns Effect from a Single Image”, SIGGRAPH Asia **2019**



Color Grading

– Kang et al. "High dynamic range video", *SIGGRAPH*, **2003**

– Bonneel et al, “Example-based Video Color Grading”, SIGGRAPH Asia, **2013**

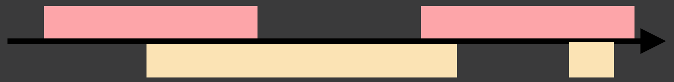


Timeline Editing

– Berthouzoz et al, “Tools for placing cuts and transitions in interview video ”, SIGGRAPH **2012**

– Truong et al. “QuickCut: An Interactive Tool for Editing Narrated Video” UIST **2016**

– Leake et al, “Computational video editing for dialogue-driven scenes”, SIGGRAPH **2017**



Sound

– Rubin et al, “Content-based tools for editing audio stories ”, UIST **2013**

– Rubin et al, “Generating emotionally relevant musical scores for audio stories”, UIST **2014**

What Makes the Cinematic Look

Image Quality

Focus

Lighting

Image Quality Focus Lighting

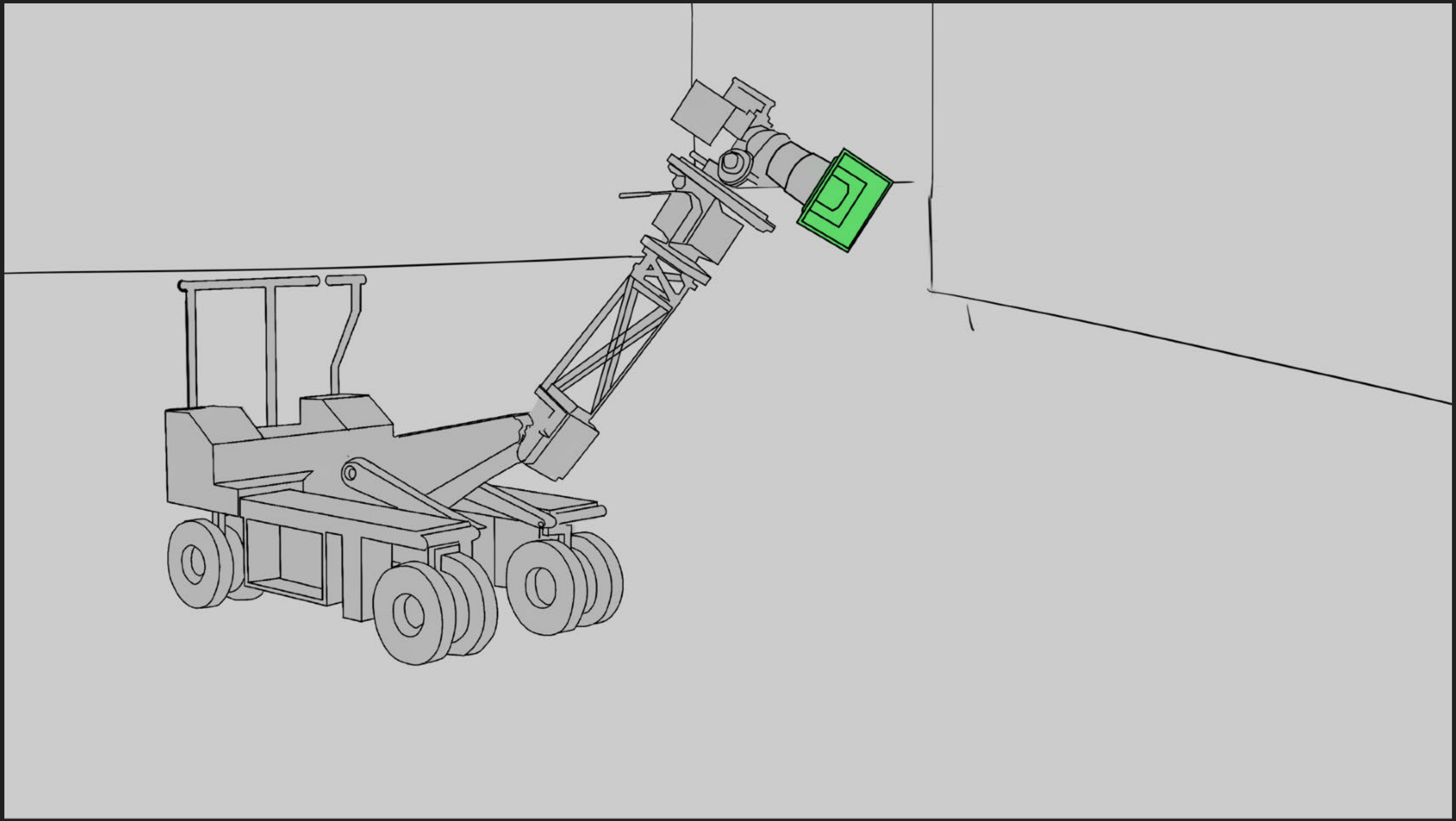


Image Quality Focus Lighting

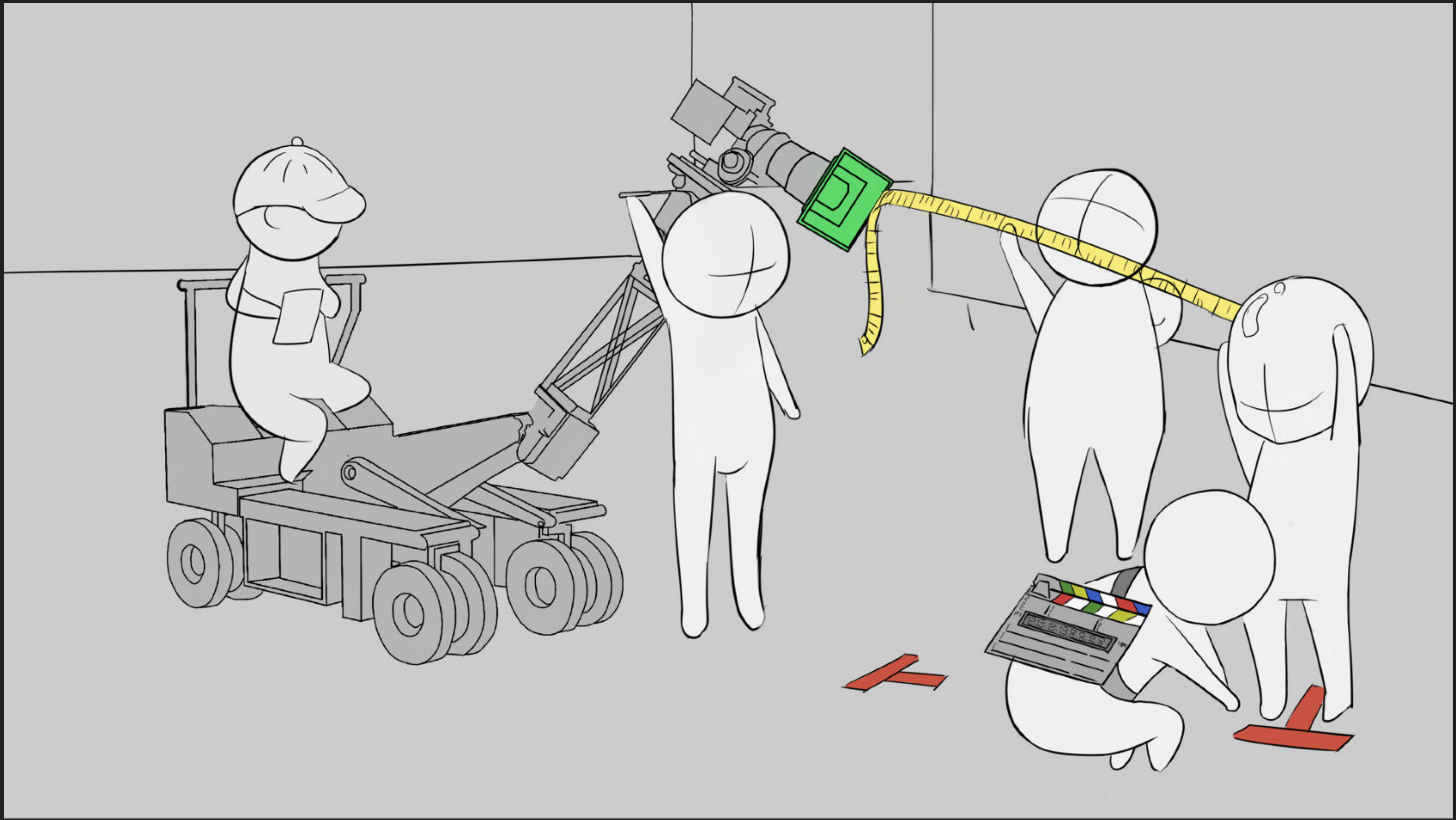


Image Quality   Focus   Lighting

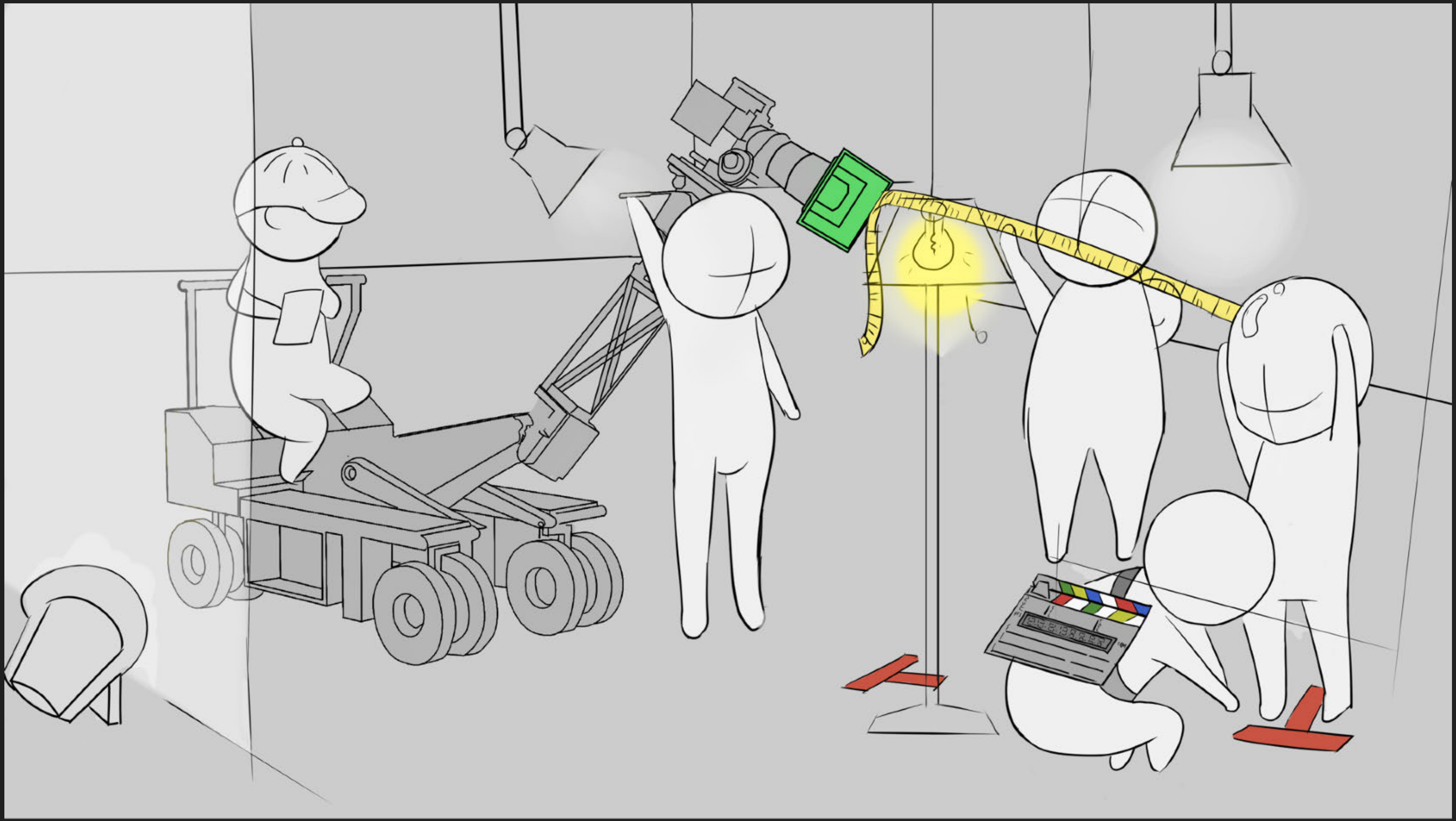
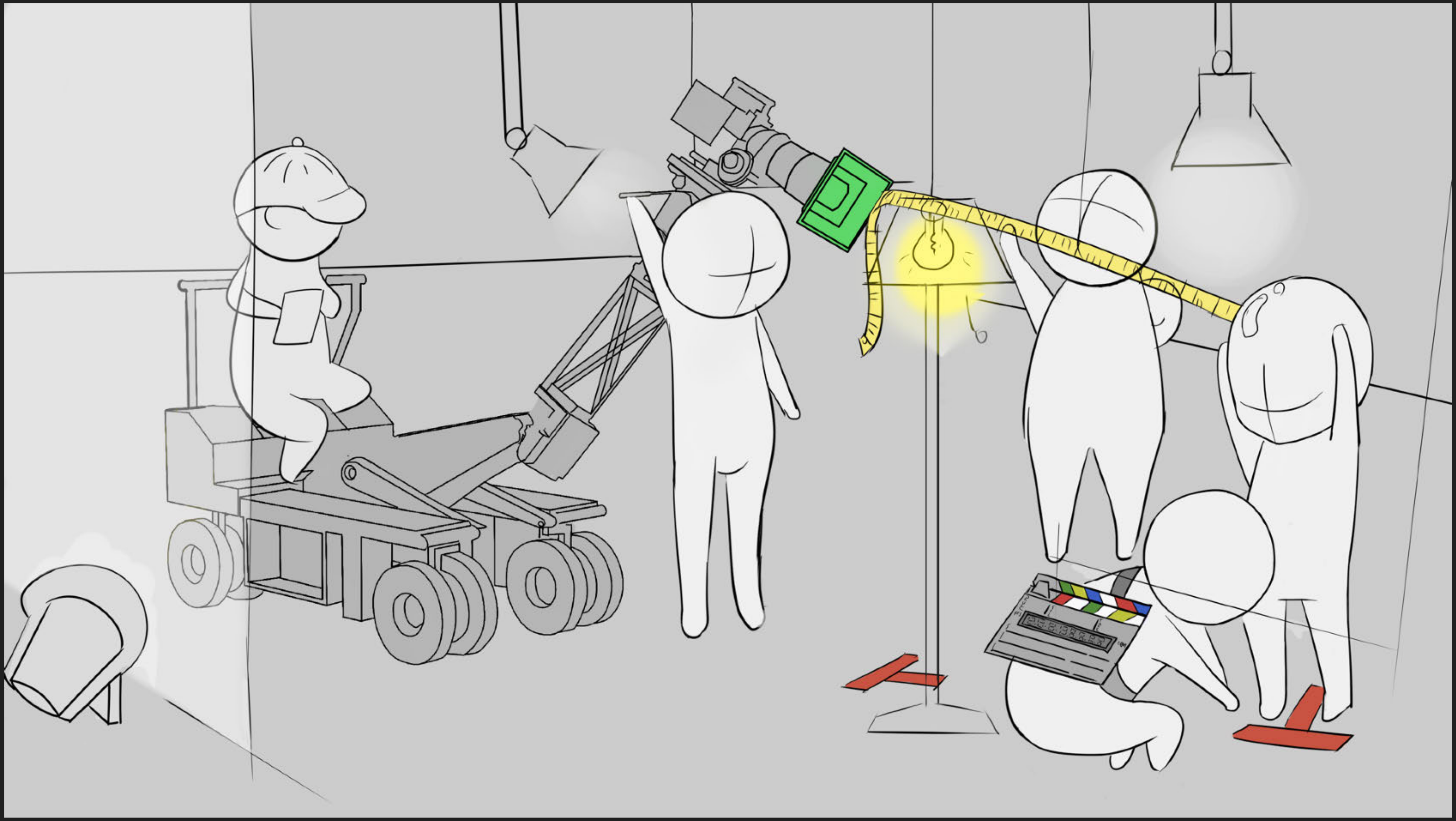


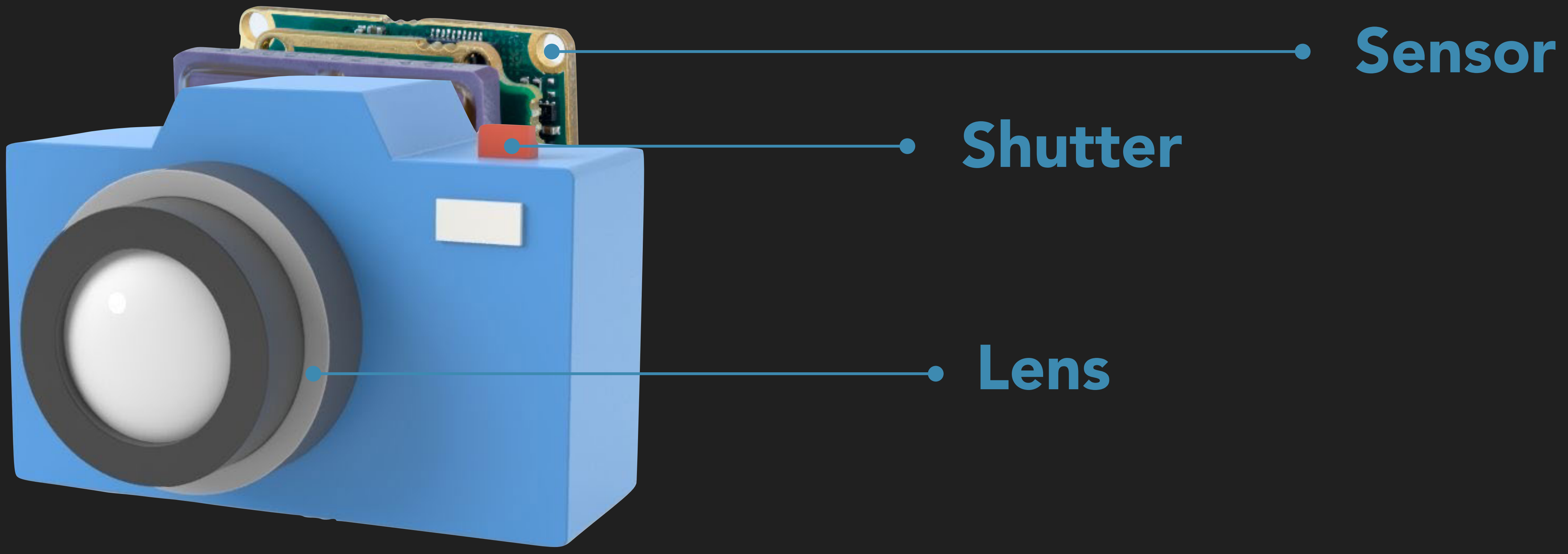
Image Quality

Focus

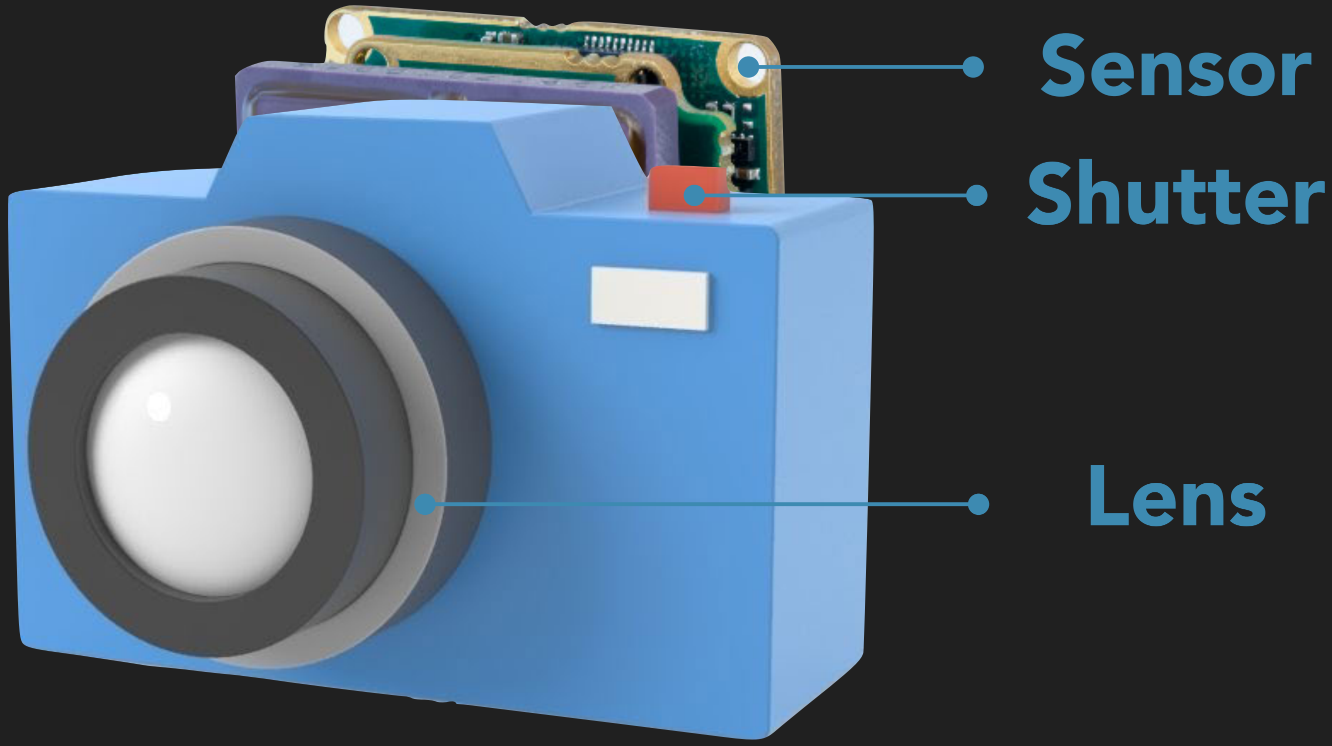
Lighting



What We Have in a Casual Imaging Device



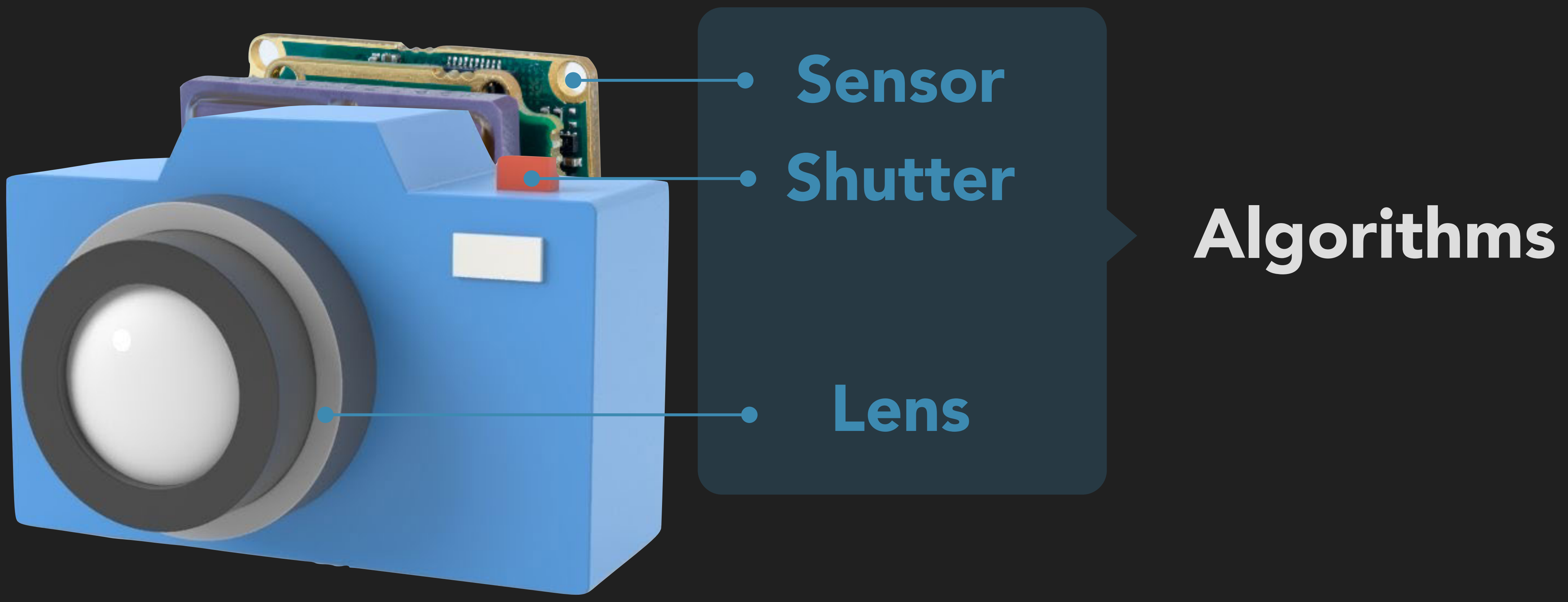
Lens, Shutter, Sensor



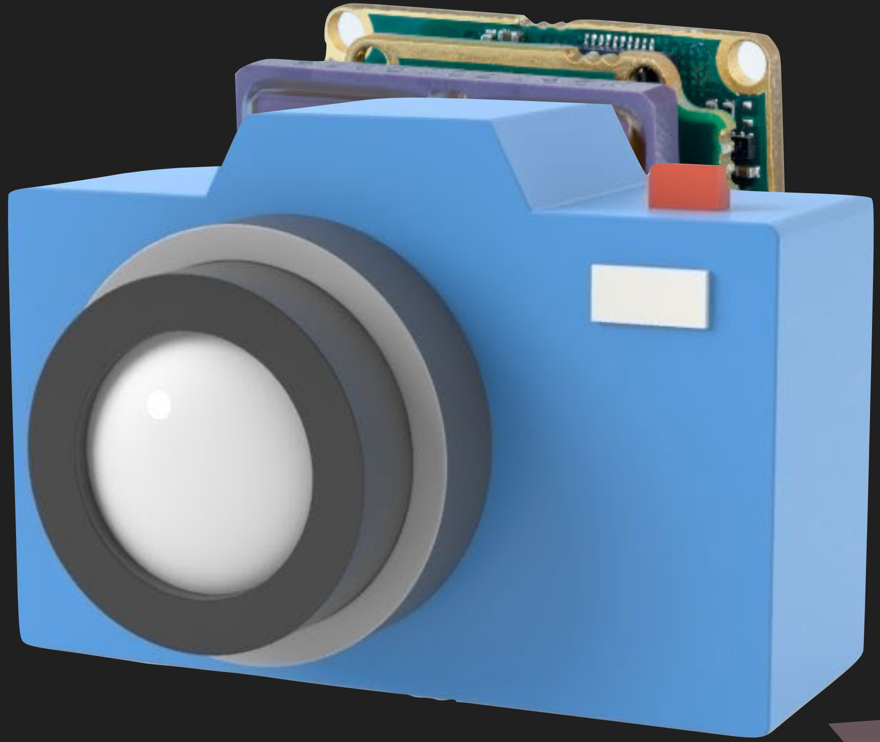
Film lighting

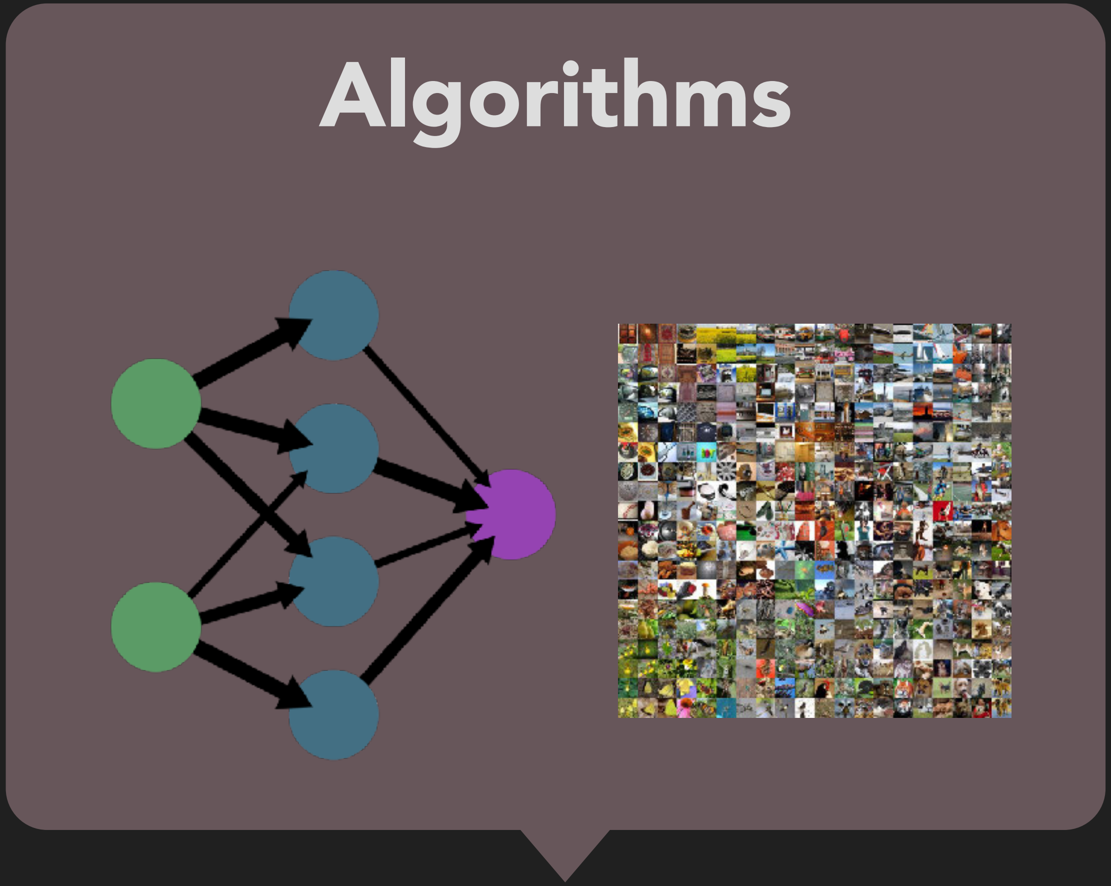
Custom camera system for making “The Irishman”

Lens, Shutter, Sensor, and Algorithms



Connecting Imaging and Context





Context

Connecting Imaging and Context

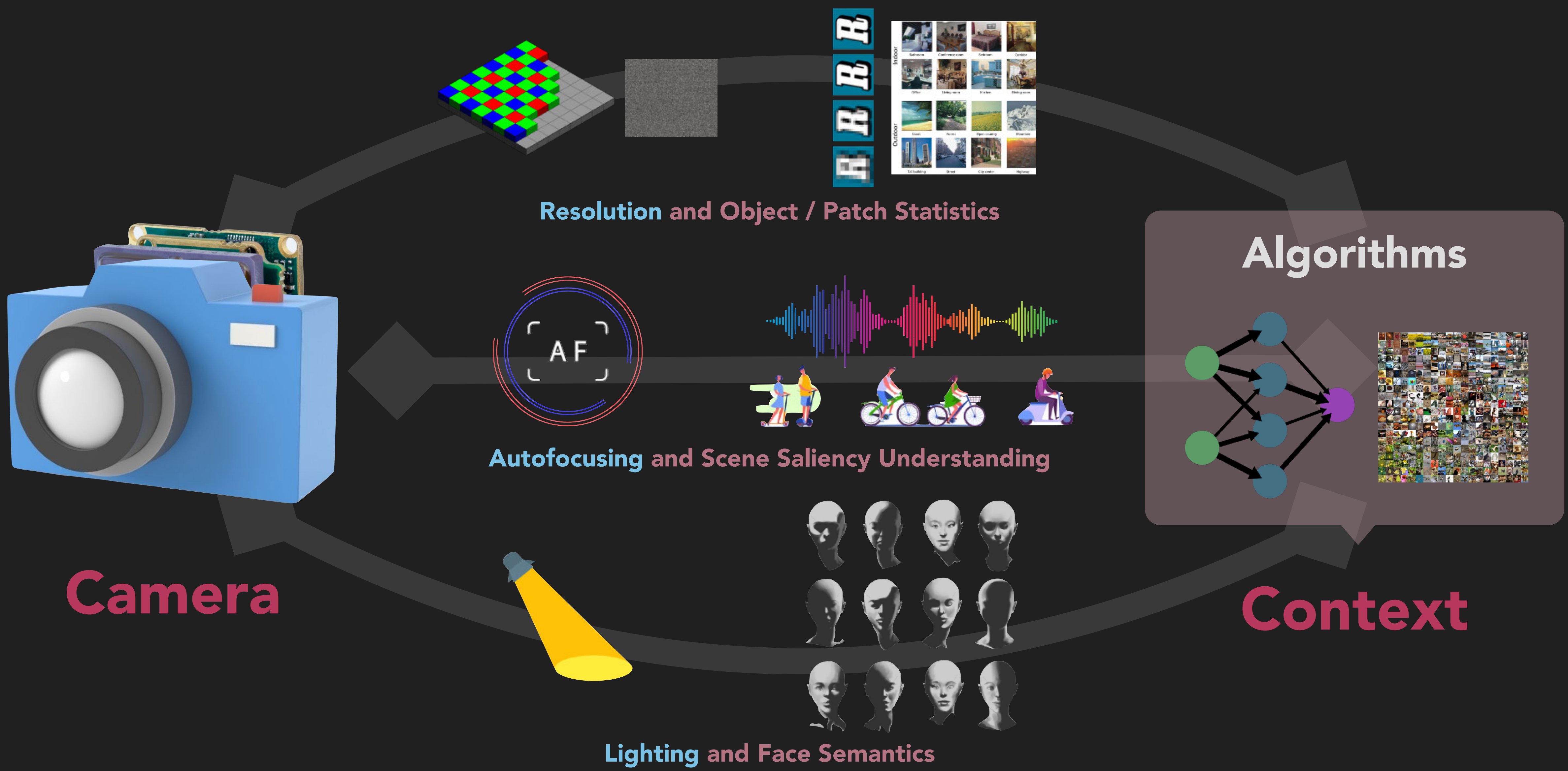


Image Quality

Zoom To Learn, Learn To Zoom

Zhang et al, CVPR 2019

- Learning From Raw Sensor
- Image Super-Resolution

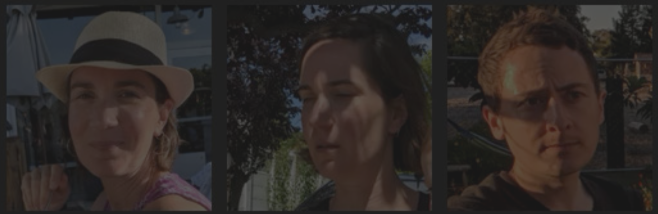
Focus

Synthetic Defocus and Look-Ahead Autofocus for Causal Videography

Zhang et al, SIGGRAPH 2019

- Video Synthetic Defocus
- 'Future' Scene Understanding

Lighting



Portrait Shadow Manipulation

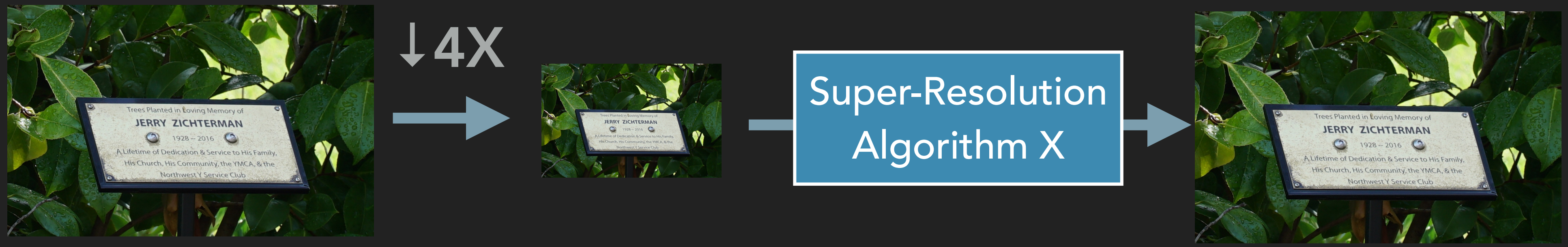
Zhang et al, SIGGRAPH 2020

- Shadow Editing Post-Capture
- Embedding Professional Lighting Principals

Super-Resolution — Digital Zoom

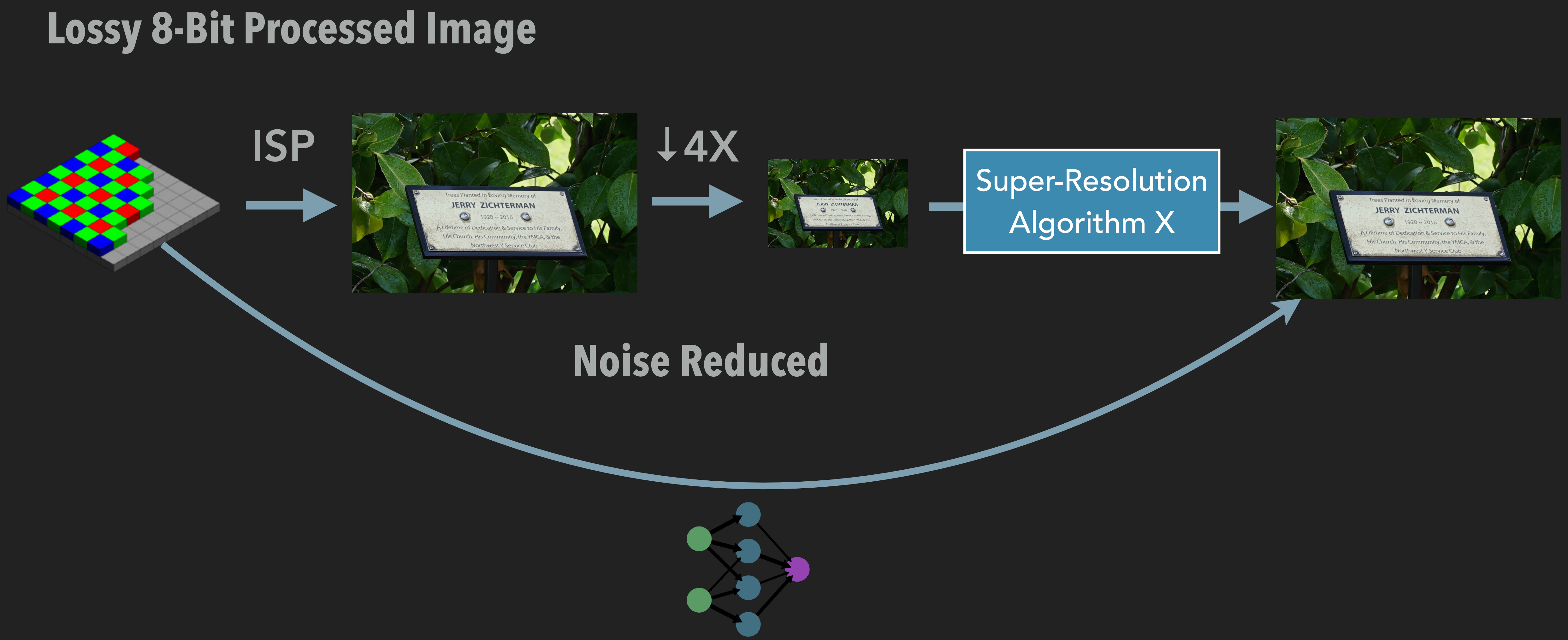
Super-Resolution — Digital Zoom

Problems in Synthetic Super-Resolution Setups



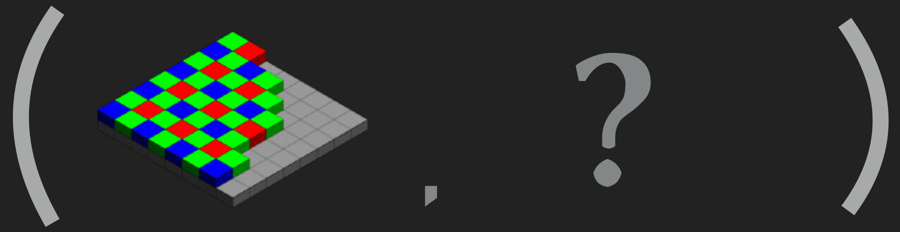
Noise Reduced

Problems in Synthetic Super-Resolution Setups

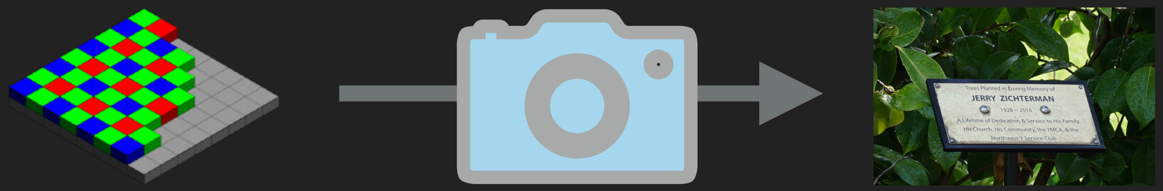


From Raw Sensor to Super-Resolved RGB

How Do We Get Ground Truth for Raw Sensor Data?



