



SIGGRAPH THINK
2020 [S2020.SIGGRAPH.ORG](https://s2020.siggraph.org) BEYOND

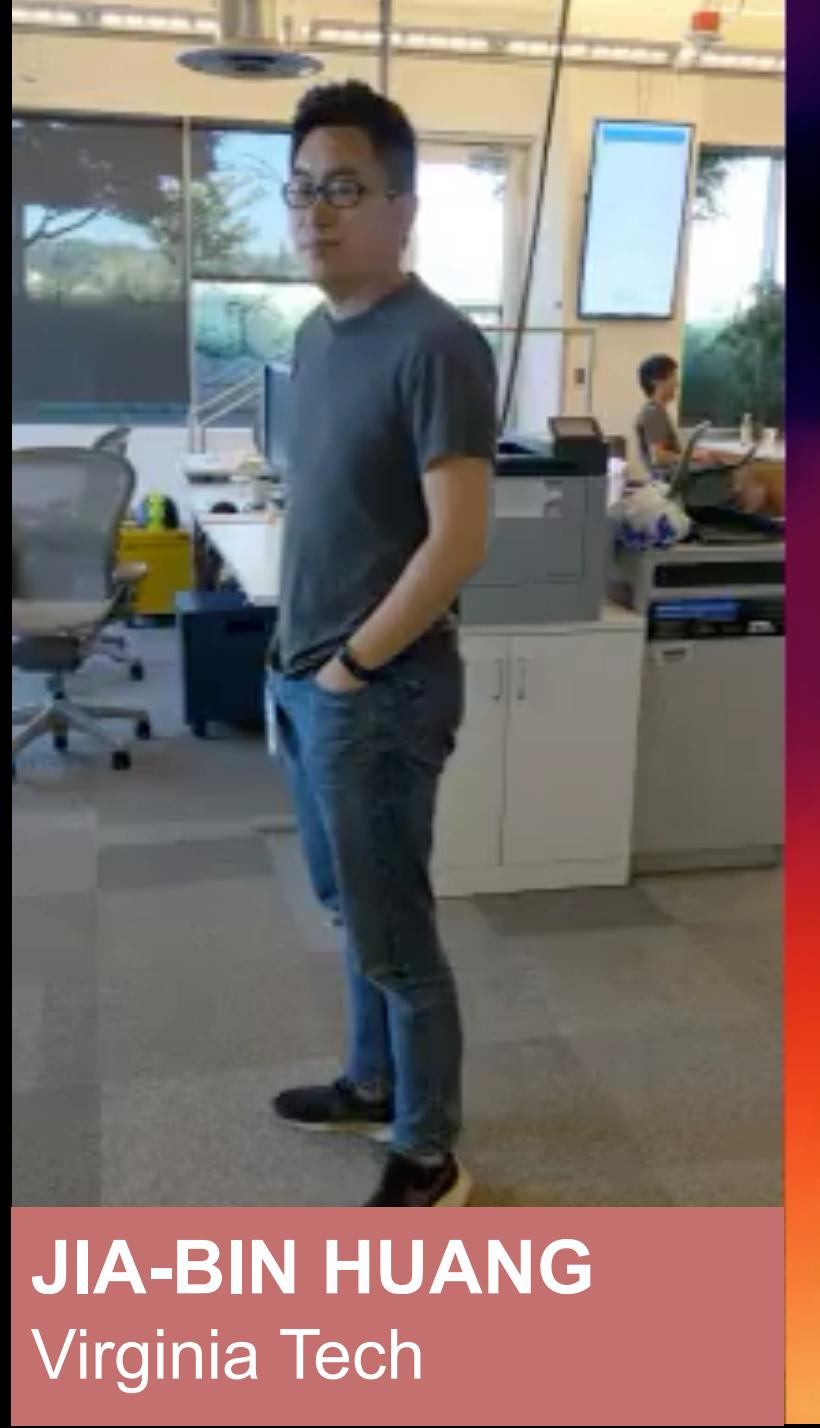
CONSISTENT VIDEO DEPTH ESTIMATION

Xuan Luo¹, Jia-Bin Huang², Richard Szeliski³,
Kevin Matzen³, Johannes Kopf³

¹*University of Washington*, ²*Virginia Tech*, ³*Facebook*



XUAN LUO
University of Washington



JIA-BIN HUANG
Virginia Tech



RICHARD SZELISKI
Facebook



KEVIN MATZEN
Facebook

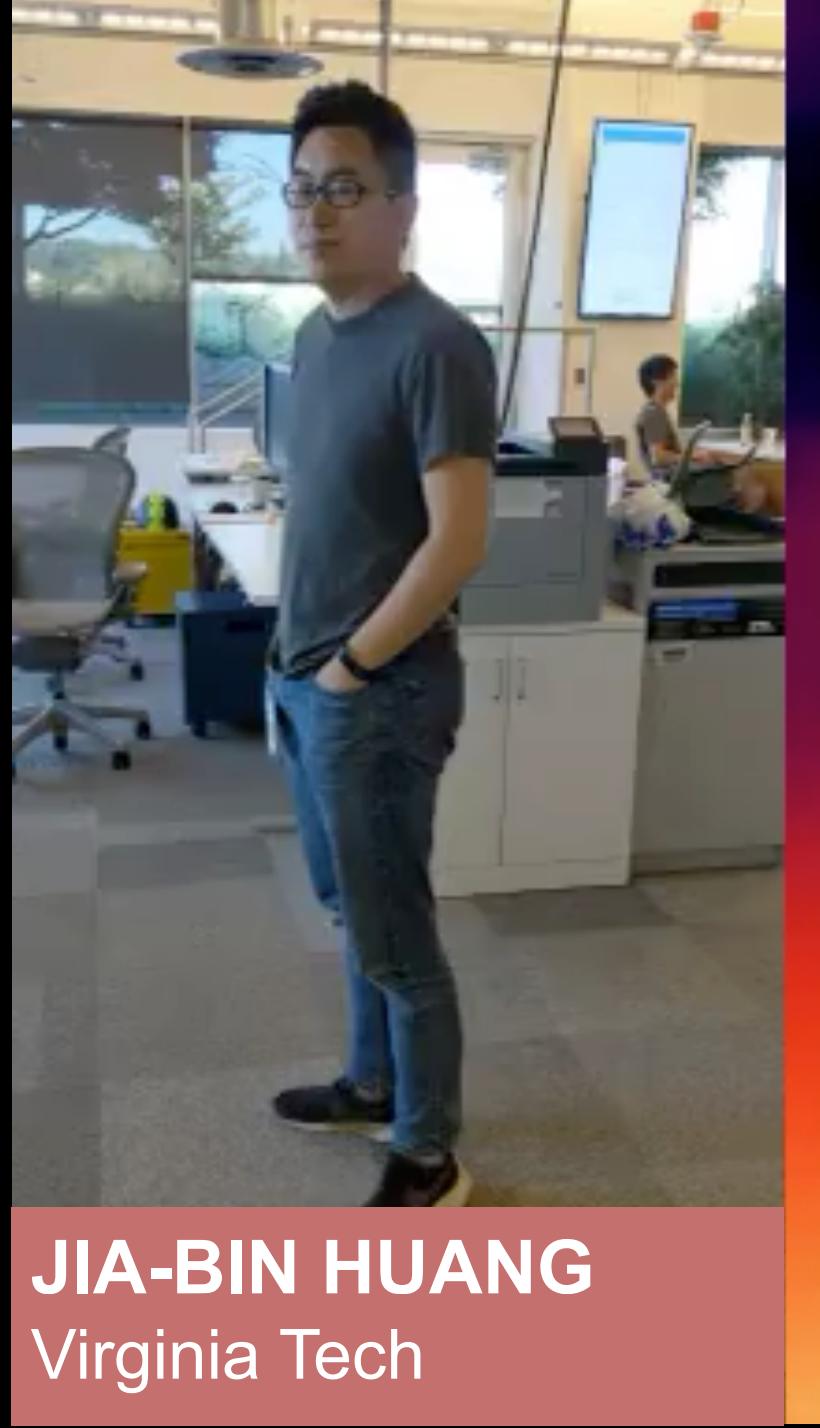


JOHANNES KOPF
Facebook





XUAN LUO