How Do We Do Better Than the Built-in Camera ISP?



The Upper Bound of Computational Zoom Is Optical Zoom

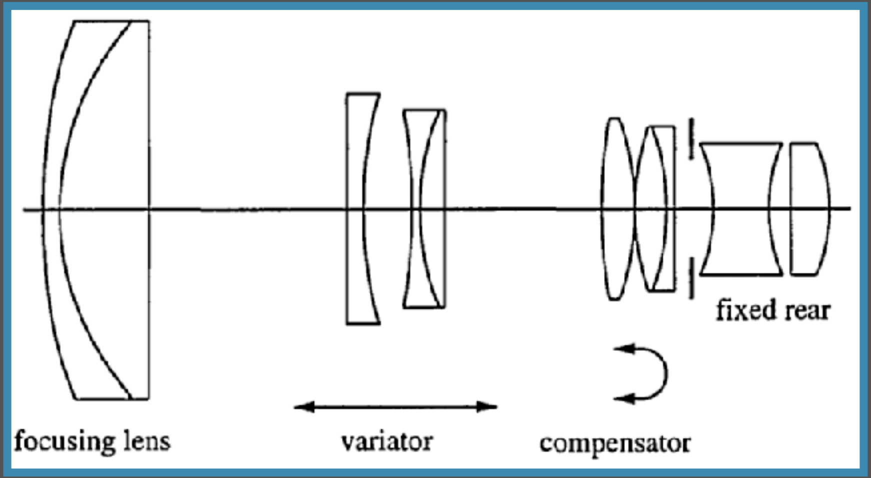
Vid 1

G.T. Capture With Optical Zoom

G.T. Capture With a Zoom Lens

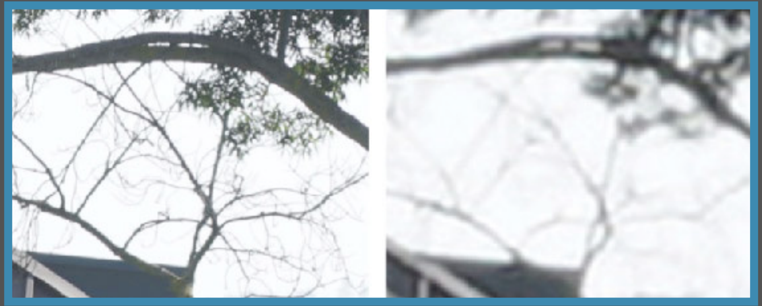
Sources of Slight Mis-Alignment

Change in Effective Camera Center



Willson, Reg G., and Steven A. Shafer. "What is the center of the image?" JOSA 1994

Aligning Different Resolution



A Loss Function for Slightly Unaligned Data Pairs

Image Enhancement

pixel-wise aligned

Contextual Style Transfer

contextually aligned

[Mechrez et al ECCV 2018]

Image Translation

unaligned

w_s

$$\mathbb{D}_{p_i,q_j} \begin{cases} (X_{pi},\ Y_{pi}) & \text{VGG Features} \\ (X_{qi},\ Y_{qi}) & \text{VGG Features} \end{cases}$$

Target Source

$$L(P, Q) = \frac{1}{N}\sum_{i}^{N} \min_{j=1,\dots,M} \left(\mathbb{D}_{p_i,q_j} + w_s \mathbb{D}_{p_i,q_j}\right)$$

Encoding Position Into Feature Vector

4X Results

Bicubic Upsampling

Contextual Loss

Contextual Bilateral Loss (CoBi)

Input (Bicubic Upsample)

ESRGAN


NO PERMIT
PARKING
NO PARKING
2:30 AM TO 6:00 AM
3 HOUR PARKING
6:00 AM TO 6:00 PM
UNRESTRICTED
6:00 PM TO 2:30 AM

Ours 4X


NO PERMIT
PARKING
NO PARKING
2:30 AM TO 6:00 AM
3 HOUR PARKING
6:00 AM TO 6:00 PM
UNRESTRICTED
6:00 PM TO 2:30 AM

4X Results

4X Results

4X Results


NO PARKING
8am - 12noon
1st
THURSDAYS FOR
STREET SWEEPING
Christie AVE
STOP

4X
Results


Lease
Colliers
Ken Meyersieck
510 433 5802
Aileen Dolby
510 433 5815

4X Results

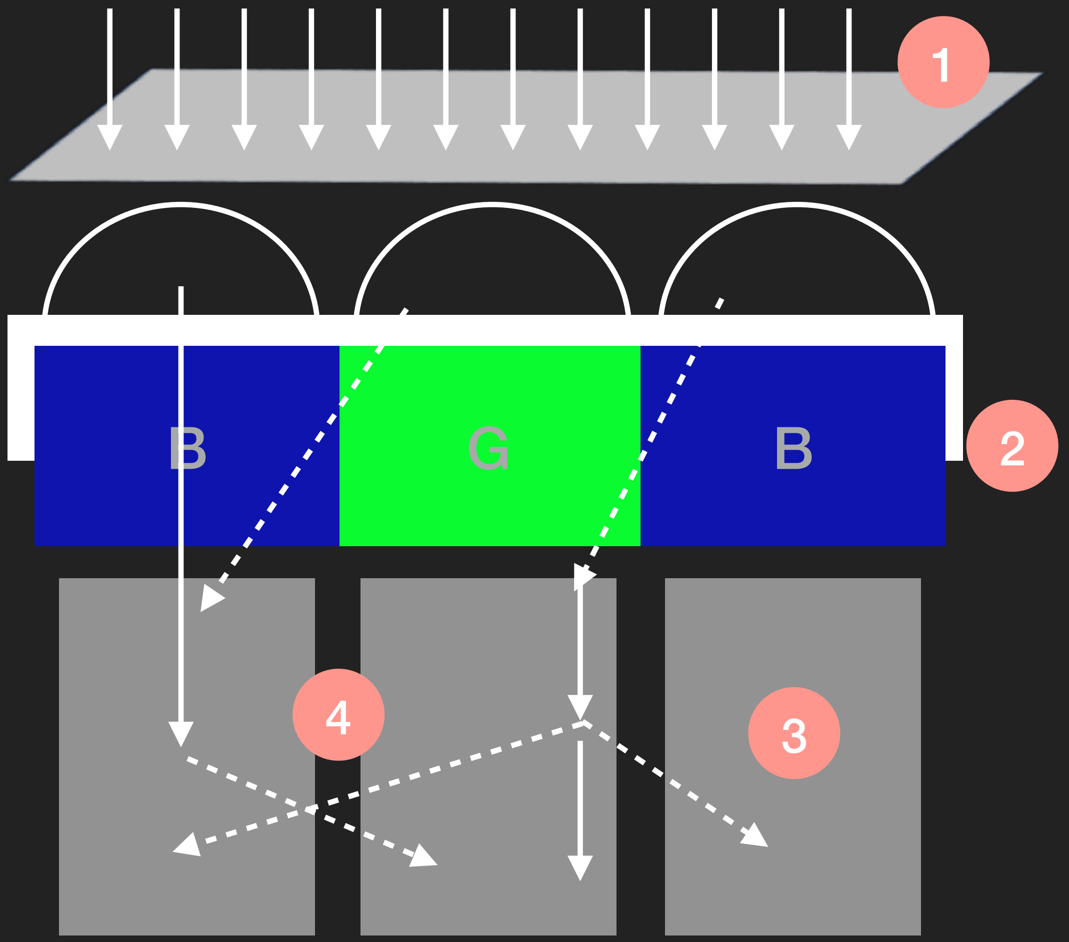

Lease
Colliers
Ken Meyersieck
510 433 5802
Aileen Dolby
510 433 5815

Input (Bicubic Upsample)

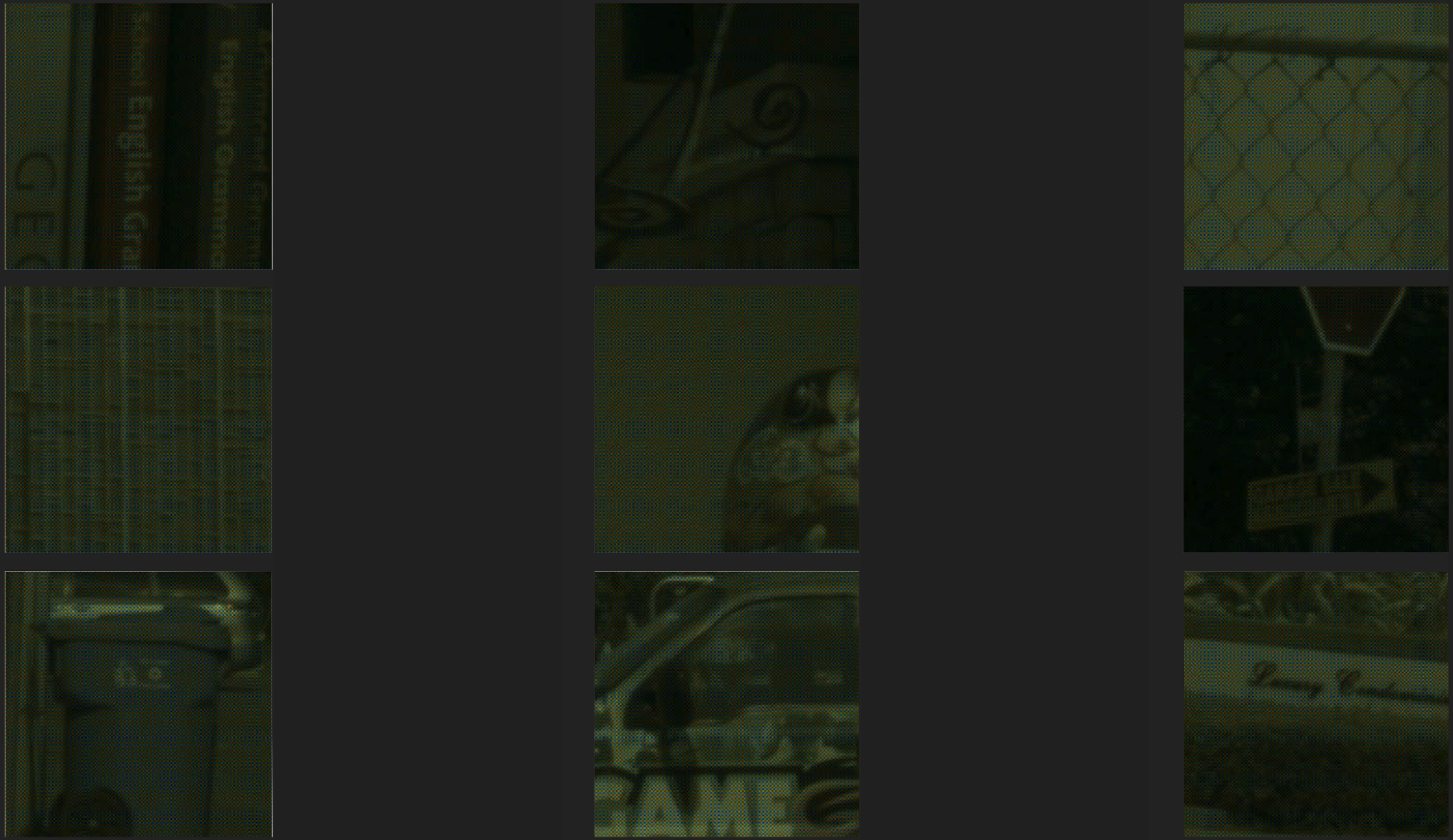
Train With Synthetic Sensor

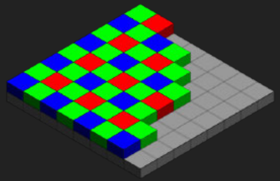
Train With Real Sensor (Ours)

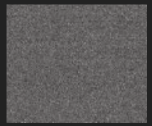
Raw Noise Is Difficult To Model

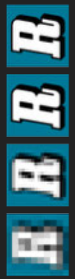


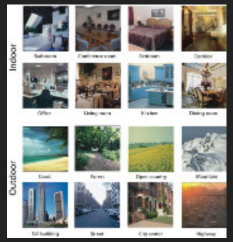
Features	Synthetic	Real
1 AA Filter	No	Yes / No
2 Bit Depth	8	12-14
3 Crosstalk	No	Yes
4 Fill Factor	100%	<100%



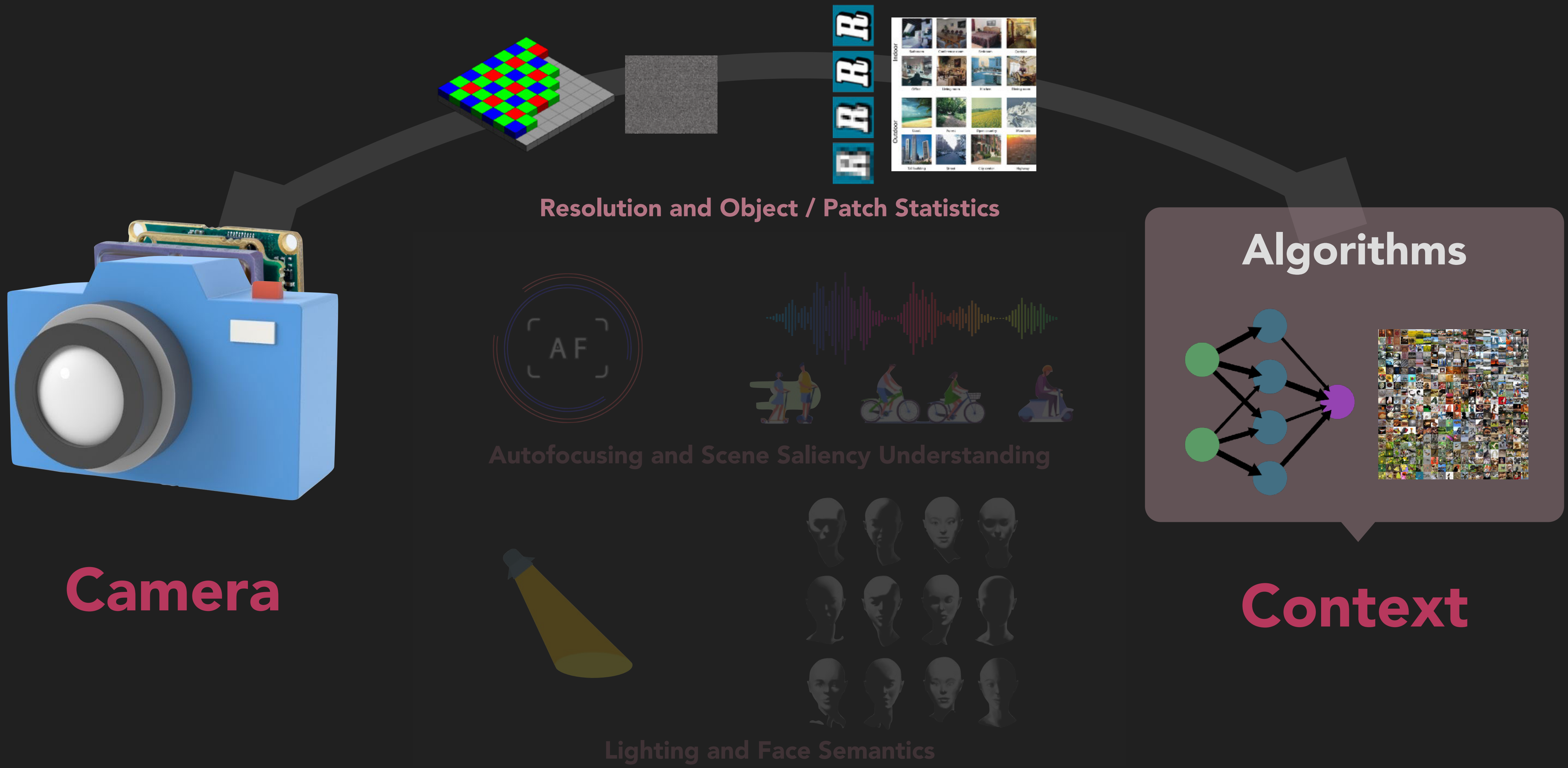


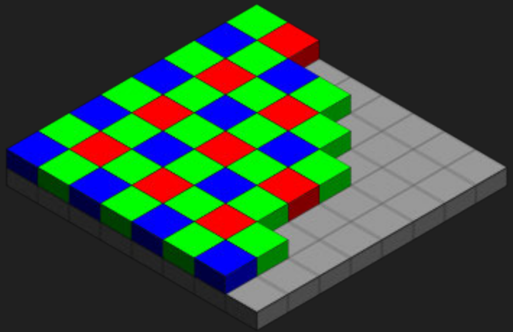


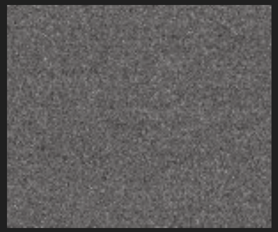




Context-Aware Casual Imaging







Sensor: 35.8 x 23.9 mm

Sensor: 5.12 x 3.84 mm

Image Quality

Zoom To Learn, Learn To Zoom

Zhang et al, CVPR 2019

- Learning From Raw Sensor
- Image Super-Resolution

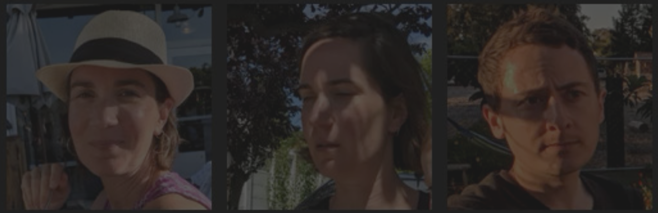
Focus

Synthetic Defocus and Look-Ahead Autofocus for Causal Videography

Zhang et al, SIGGRAPH 2019

- Video Synthetic Defocus
- 'Future' Scene Understanding

Lighting



Portrait Shadow Manipulation

Zhang et al, SIGGRAPH 2020

- Shadow Editing Post-Capture
- Embedding Professional Lighting Principals

Cinematic Focus is Appealing but Challenging

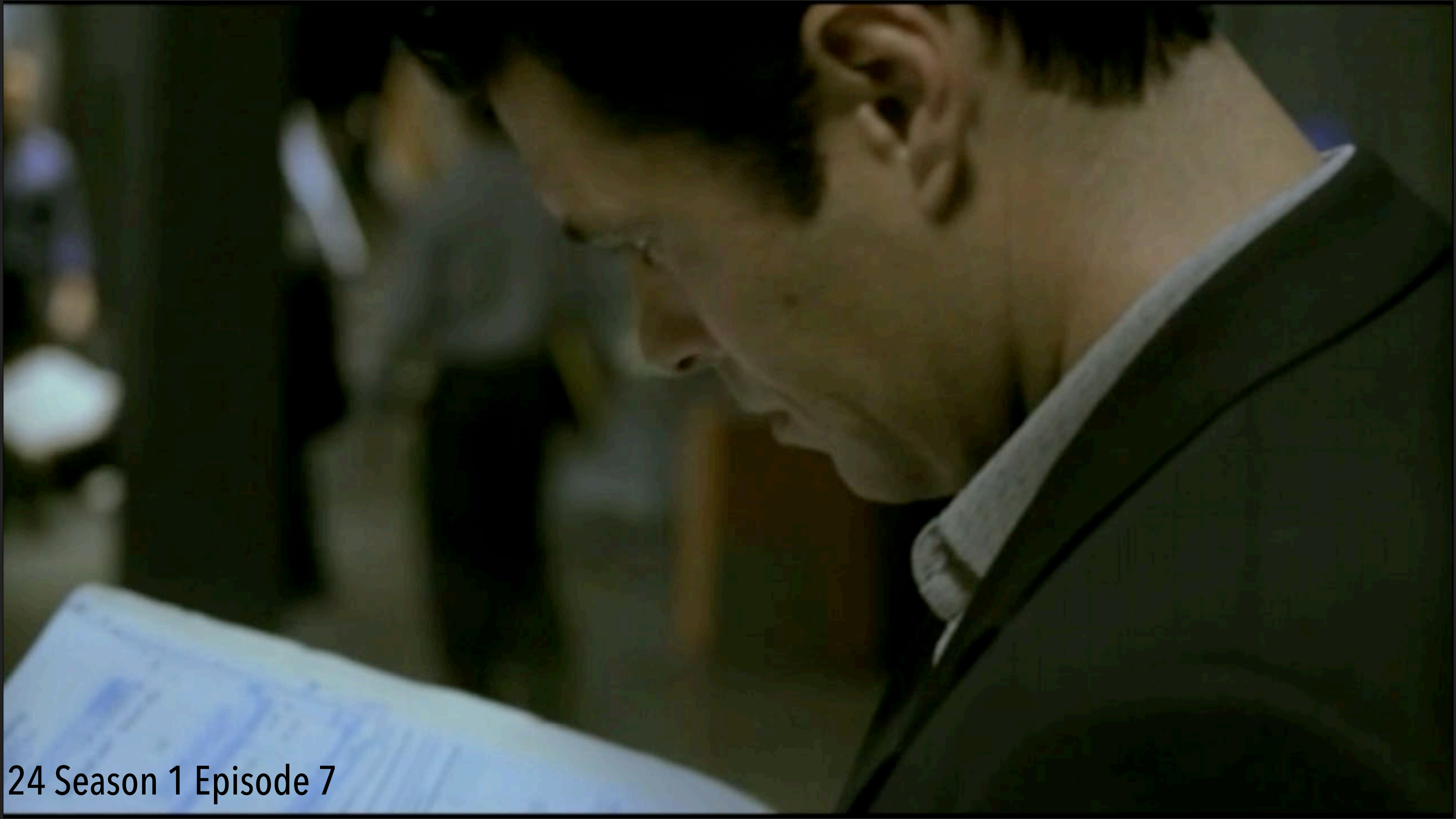
12 years a slave

Shallow Depth of Field (DOF) Is Visually Appealing

12 years a slave

Vid 2

Focus Guides the Viewer's Gaze



Vid 3

Casual Videos Rely on Autofocus, Always Have Focus Errors

Video credit: Ren Ng

Small Phone Cameras Don't Give Shallow DOF

Video credit: Ren Ng

OUR GOAL

Cinematic Focusing for Casual Video.

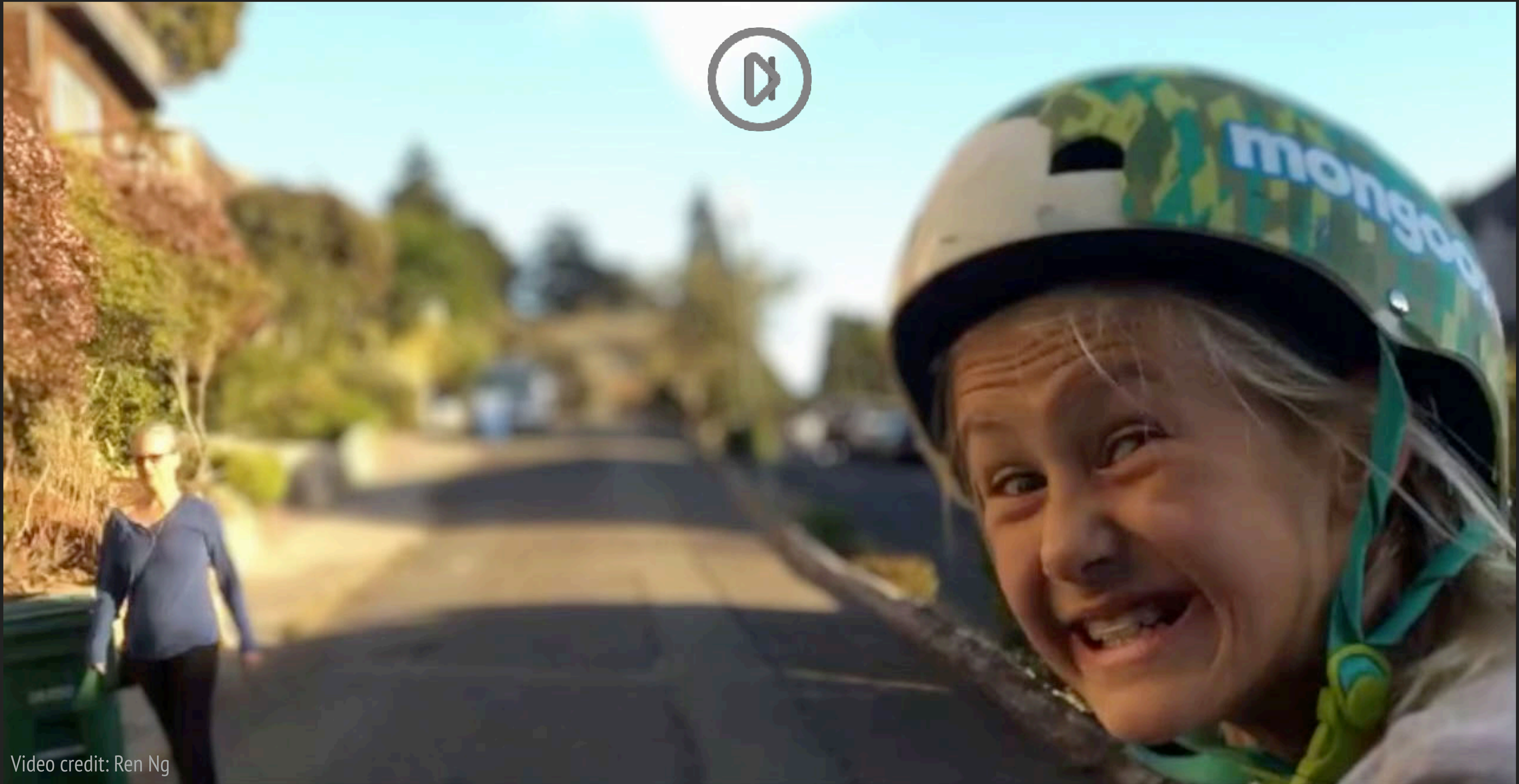
Vid 4

Input Casual Video

Video credit: Ren Ng

Vid 5

Output Autofocus Video From our System



Good Video Autofocus Needs Future Information

- We identify that good video autofocus needs information about the future
- We make that future information available to video autofocus for the first time
- We built a prototype system to demonstrate the potential of this approach

RVR

Making Videos Refocusable

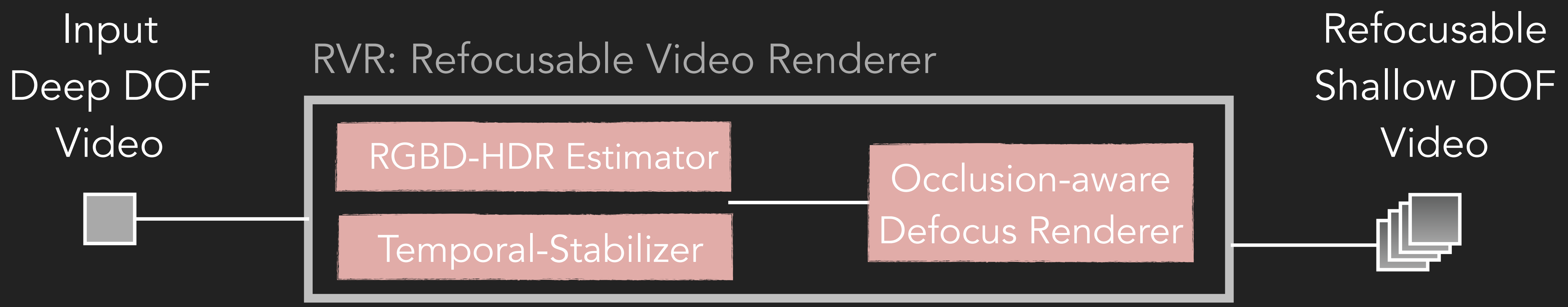
- Synthesizing shallow DOF from deep DOF videos

LAAF

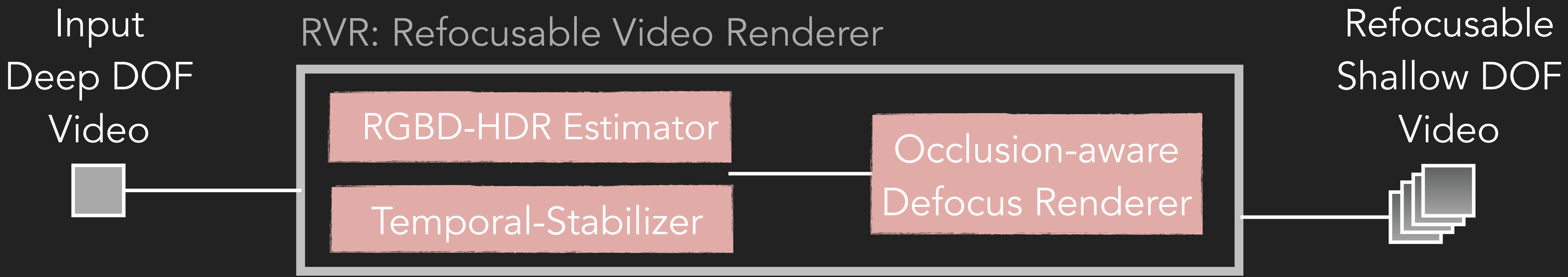
Look-Ahead Autofocus

- A class of autofocus algorithms that looks at future frames

RVR Making Videos Refocusable

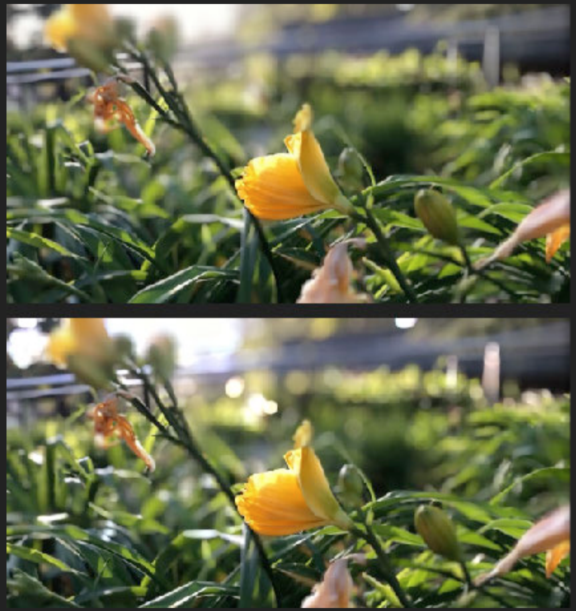


RVR Making Videos Refocusable



APERTURE DATASET

HDR RECOVERY



TEMPORAL COHERENCY

Vid 6

Shallow DOF Makes Casual Videos Appealing

When and Where To Focus?

Vid 7

RVR-Only: Shallow DOF From RVR Gives Wrong Focus!

RVR

Making Videos Refocusable

- Synthesizing shallow DOF from deep DOF videos

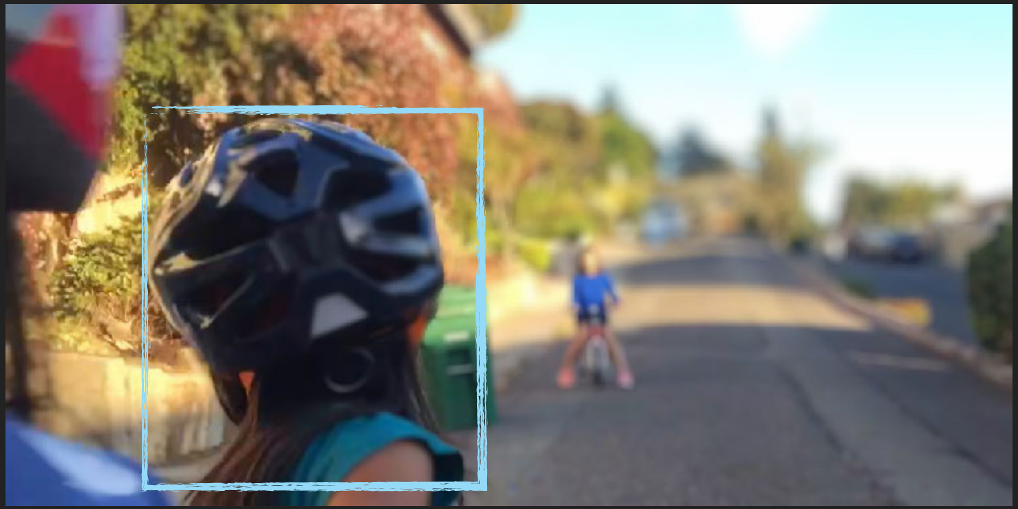
LAAF

Look-Ahead Autofocus

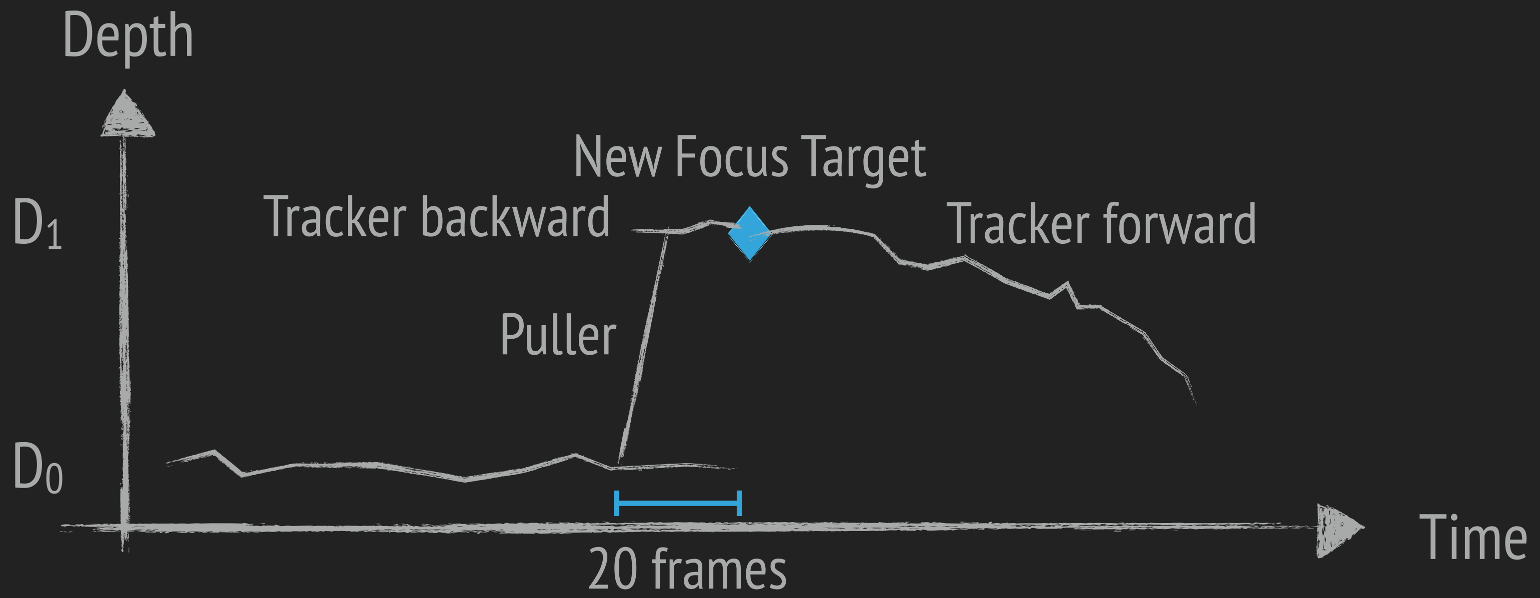
- A class of autofocus algorithms that looks at future frames

New Focus Target + Tracking + Focus Puller

Previous Focus Target



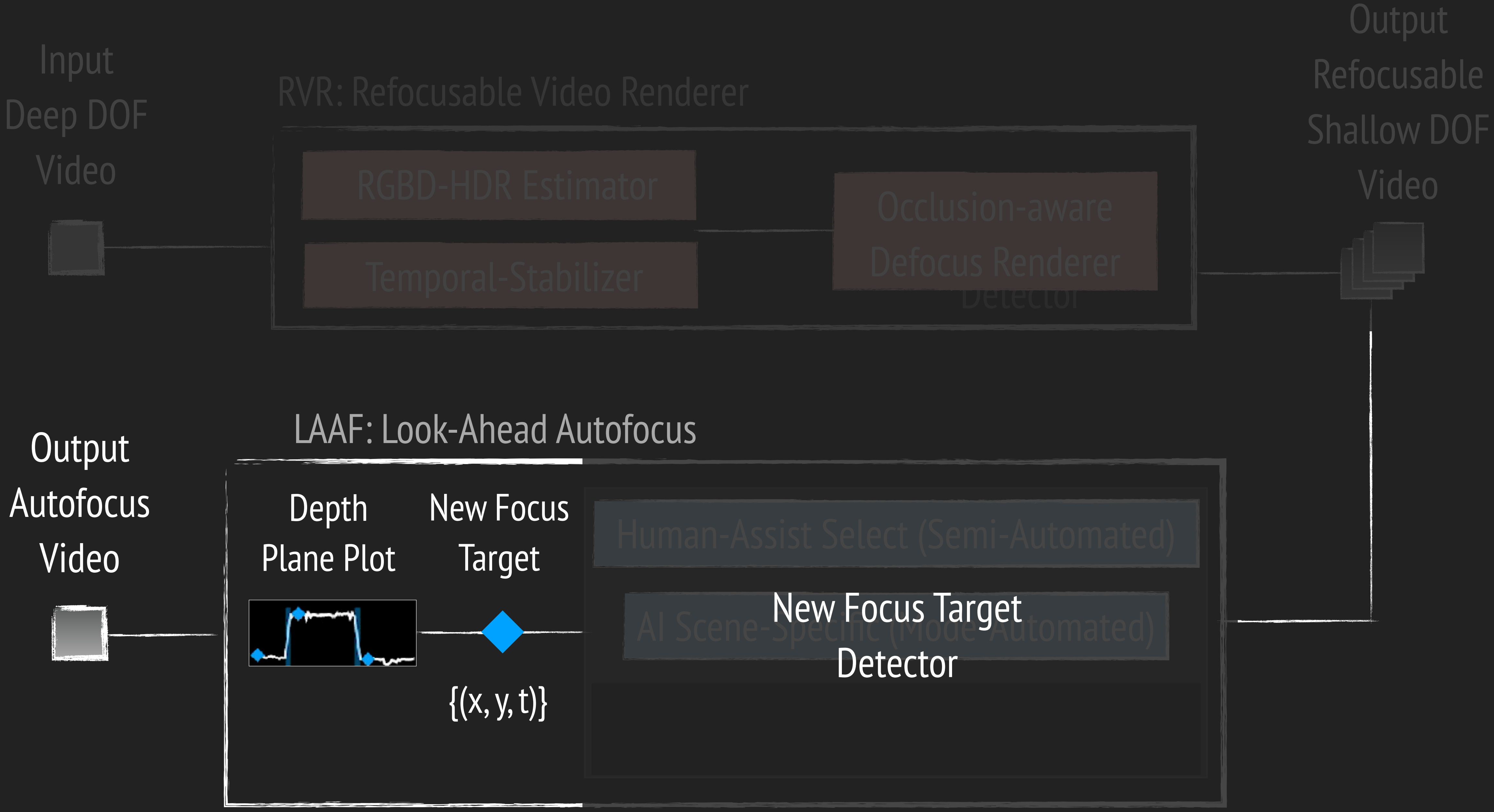
New Focus Target



Focus Puller

Previous Focus Target (baby) → New Focus Target (mom)

LAAF Tells When and Where To Focus



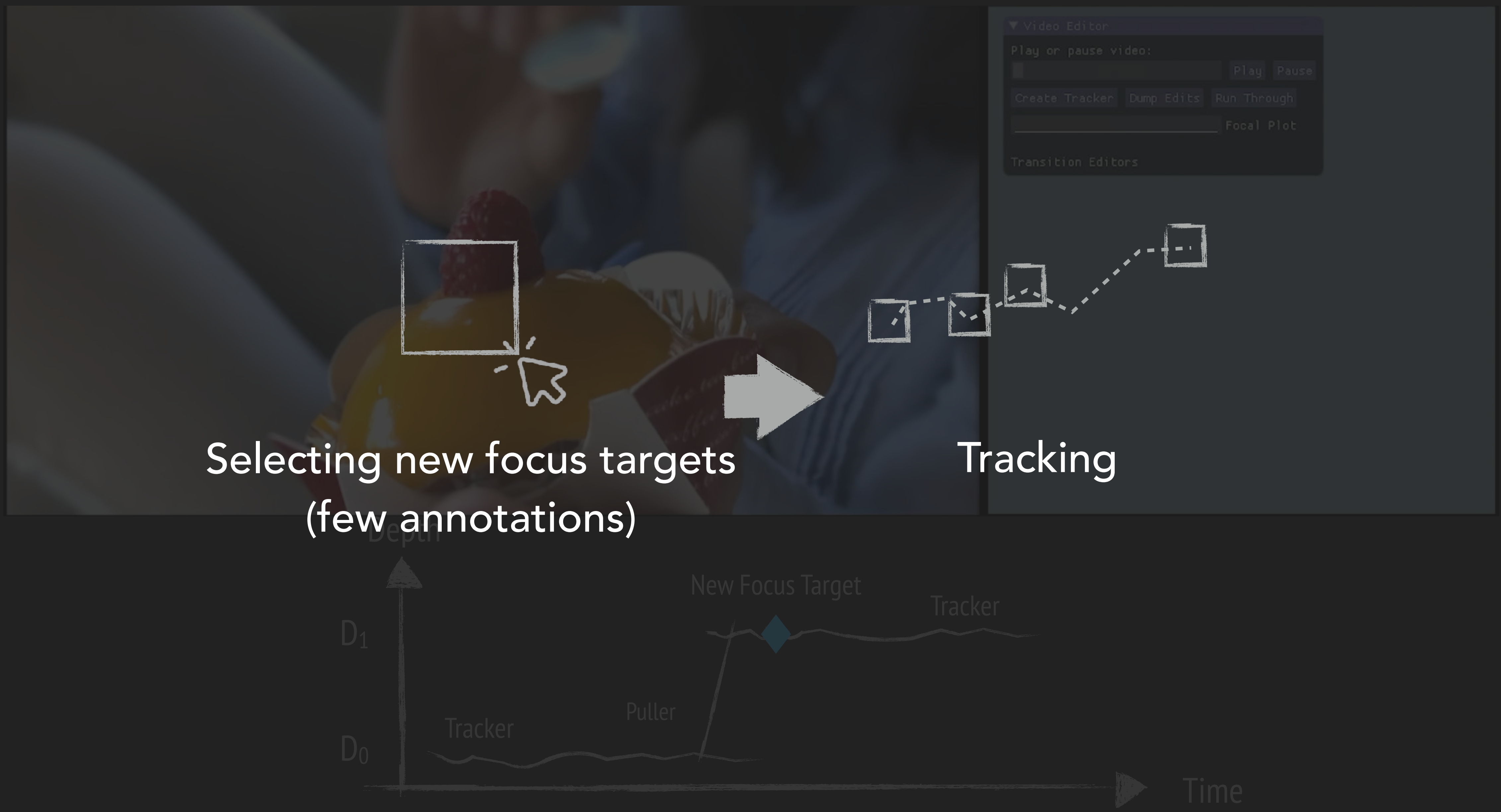
New Focus Target Detector

Human-Assisted Semi-Autofocus

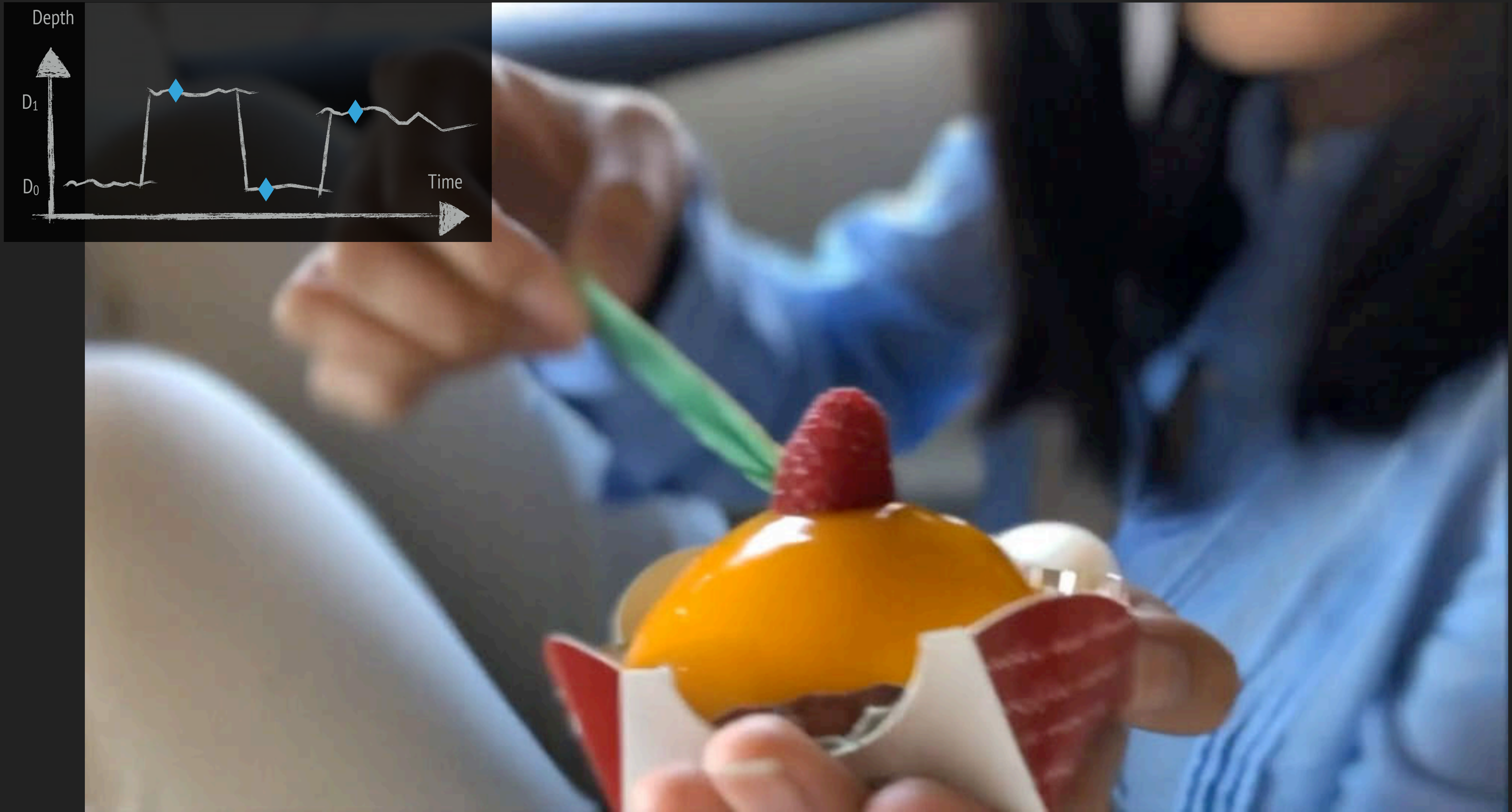
Scene-Specific Autofocus

Vid 8

GUI + Tracking To Get New Focus Targets



Autofocus Result With LAAF-GUI



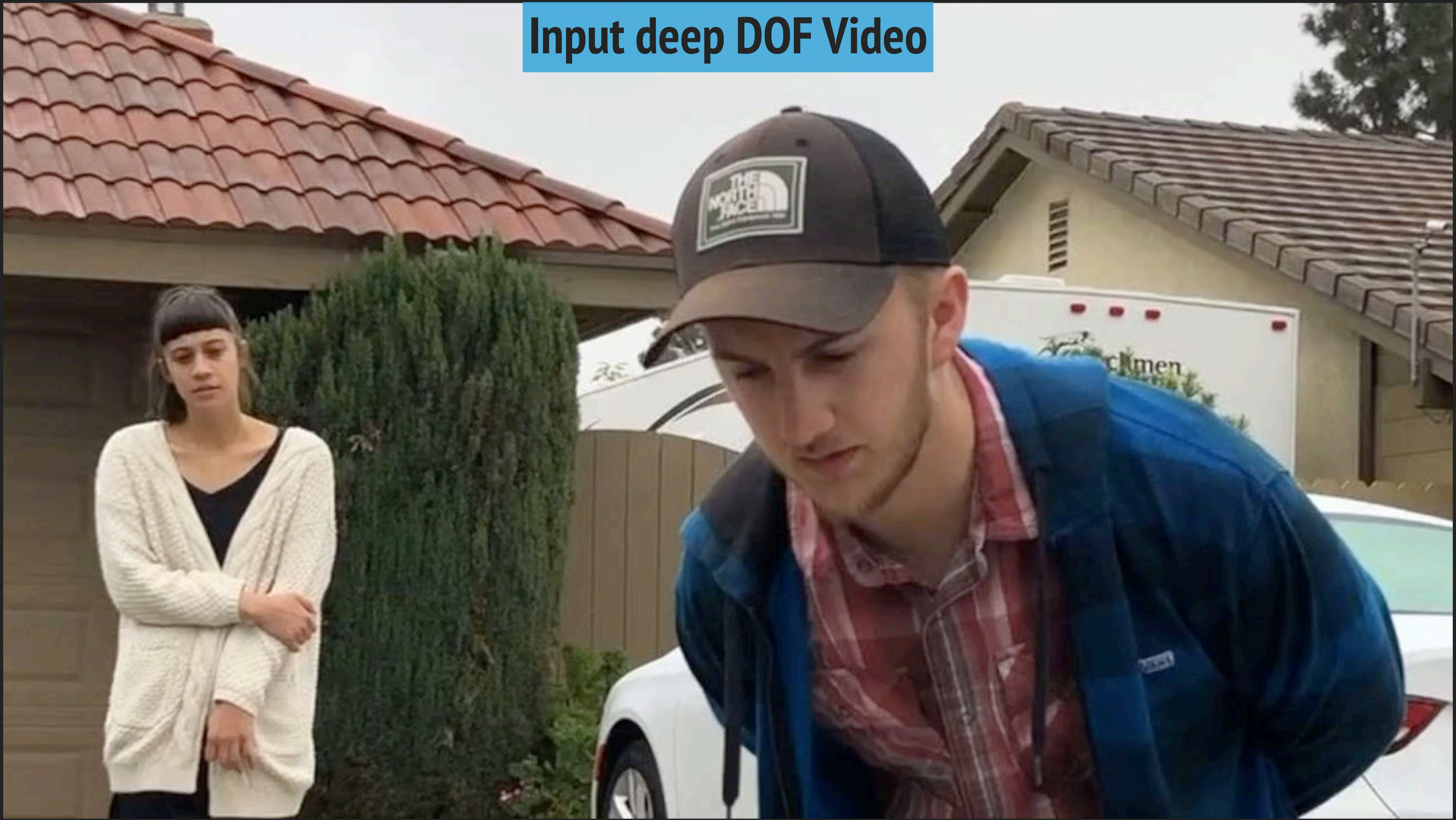
New Focus Target Detector

Human-Assisted Semi-Autofocus

Scene-Specific Autofocus

Audio-Driven LAAF Example — Come See the Spider?

Vid 9



Vid 10

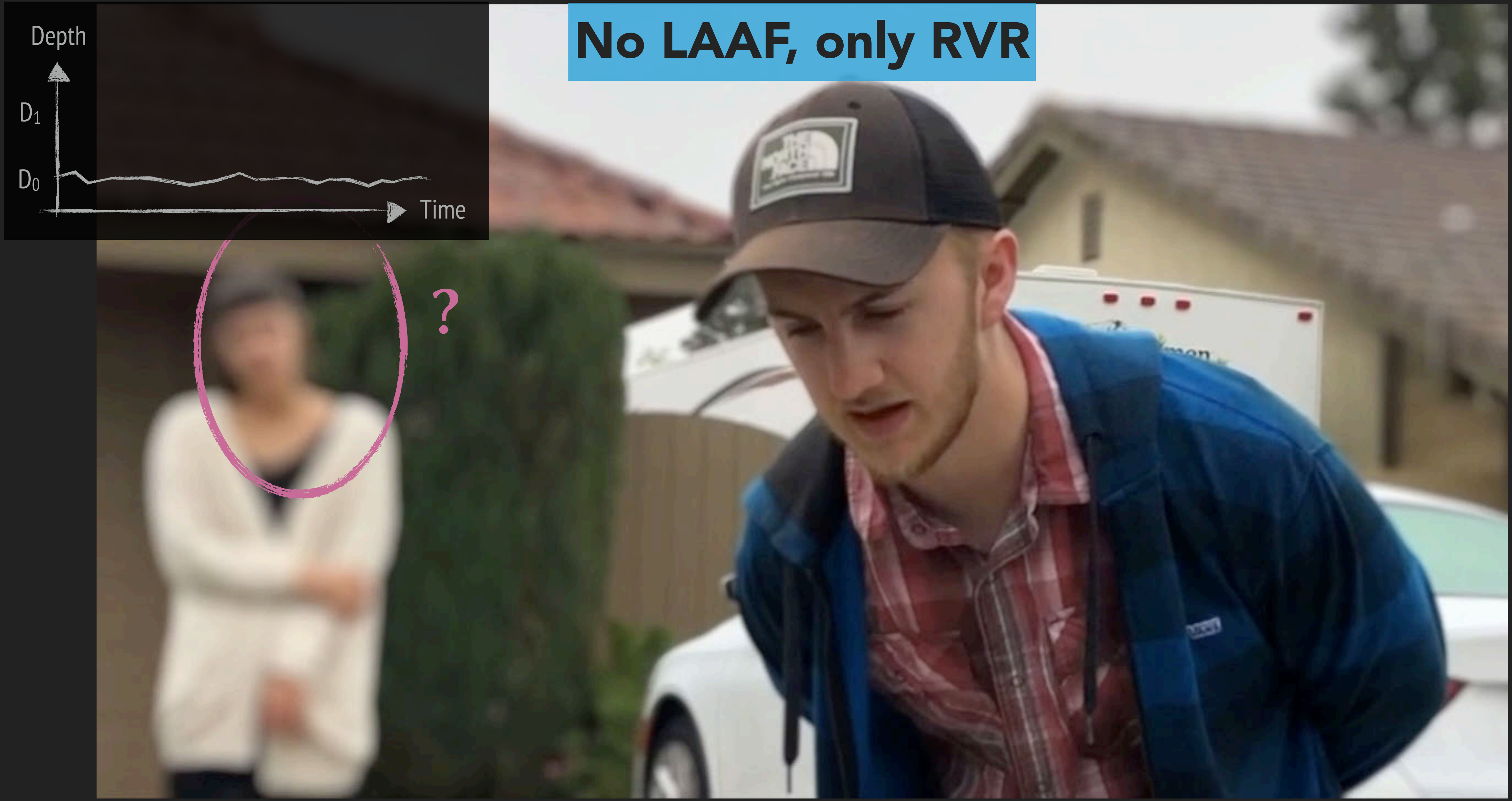
Scene-Specific Heatmap Identifies New Focus Targets

Owens et al, "*Audio-visual scene analysis with self-supervised multisensory features*", ECCV 2018

Autofocus Output With Audio-Aware LAAF

Vid 12

What if We Use RVR to Naïvely Apply Shallow DOF ?



Mode-Automatic: Scene-Specific Autofocus

Set focus right before

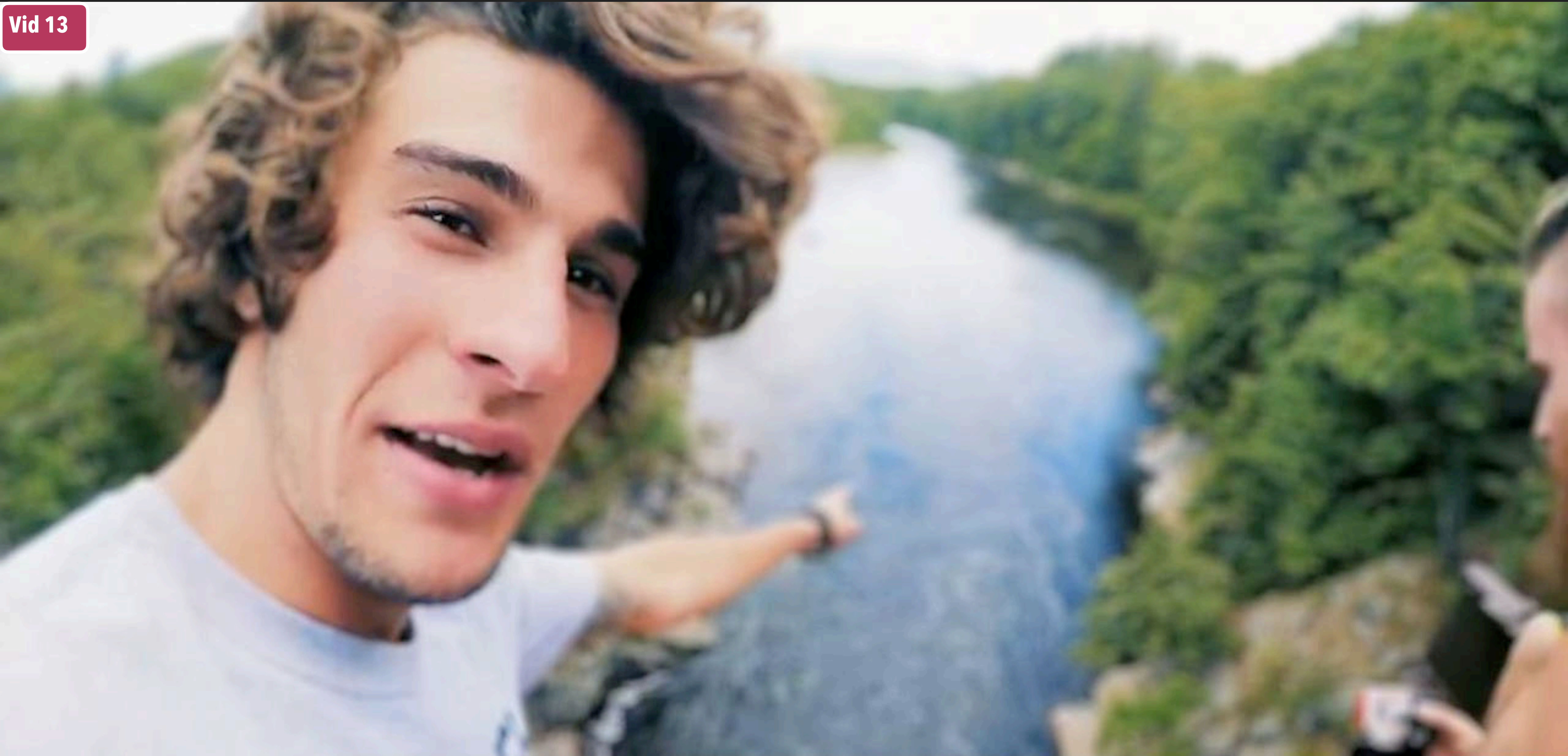
one speaks

an action happens

Audio-aware RVR-LAAF

Action-aware RVR-LAAF

LAAF Makes Casual Videos Contextually Meaningful



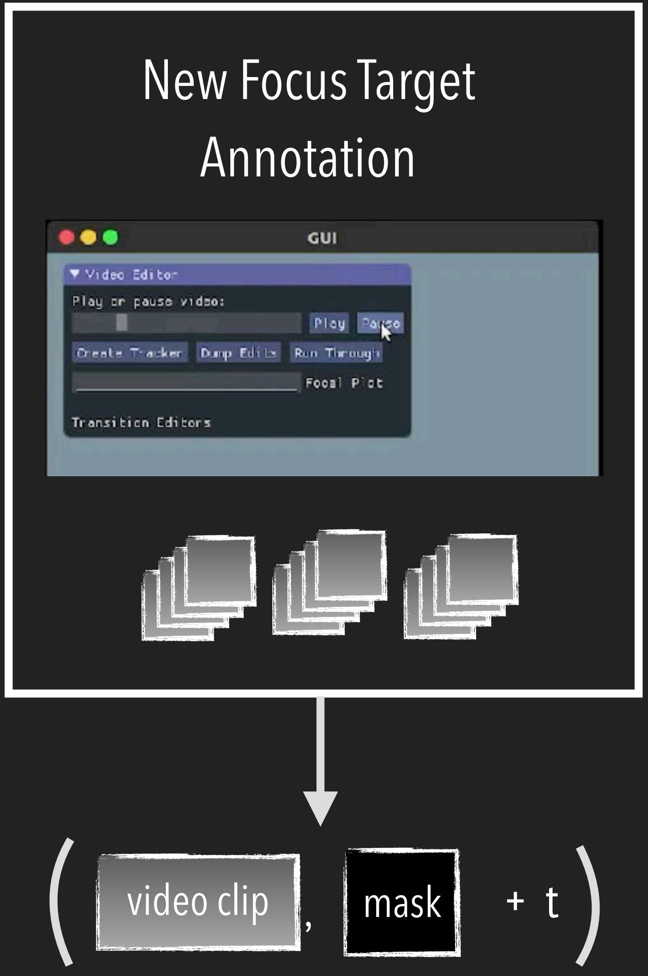
New Focus Target Detector

Human-Assisted Semi-Autofocus

Scene-Specific Autofocus

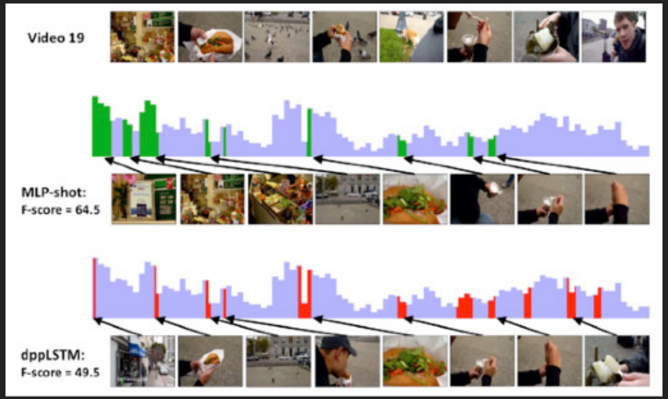
Video Saliency as New Focus Targets

Make focusing fully automatic



Data-driven Autofocus

Video Summarization



New video saliency dataset

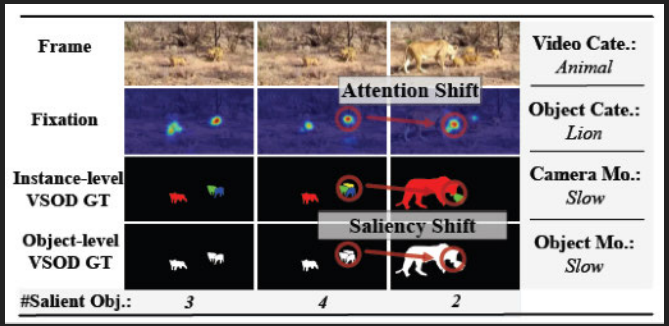
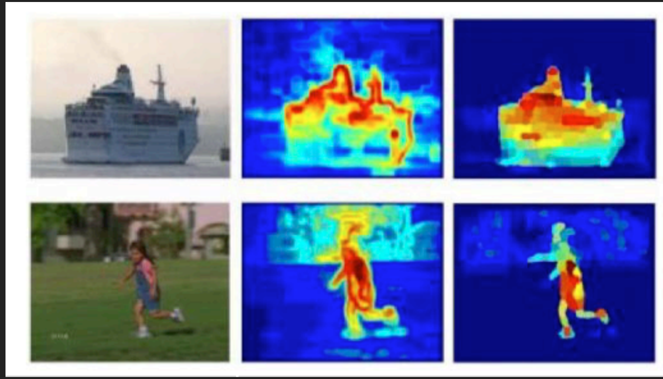
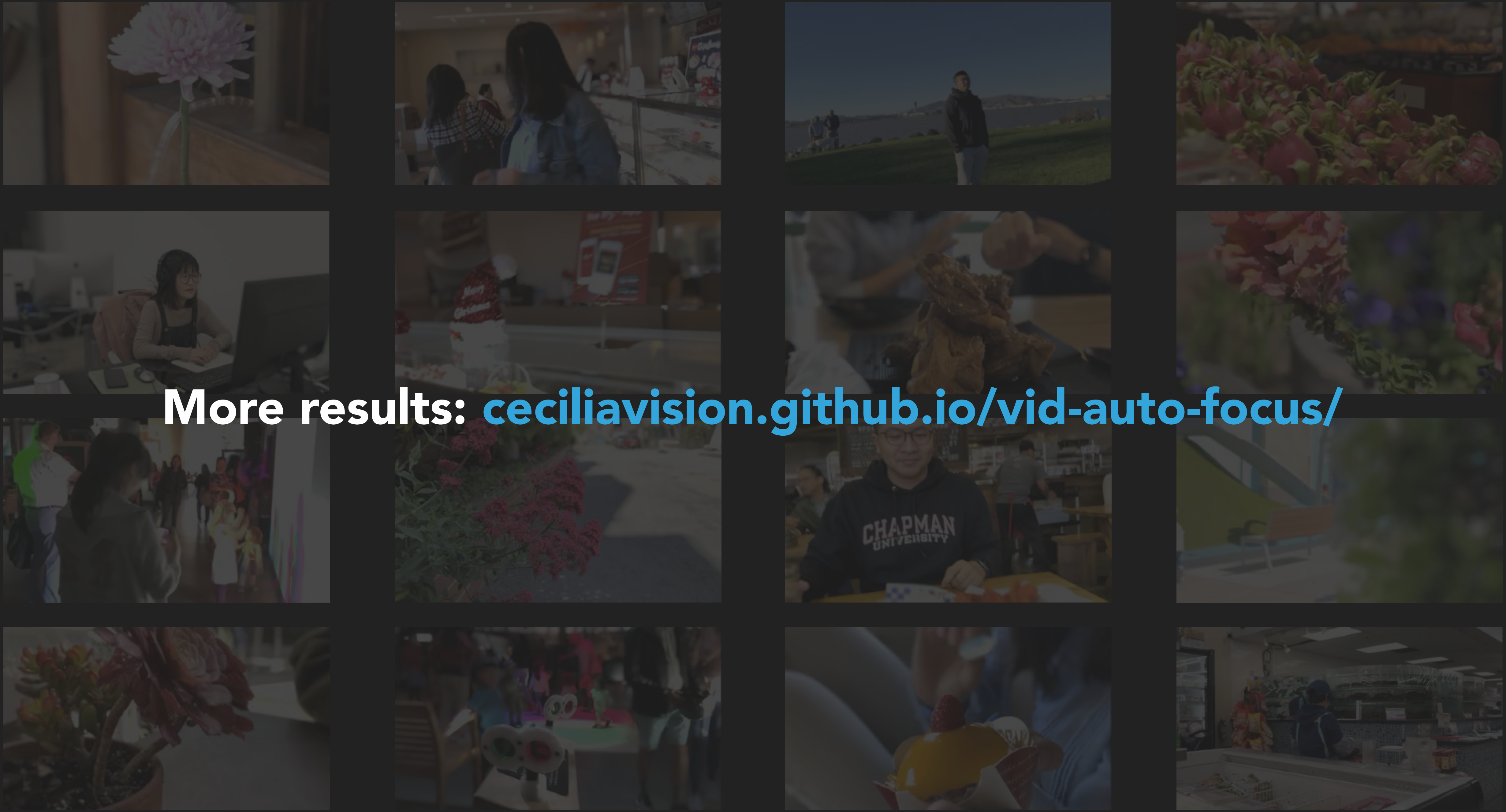