University of Washington



JIA-BIN HUANG
Virginia Tech



RICHARD SZELISKI
Facebook



KEVIN MATZEN
Facebook



JOHANNES KOPF
Facebook

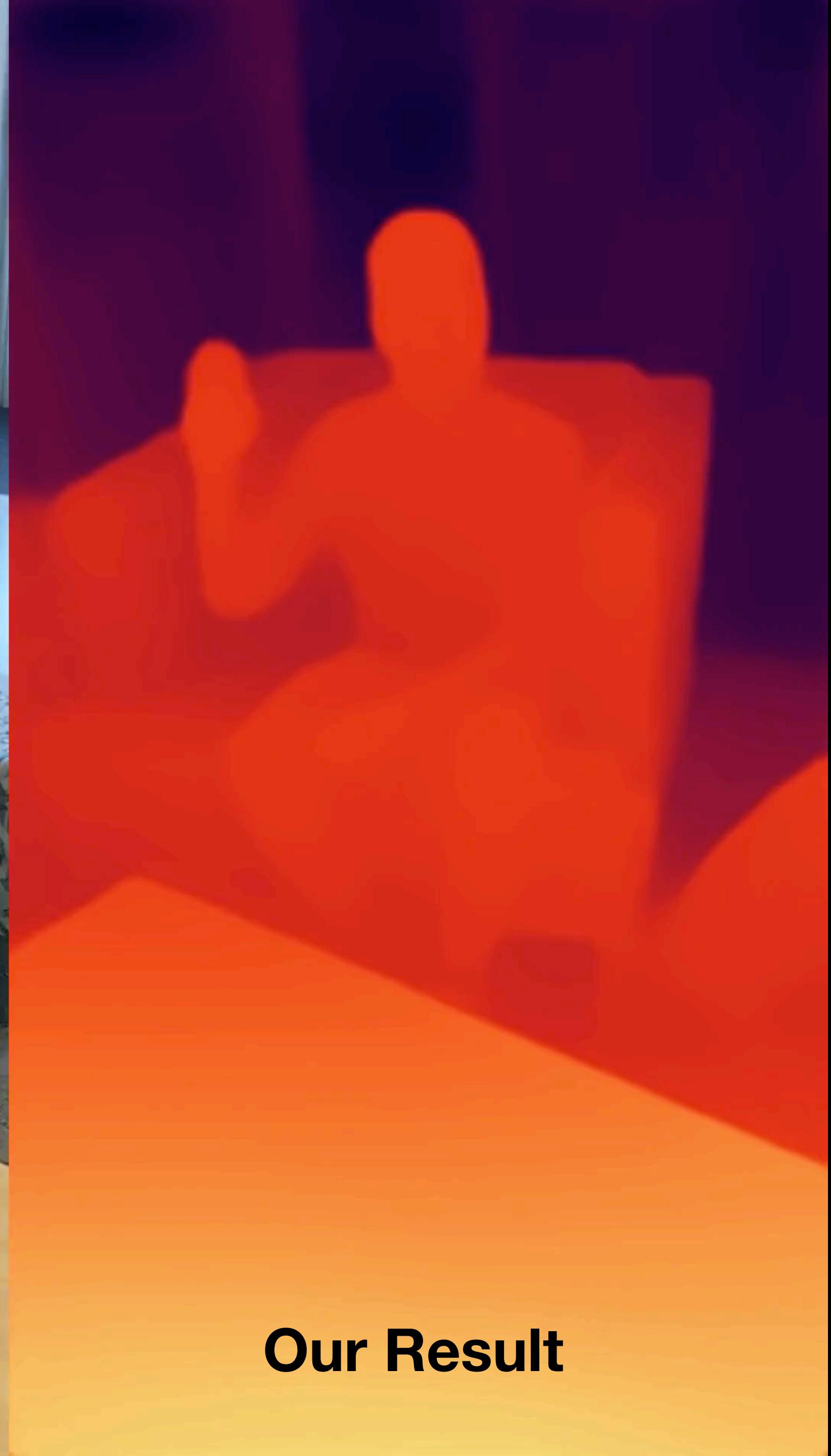








Input



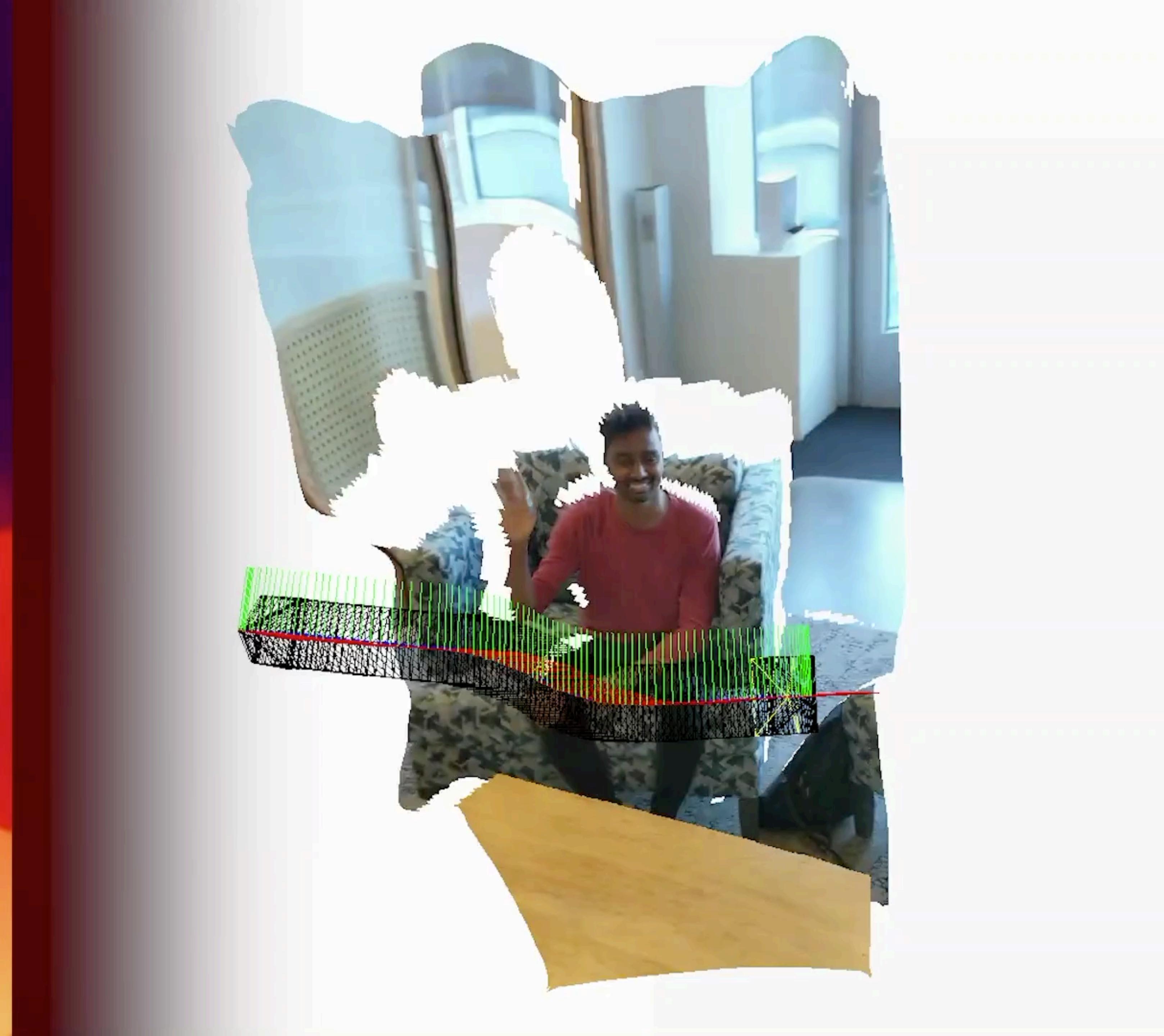
Our Result

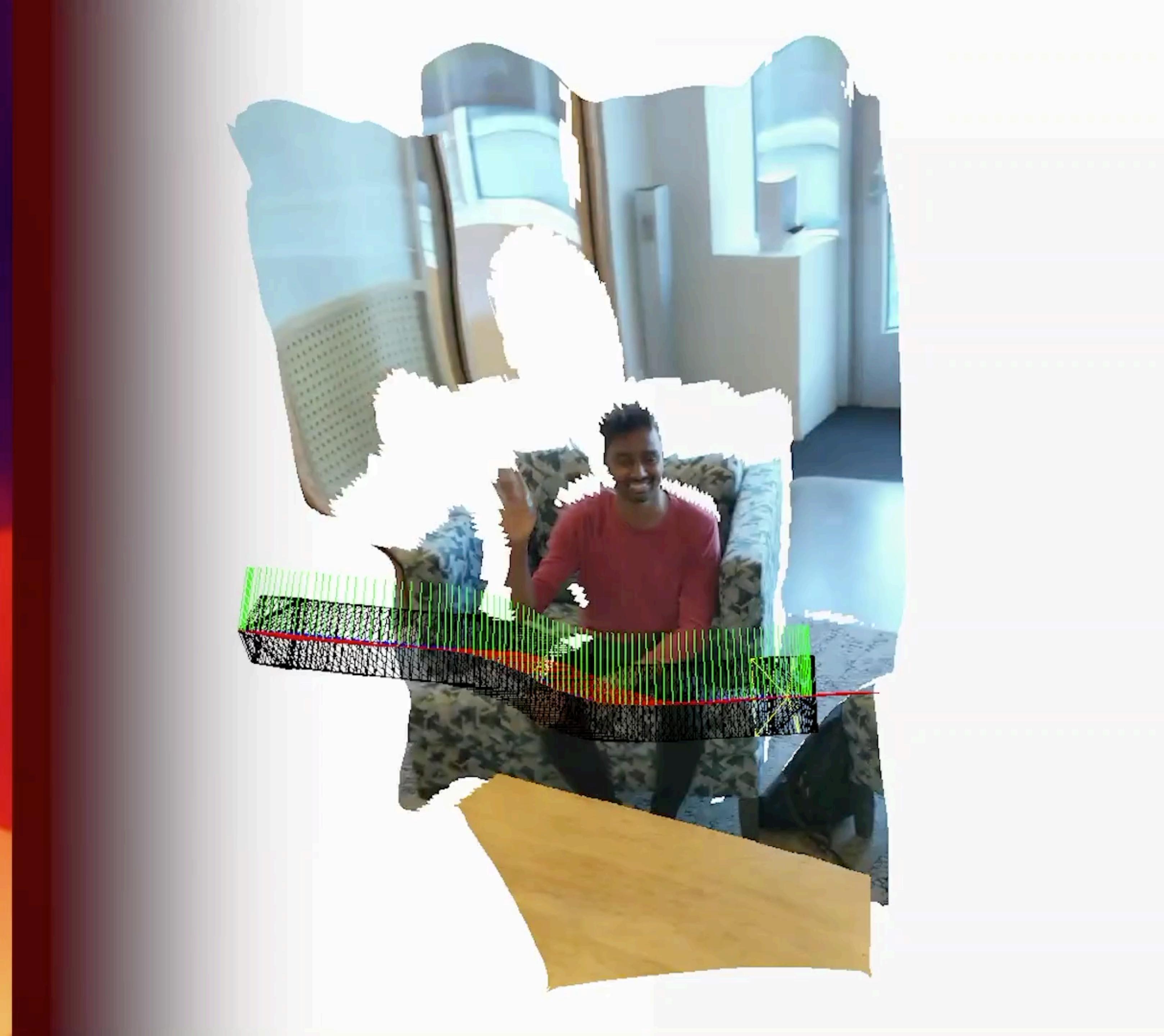


Input



Our Result

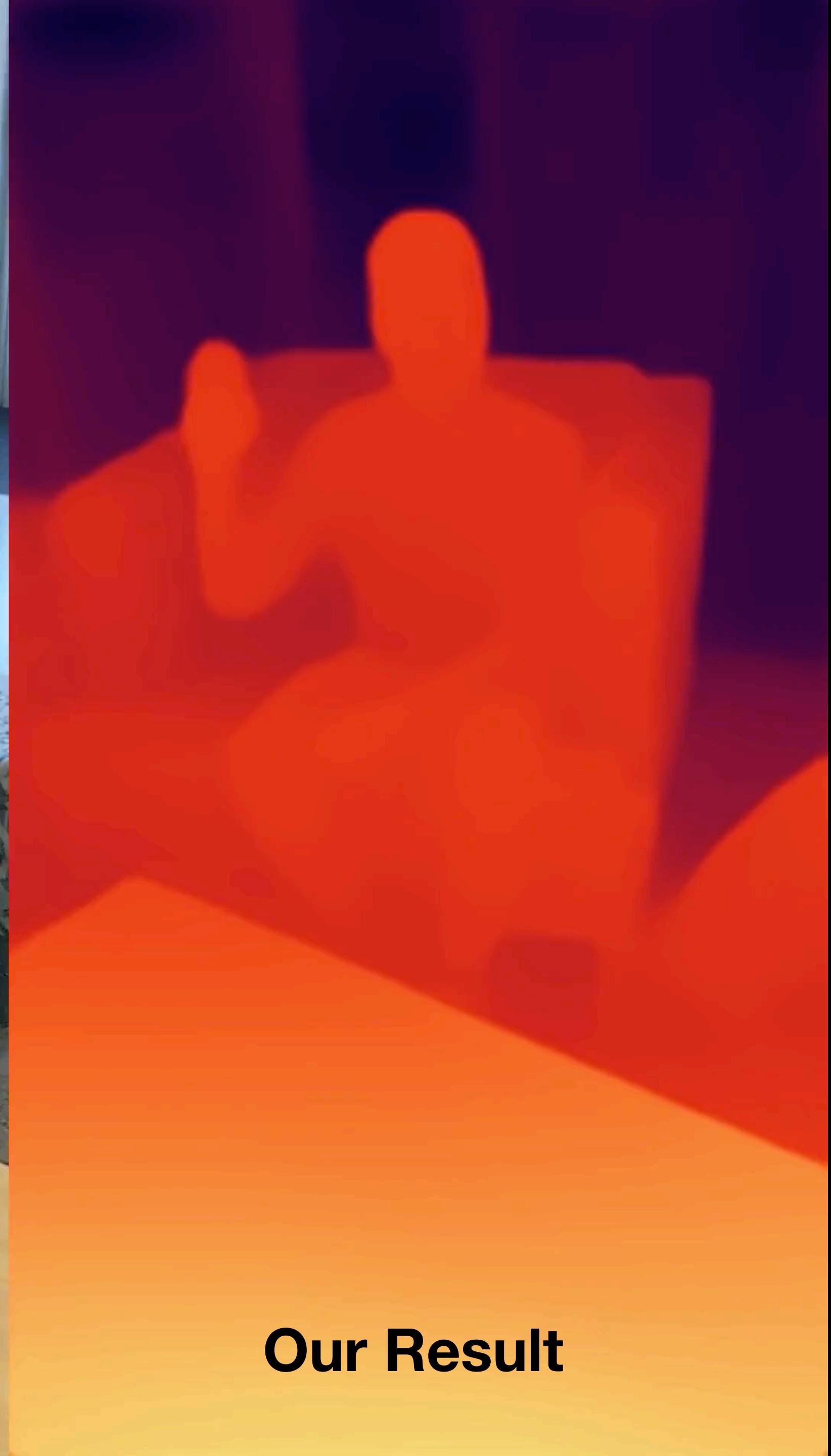








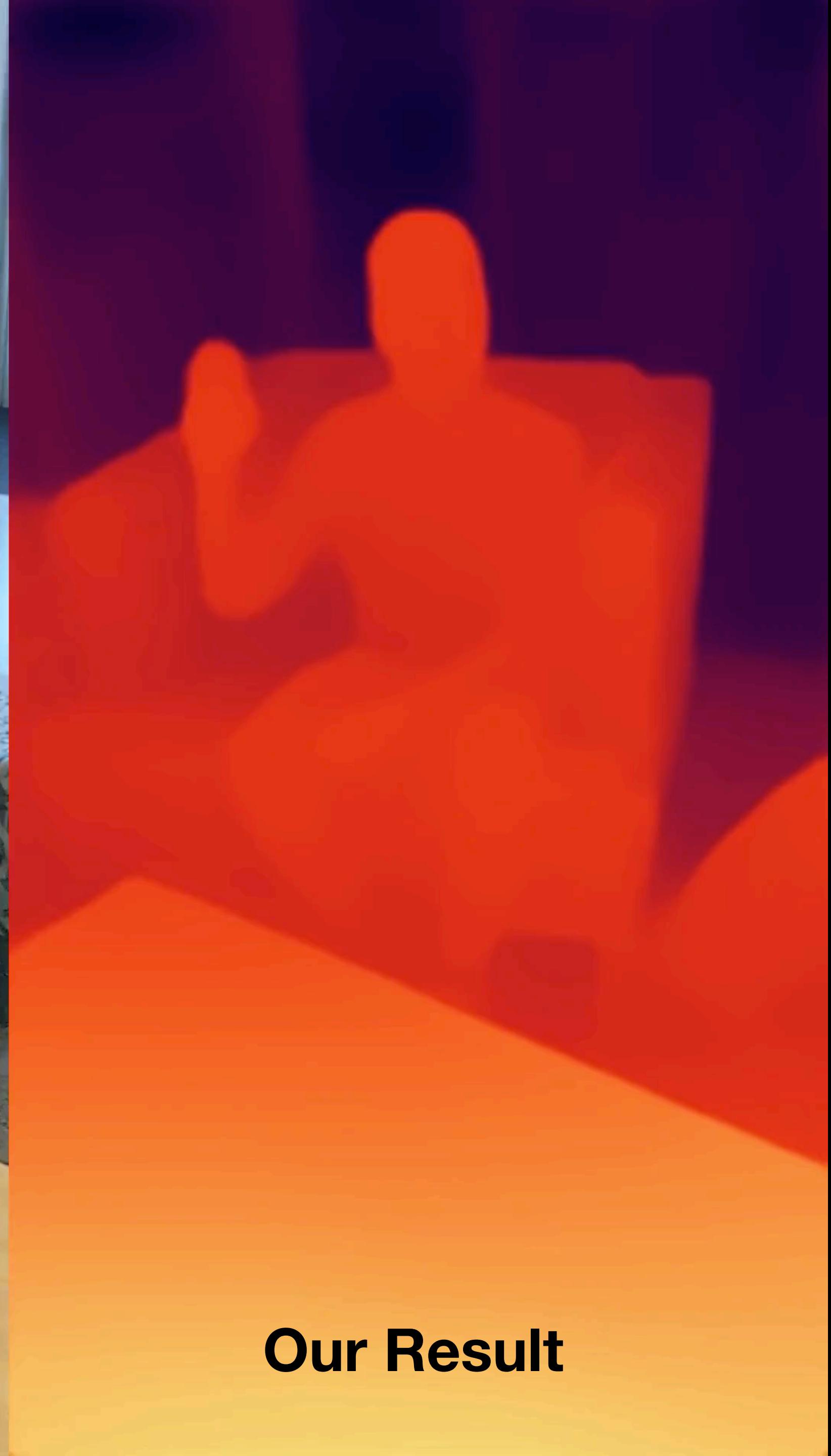
Input



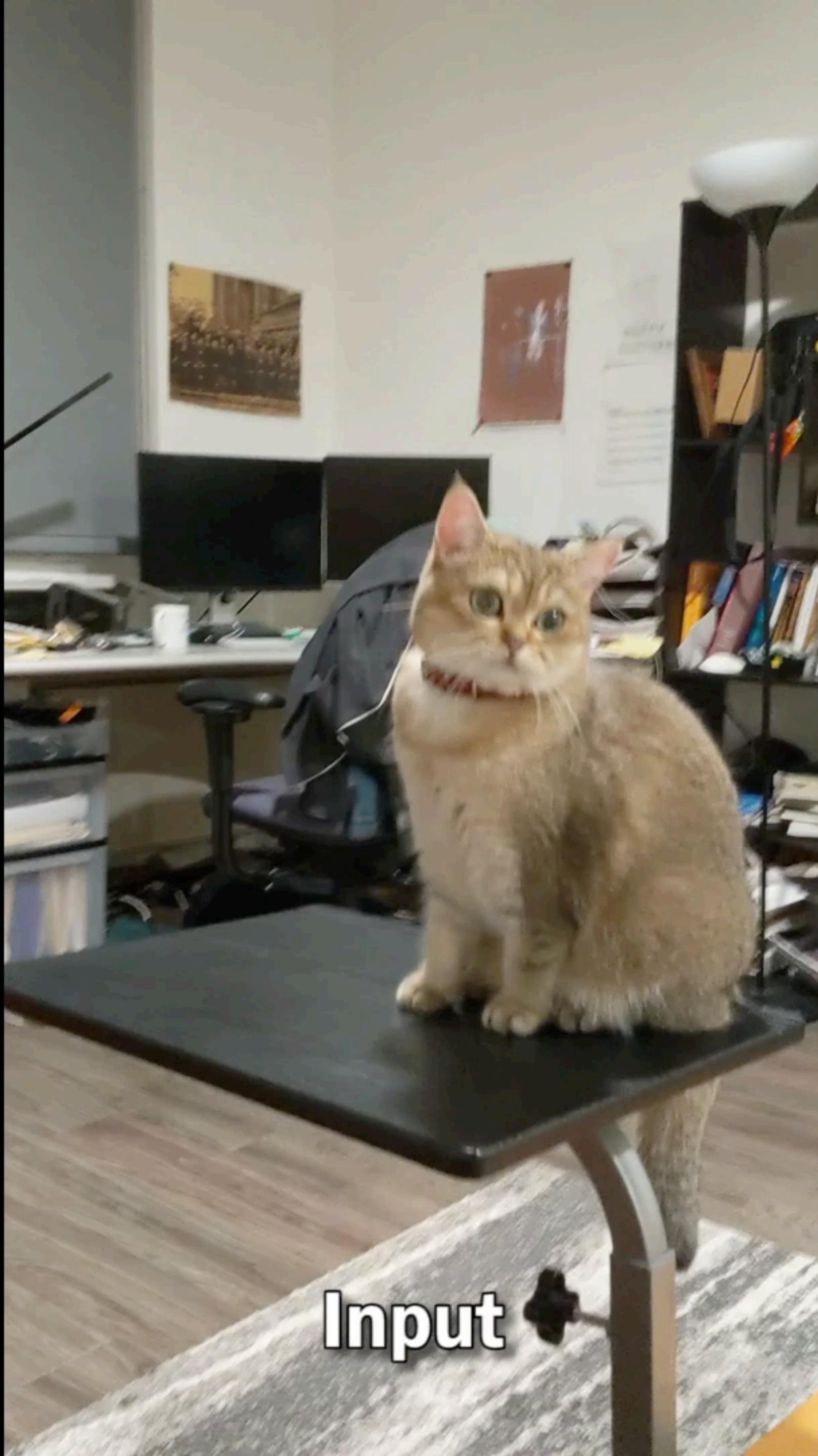
Our Result



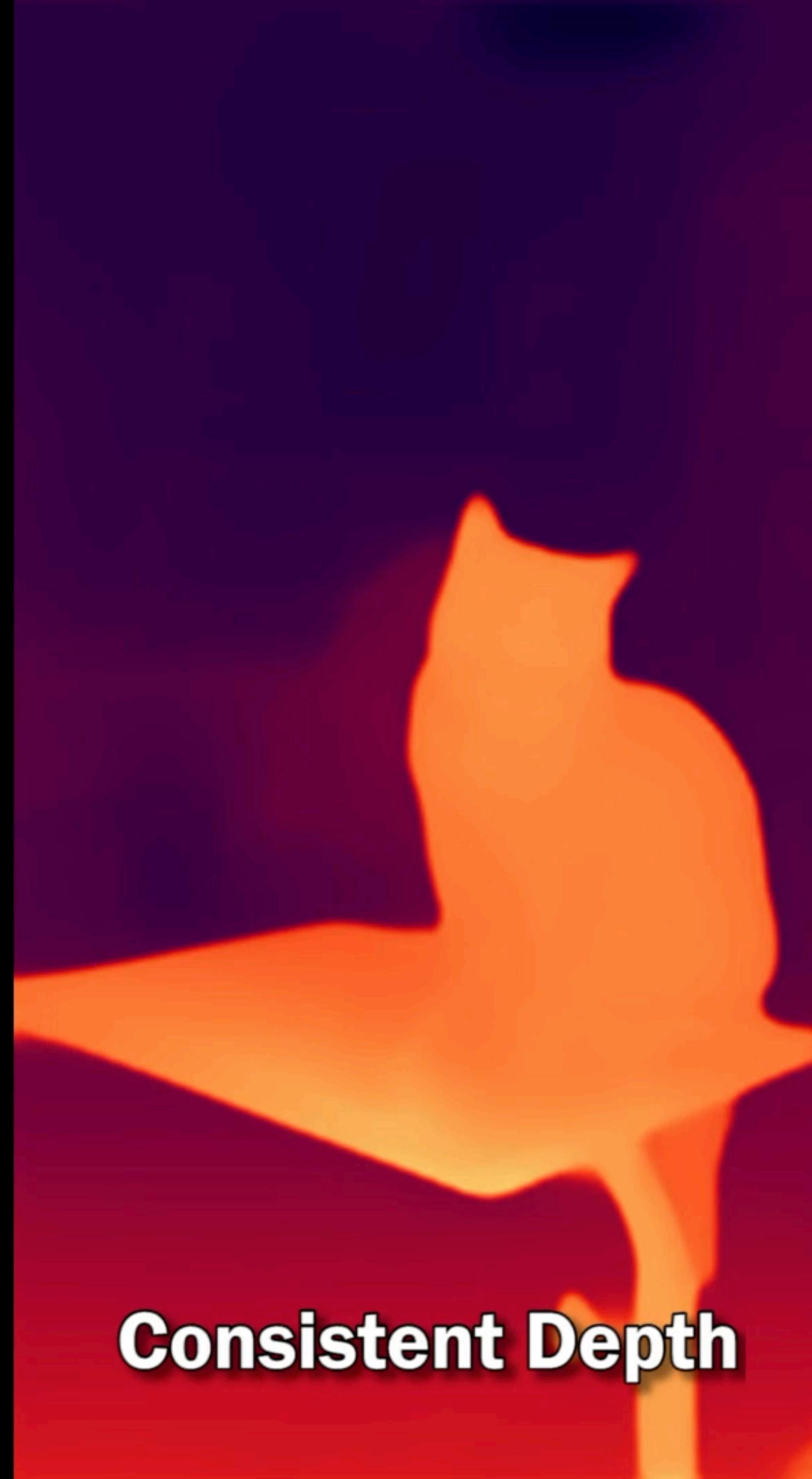
Input



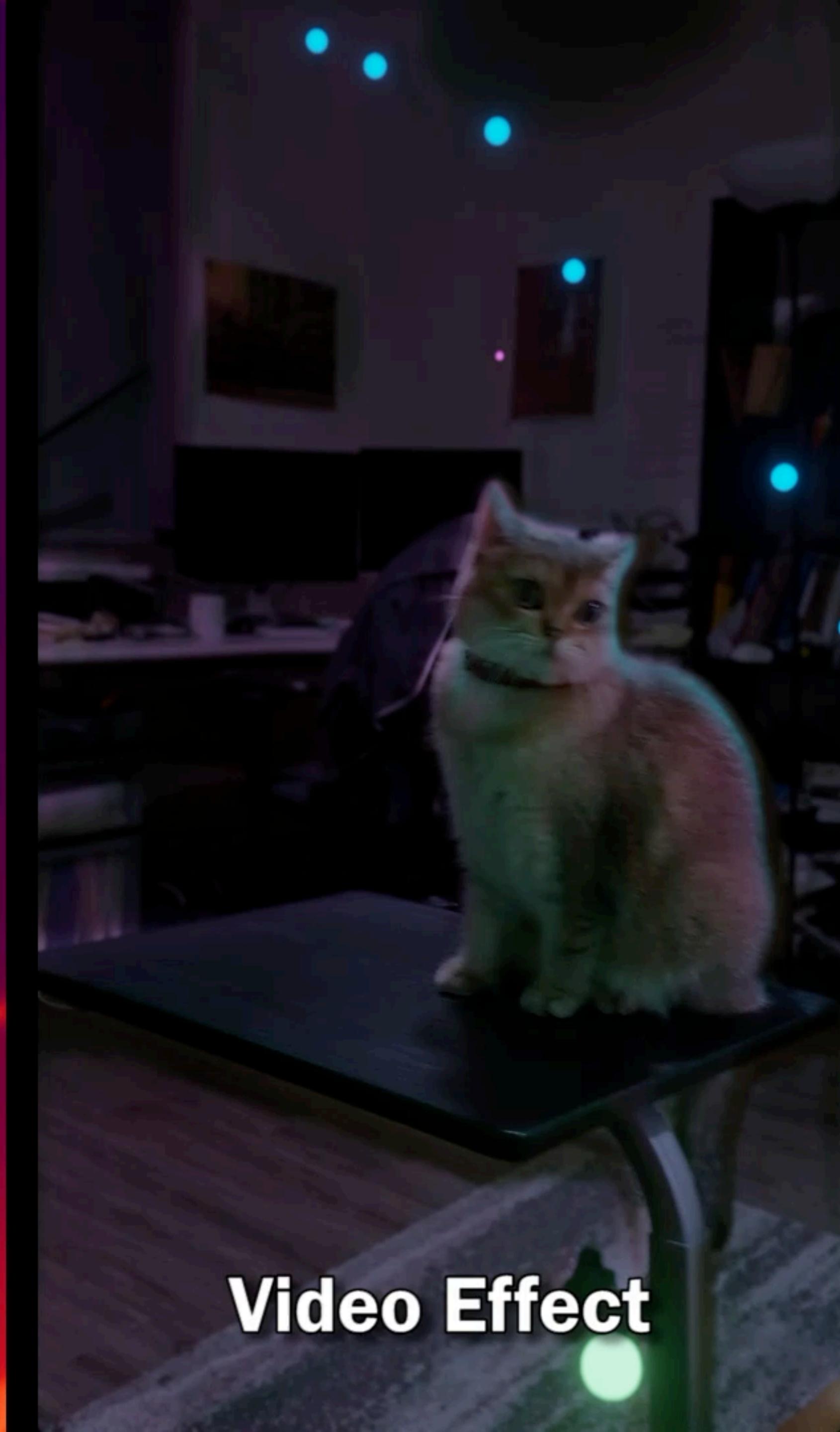
Our Result



Input



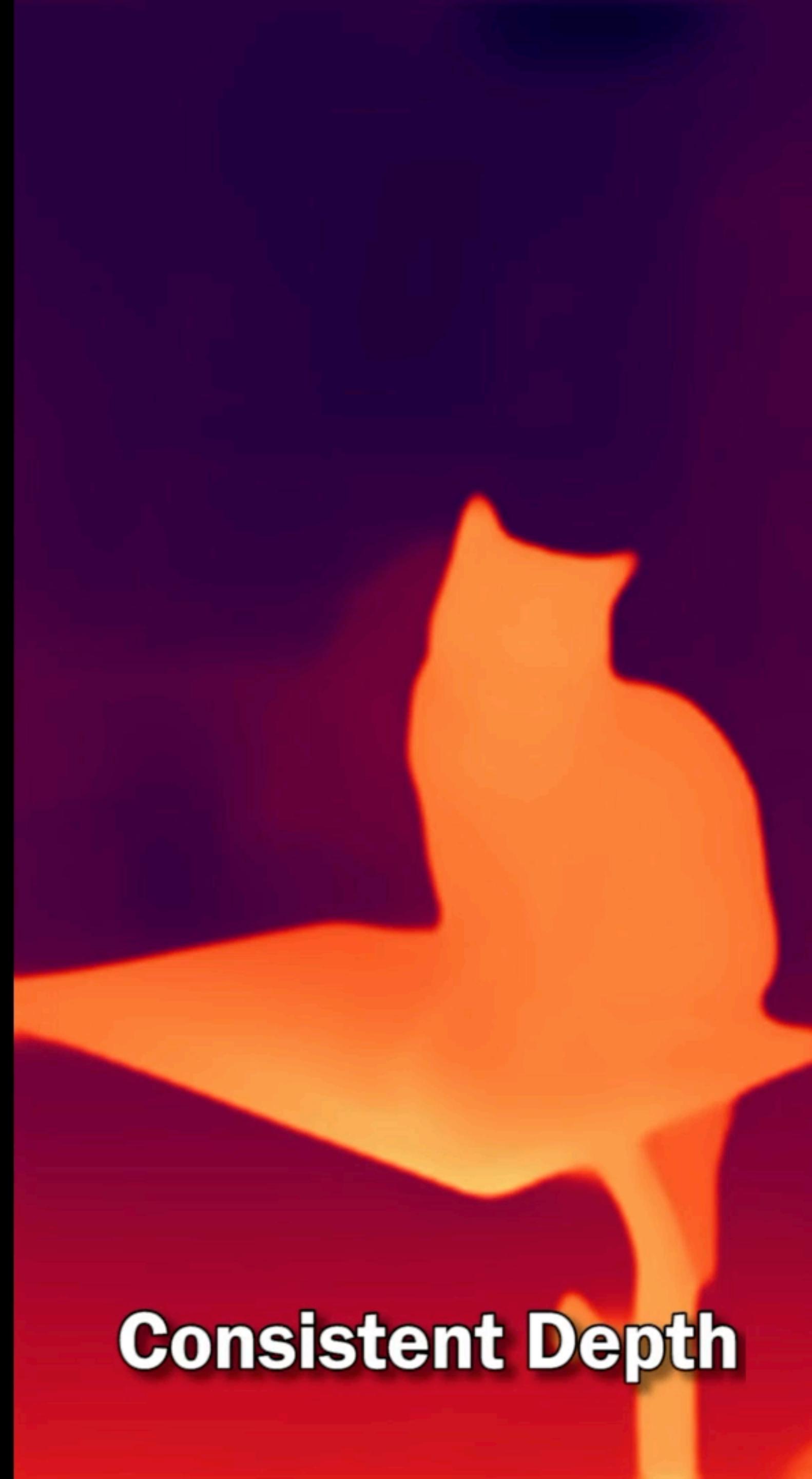
Consistent Depth



Video Effect



Input



Consistent Depth



Video Effect





Structure from Motion

Dense

N



COLMAP
Schonberger et al. 2016

Structure from Motion

Dense

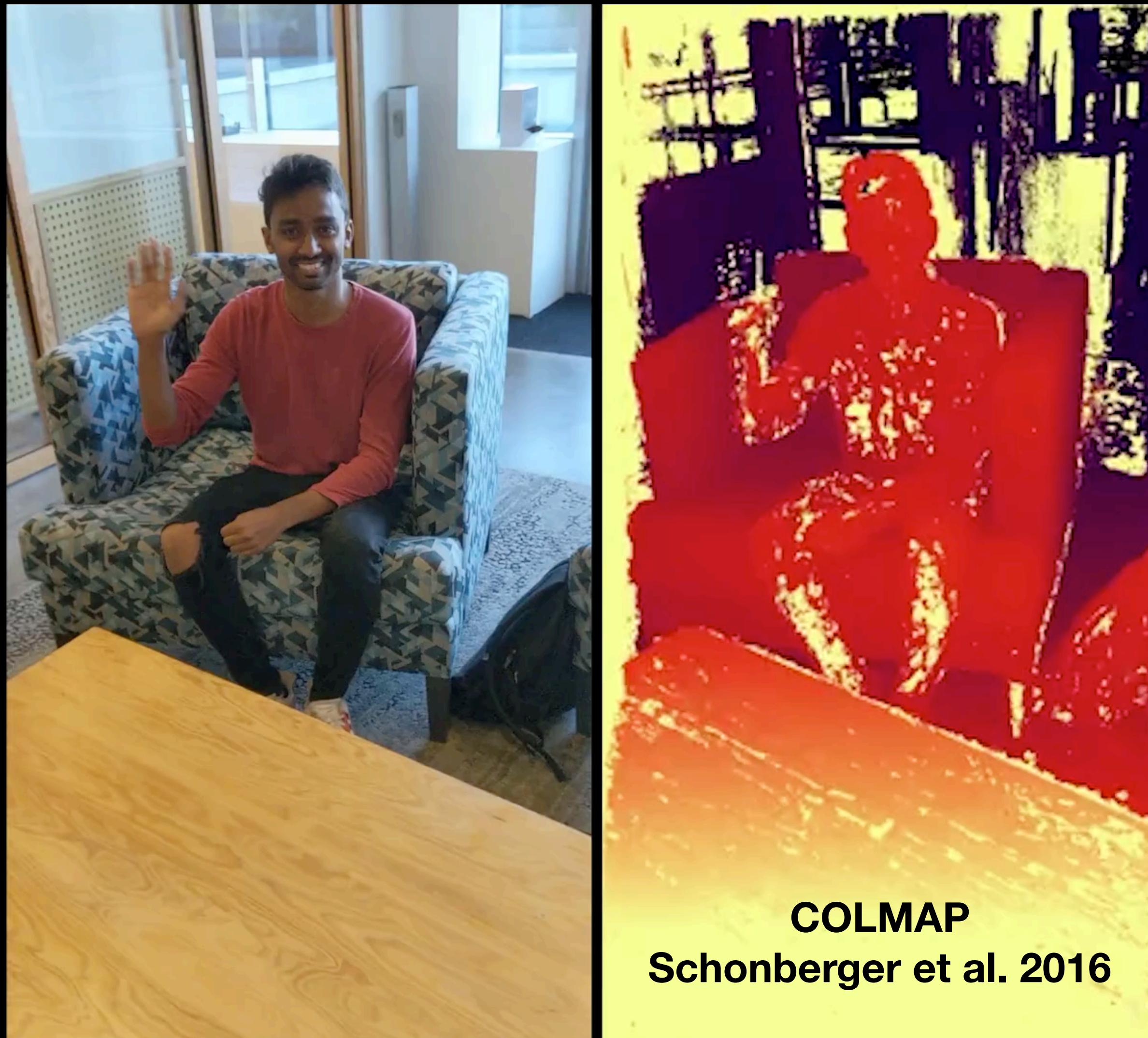
N



COLMAP
Schonberger et al. 2016

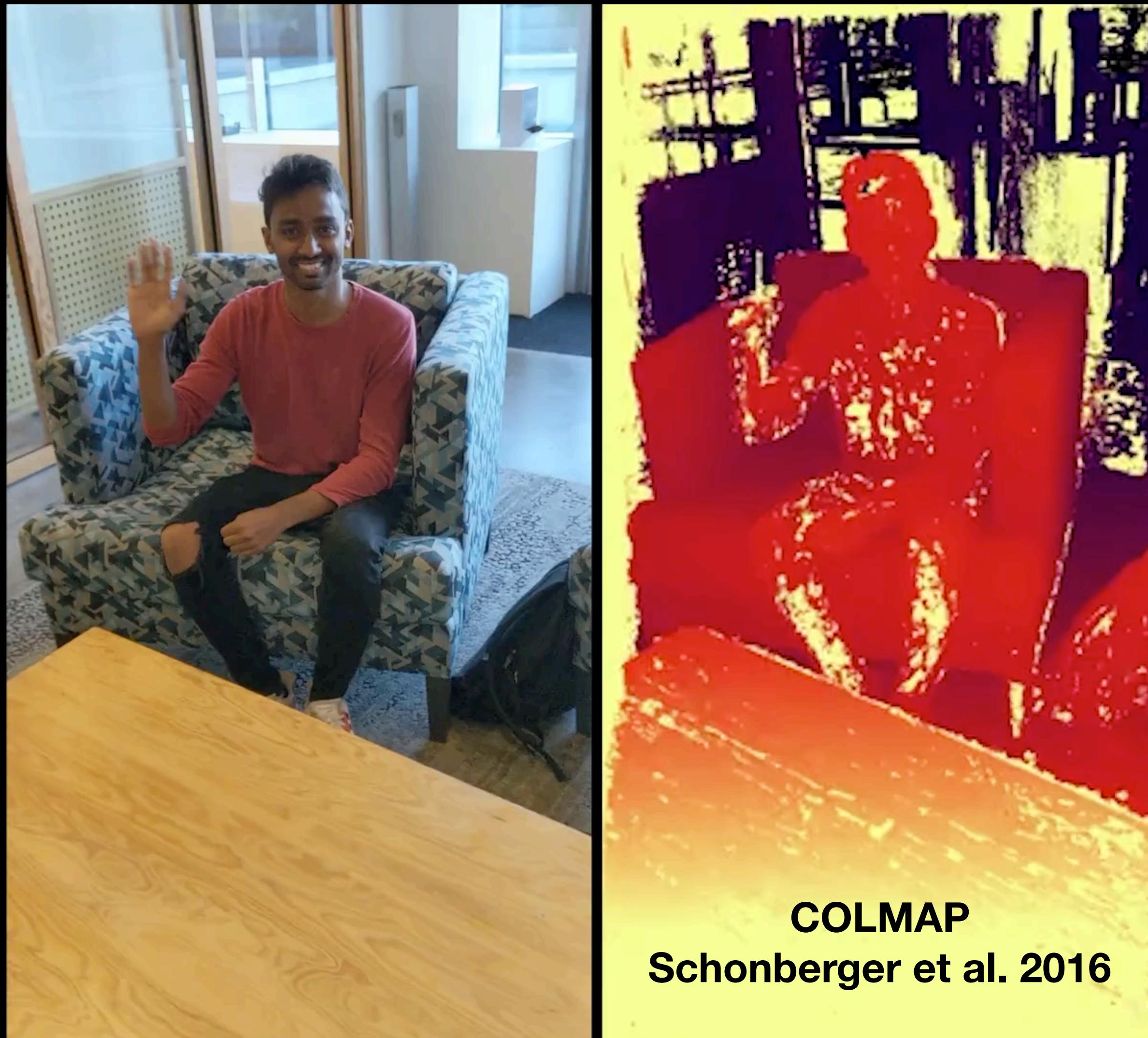
Multi-View Stereo (MVS)

Dense	Semi
Dynamics	N



Multi-View Stereo (MVS)