Image and Video Saliency





More results: ceciliavision.github.io/vid-auto-focus/

Context-Aware Casual Imaging

Image Quality

Zoom To Learn, Learn To Zoom

Zhang et al, CVPR 2019

- Learning From Raw Sensor
- Image Super-Resolution

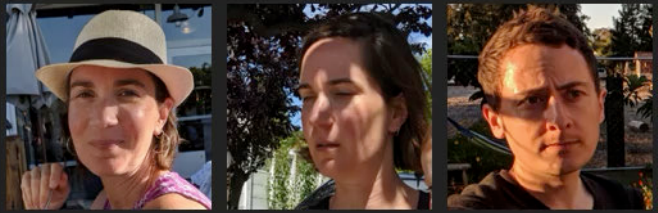
Focus

Synthetic Defocus and Look-Ahead Autofocus for Causal Videography

Zhang et al, SIGGRAPH 2019

- Video Synthetic Defocus
- ‘Future’ Scene Understanding

Lighting

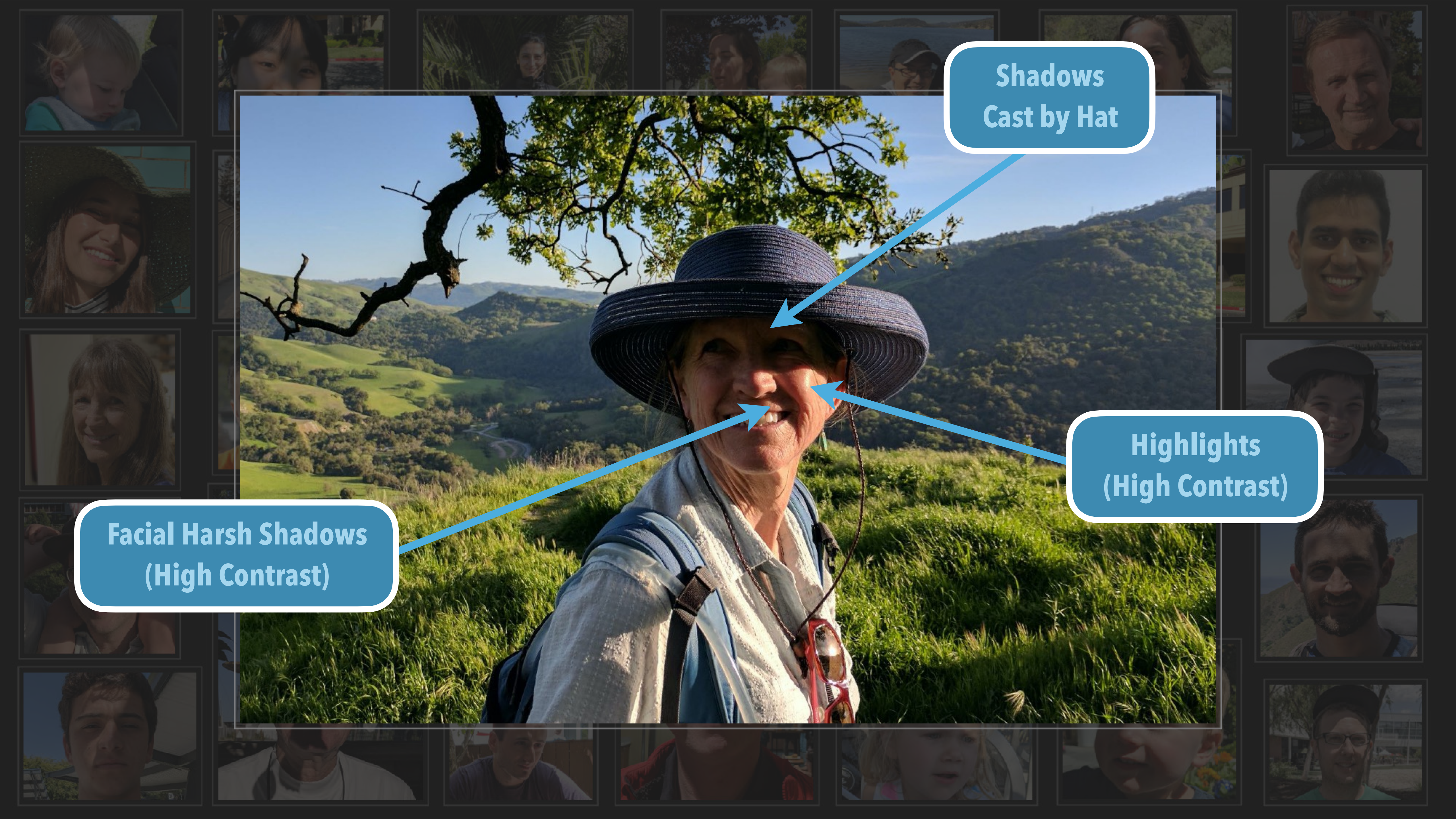


Portrait Shadow Manipulation

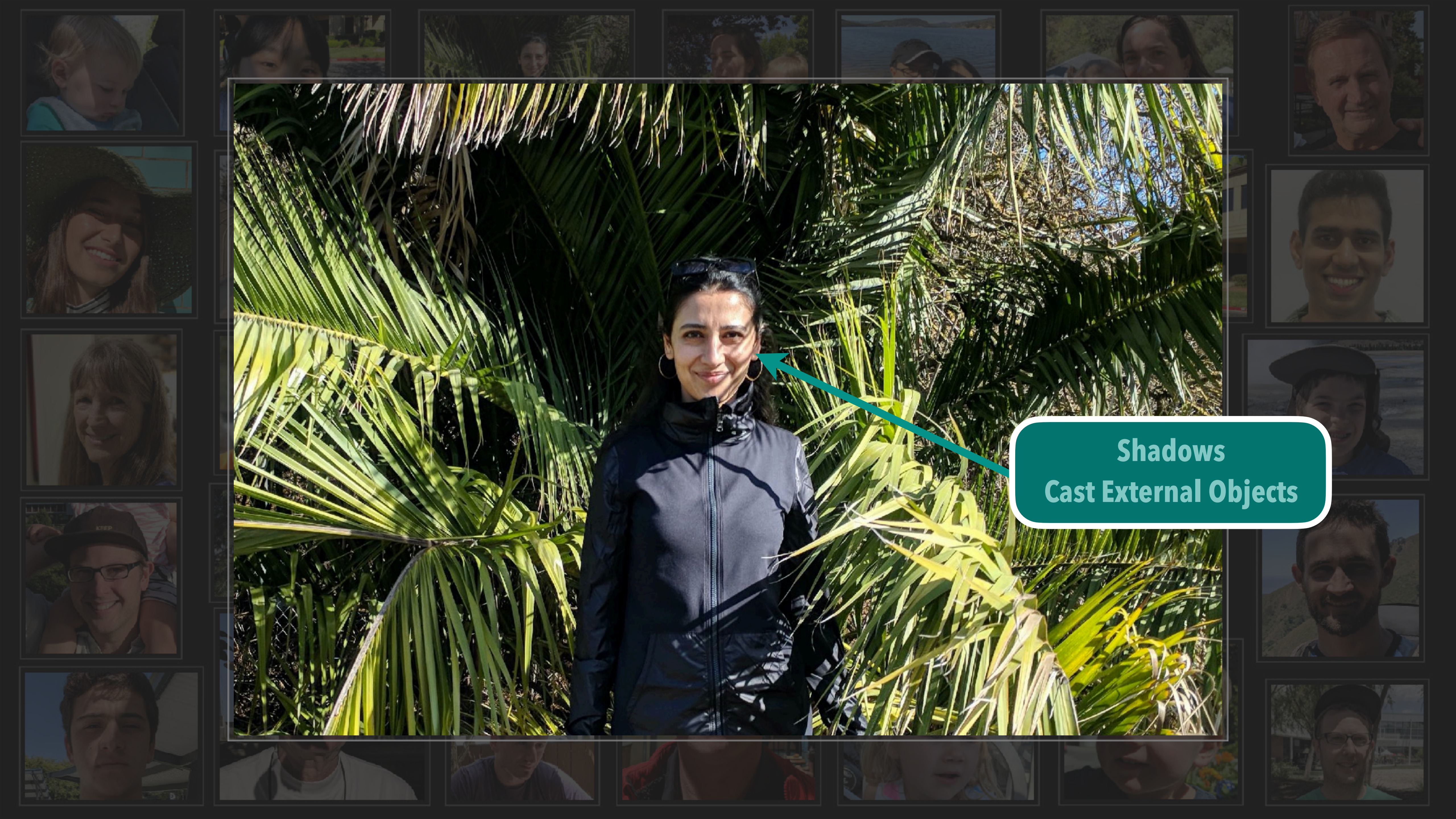
Zhang et al, SIGGRAPH 2020

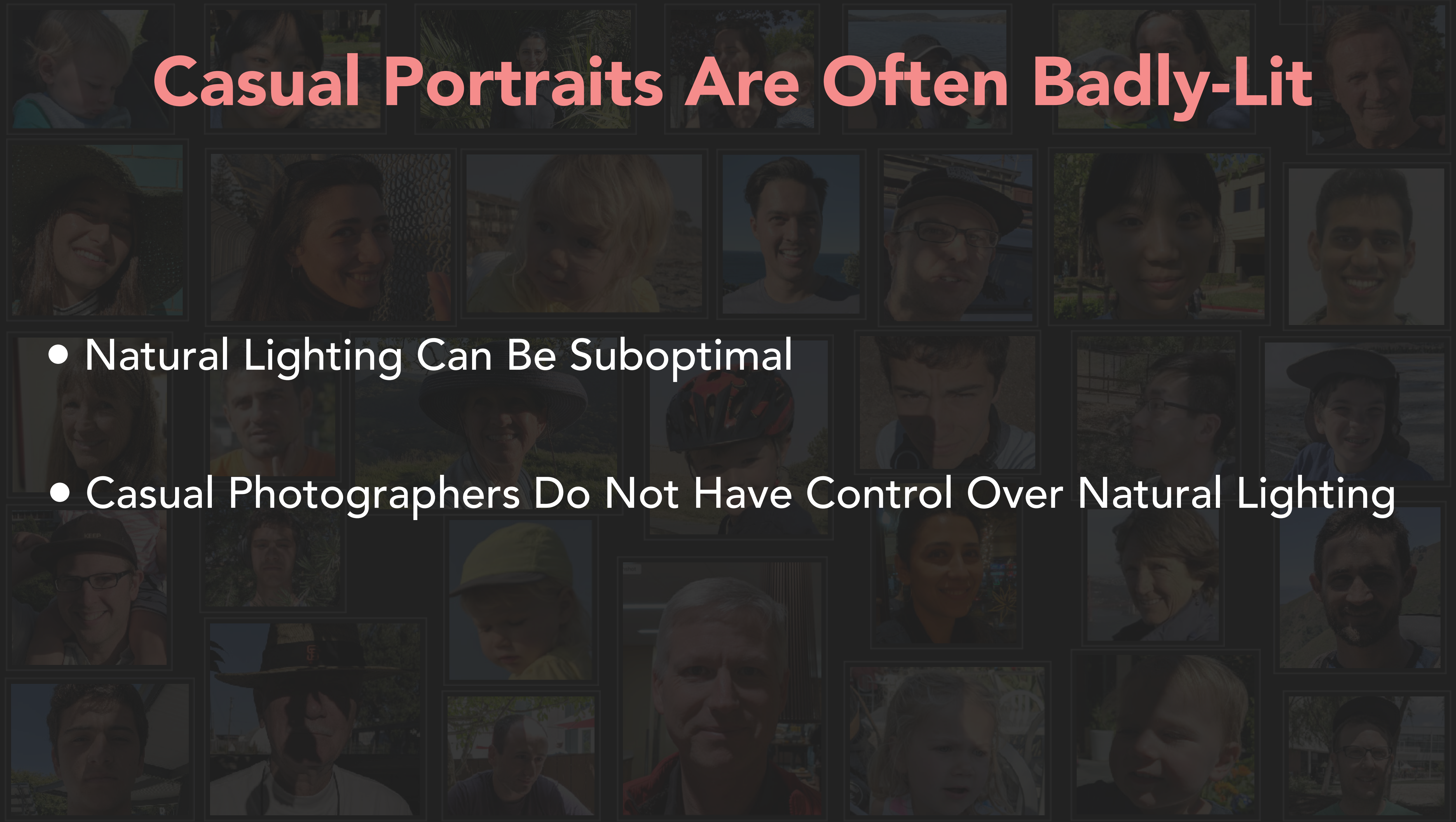
- Shadow Editing Post-Capture
- Embedding Professional Lighting Principals


KEEP


Shadows
Cast by Hat
Highlights
(High Contrast)
Facial Harsh Shadows
(High Contrast)


KEEP


Shadows
Cast External Objects



Casual Portraits Are Often Badly-Lit

- Natural Lighting Can Be Suboptimal
- Casual Photographers Do Not Have Control Over Natural Lighting

Two Sources of Portrait Shadows

- Facial Features Cast High Contrast Shadows
- External Objects Cast Randomly Shaped Shadows


KEEP

Portrait Shadow By Two Types

Foreign Shadow:

cast by external objects and hats

Facial Shadow:

cast by facial features

We Enable Shadow-Lighting Editing

Foreign Shadow

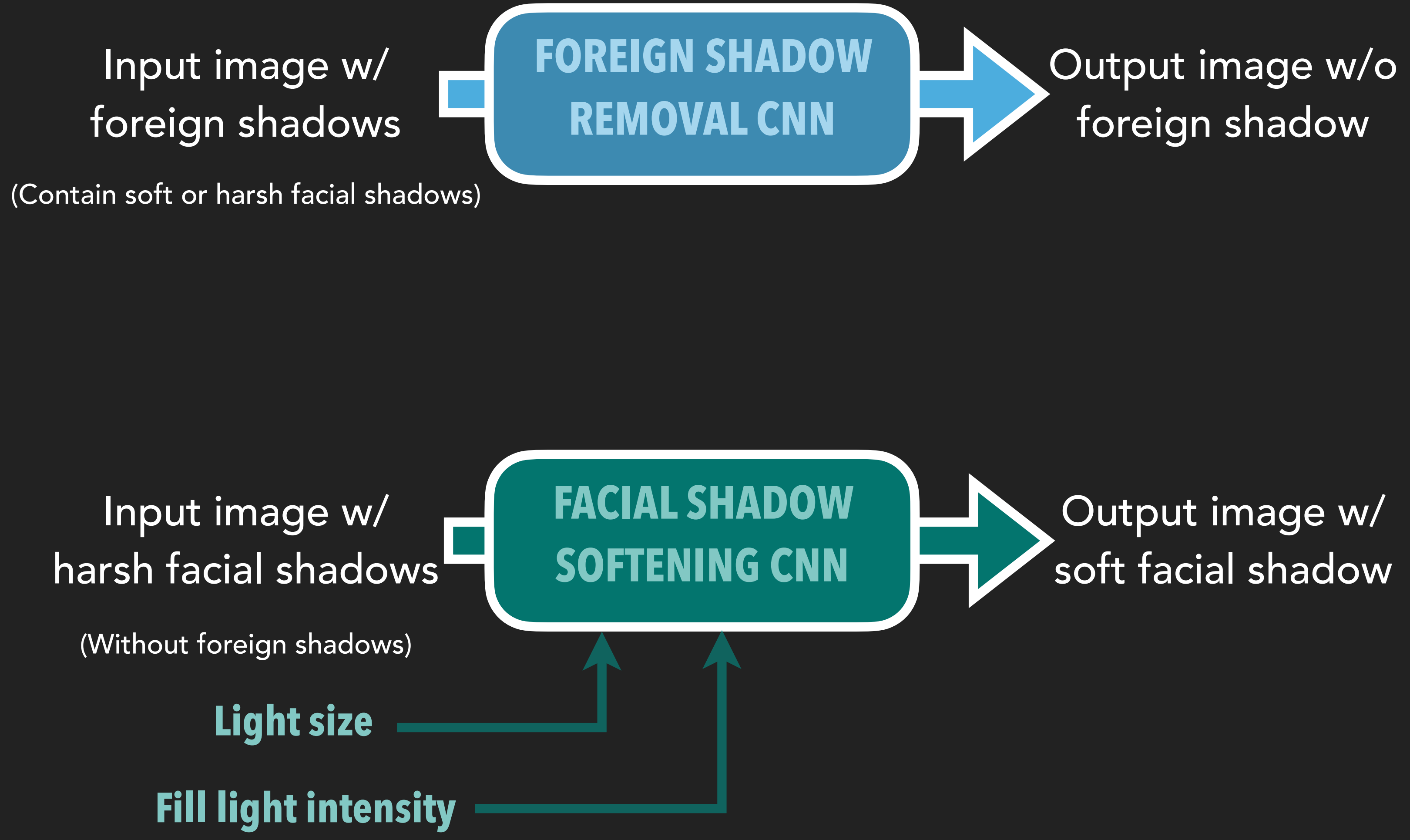
Facial Shadow

Lighting Ratio

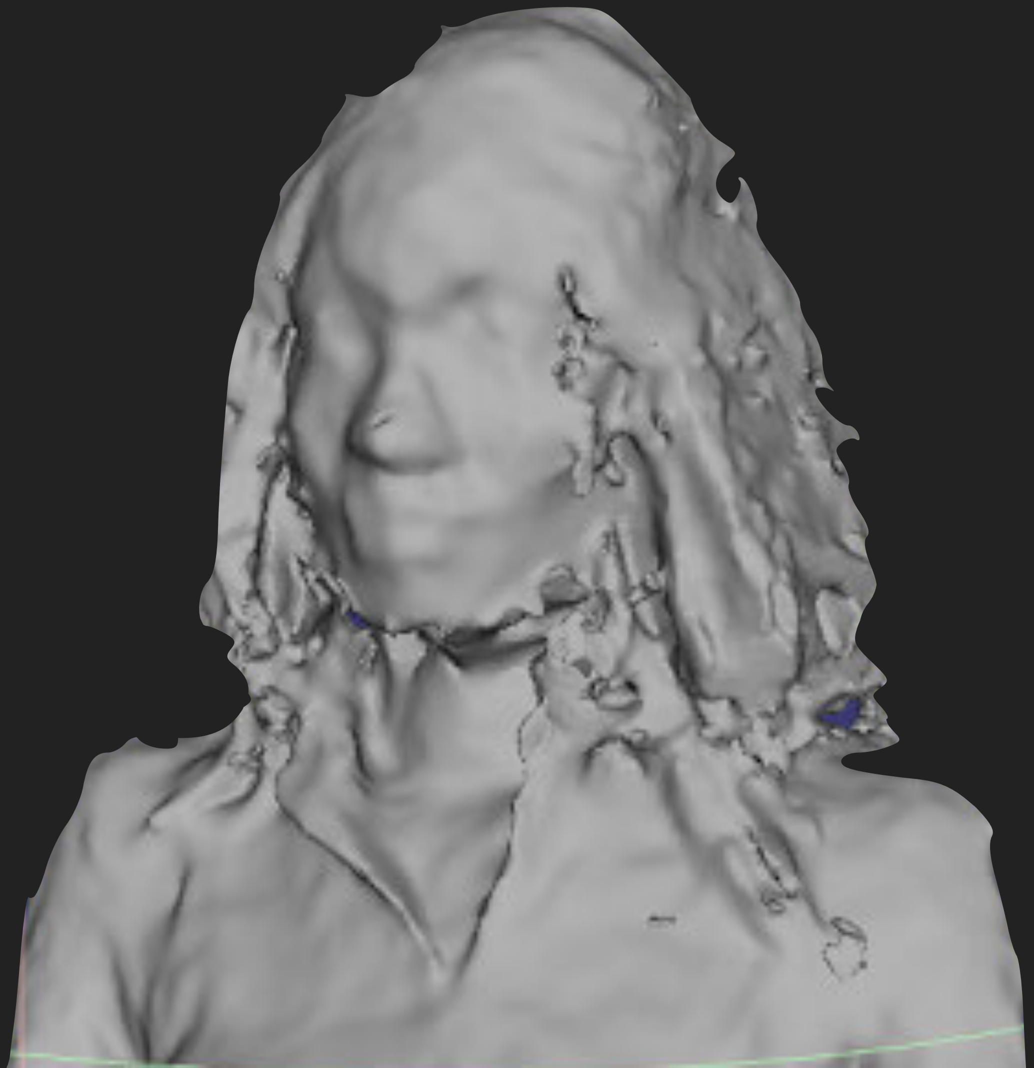
Badly-Lit Portrait Shadow Can Be Enhanced Using Knowledge of Good Lighting

- We identify the two types of portrait shadows and edit them differently according to different lighting principles
- We embed good lighting principles into the system using machine learning and synthesized data
- We show facial shadow manipulation benefits from symmetry modeling
- We demonstrate the system on real-world casually taken portrait photos