Dense	Semi
Dynamics	N



Single-frame Depth Estimation

Dense	Y
Dynamics	Y
Temporal Coherence	N
Geometric Consistency	N



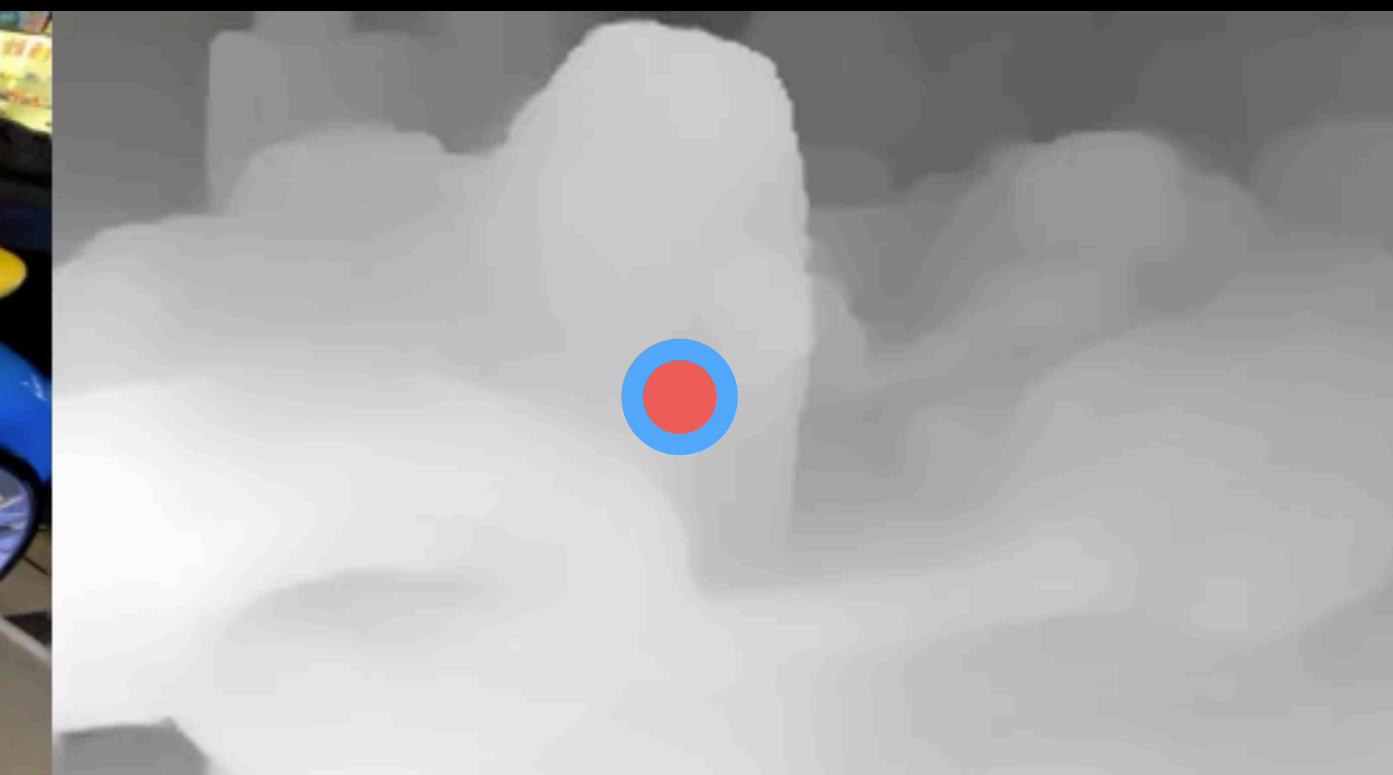
Single-frame Depth Estimation

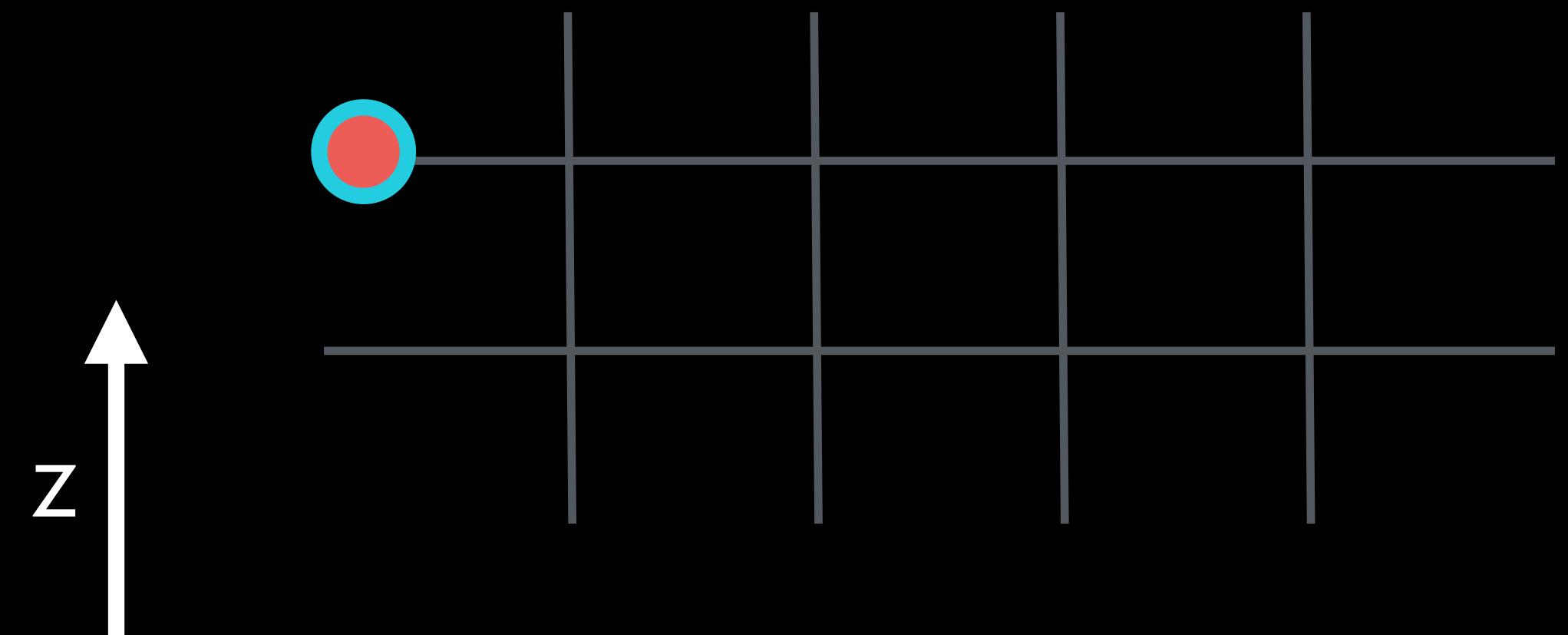
Dense	Y
Dynamics	Y
Temporal Coherence	N
Geometric Consistency	N

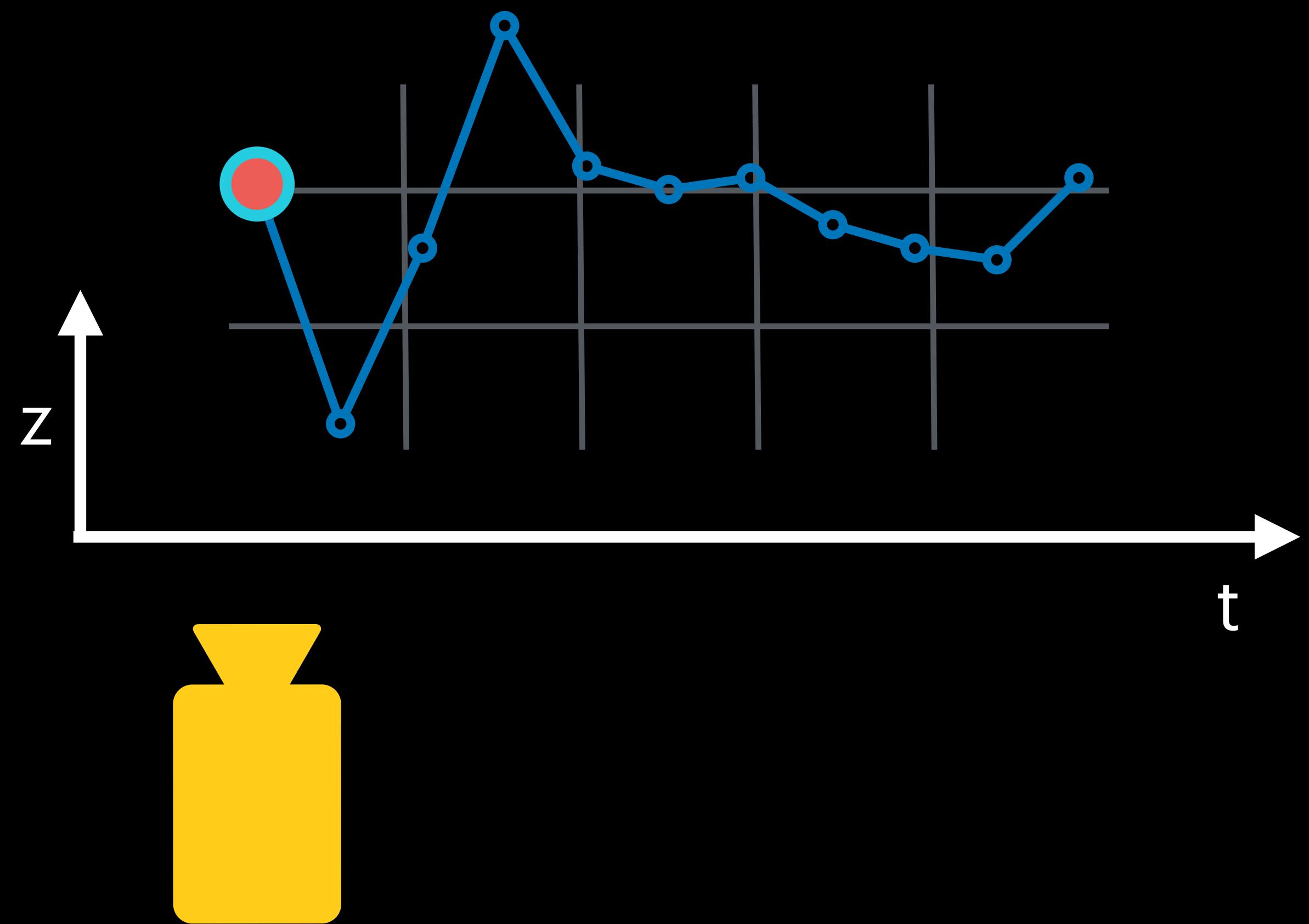


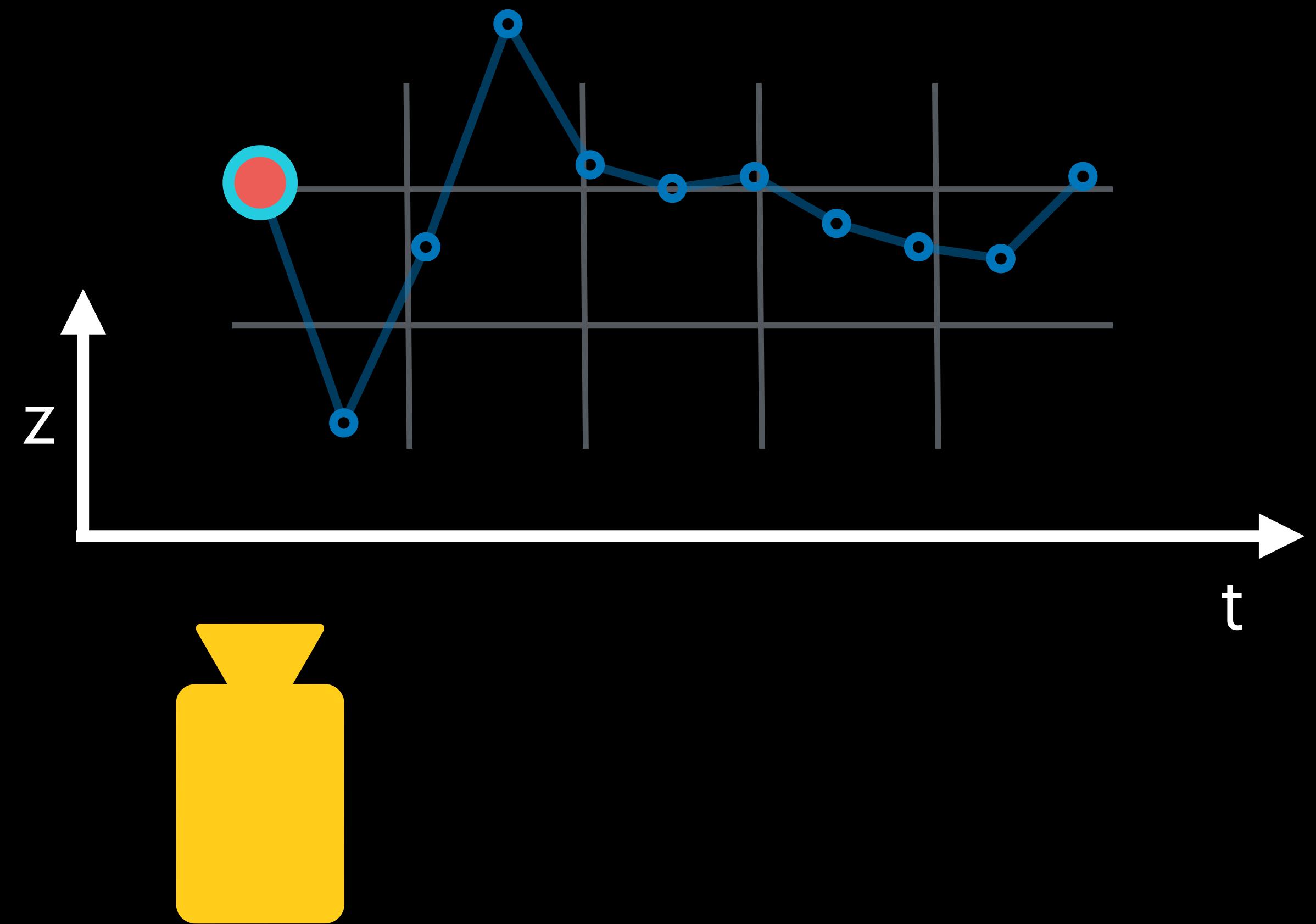


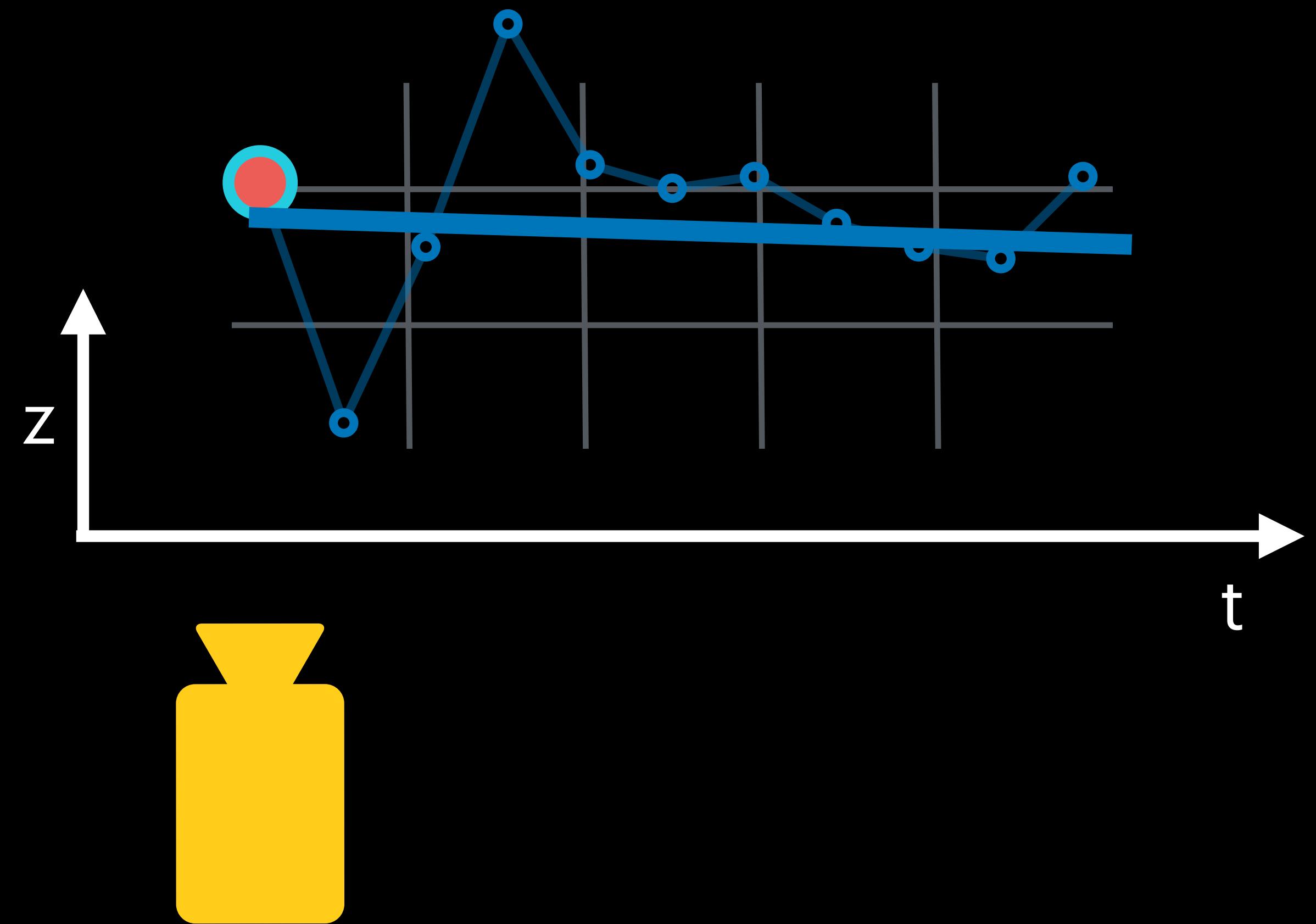


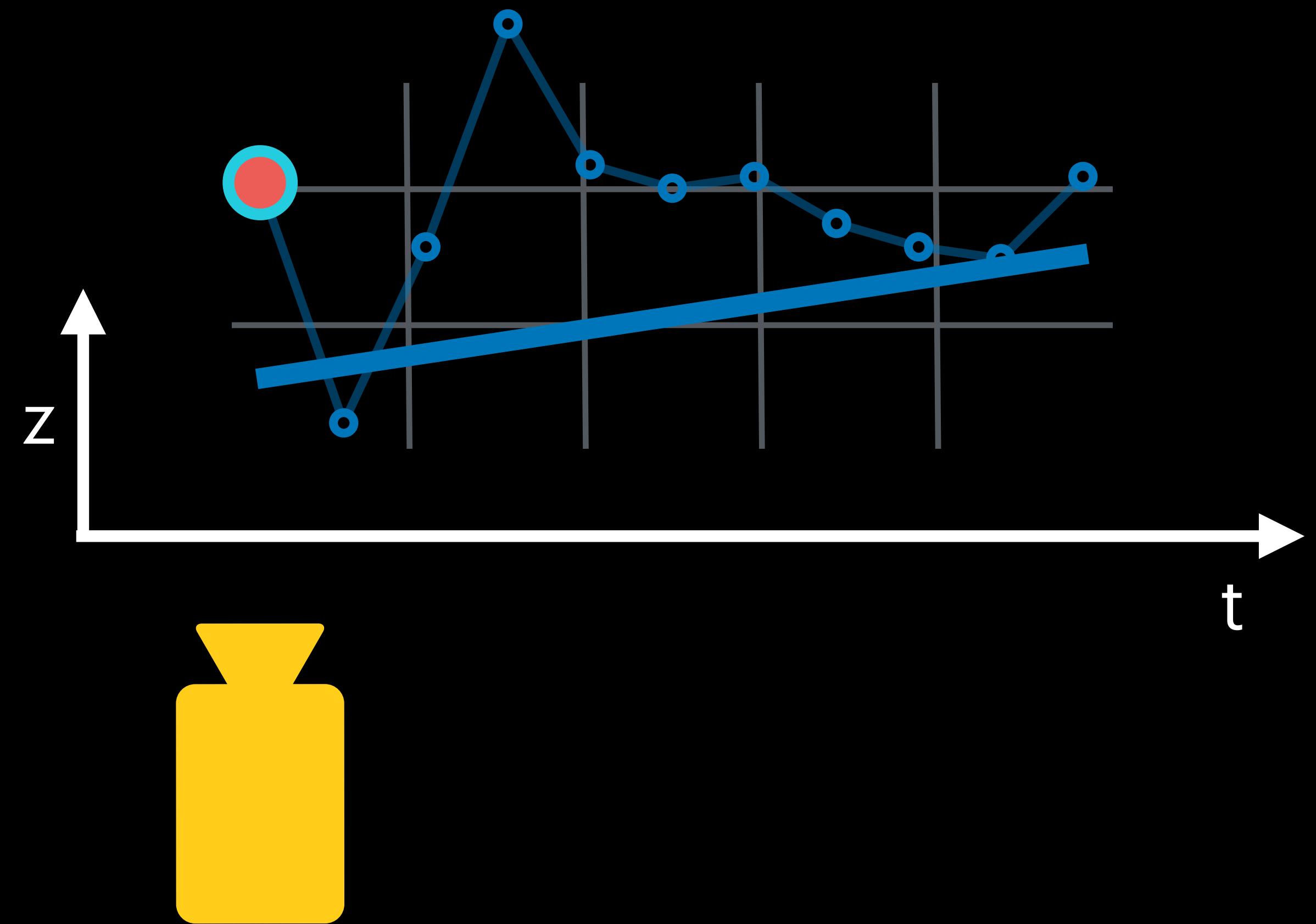