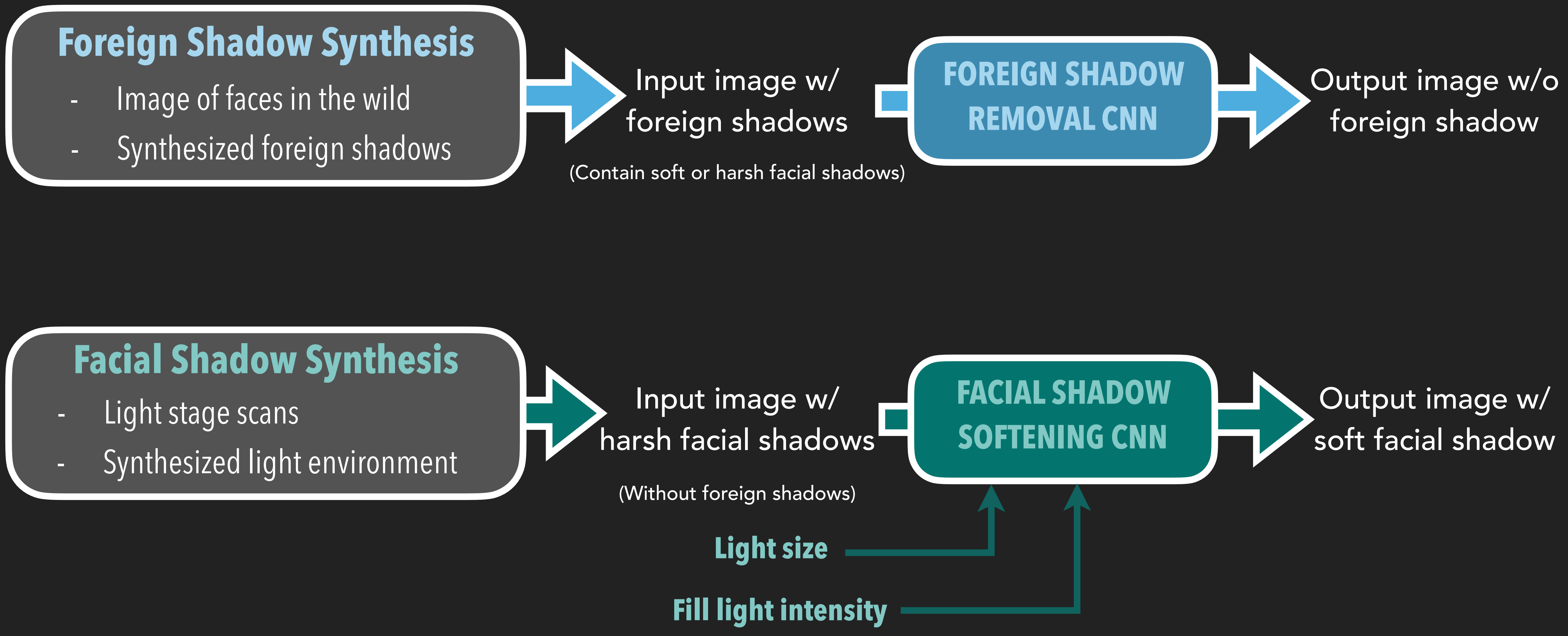
Two Models: Input and Output



3D Data Is Noisy and Non-Scalable



Two Data Synthesis Models



Foreign Shadow Removal Model and Controllable Facial Shadow Softening Model

Foreign Shadow Synthesis

- Image of faces in the wild
- Synthesized foreign shadows

Input image w/ foreign shadows

FOREIGN SHADOW REMOVAL CNN

Output image w/o foreign shadow

Facial Shadow Synthesis

- Light stage scans
- Synthesized light environment

Input image w/ harsh facial shadows

FACIAL SHADOW SOFTENING CNN

Output image w/ soft facial shadow

Light size

Fill light intensity

Foreign Shadow Synthesis Model

$$I = I_l \circ (1 - M) + I_s \circ M$$

Synthesized input

Lit-image

Shadow-image

with color variation

Shadow mask

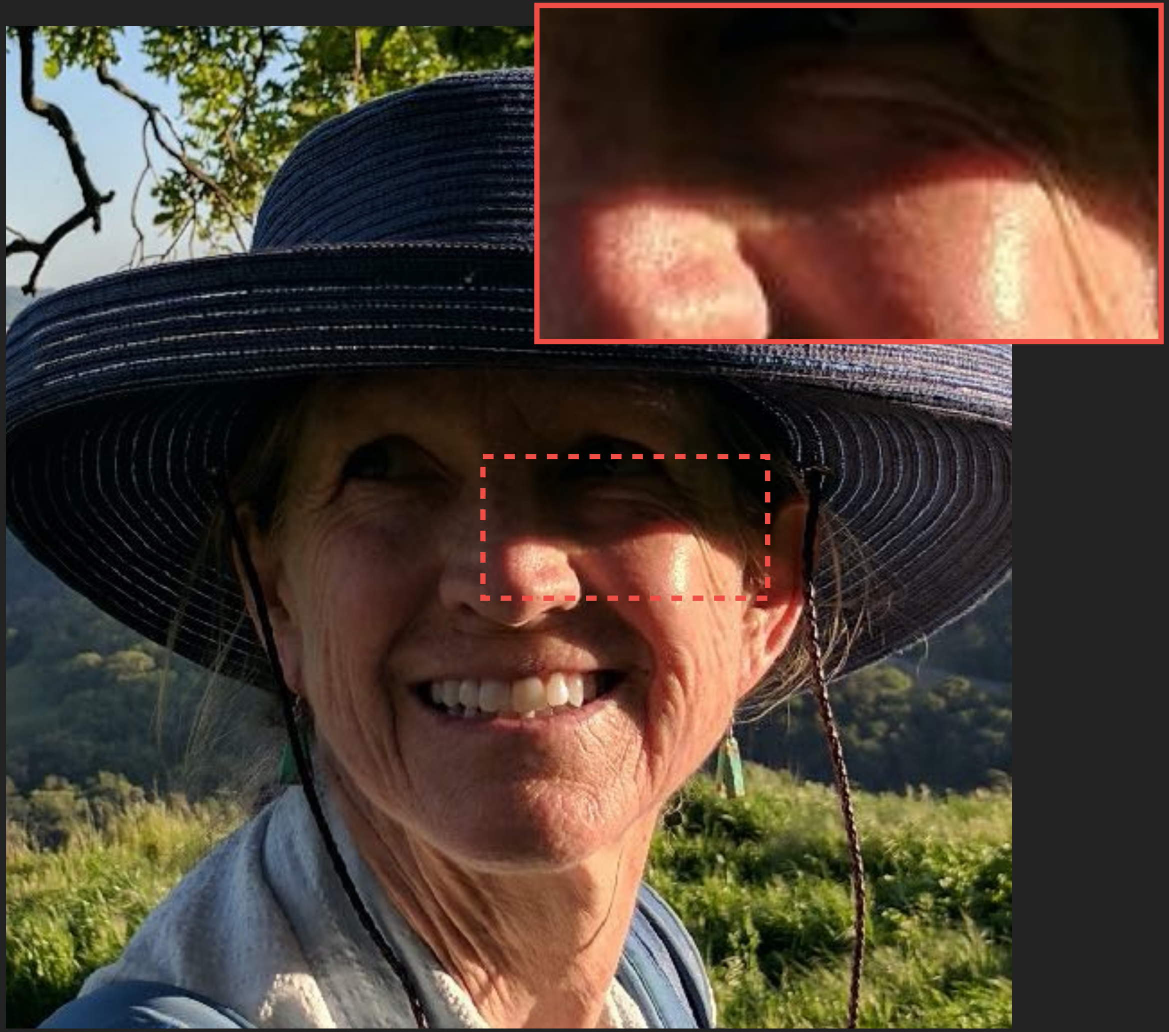
with subsurface scattering

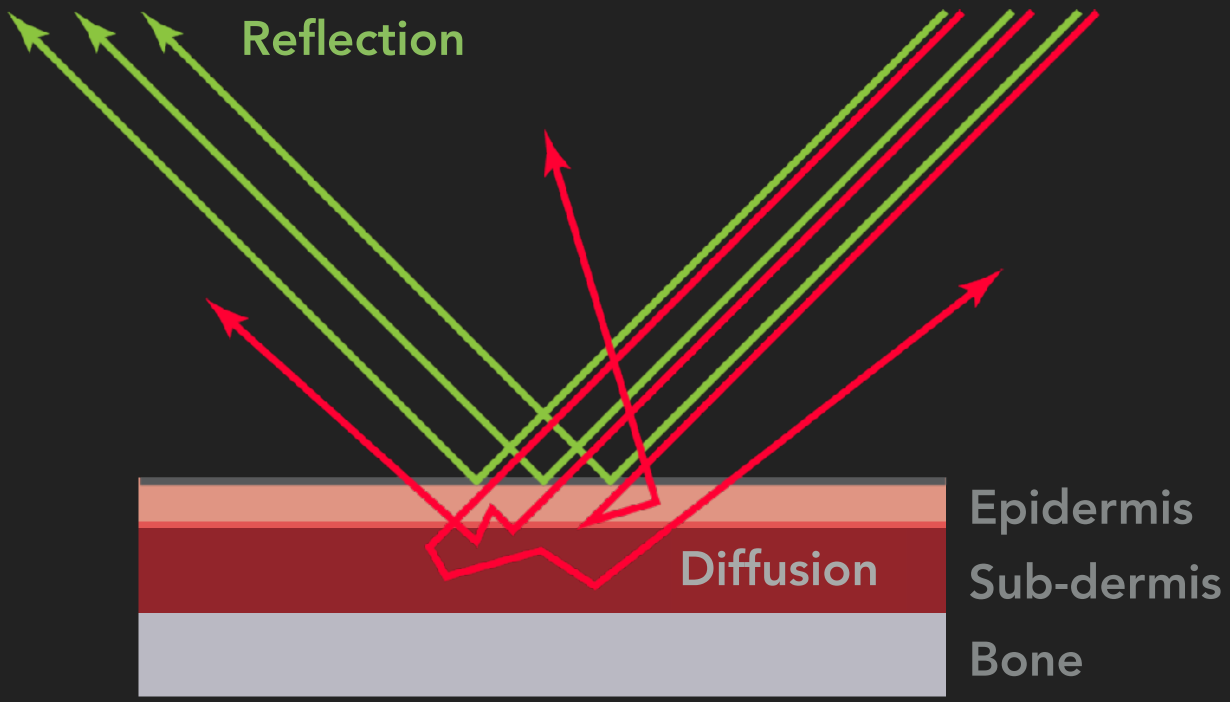
with spatial intensity variation

with spatially-varying blur

How Does Shadow Appear On Faces

1. Subsurface Scattering Approximation

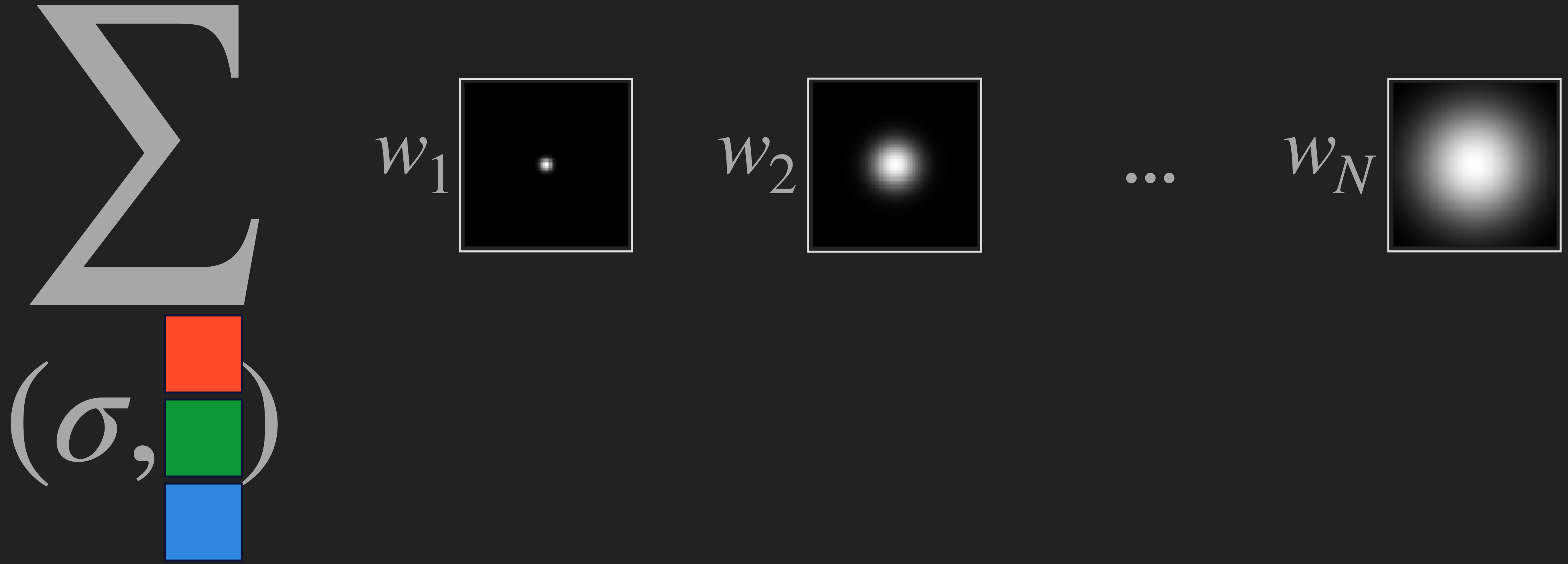




Subsurface Scattering on Skin

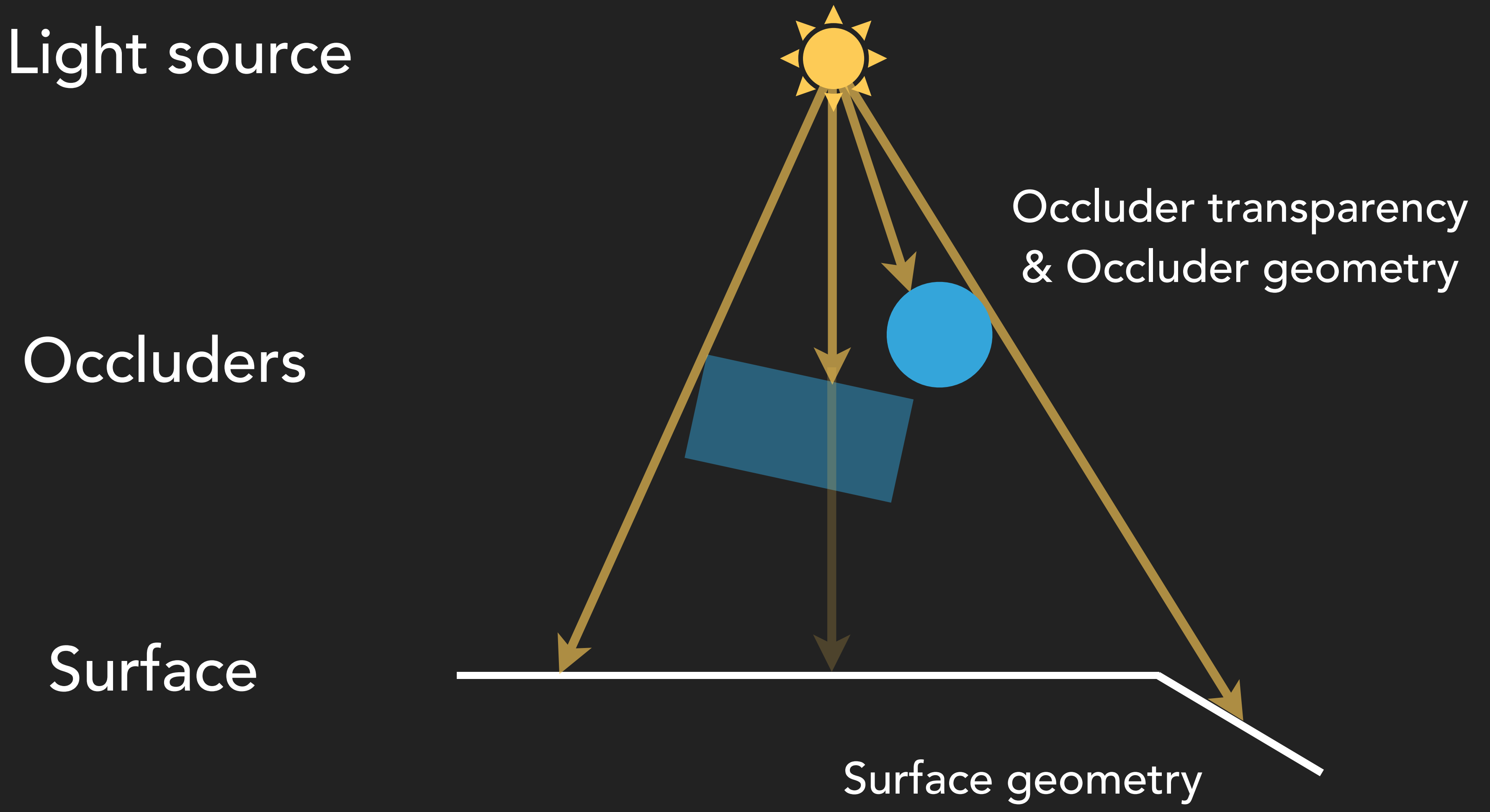
How Does Shadow Appear On Faces

1. Subsurface Scattering Approximation



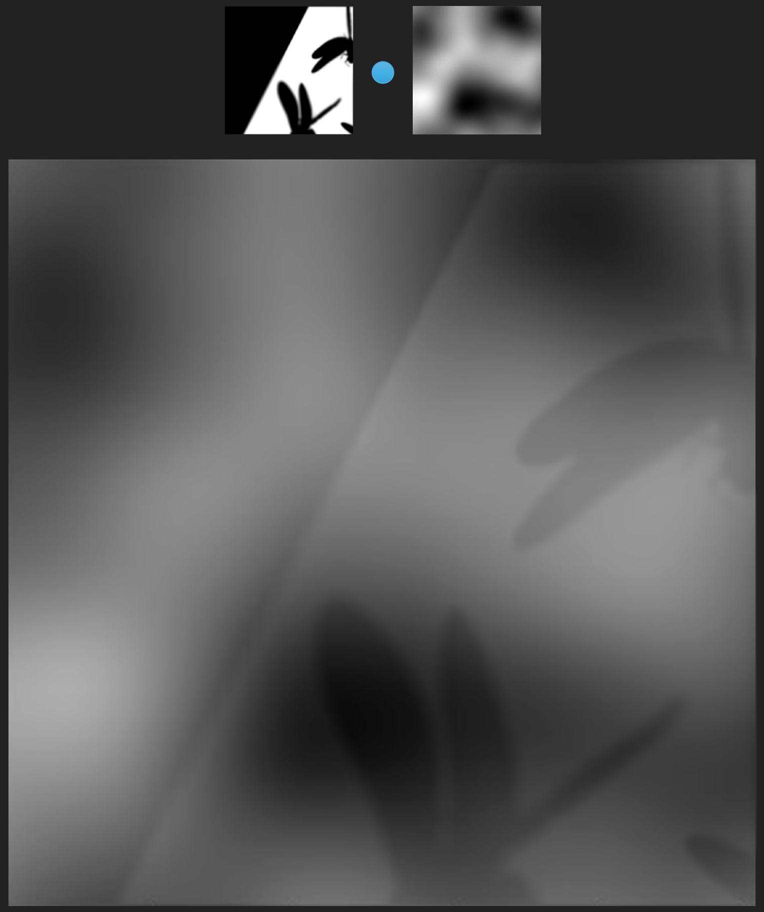
How Does Shadow Appear On Faces

2. Spatially-Varying Appearance



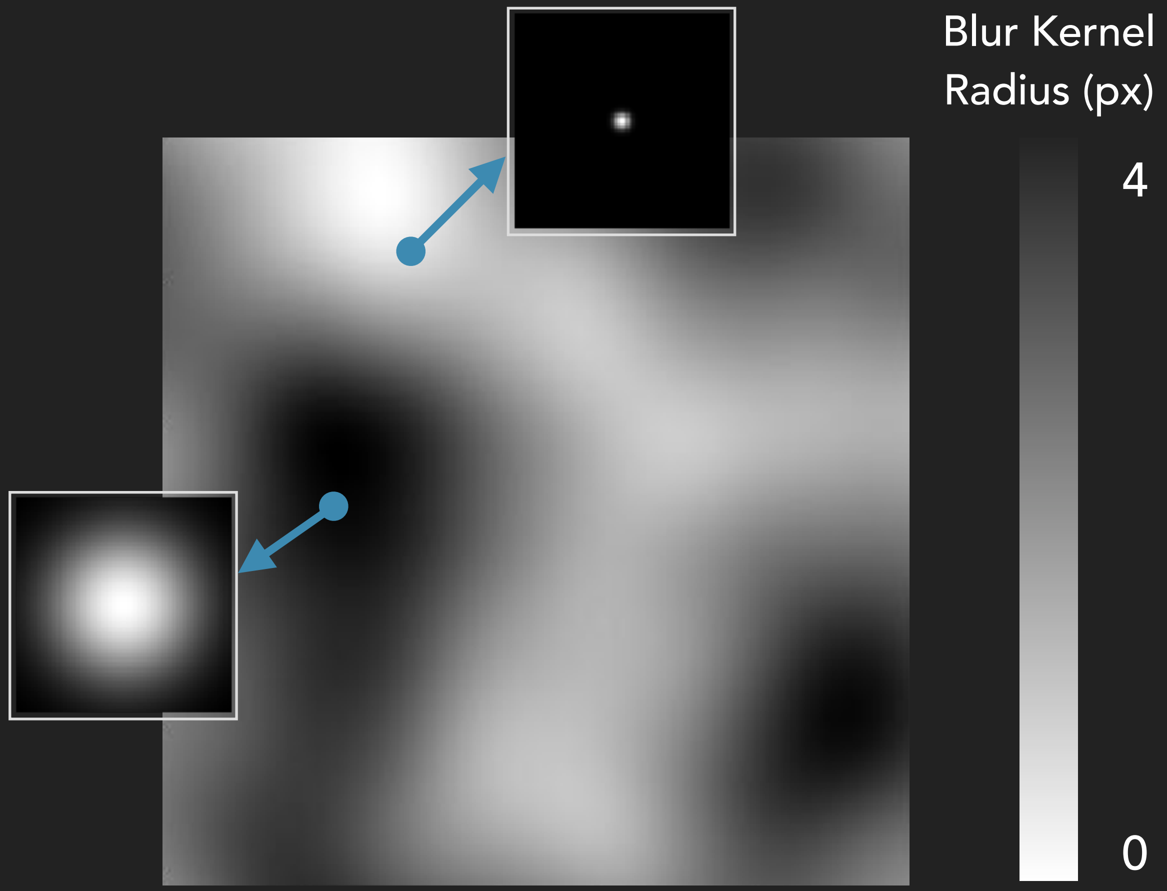
How Does Shadow Appear On Faces

Spatially-Varying Blur



Per-pixel Intensity Map modeled by Perlin Noise

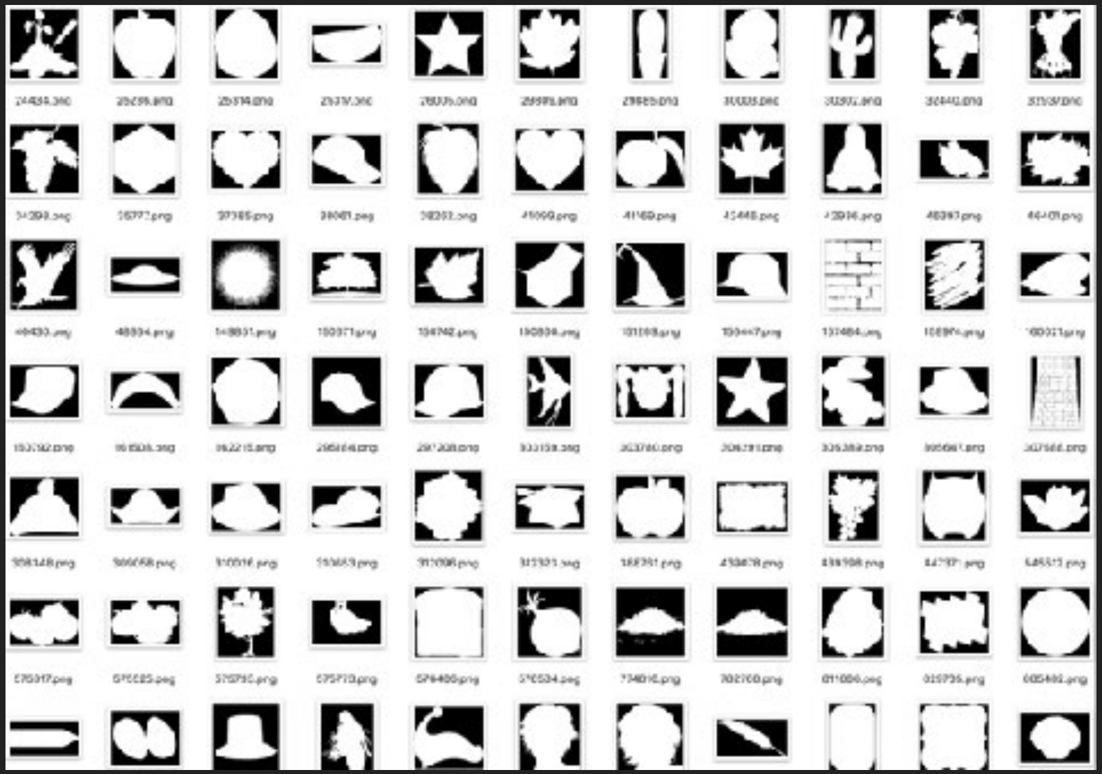
Spatially-Varying Intensity



Per-pixel Blur Map modeled by Perlin Noise

How Does Shadow Appear On Faces

3. Shape Variation



Source 1: geometry randomly sampled from a collection of silhouette images

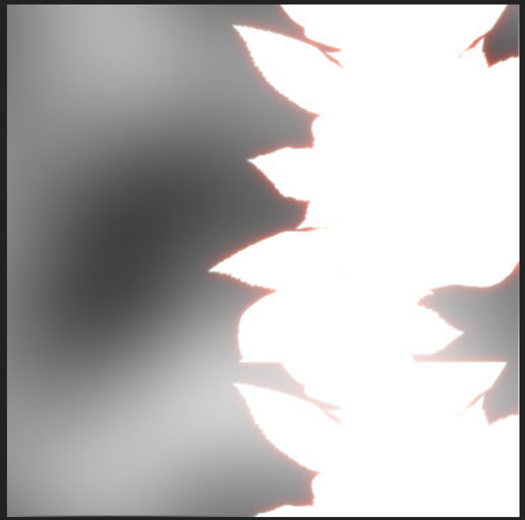
Source 2: geometry randomly generated using Perlin noise

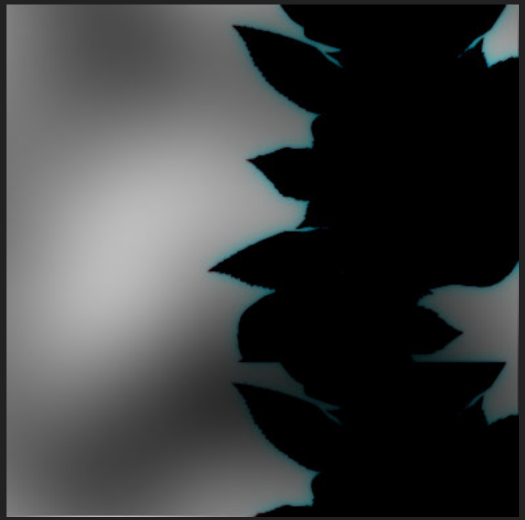
$$I = I_l \circ (1 - M) + I_s \circ M$$

with color variation

4. Color Variation

$$I = I_l \circ (1 - M) + I_s \circ M$$





with color variation

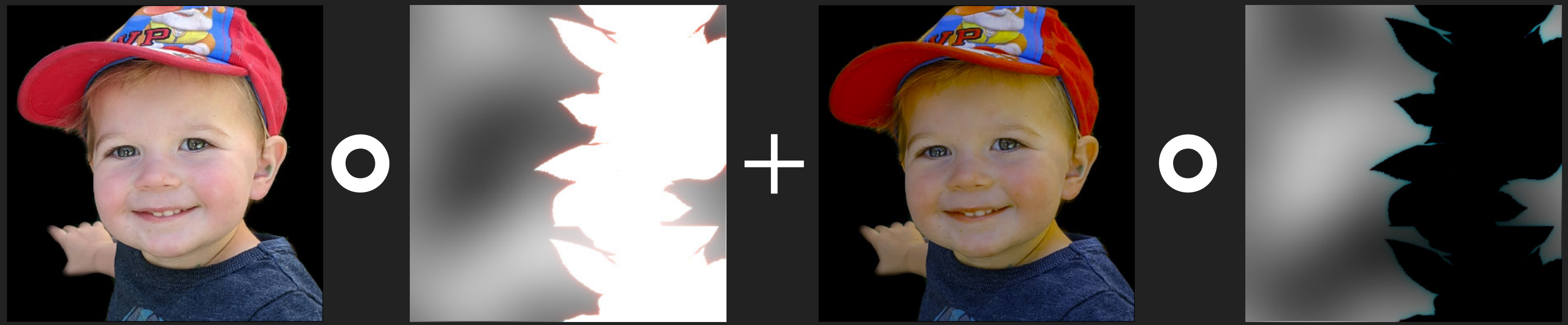
with spatial intensity variation

with spatially-varying blur

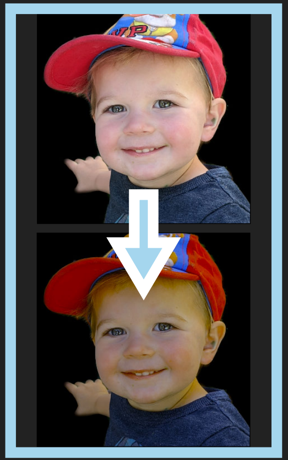
with subsurface scattering approximation

$$I = I_l \circ (1 - M) + I_s \circ M$$

Synthesized image
with foreign shadow

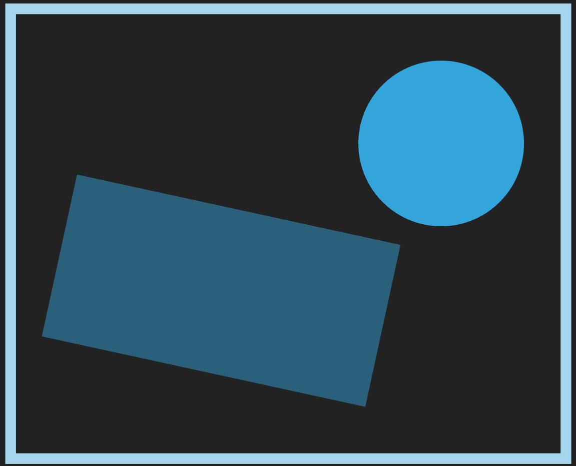


Generate A "Super-Set" Of Foreign Shadow



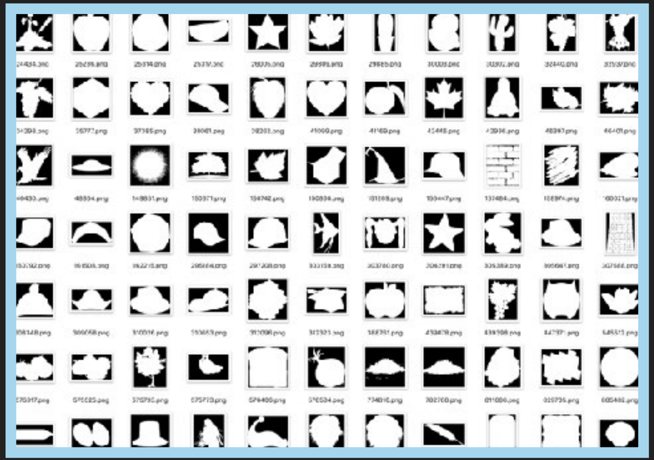
color jitter matrix

subsurface scattering approximation weights



spatial intensity variance

spatially-varying blur size



random shape sampling

Example Foreign Shadow Removal Training Data
— Input Image With Foreign Shadow

Example Foreign Shadow Removal Training Data
— Ground Truth Image With No Foreign Shadow

Example Foreign Shadow Removal Training Data
— Ground Truth Image With No Foreign Shadow

500K 'in-the-wild' images of
5000 unique subjects

Full synthesis code and demo are now online: eecs.berkeley.edu/~cecilia77/project-pages/portrait

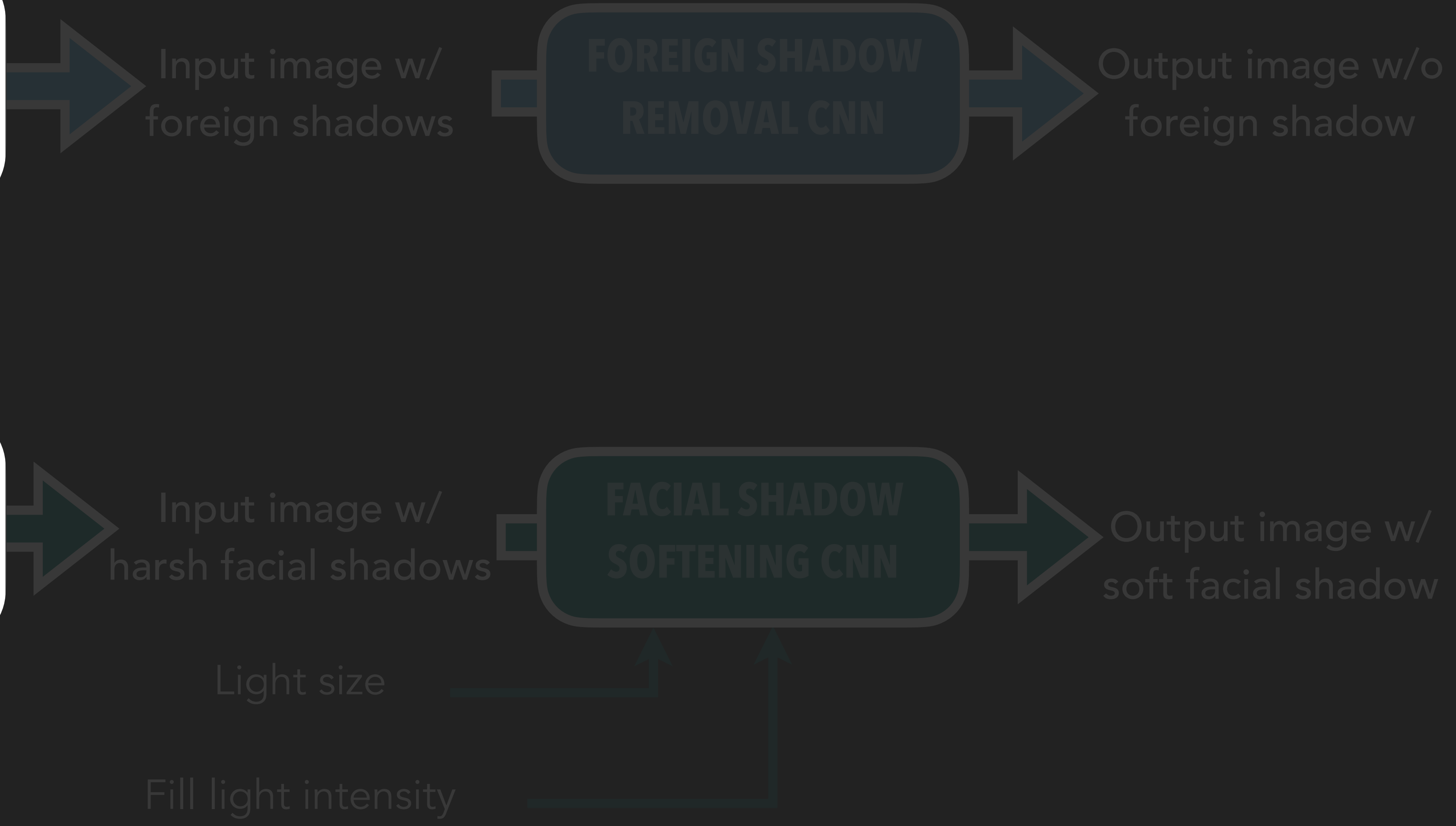
Foreign Shadow Removal Model and Controllable Facial Shadow Softening Model

Foreign Shadow Synthesis

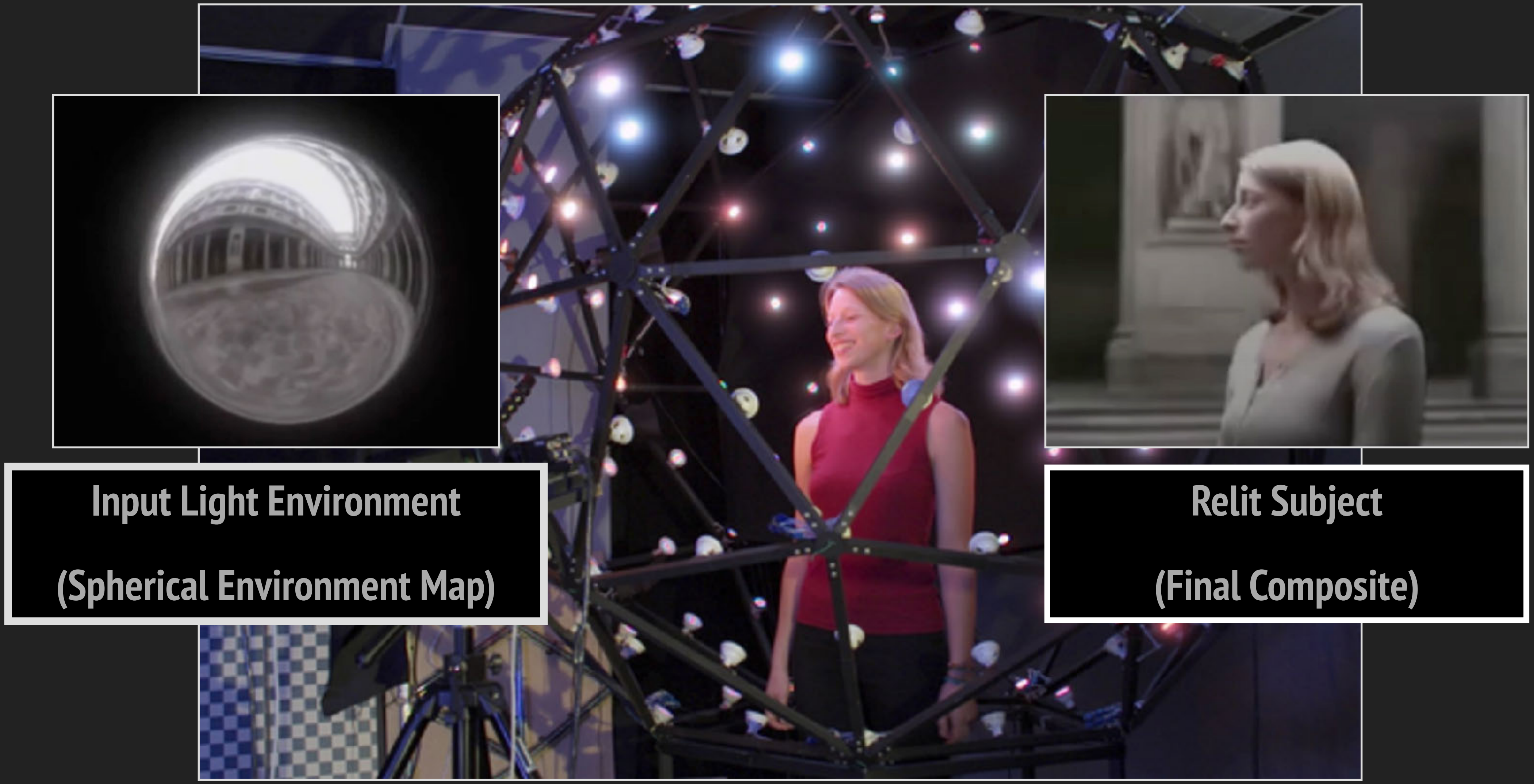
- Image of faces in the wild
- Synthesized foreign shadows

Facial Shadow Synthesis

- Light stage scans
- Synthesized light environment

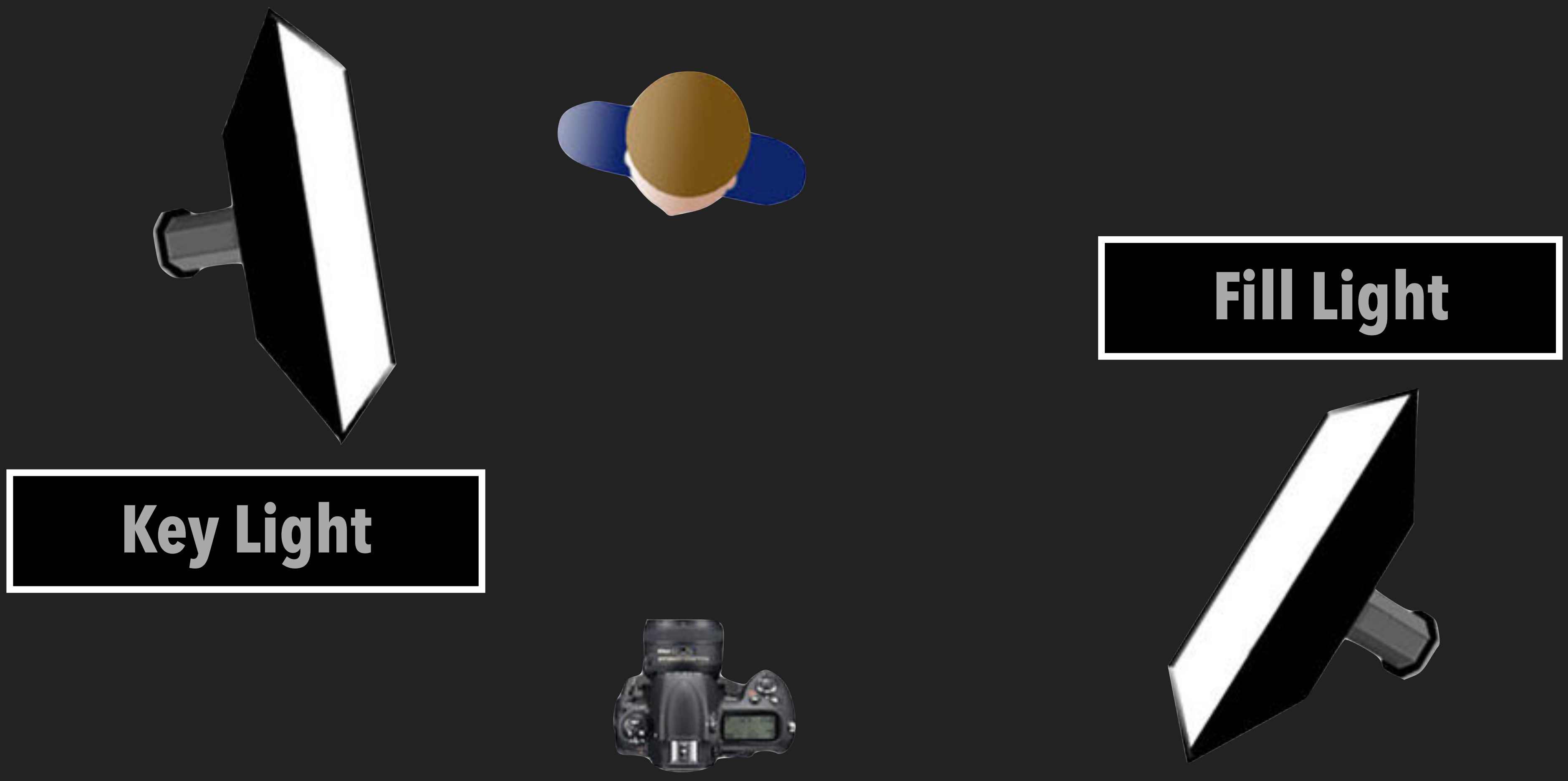


The Light Stage



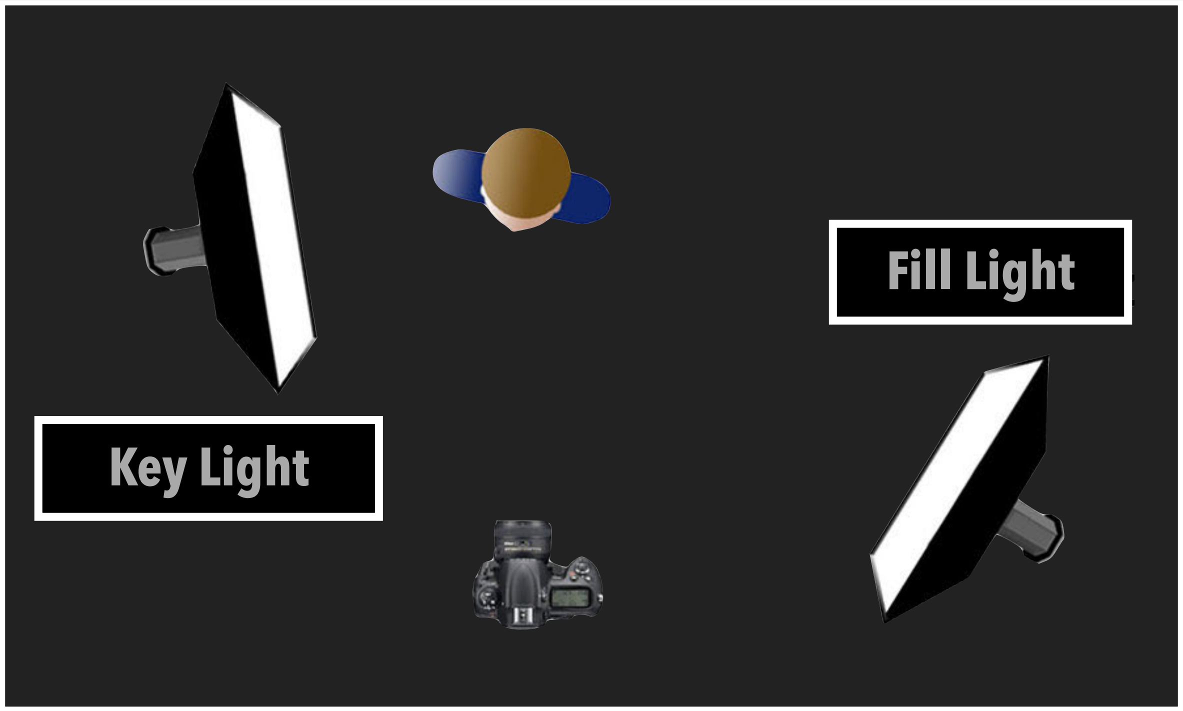
Debevec, Paul, et al. "A Lighting Reproduction Approach to Live-Action Compositing." ACM Transactions on Graphics (TOG) 21.3 (2002): 547-556.

Studio Lighting Principals



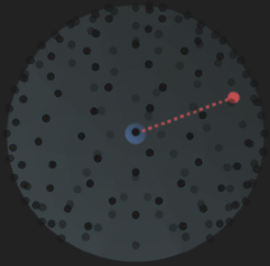
- Ambient Light
- Hair / Rim Light
- Kicker
- Multiple Key
- ...

Studio Lighting To Light Stage



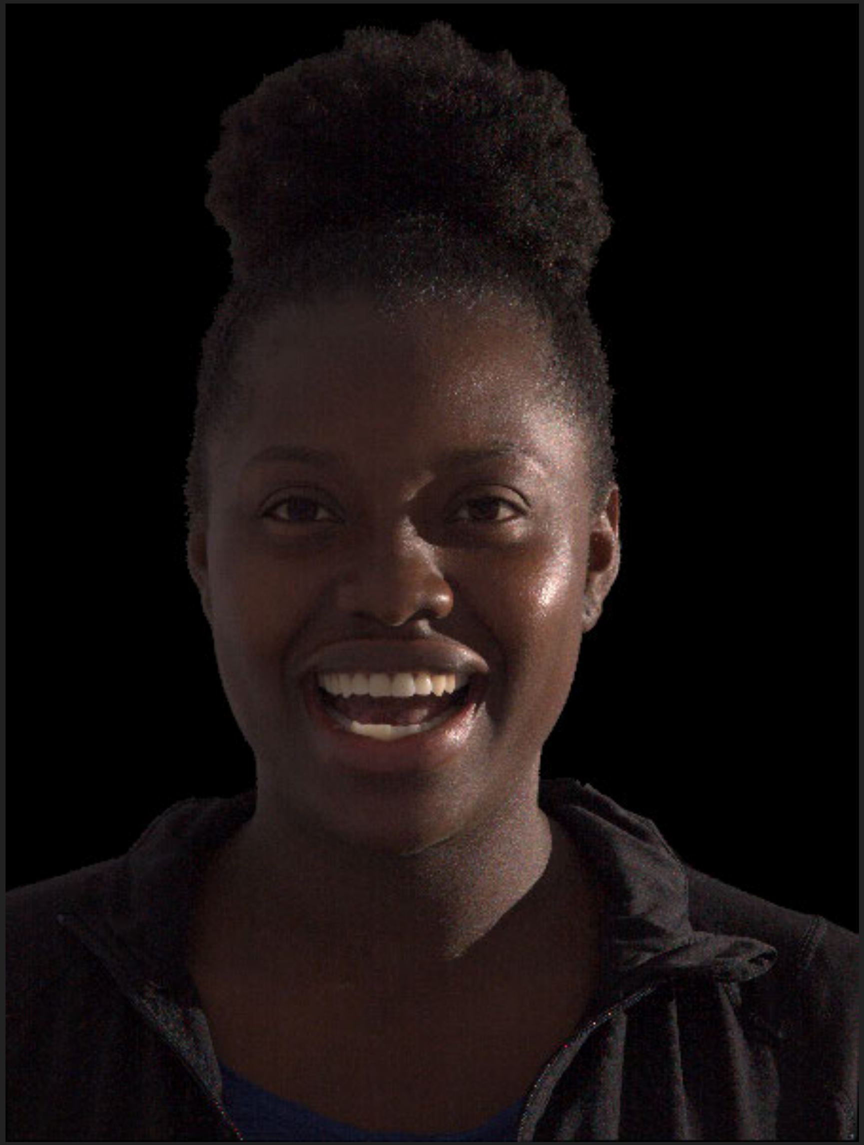
Synthetic studio lighting environment

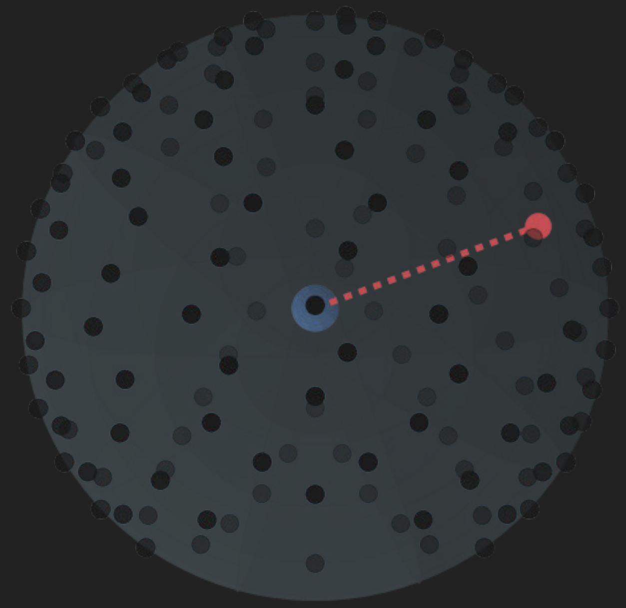
Light stage representation



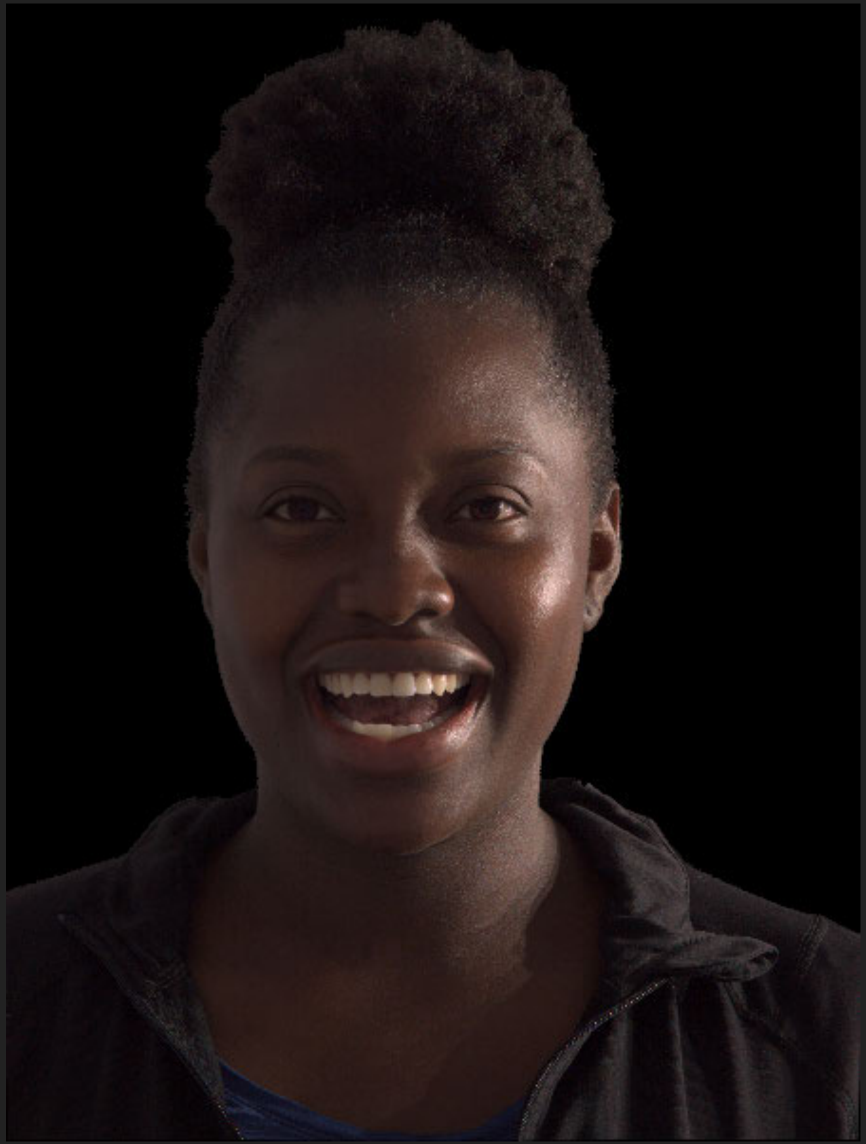
Simplified illustration

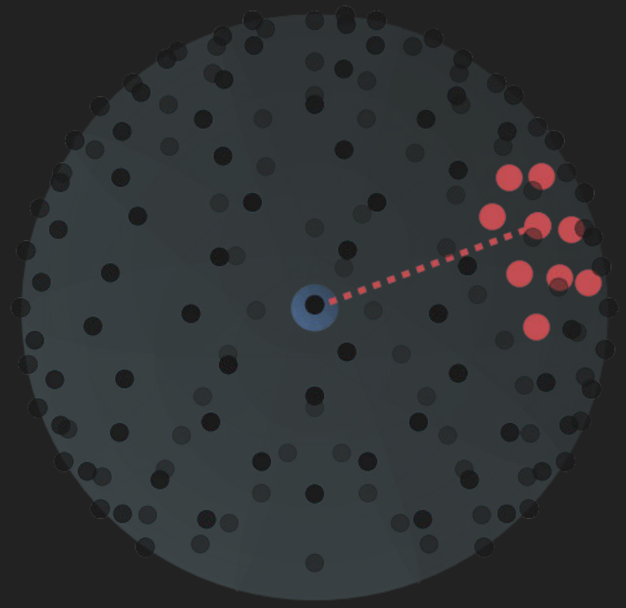
Using Light Stage Scan To Synthesize Facial Shadow





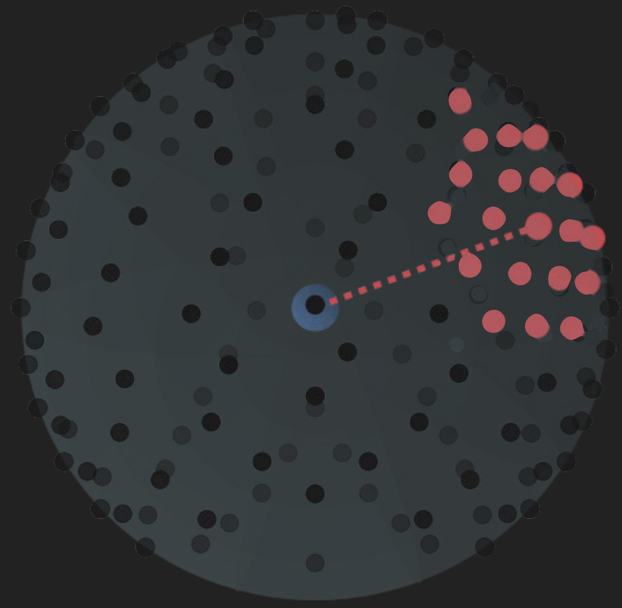
Using Light Stage Scan To Synthesize Facial Shadow





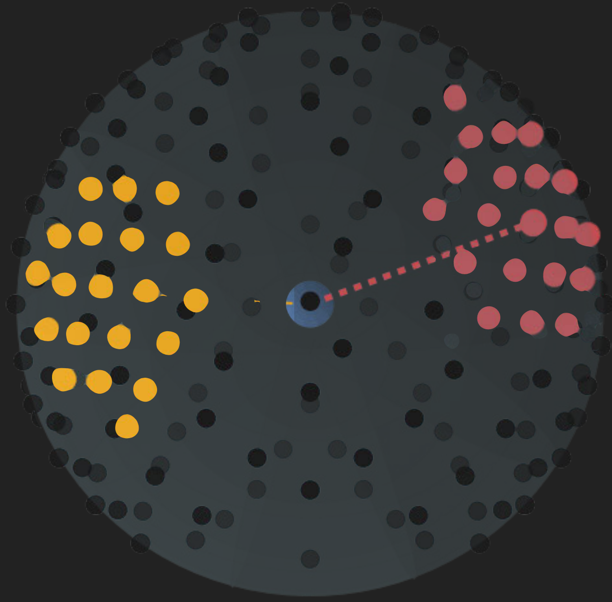
- Key light softened with light size 10

Using Light Stage Scan To Synthesize Facial Shadow



- Key light softened with light size 40

Using Light Stage Scan To Synthesize Facial Shadow

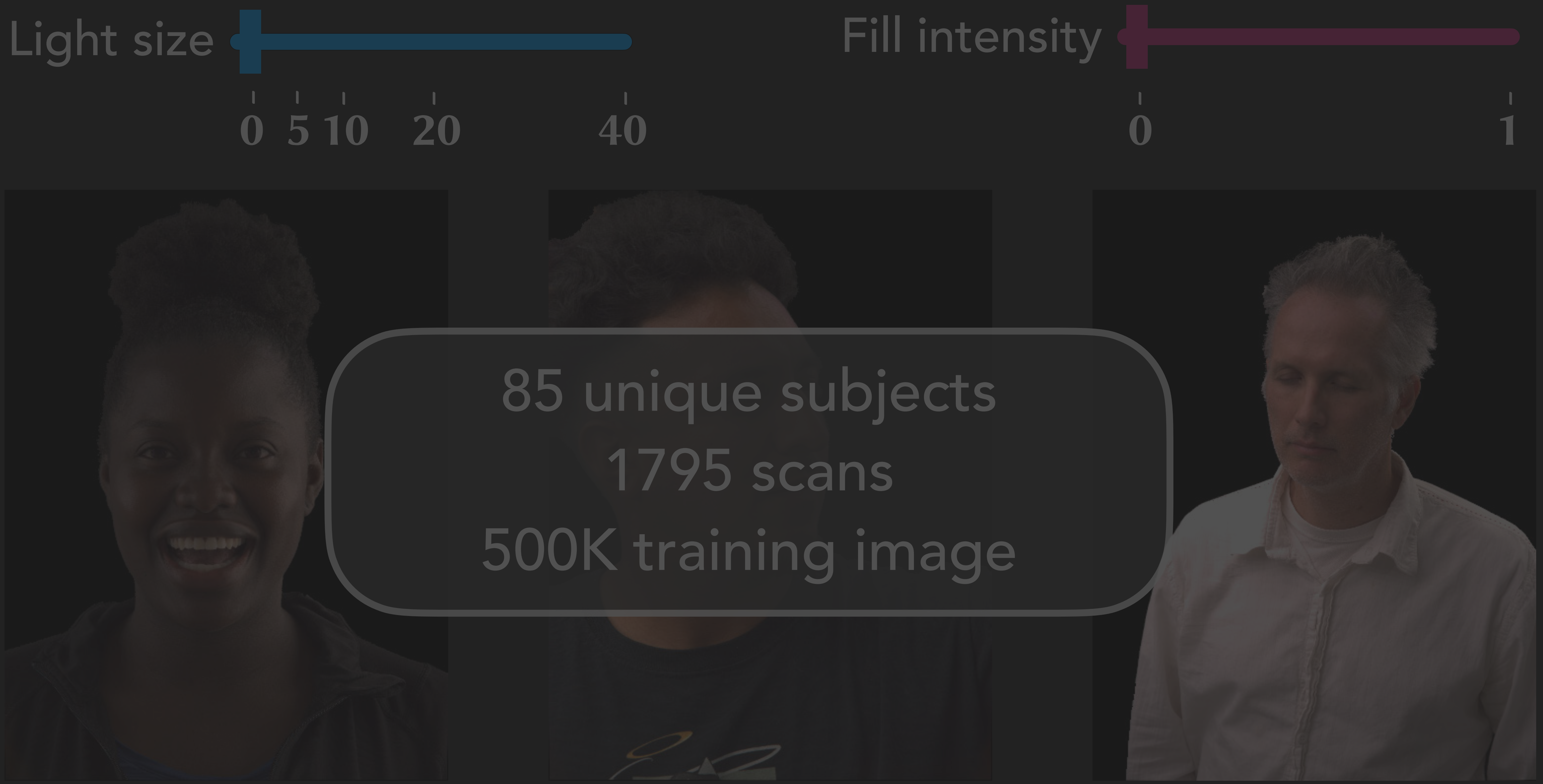


• Key light softened with light size 40

+

• Fill light max intensity*

*max intensity set to 1/10 of key light intensity



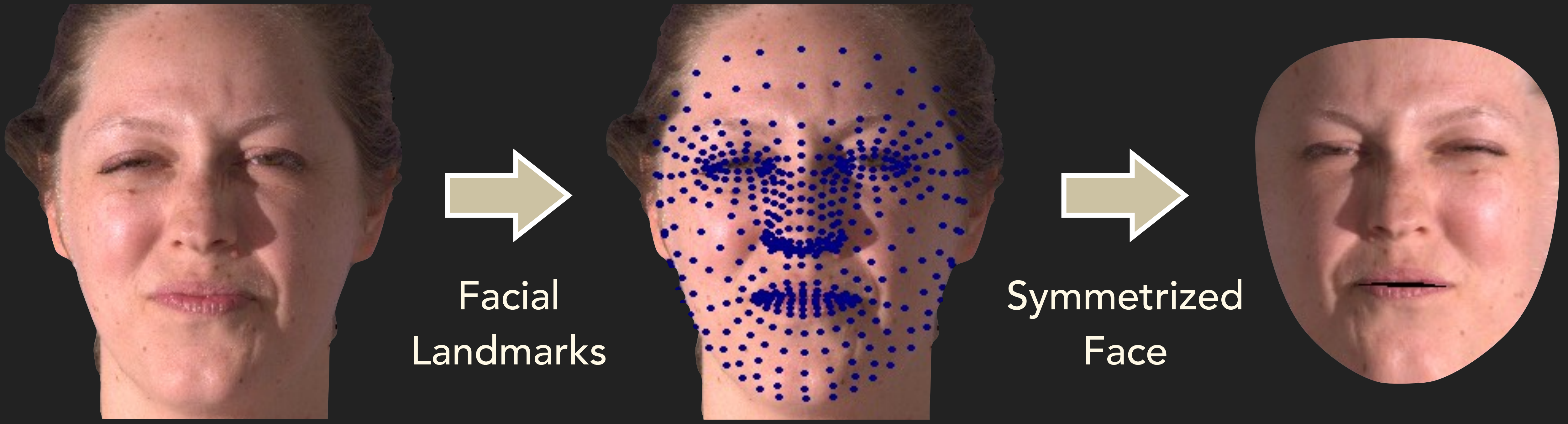
Examples of synthesized facial shadow data

Modeling Facial Symmetry

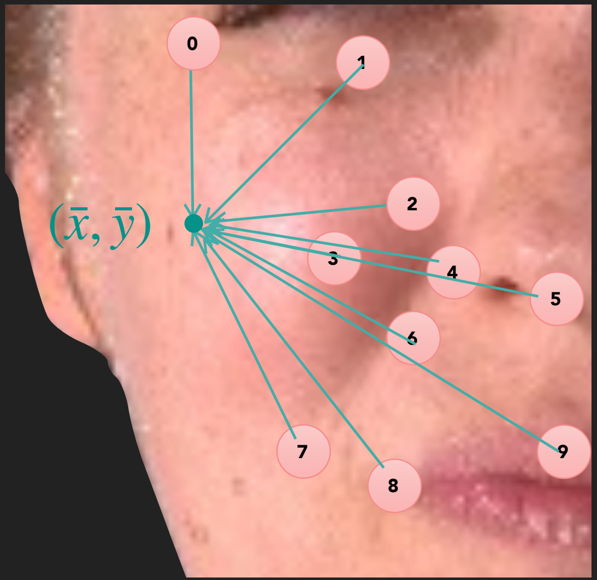
Modeling Facial Symmetry

- Human faces tend to be symmetric
- Facial shadow cast upon a face is likely not symmetric

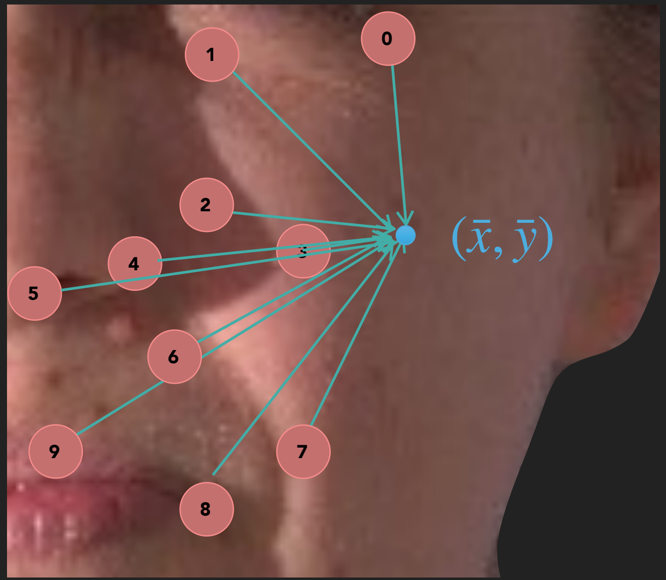
Synthesizing Symmetrized Face



Synthesizing Symmetrized Face



Facial Landmarks (u_i, v_i)



Facial Landmark Pairs $(u_{\bar{i}}, v_{\bar{i}})$

$$\bar{I} = I\left(\sum_j \mathrm{W}_{i,j} u_{\bar{j}}, \sum_j \mathrm{W}_{i,j} v_{\bar{j}}\right)$$

Input Faces

Input Faces and Their 'Mirrored' Pairs

Model and Training

- GridNet[1] architecture for both foreign shadow removal and facial shadow softening model
- Facial softening model is conditioned on the light size and fill intensity
- $I_{out} = I_{in} \cdot A + B$

[1] Fourure, Damien, et al. "Residual Conv-Deconv Grid Network for Semantic Segmentation." *BMVC 2017*.

Results — Foreign Shadow Removal

- Evaluation dataset: from high-frame rate video capture
- Test dataset: In-the-wild portrait photos

High Frame-Rate (60fps) Video For Evaluation

Multiple image pairs from a single video capture

Vid 15

Examples of Captured Videos

- 8 subjects
- 100 evaluation image pairs

Evaluation and test dataset are now online: eecs.berkeley.edu/~cecilia77/project-pages/portrait

Results on Foreign Shadow Removal

Cun et al. [1]

Ours

[1] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. 2020. Towards Ghost-free Shadow Removal via Dual Hierarchical Aggregation Network and Shadow Matting GAN. AAAI.

Results on Foreign Shadow Removal

Cun et al. [1]

Ours

Ground Truth

[1] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. 2020. Towards Ghost-free Shadow Removal via Dual Hierarchical Aggregation Network and Shadow Matting GAN. AAAI.

Results on Foreign Shadow Removal

Cun et al. [1]

Ours

[1] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. 2020. Towards Ghost-free Shadow Removal via Dual Hierarchical Aggregation Network and Shadow Matting GAN. AAAI.

Results on Foreign Shadow Removal

Cun et al. [1]

Ours

Ground Truth

[1] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. 2020. Towards Ghost-free Shadow Removal via Dual Hierarchical Aggregation Network and Shadow Matting GAN. AAAI.

Results on Foreign Shadow Removal

Baseline Removal Model	Rendered Test Set			Real Test Set		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
Input Image	20.657	0.807	0.206	19.671	0.766	0.115
[Guo et al. 2012]	19.170	0.699	0.359	15.939	0.593	0.269
[Hu et al. 2019]	20.895	0.742	0.238	18.956	0.699	0.148
[Cun et al. 2020]	22.405	0.845	0.173	19.386	0.722	0.133
Ours	<u>29.814</u>	<u>0.926</u>	<u>0.054</u>	<u>23.816</u>	<u>0.782</u>	<u>0.074</u>



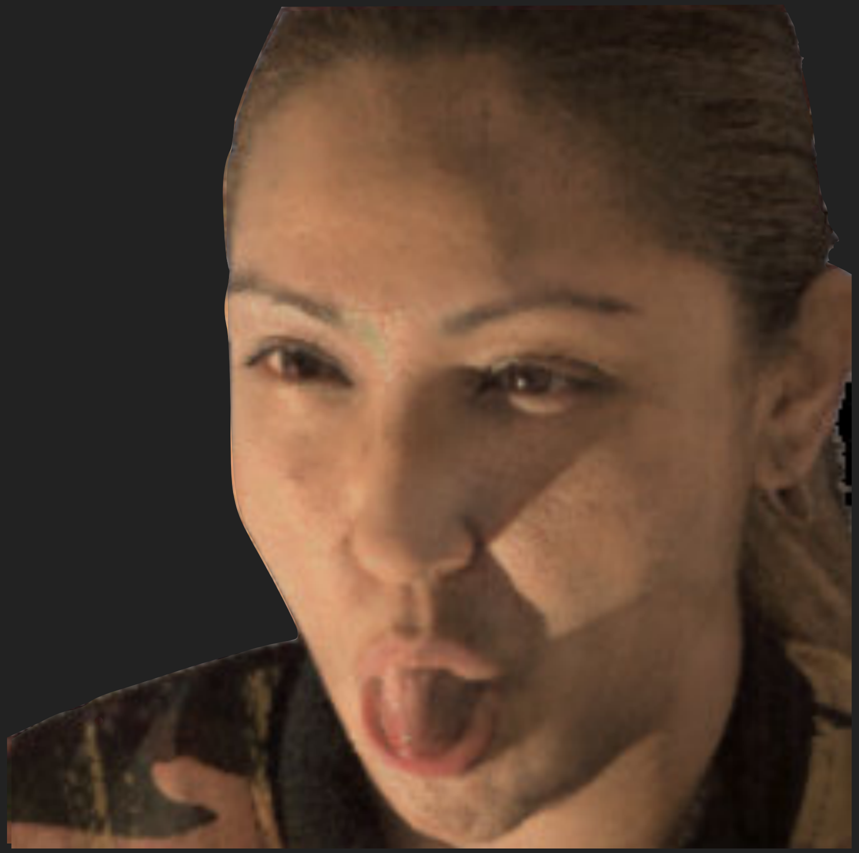




Results — Facial Shadow Softening

- Evaluation dataset: from Light Stage rendering
- Test dataset: In-the-wild portrait photos

Results on Facial Shadow Softening

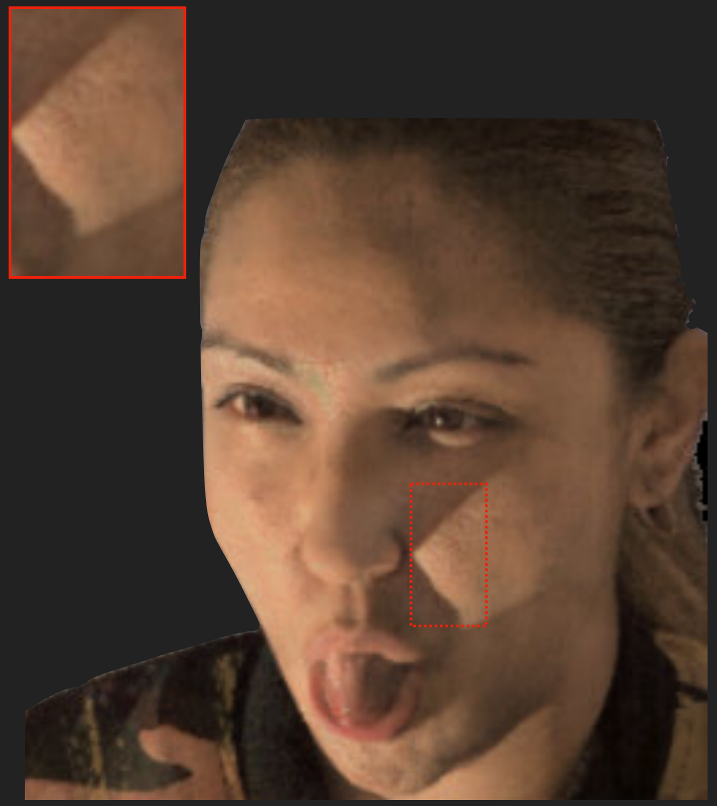


Sun et al. [1]

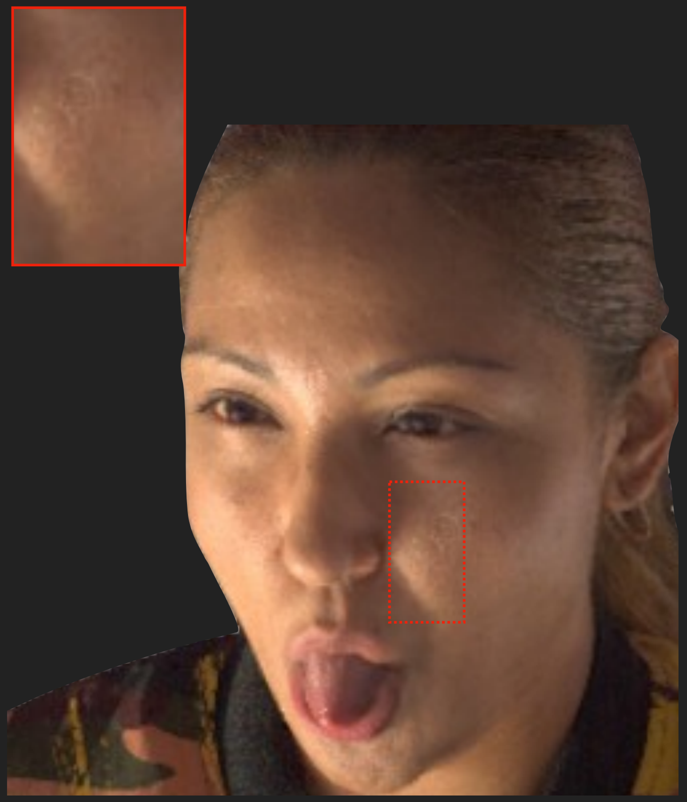
Ours

[1] Sun et al. "Single Image Portrait Relighting" SIGGRAPH 2019

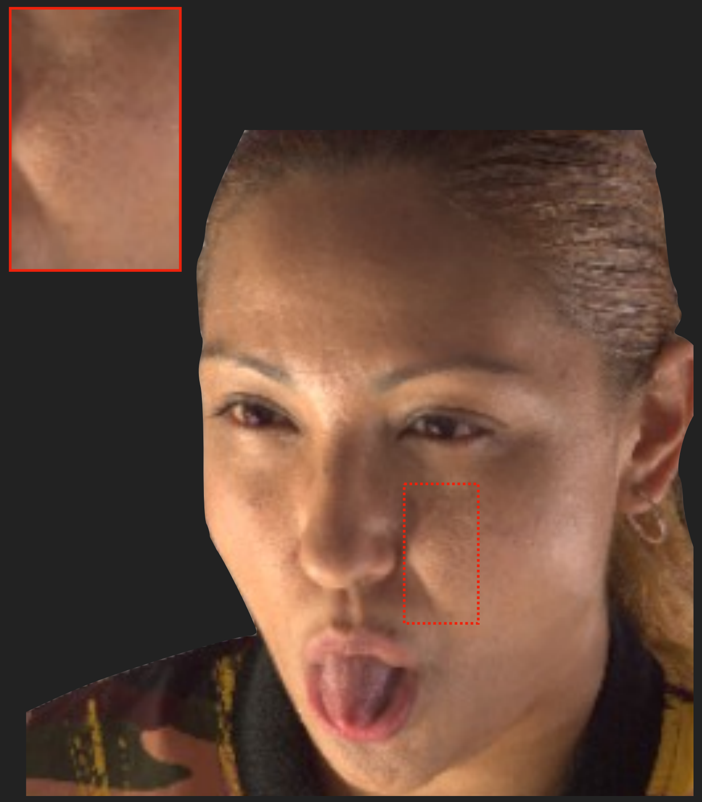
Results on Facial Shadow Softening



Sun et al. [1]
(blur environment map and relight)



Ours



Ground Truth

[1] Sun et al. "Single Image Portrait Relighting" SIGGRAPH 2019

Results on Facial Shadow Softening

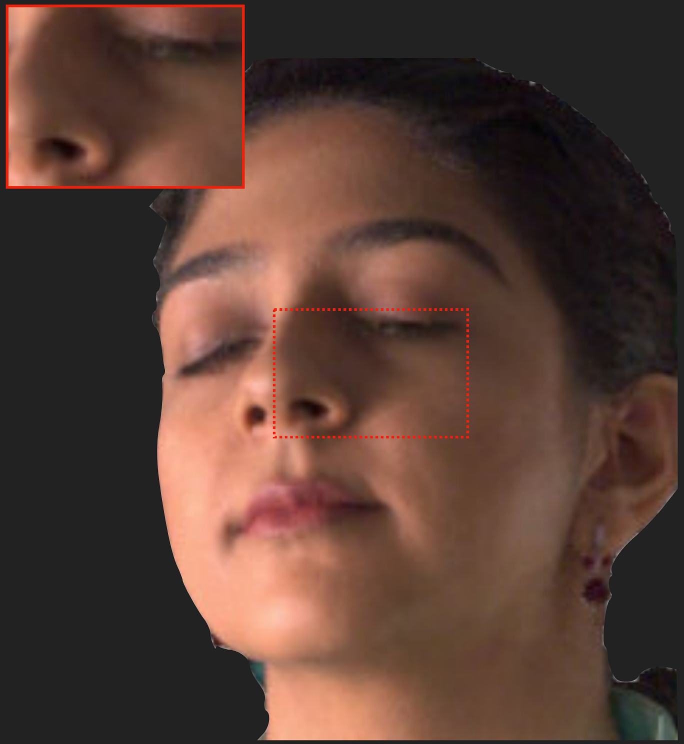
Sun et al. [1]

Ours

[1] Sun et al. "Single Image Portrait Relighting" SIGGRAPH 2019

Results on Facial Shadow Softening

Sun et al. [1]
(blur environment map and relight)



Ours

Ground Truth

[1] Sun et al. "Single Image Portrait Relighting" SIGGRAPH 2019

Results on Facial Shadow Softening

Baseline Shadow Reduction Model	PSNR↑	SSIM↑	LPIPS↓
PR-net [Sun et al. 2019]	21.639	0.709	0.152
Ours w/o Symmetry	24.232	0.826	0.065
Ours w/ Symmetry	<u>26.740</u>	<u>0.914</u>	<u>0.054</u>

Results on facial shadow softening

Light size

40 -

20 -

10-

0 -

Fill intensity

1* -

0.5 -

0.25-

0 -

*1 corresponds to the max fill light intensity

Results on facial shadow softening

Before

After

Results on facial shadow softening

Light size

40 -

20 -

10-

0 -

Output light size 40, fill 0.25

Fill intensity

1* -

0.5 -

0.25-

0 -

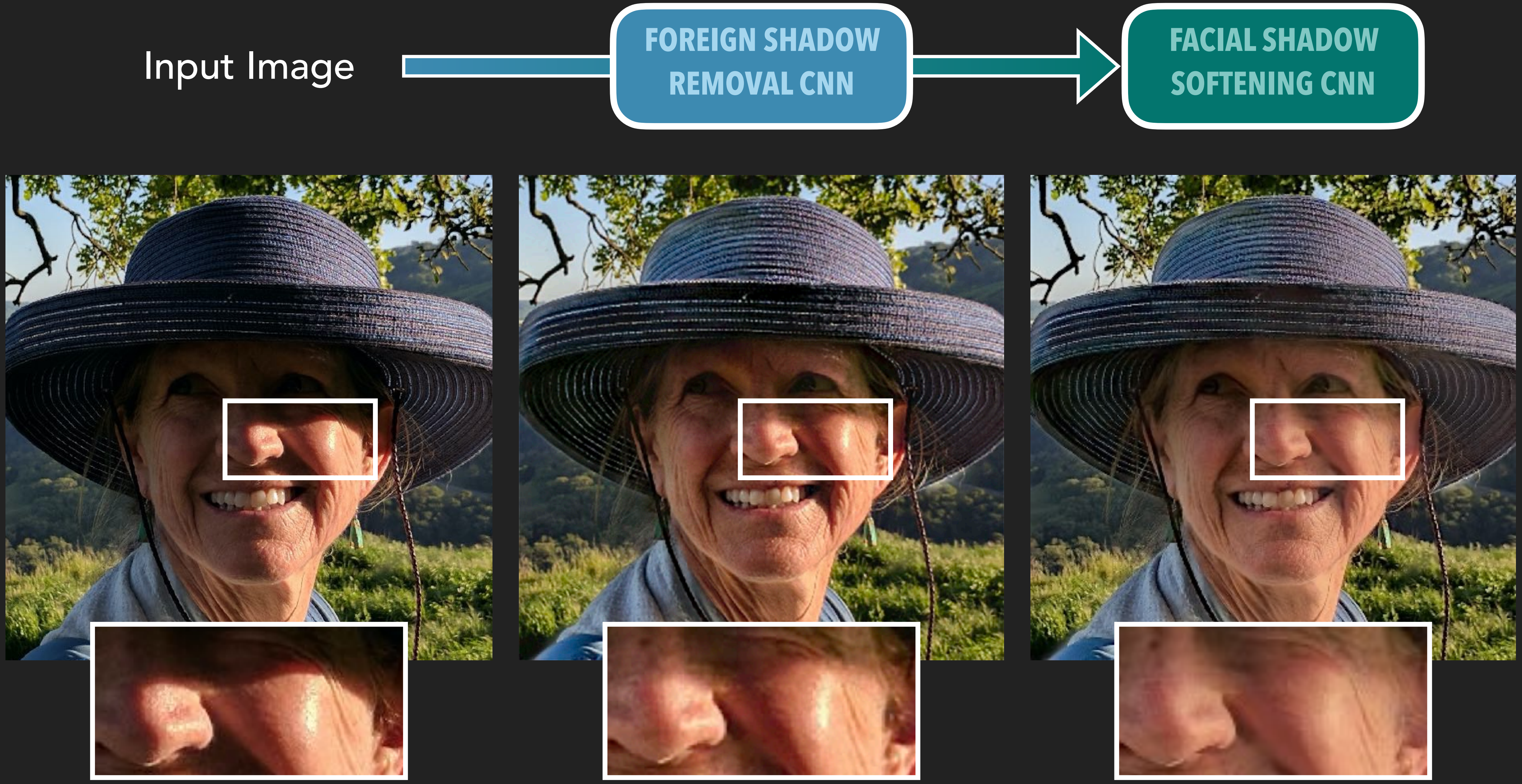
*1 corresponds to the max fill light intensity