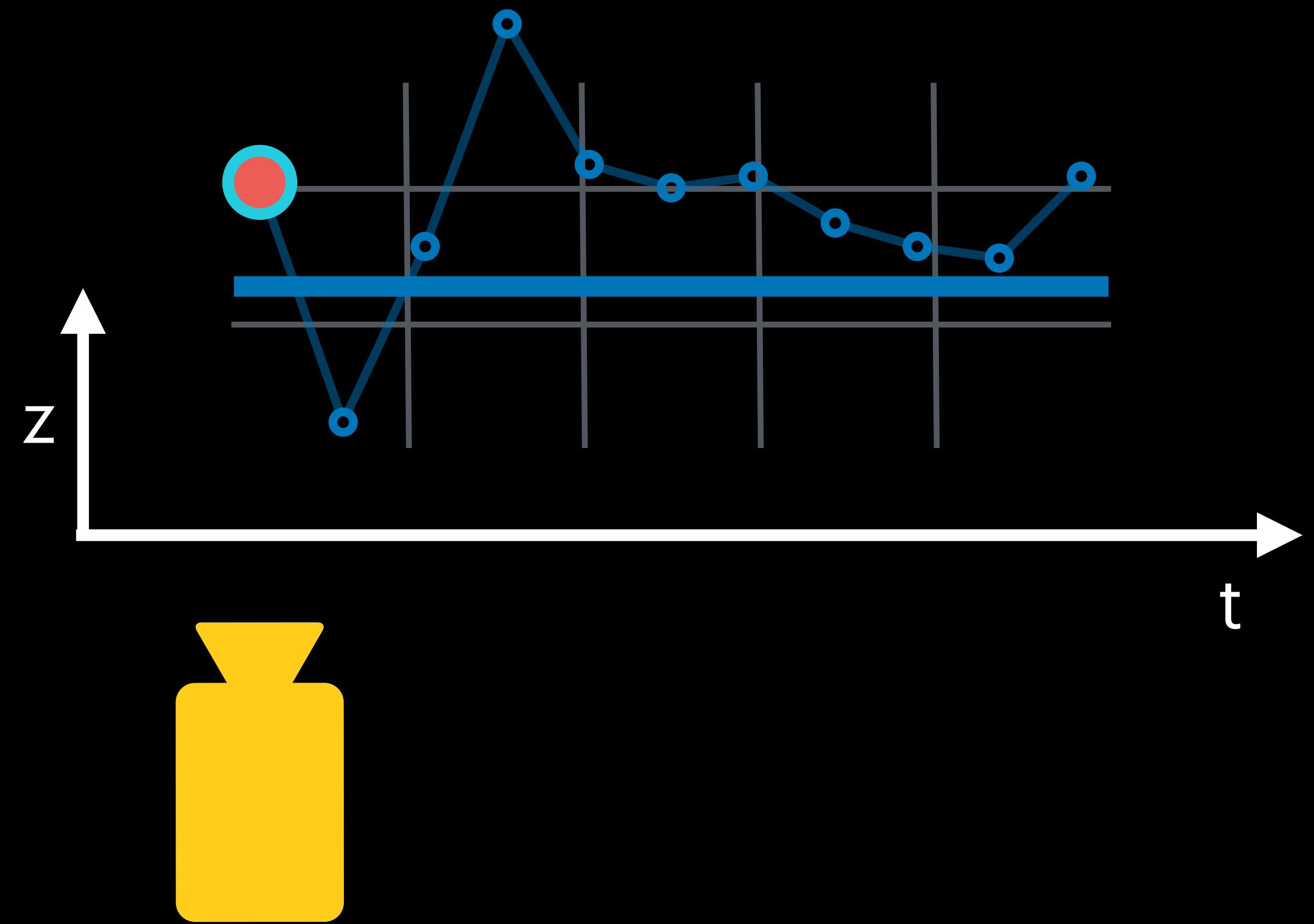






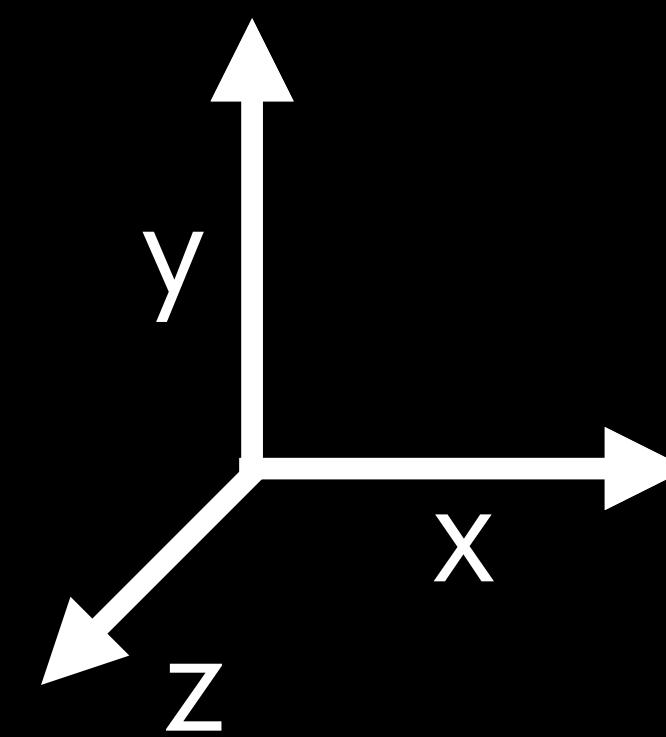
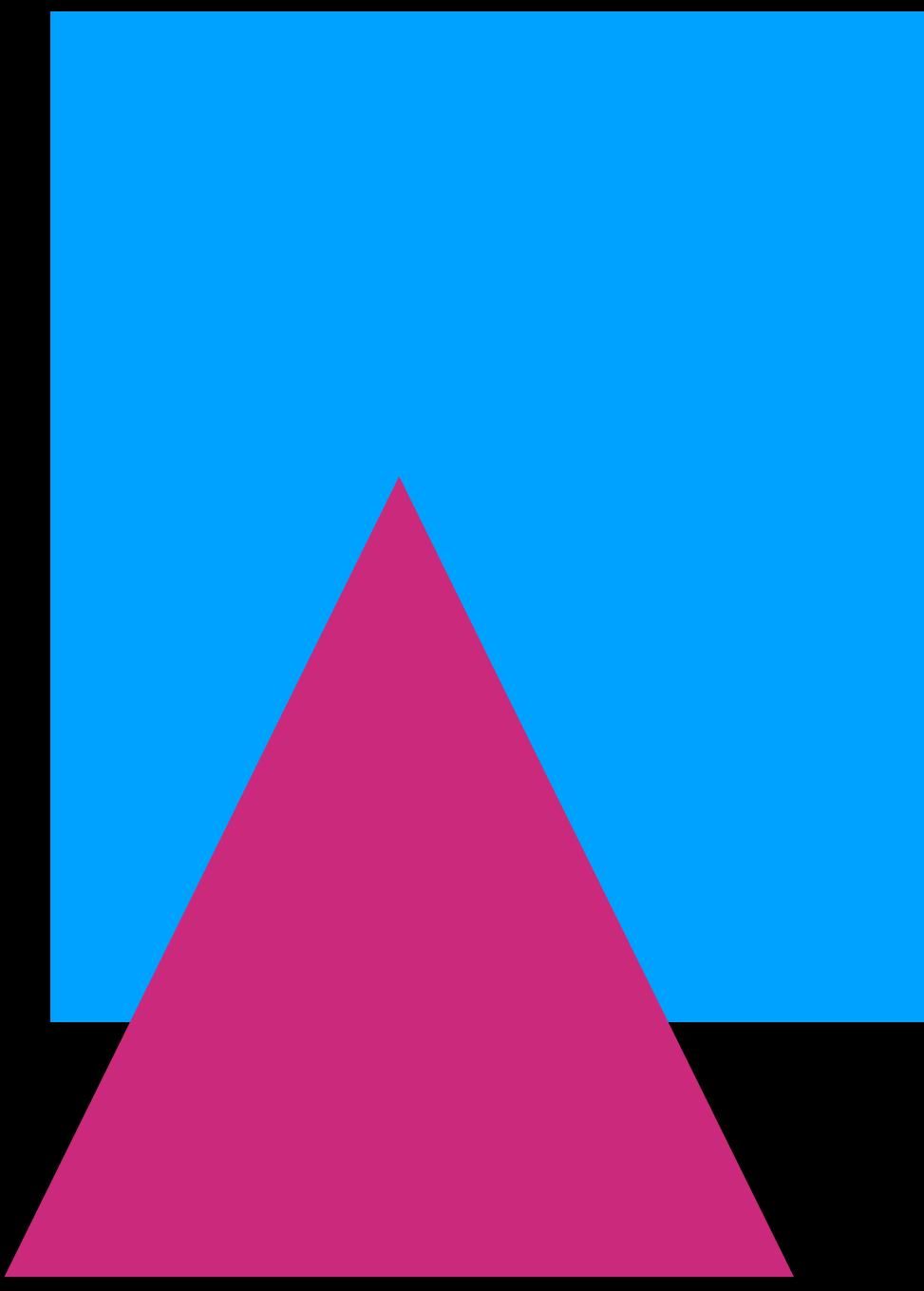