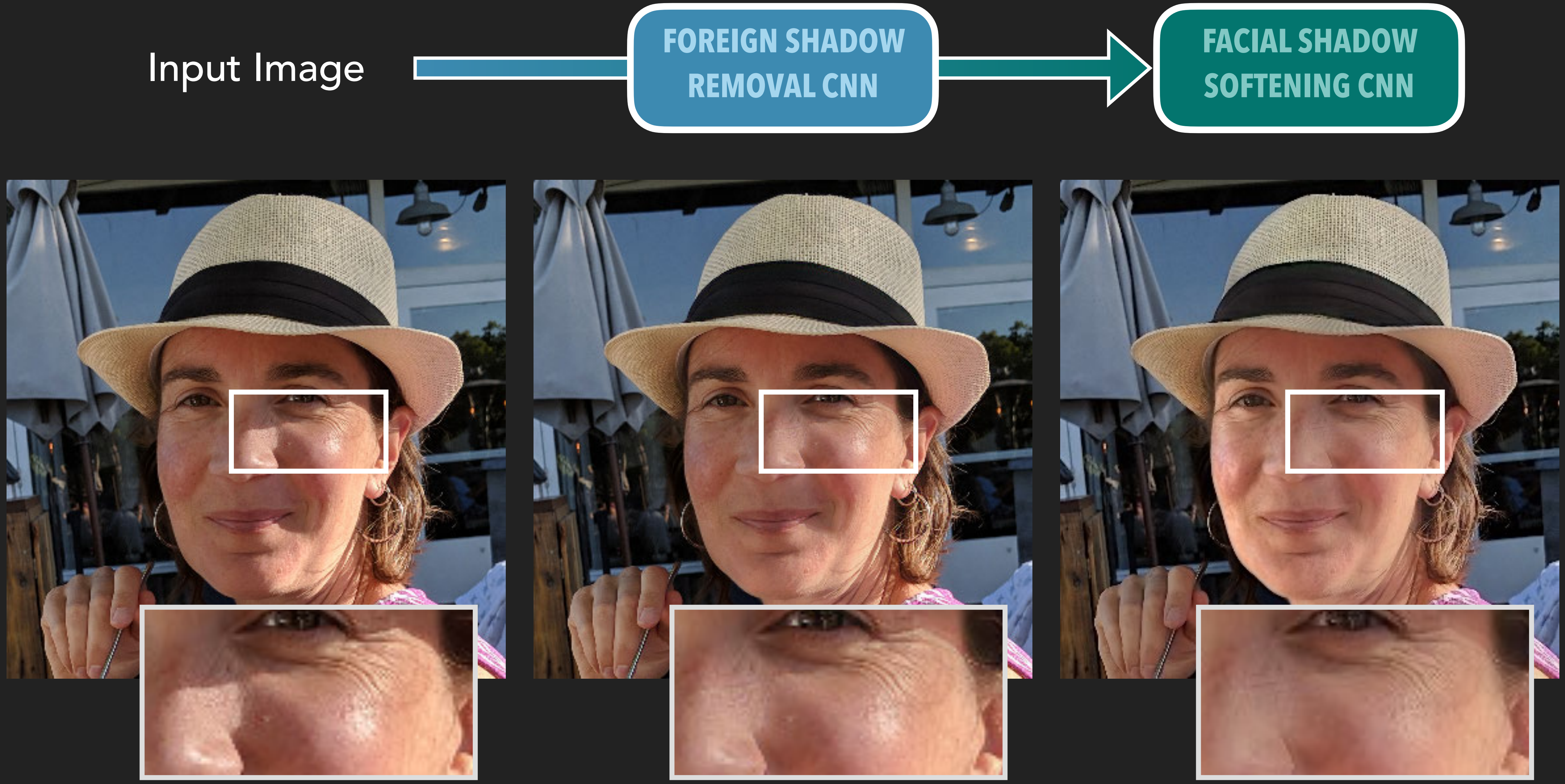
Results on facial shadow softening

Before

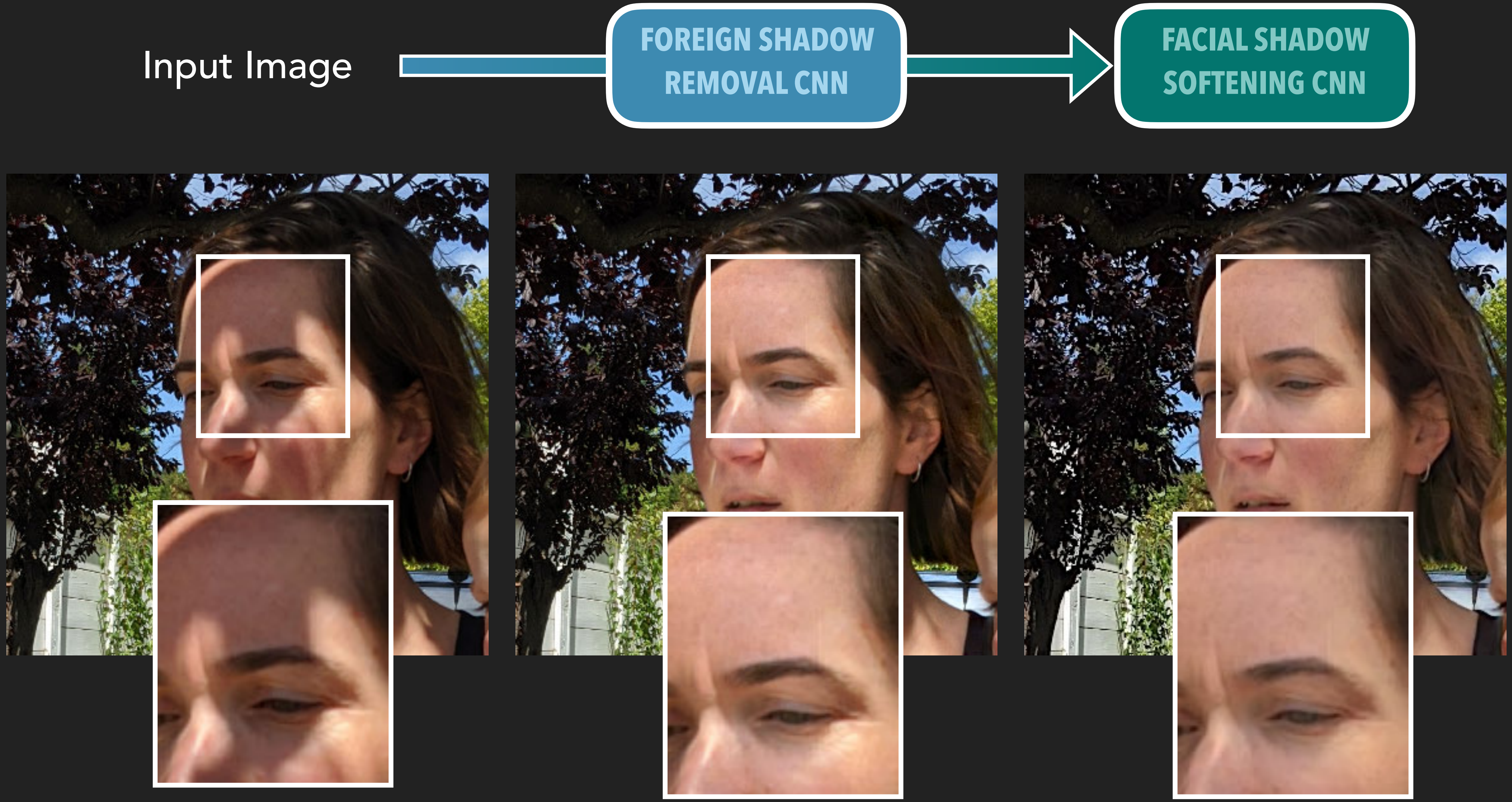
After

Results — 2-Stage Portrait Enhancement

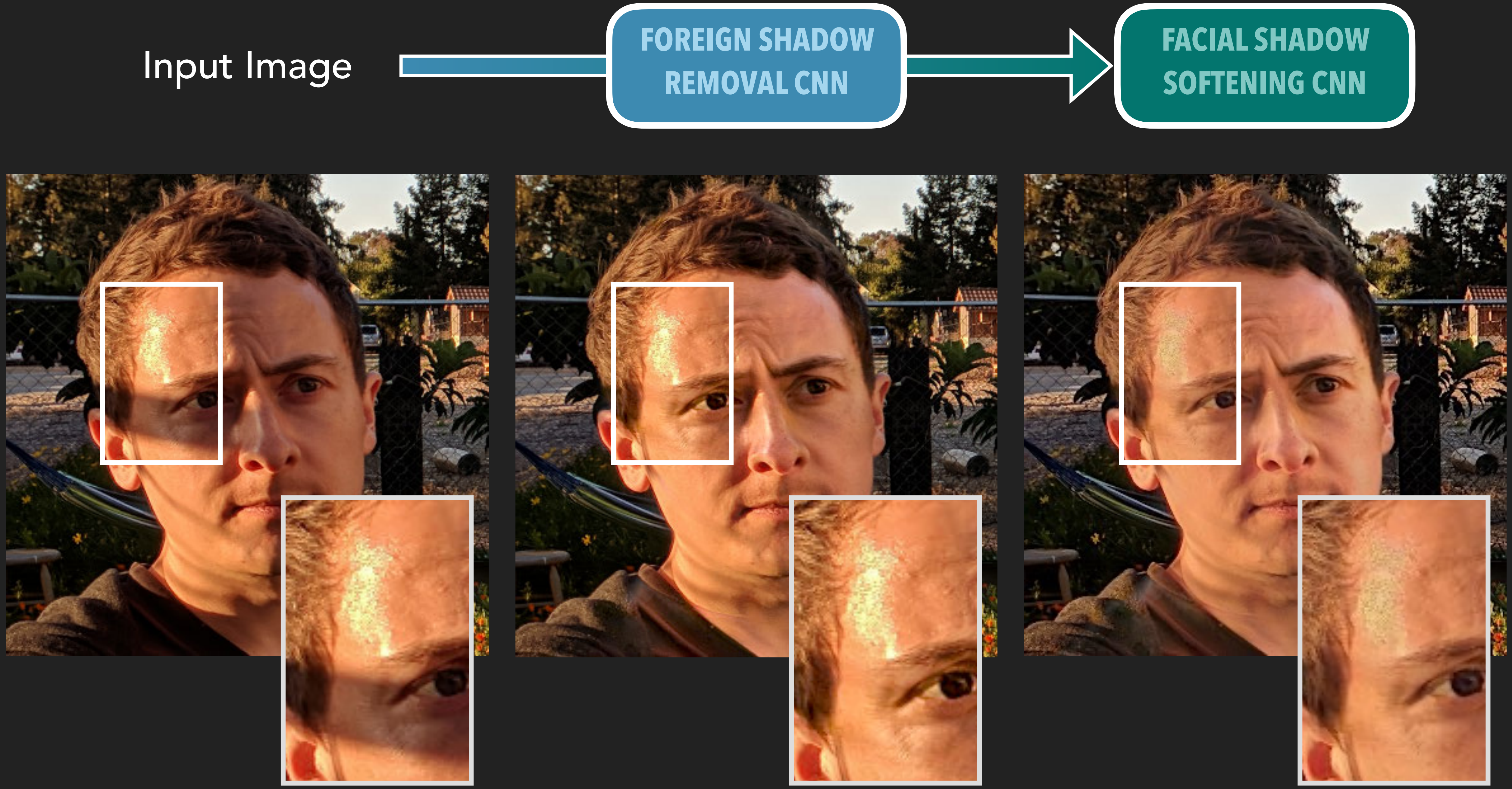

Input Image
FOREIGN SHADOW REMOVAL CNN
FACIAL SHADOW SOFTENING CNN



Input Image
FOREIGN SHADOW REMOVAL CNN
FACIAL SHADOW SOFTENING CNN



Input Image
FOREIGN SHADOW REMOVAL CNN
FACIAL SHADOW SOFTENING CNN


Input Image
FOREIGN SHADOW REMOVAL CNN
FACIAL SHADOW SOFTENING CNN



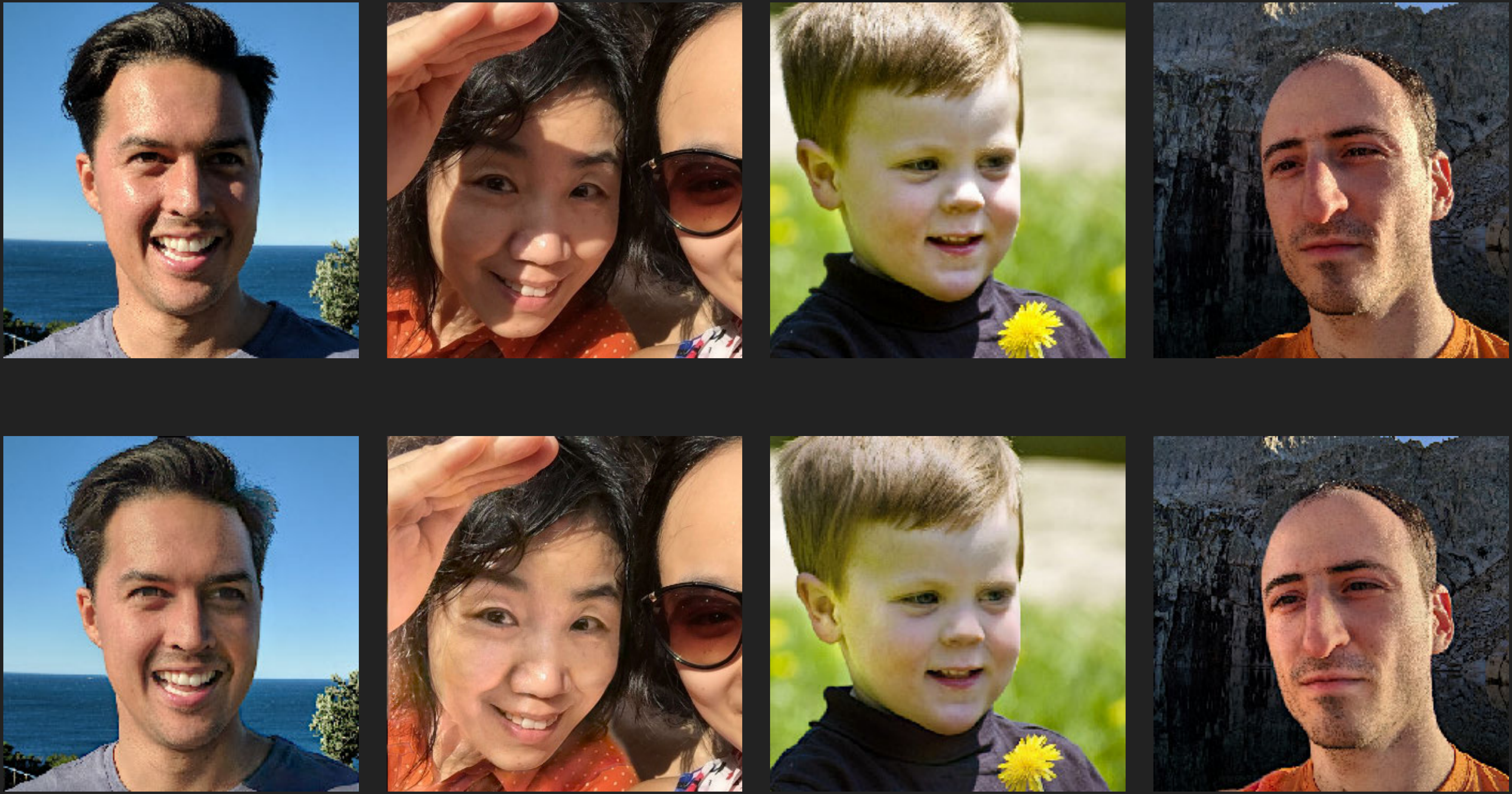

More Results on Portrait Enhancement
















KEEP


KEEP
KEEP

Context-Aware Casual Imaging

Image Quality

Zoom To Learn, Learn To Zoom

Zhang et al, CVPR 2019

- Learning From Raw Sensor
- Image Super-Resolution

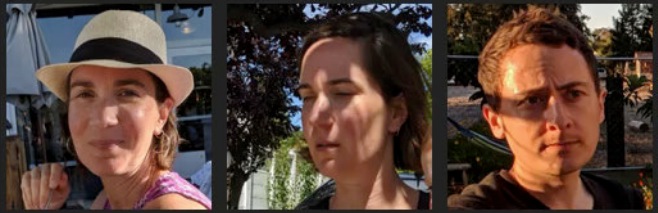
Focus

Synthetic Defocus and Look-Ahead Autofocus for Causal Videography

Zhang et al, SIGGRAPH 2019

- Video Synthetic Defocus
- ‘Future’ Scene Understanding

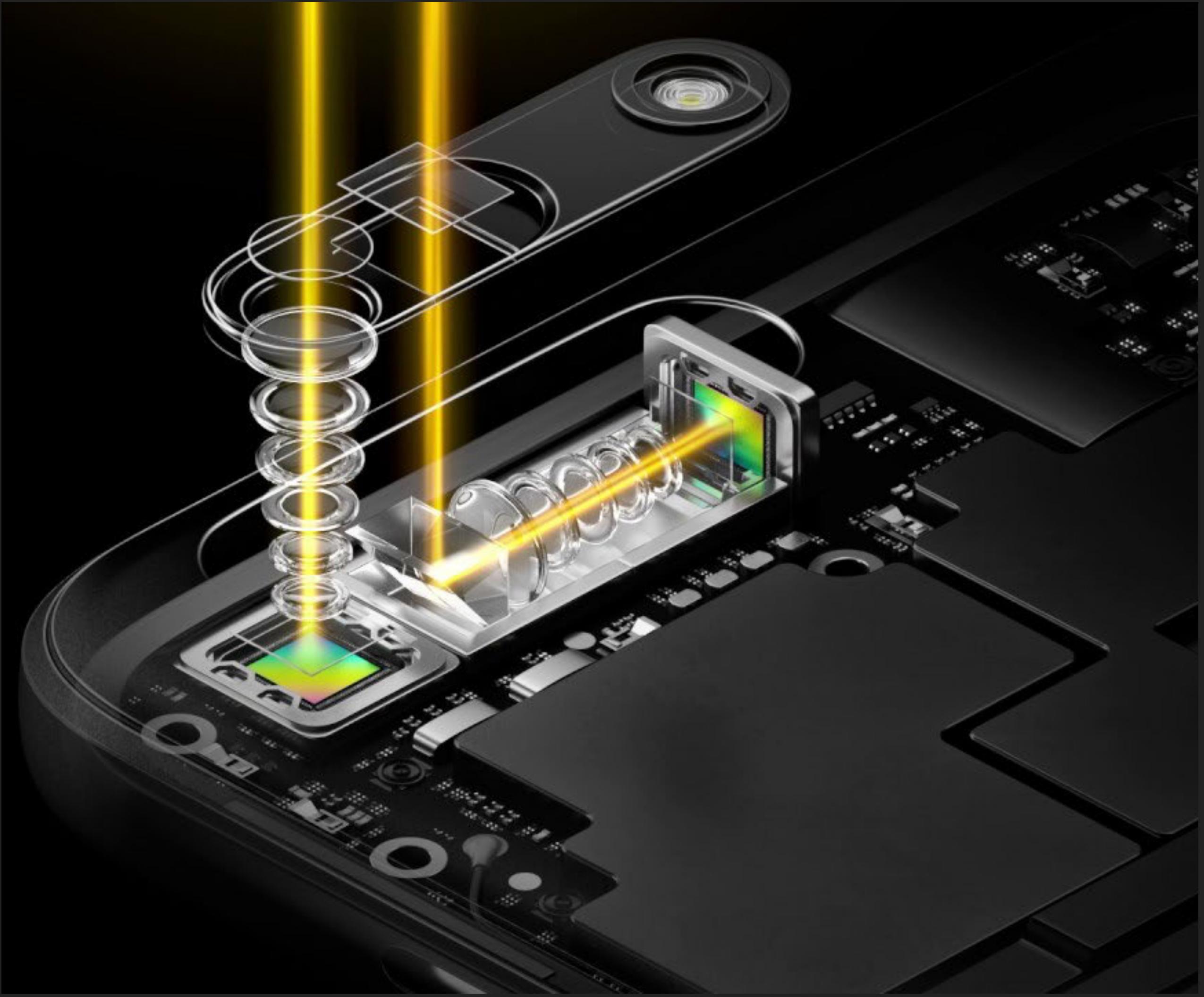
Lighting



Portrait Shadow Manipulation

Zhang et al, SIGGRAPH 2020

- Shadow Editing Post-Capture
- Embedding Professional Lighting Principals



- Multi-Camera Systems for Cinematic Storytelling
- Generative Models for Photo Editing Systems
- Recent Advancement in Imaging Devices

What's Next?

Image Quality

Zoom To Learn, Learn To Zoom

Zhang et al, CVPR 2019

Single Image Reflection Removal With Perceptual Losses

Zhang et al, CVPR 2018

Learned Dual-View Reflection Removal

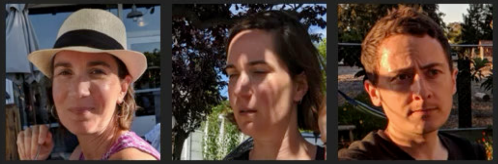
Niklaus et al

Focus

Synthetic defocus and look-ahead autofocus for causal videography

Zhang et al, SIGGRAPH 2019

Lighting

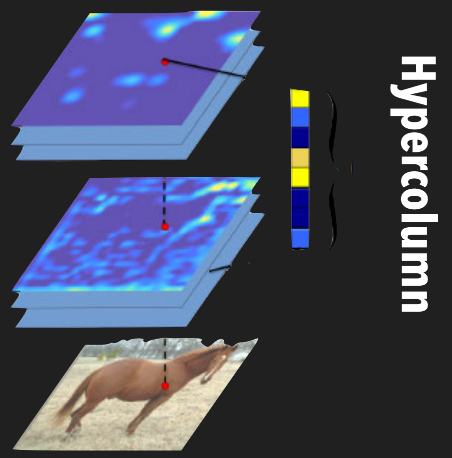


Portrait Shadow Manipulation

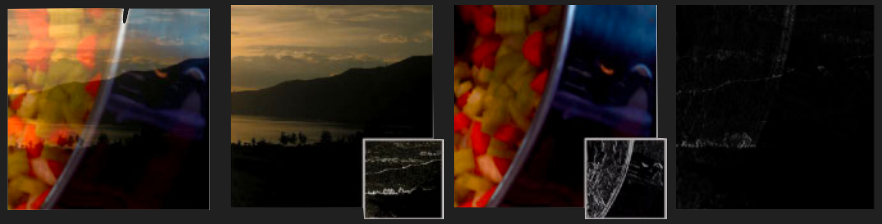
Zhang et al, SIGGRAPH 2020

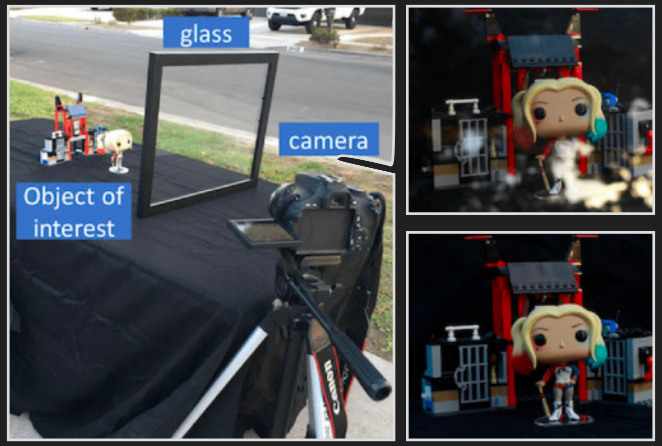
Reflection Removal

Reflection Removal With a Single Image



[CVPR2015 Hariharan et al]





- Use all-level features to bring ‘context’ into the reflection removal
- Sparsity of image gradients as a differentiable constraint
- Real-world dataset for benchmark evaluation

Reflection Removal With a Single Image

Single Image Input

Output Transmission Layer

Output Reflection Layer

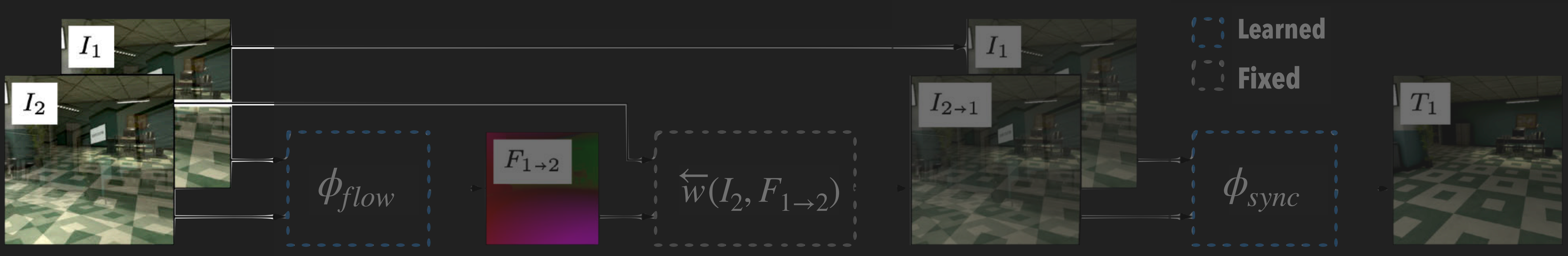
Reflection Removal With a Single Image

Failure Case: Intrinsic Ambiguity of Single-Image
Separating Transmission and Reflection

Input

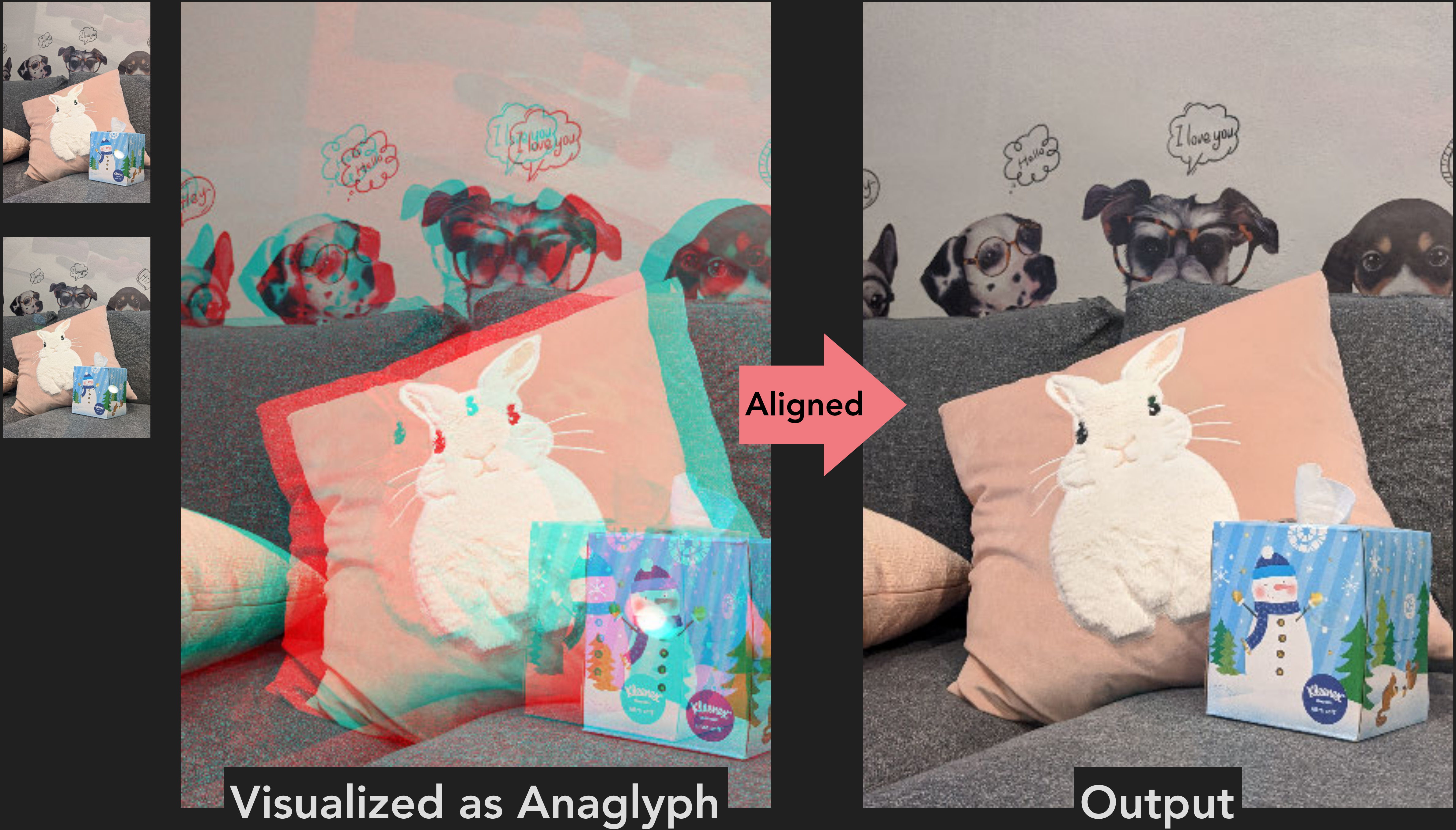
Output Transmission Layer

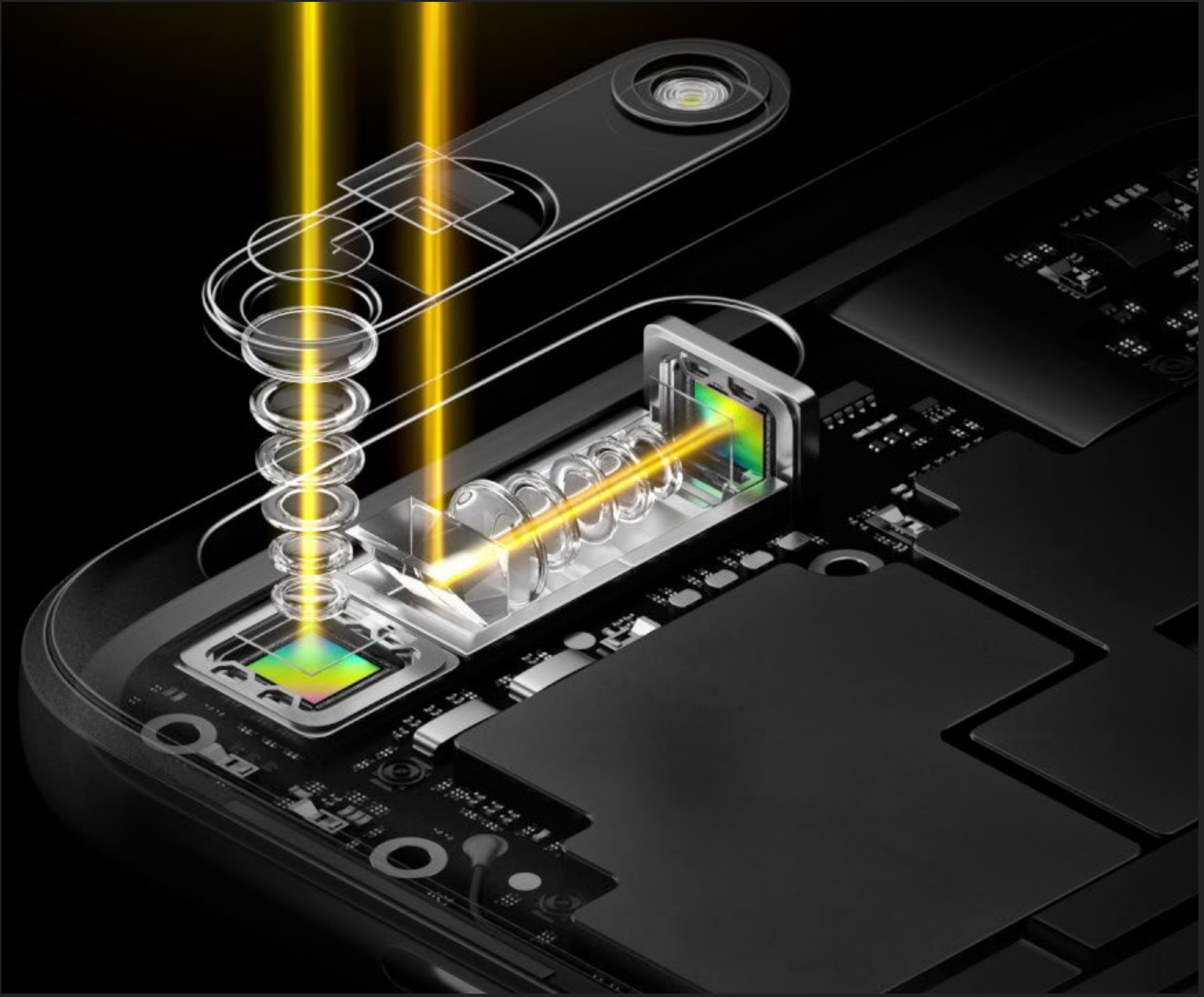
Reflection Removal With Dual-View



- Reflection-Invariant Optical Flow
- Parallax Only in Reflection Layer
- Fully Differentiable


Hello
I love you
Hi!


Visualized as Anaglyph
Aligned
Output



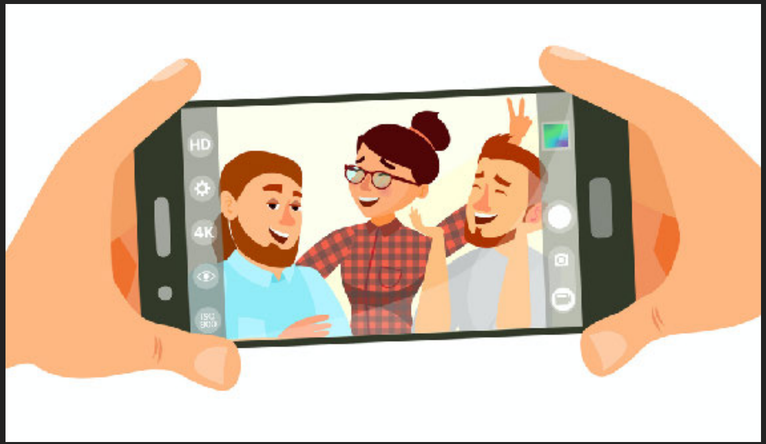
- Multi-Camera Systems for Cinematic Storytelling
- Generative Models for Photo Editing Systems
- Recent Advancement in Imaging Devices

What's Next?

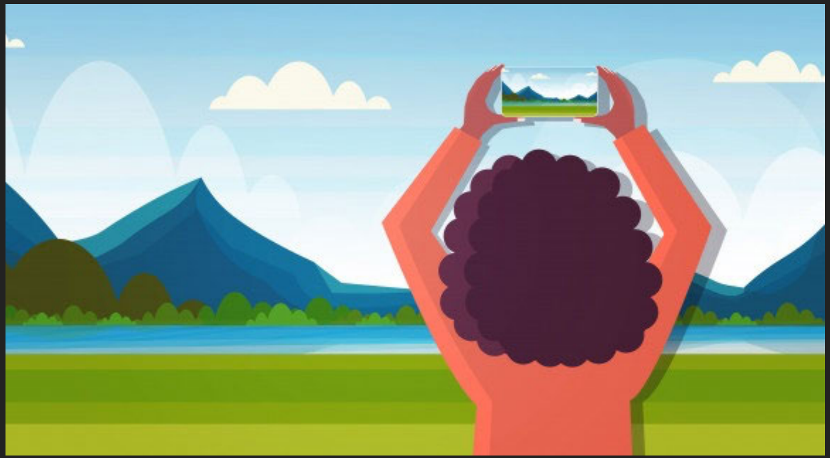
Multi-Camera System

- Most phones are equipped with >=2 cameras
  - But they hardly interact

Front Camera



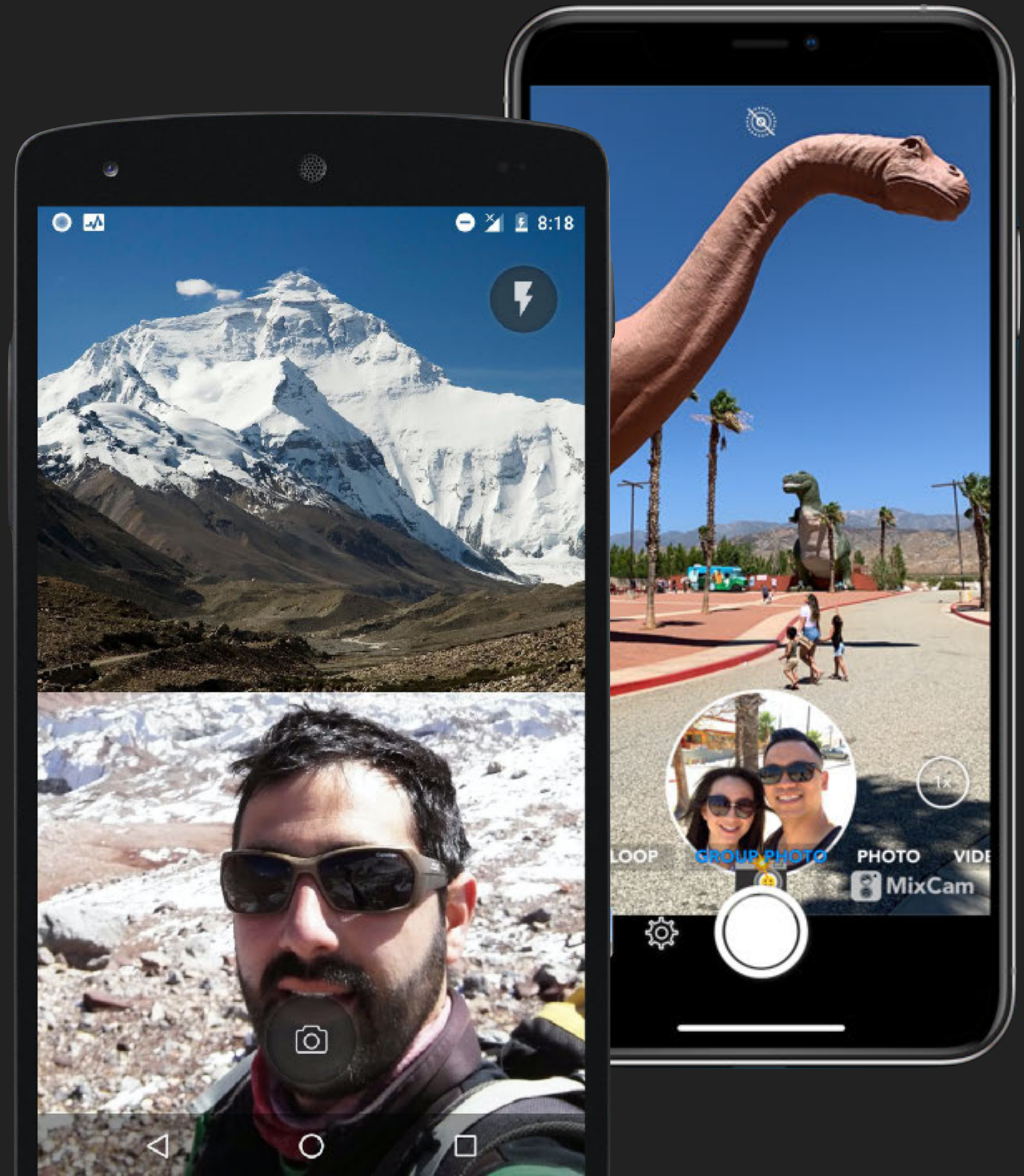
Rear Camera



Wide-Angle Camera

Multi-Camera System Brings More Context

- Additional information
  - Lighting
  - Content
  - Vlog-like storytelling

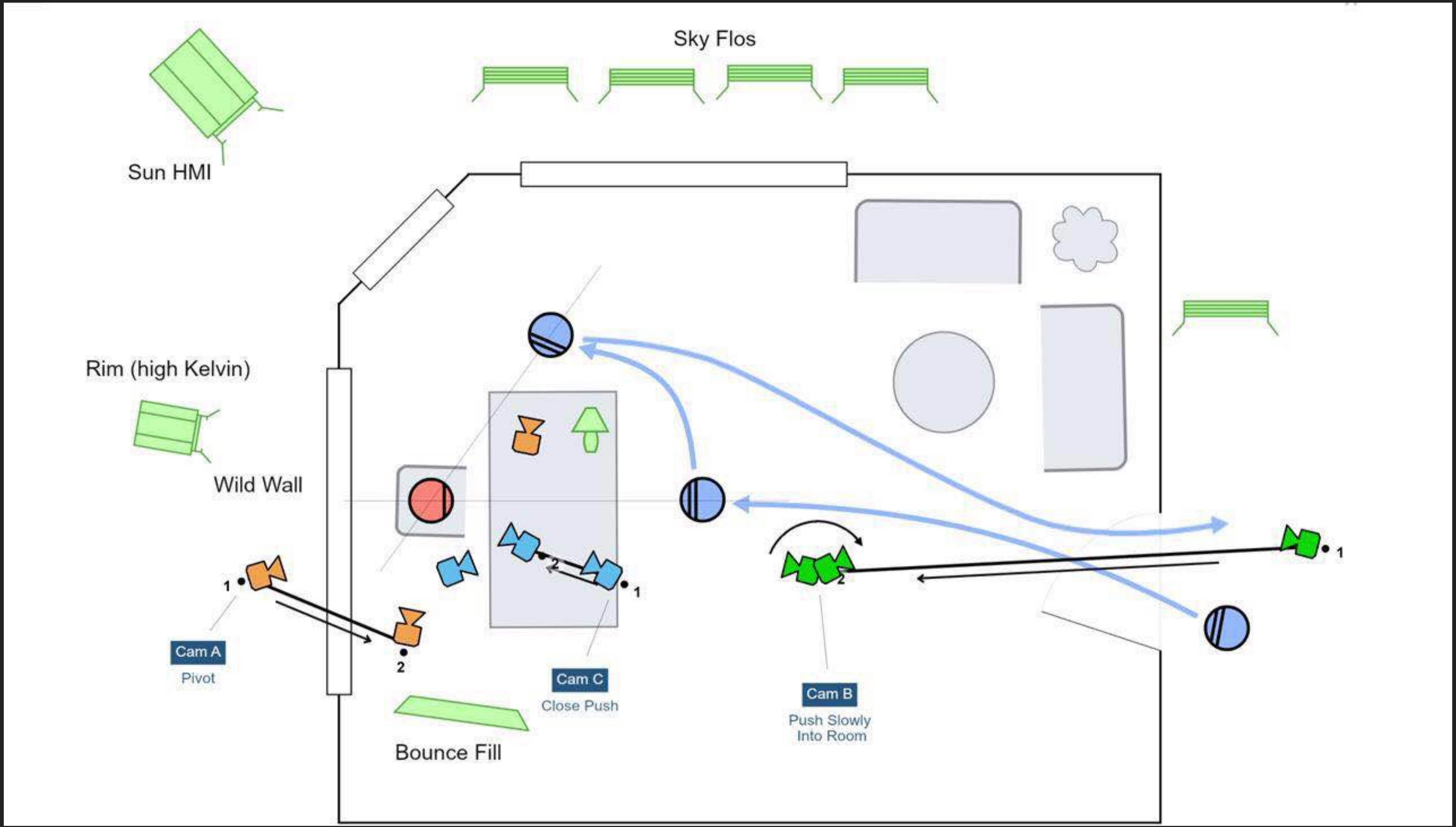


Multi-Camera for Storytelling

- 8K raw
- Wide-angle lens

Video credit: https://www.youtube.com/watch?v=YHYu_ORC7CE @ Tyler Stalman

In a Film Crew: Staging & Blocking Finds the Optimal (X, Y, Z, T)



Neural Rendering for AI Driven Photo-editing Systems

Karras, Tero, Samuli Laine, and Timo Aila. "A style-based generator architecture for generative adversarial networks."

AI Driven Photo-editing Systems *

- **Practicality**: generalize well
- **Controllability**: intuitive, high-level abstraction
- **Identity**: preserves the identity and details

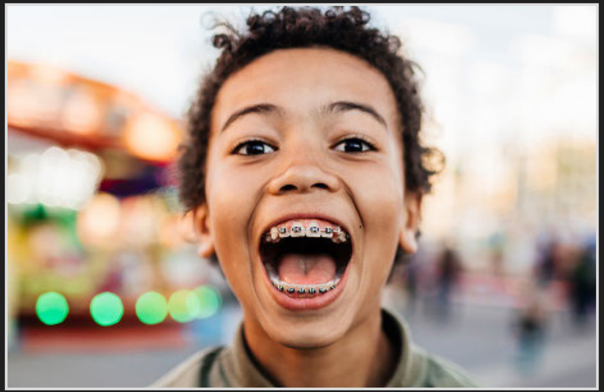
My Shot in, my Shot out

*Inspired by a conversation with Tim Brooks

Practicality: Generalization and Robustness

- Casual photos are noisy

Uncommon style

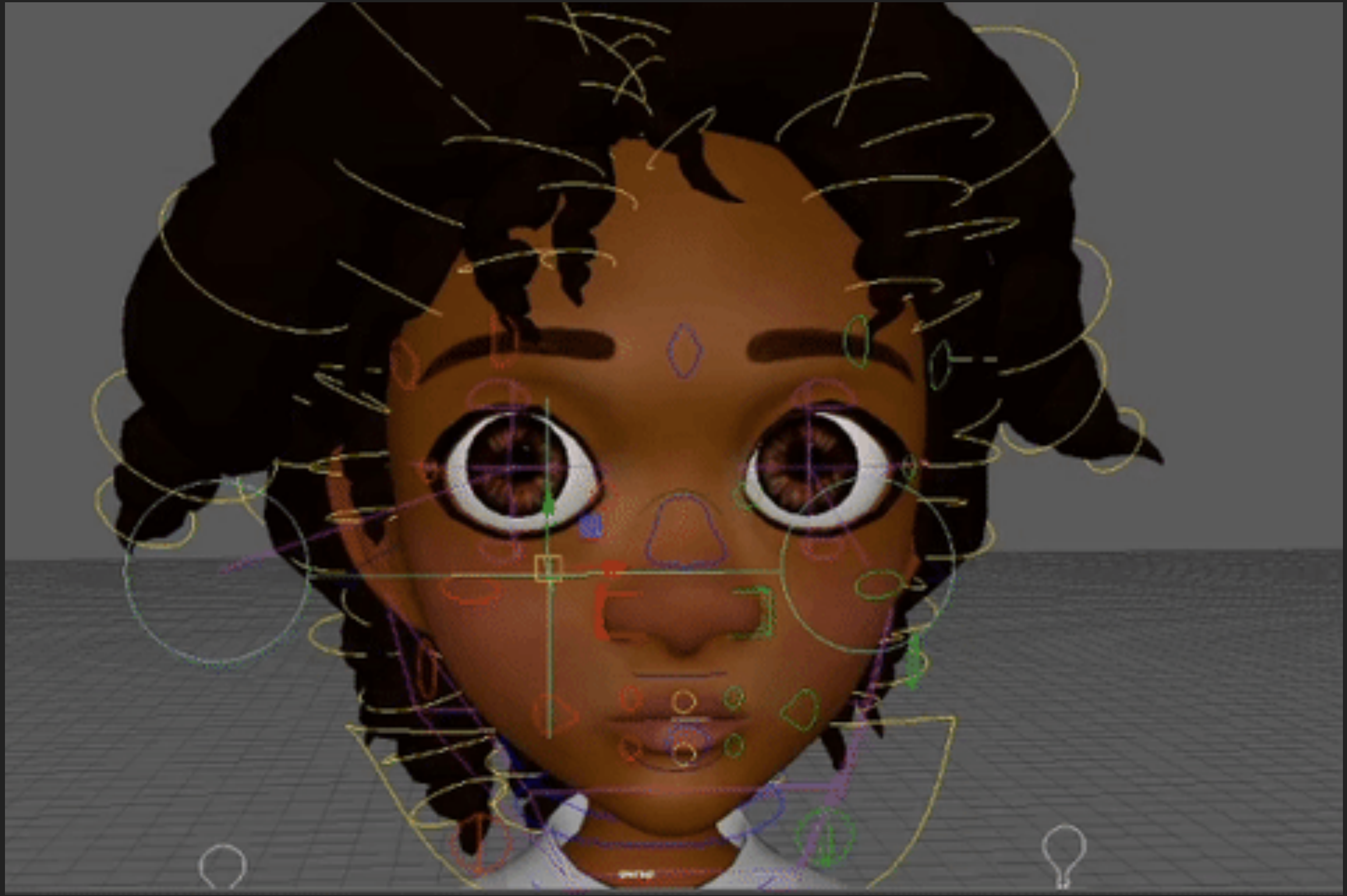


Uncommon accessories

Noisy and defocus

Pose and upper body

Controllability: Intuitive, High-Level Abstraction



Identity: Preserves the Identity and Details

[Richardson et al Arxiv 2020]
[Pidhorskyi et al CVPR 2020]

Neural Rendering for Photo-editing Systems

Controllability	Per-Pixel	Per-Segment	Per-Global-Style
Identity	Transform	Adding / Removing	Replacement / Reenactment
Photo-Editing	Enhancement	Manipulation (Resampling)	Fake

UI / UX
Data Bias
Forensics
Privacy
...

Hardware Efforts

- SPAD + 100K Burst
- AI-Based Chip & ISP
- Real-Time Eye Tracking Focusing
- 8K Raw Videos (30 mins)
- 8-Stop in-Body Stabilization ...
- Wide-Angle Lens ...
- Multi-Reflection Telephoto Lens ...
- Mobile Lidar ...

Software Efforts

- Burst Photography
- Image Warping
- Robust Temporal Alignment
- Handling Dynamic & Motion

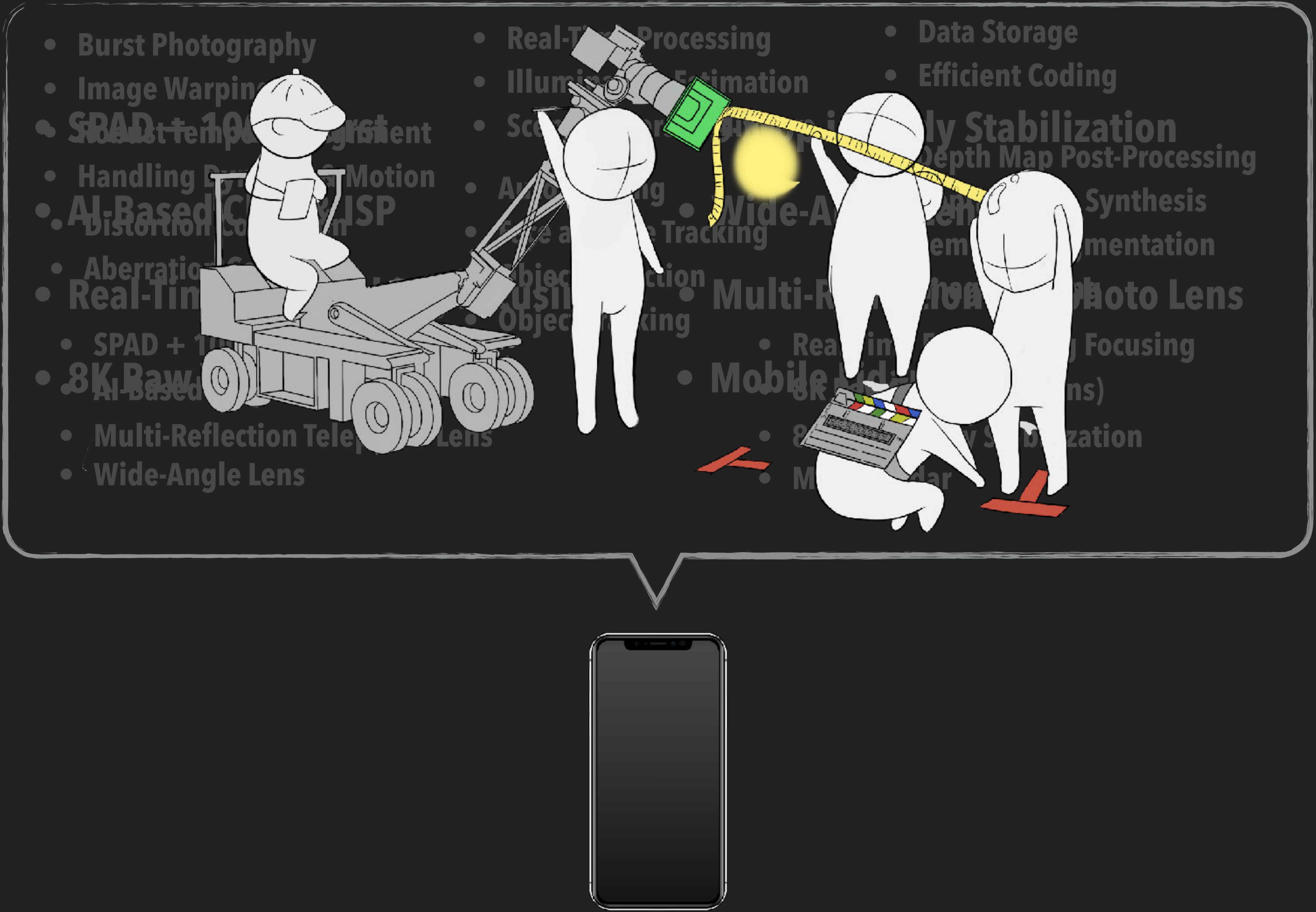
- Real-Time Processing
- Auto-X Algorithms
- Illumination/Scene Understanding

- Distortion Correction
- Aberration Correction

- Autofocusing
- Face and Eye Tracking
- Object Detection
- Object Tracking

- Data Storage
- Efficient Encoding

- Depth Map Post-Processing
- Depth-Aware Synthesis
- Semantic Segmentation
- Image Matting


Burst Photography
Data Storage
Efficient Coding
Depth Map Post-Processing
Wide-Angle Lens

A Preview of my Acknowledgment

- You will see
  - The people (and stuff) who help me survive until this moment
  - Words I've always wanted to say but didn't say
  - Photos of many cute people (and maybe a dog)
  - Many fancy (and unnecessary) Keynote animation that I never had a chance to use
  - How I see myself in 5 years (😯)

Feel free to unmute yourself ❤️

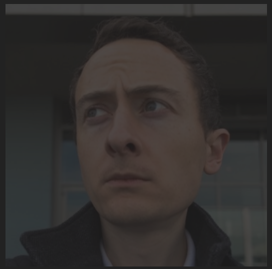
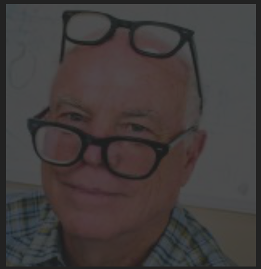
Ren

Alyosha

JRK

Marty



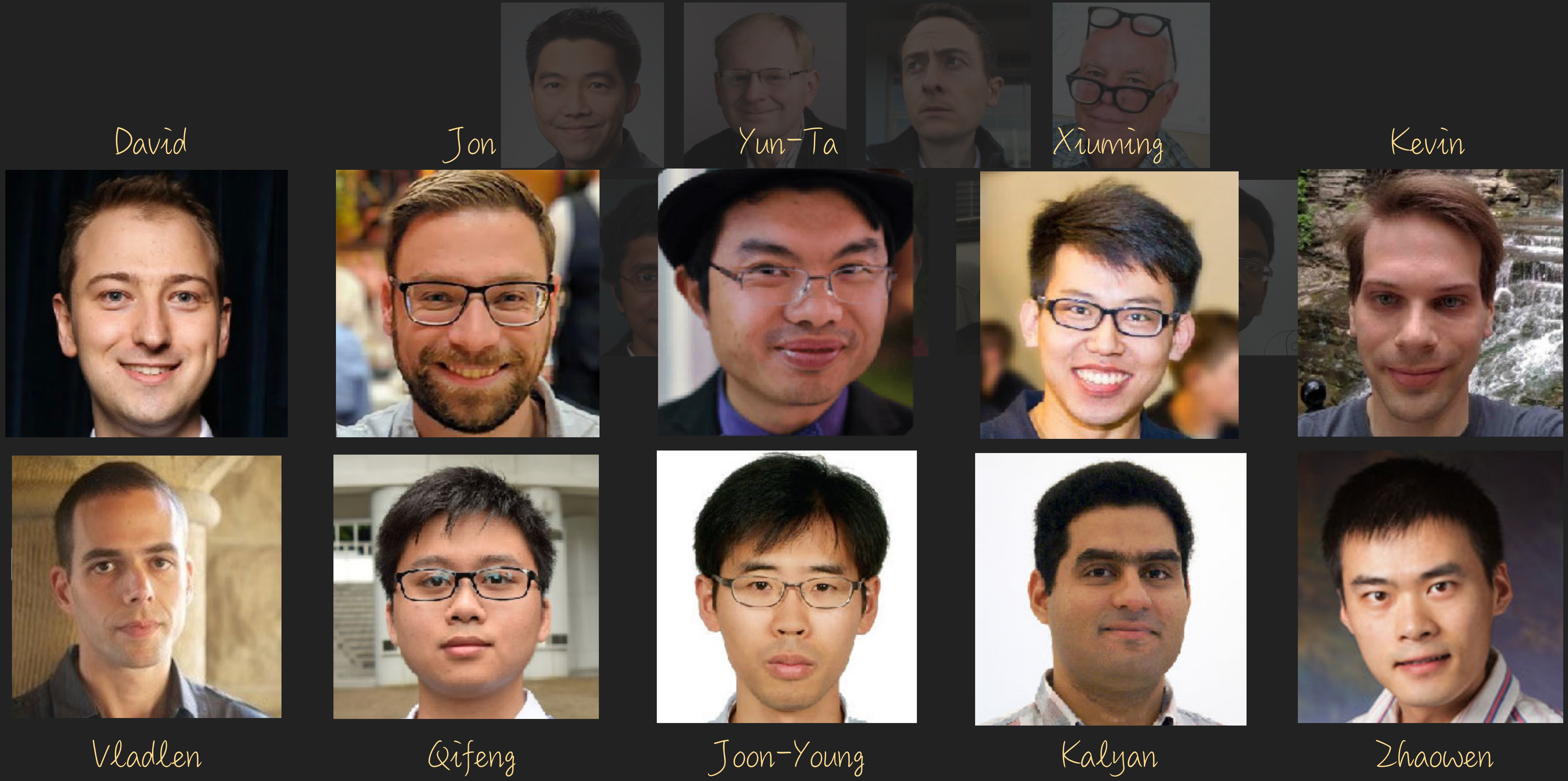
Don

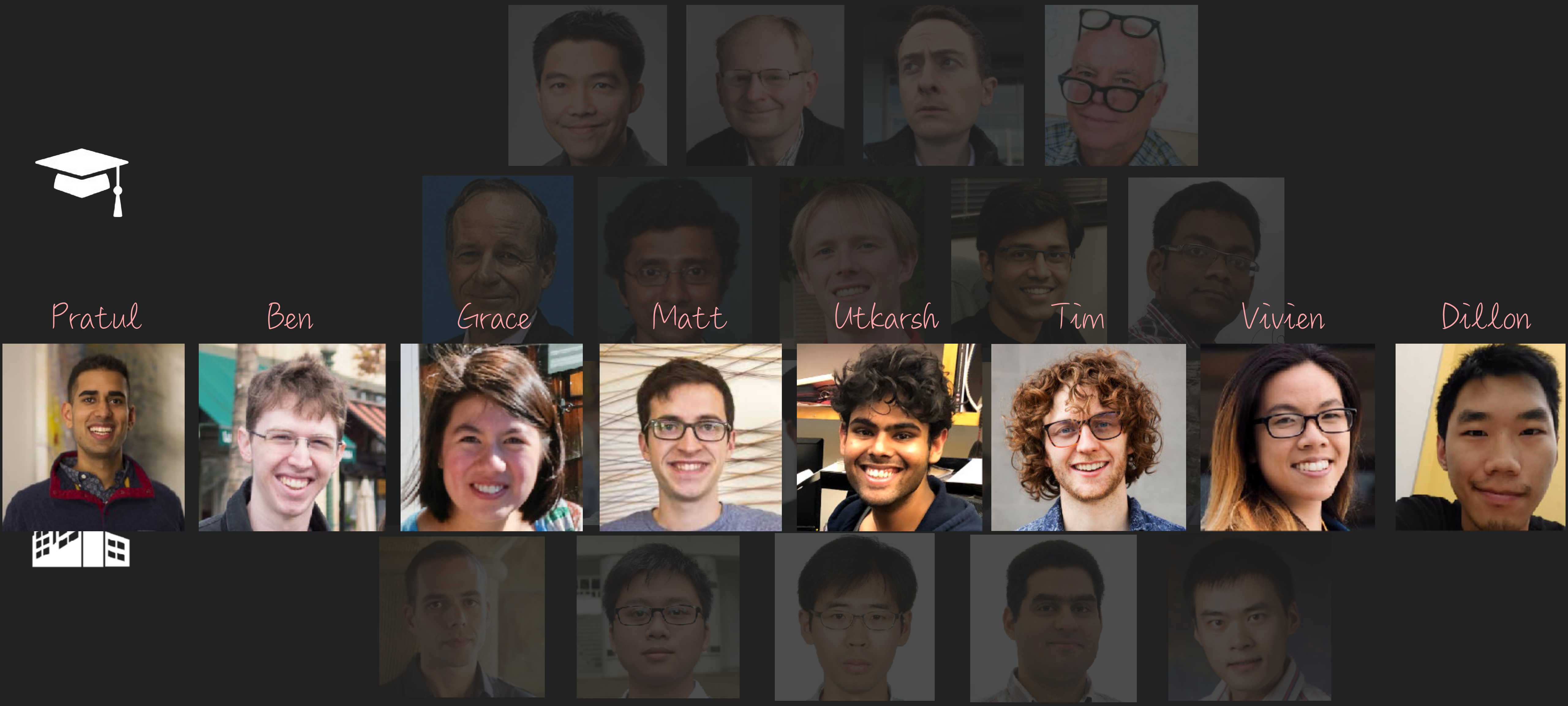
Ashok

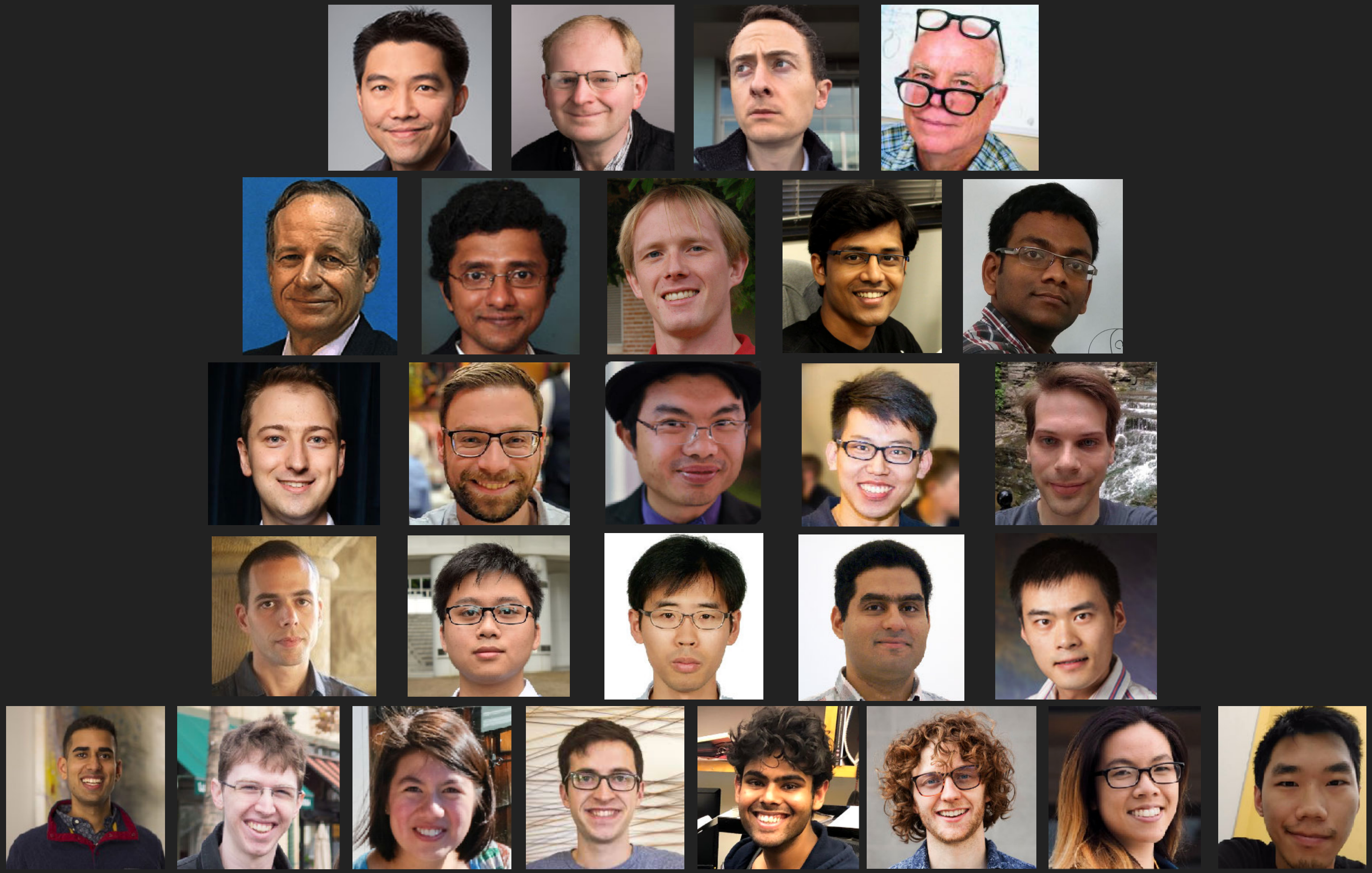
Jason

Vivek

Adithya


David
Jon
Yun-Ta
Xiuming
Kevin
Vladlen
Qifeng
Joon-Young
Kalyan
Zhaowen


Pratul
Ben
Grace
Matt
Utkarsh
Tim
Vivien
Dillon



Shirley

Audrey

Jean

Susanne

Angela

Angie


Holly
Jenny
Tinghui
Lantao
ST
QQ
Julie
Yoko
Yi
Mandy
Jun-yan
Biye
Lingqi
Daniel
Hezheng

Thanks

Specially, ...

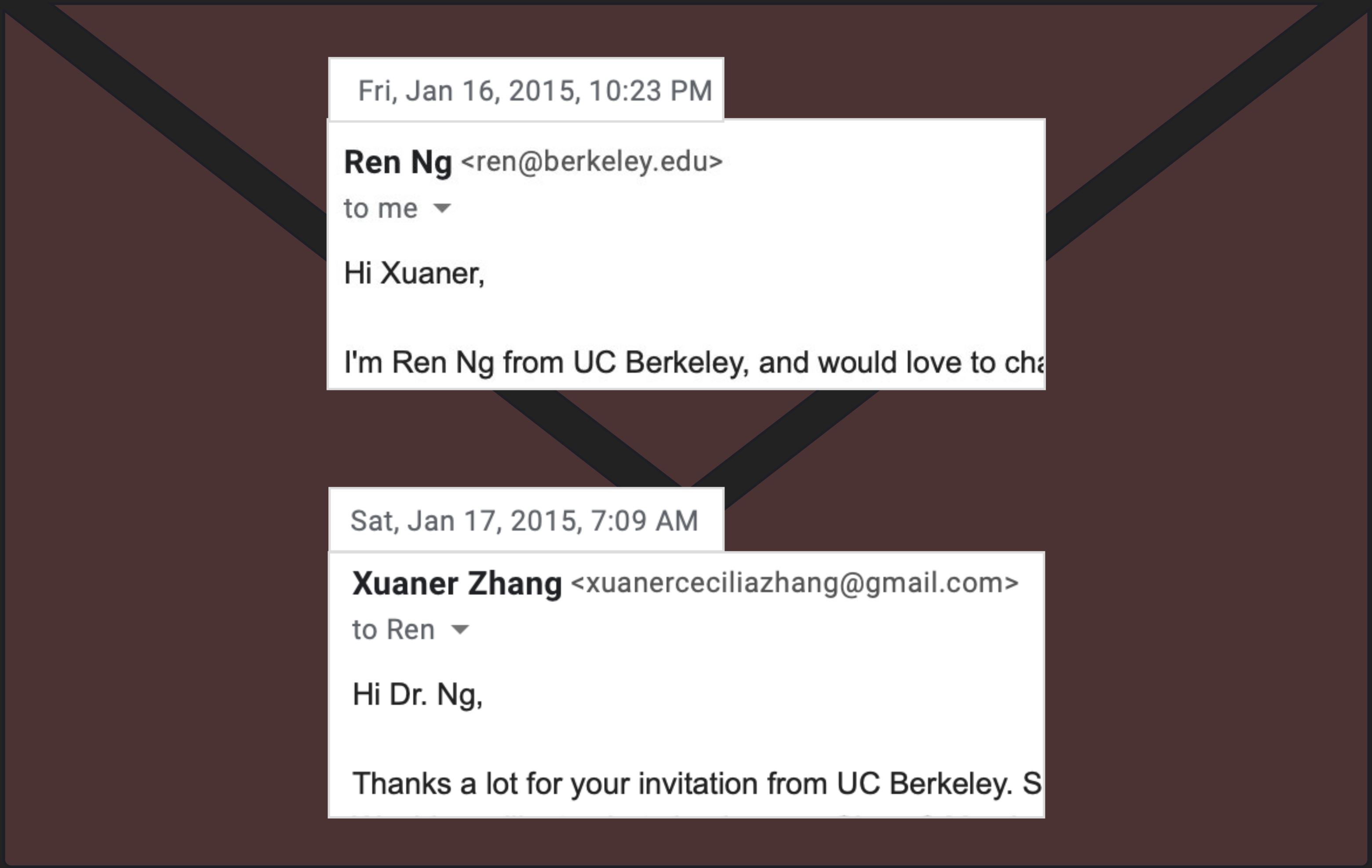
For Ren and Ren's "kids" (besides Reya and Lana)


AMD
RYZEN
THREADRIPPER

The Man Behind the Lens

Thanks Ren!

For bringing me here…


Fri, Jan 16, 2015, 10:23 PM

Ren Ng <ren@berkeley.edu>
to me

Hi Xuaner,

I'm Ren Ng from UC Berkeley, and would love to cha

Sat, Jan 17, 2015, 7:09 AM

Xuaner Zhang <xuanerceciliazhang@gmail.com>
to Ren

Hi Dr. Ng,

Thanks a lot for your invitation from UC Berkeley. S

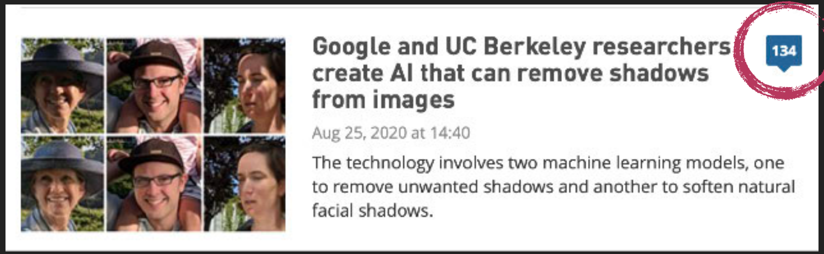
For Mom, Dad and Yoo

How Do You See Yourself in 5 Years?

DPREVIEW

PetaPixel

SLR Lounge www.slrlounge.com



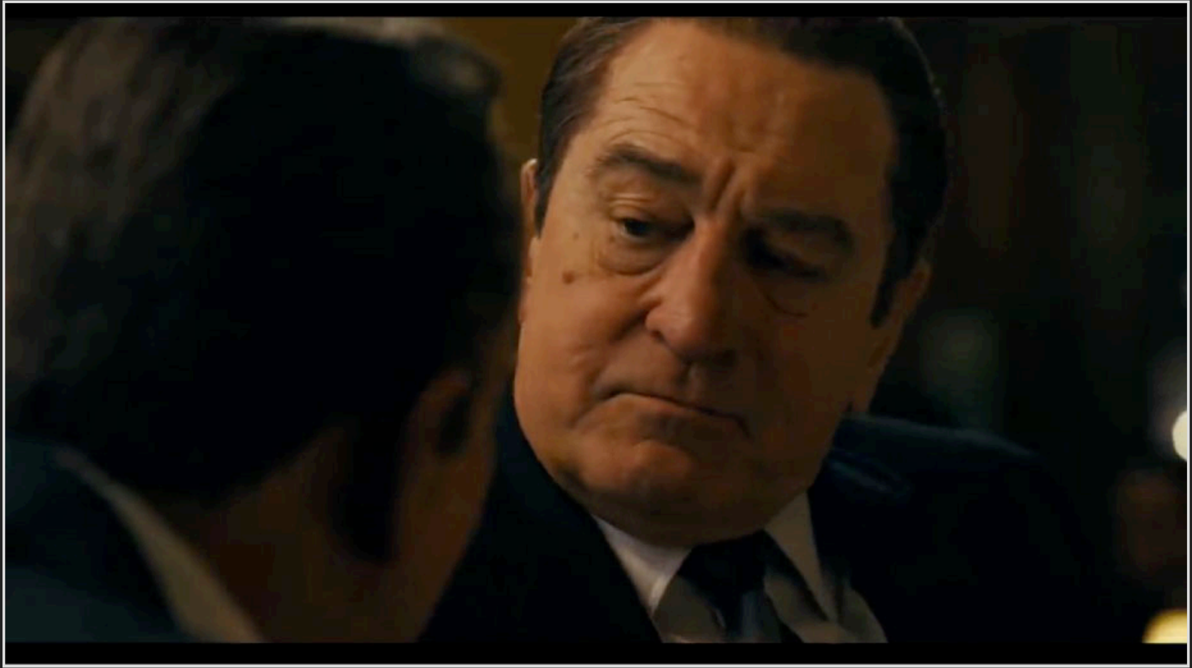
Google and UC Berkeley researchers create AI that can remove shadows from images

134

Aug 25, 2020 at 14:40

The technology involves two machine learning models, one to remove unwanted shadows and another to soften natural facial shadows.

&$!#%



"The making of The Irishman" (SIGGRAPH 2020 production session)

Thank you all!

It’s been an incredible journey!