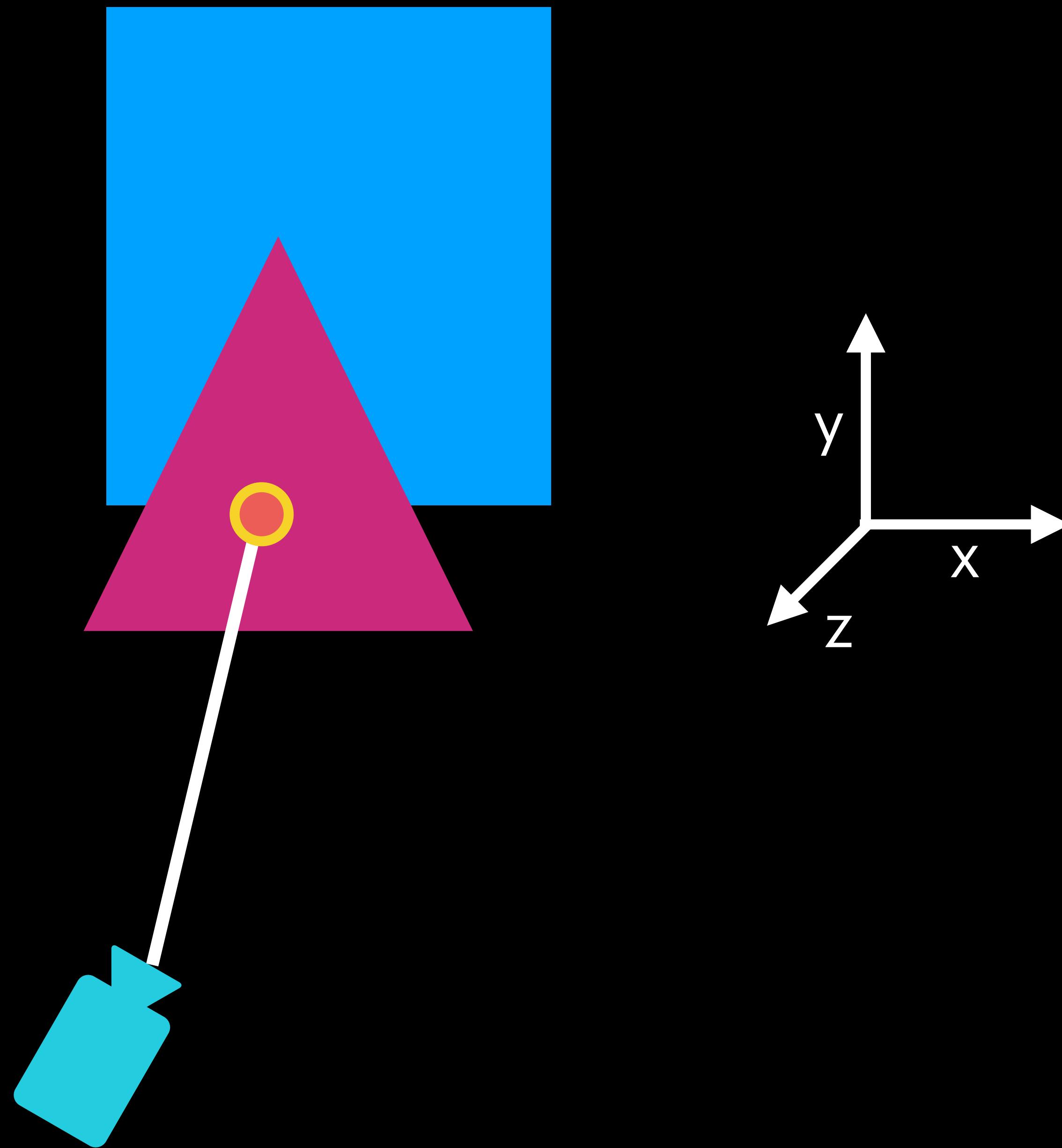


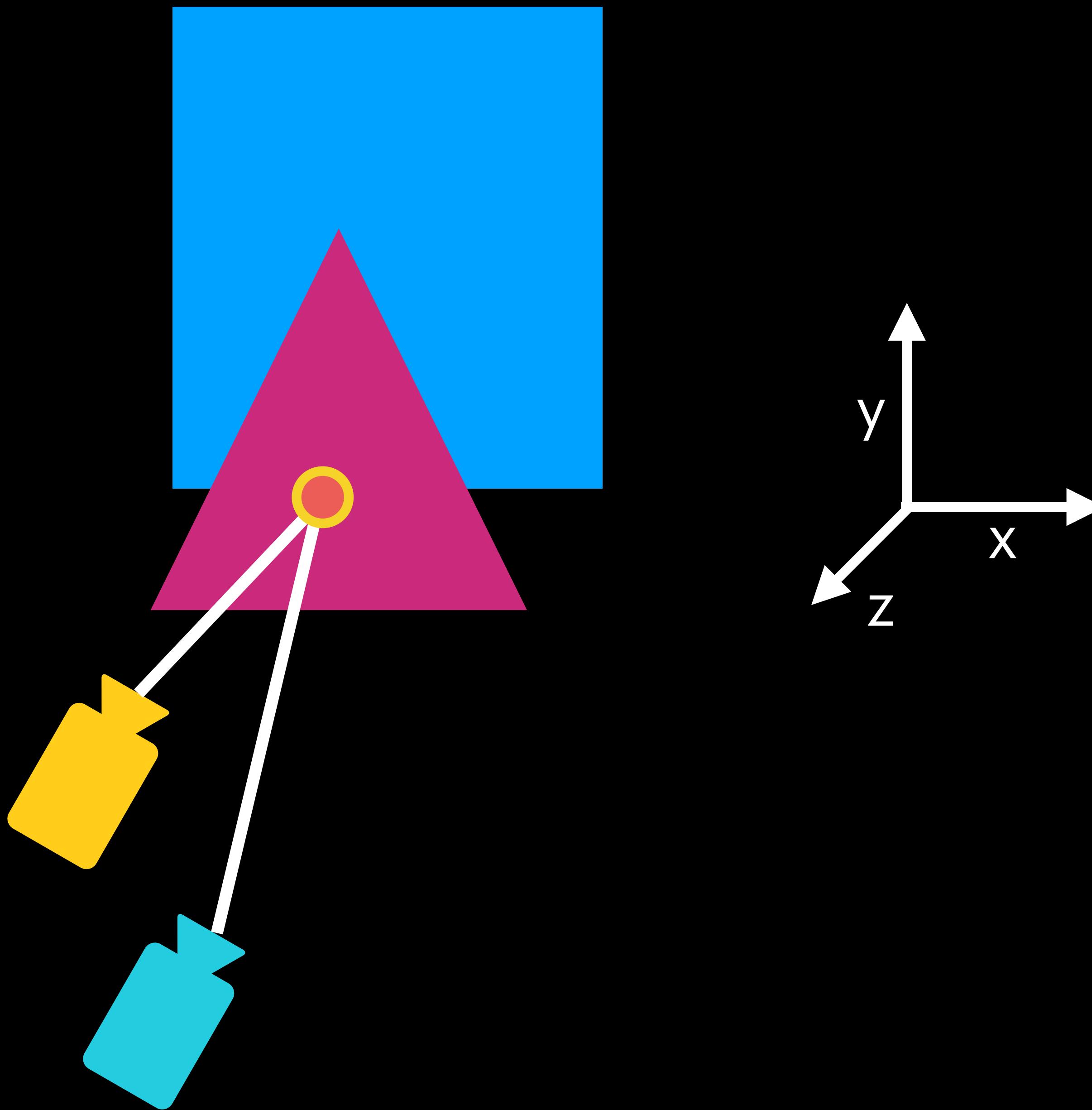


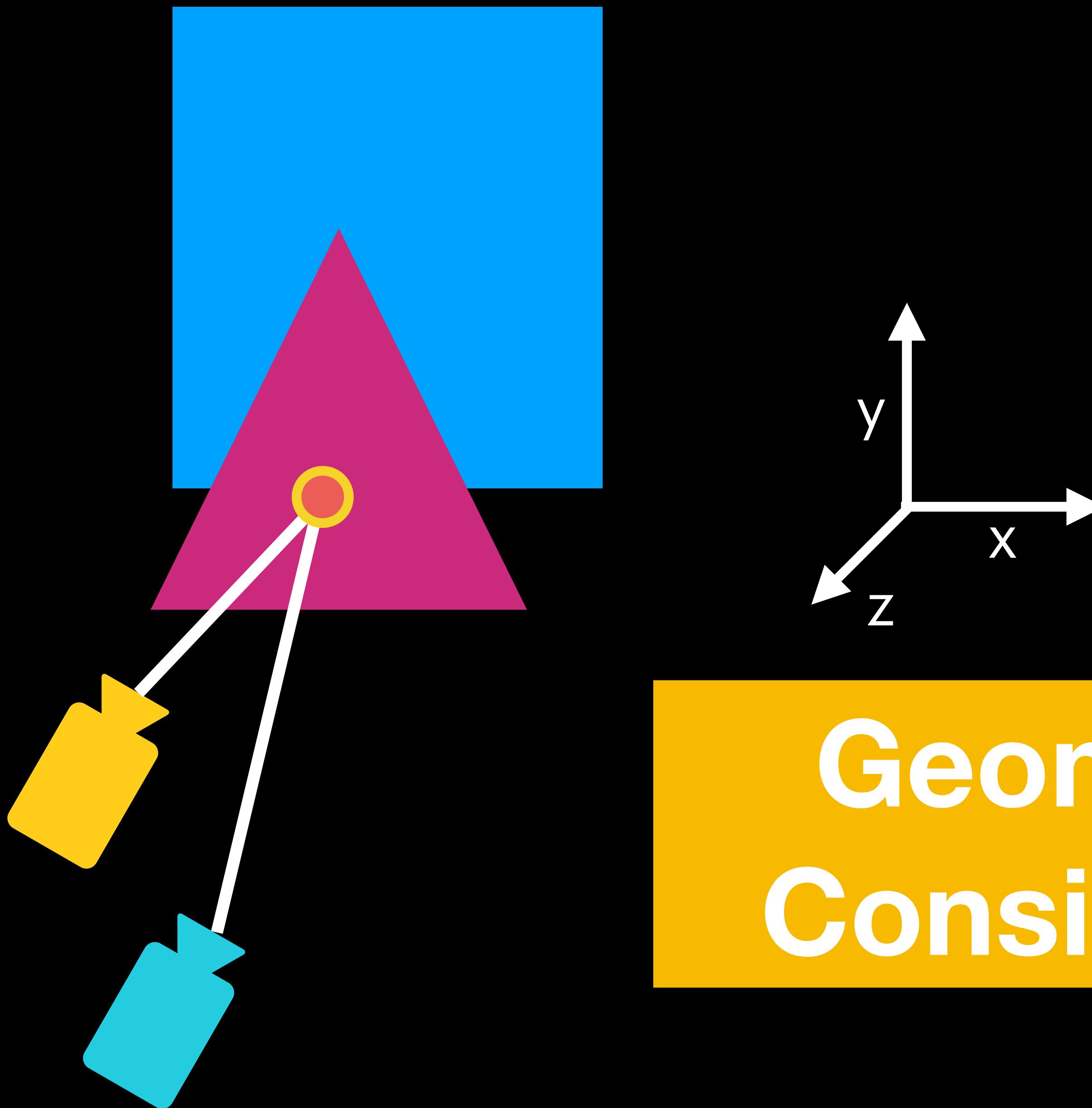




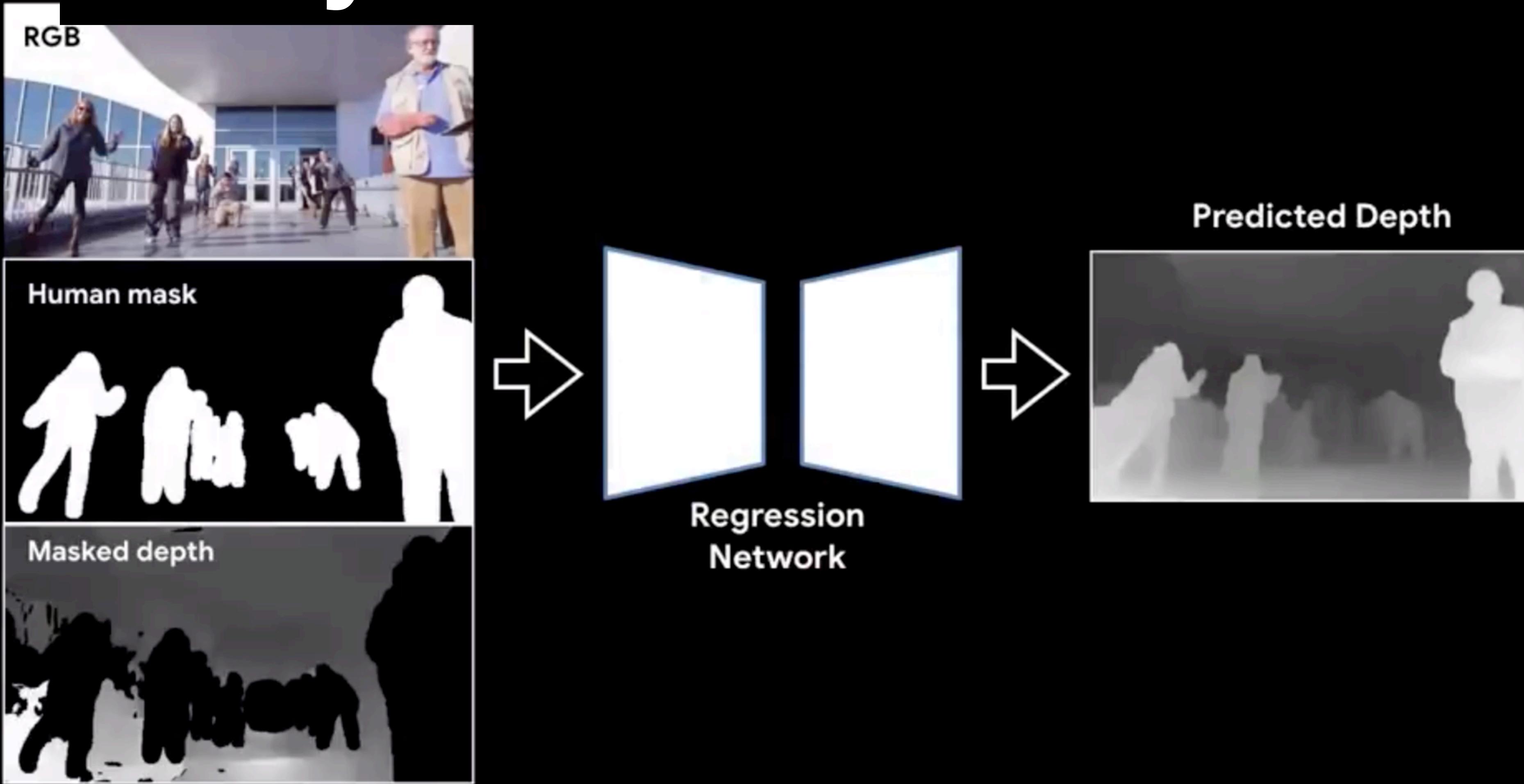








Hybrid CNN+MVS





Input



[Li et al. 2019]



Input

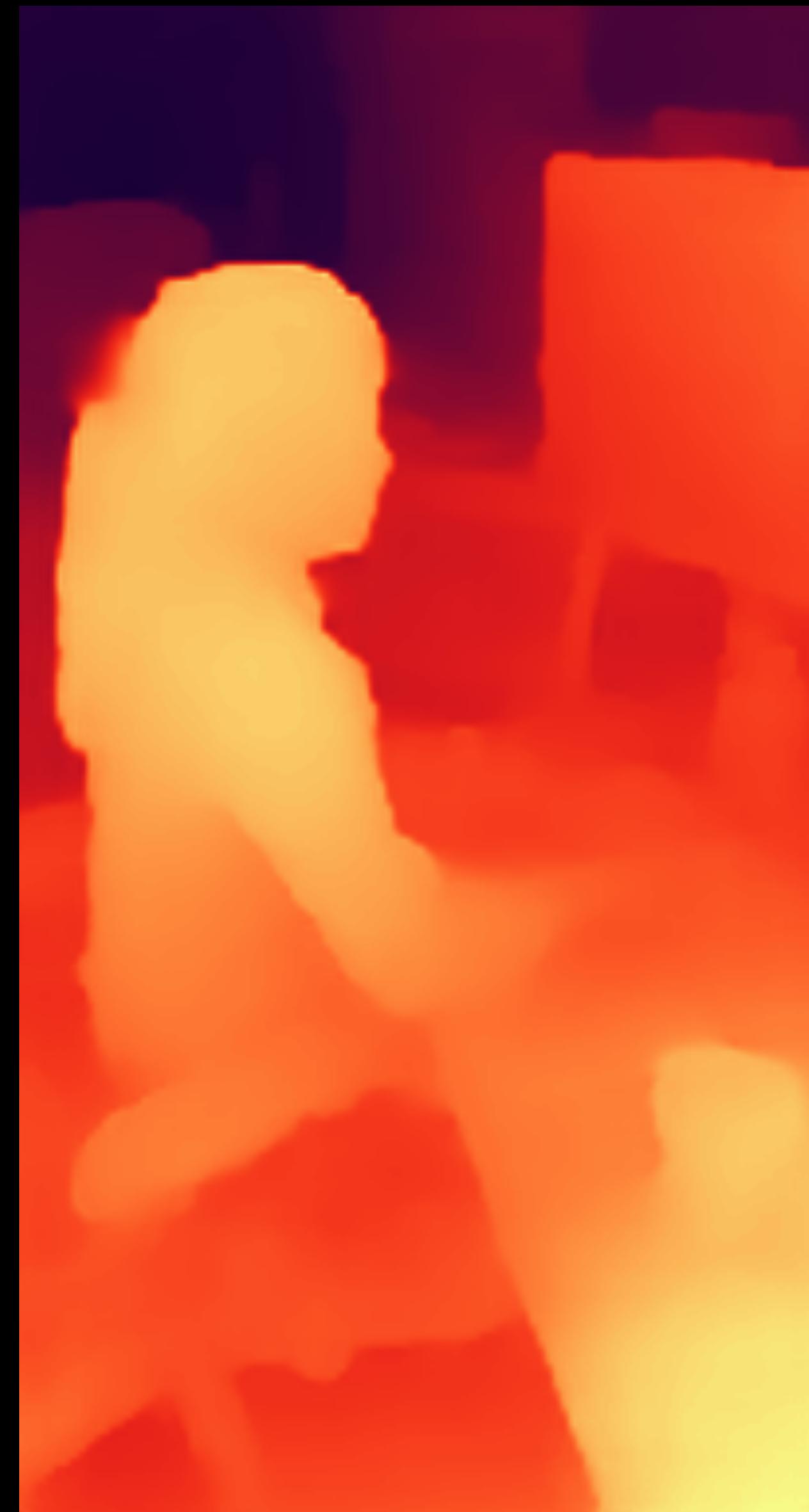
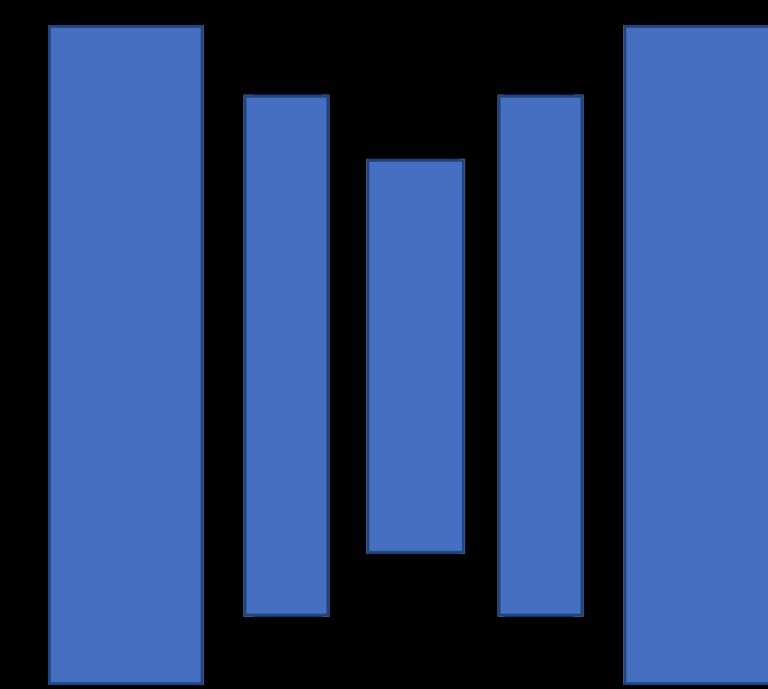
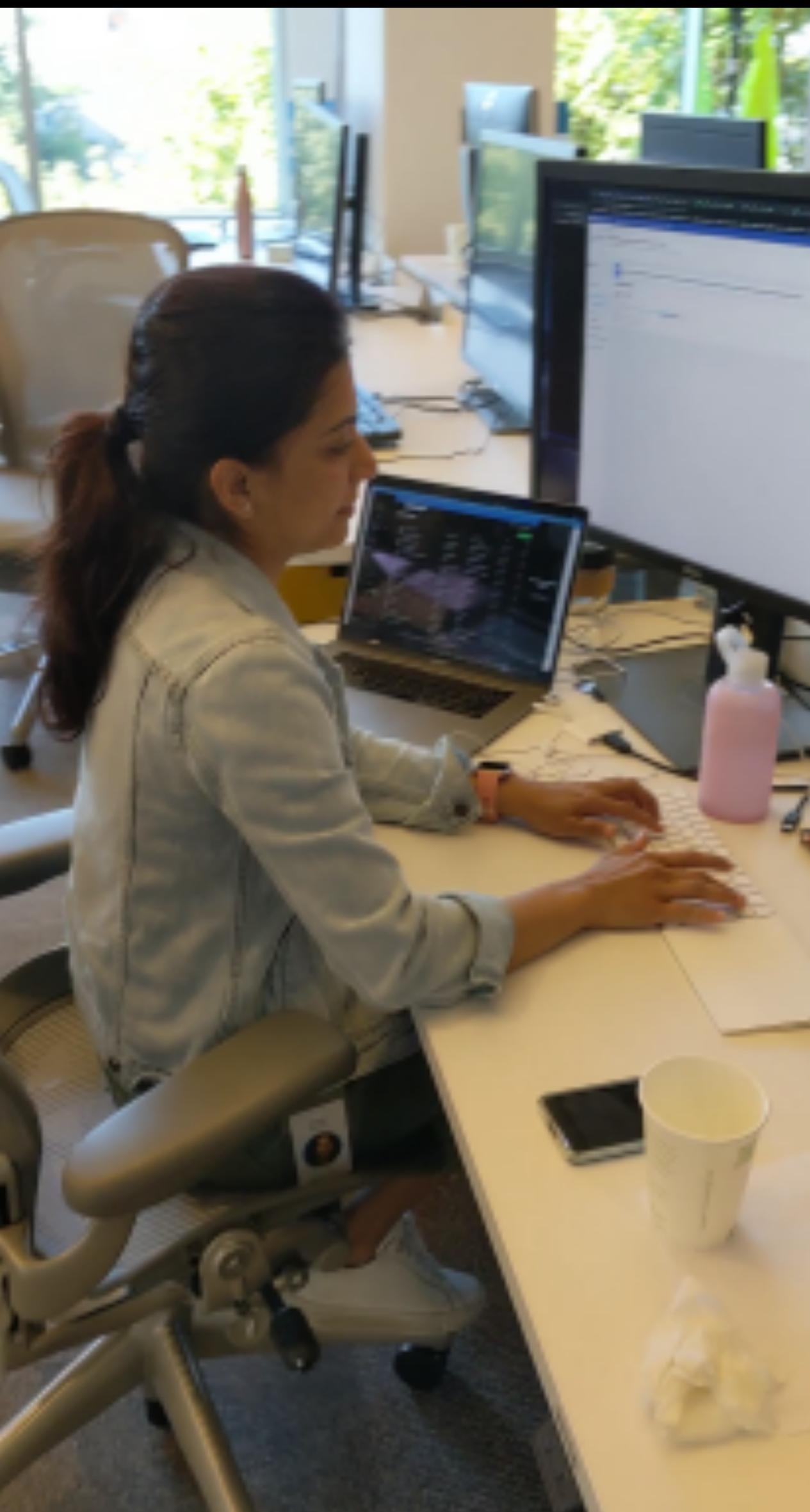


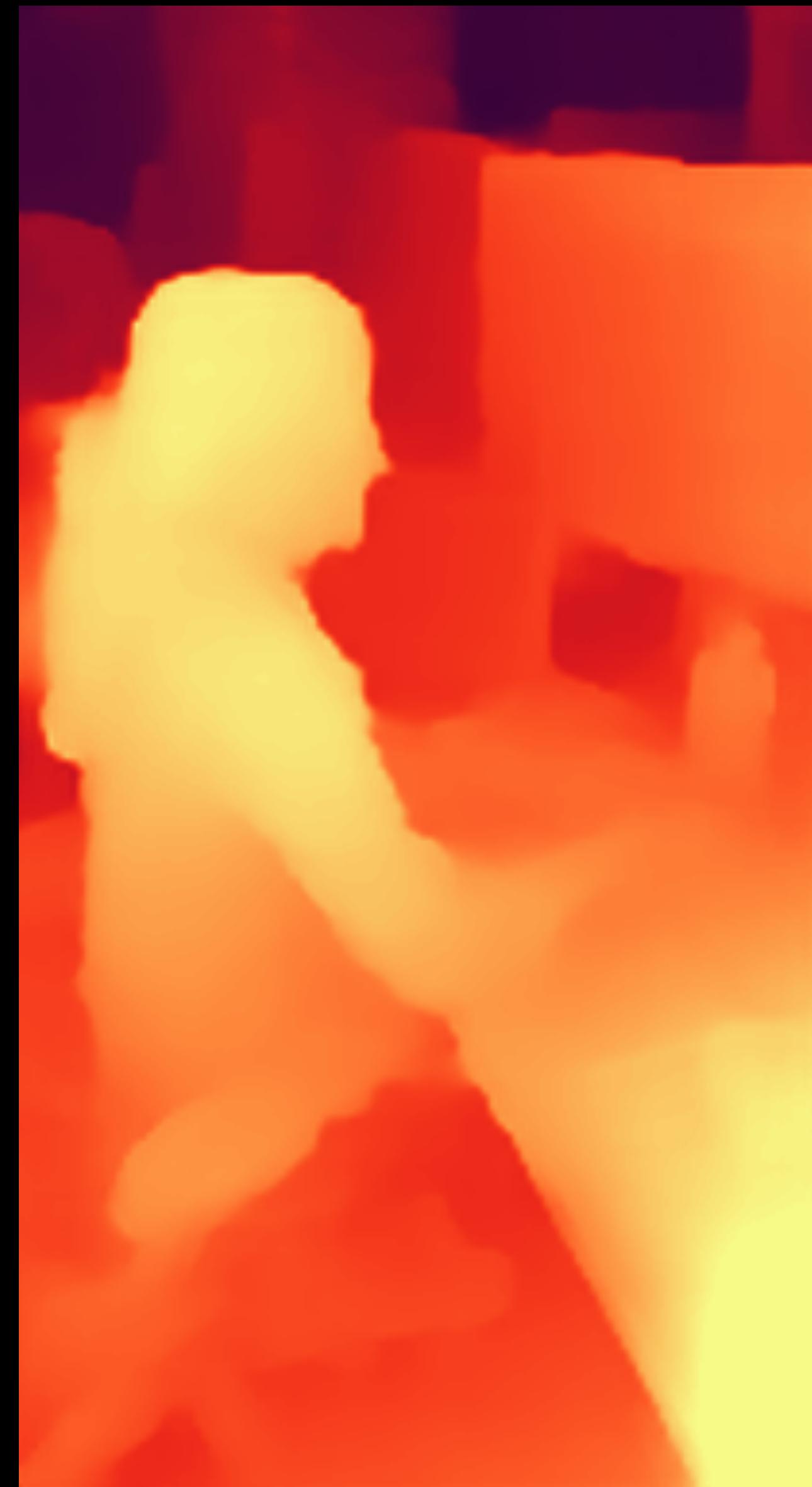
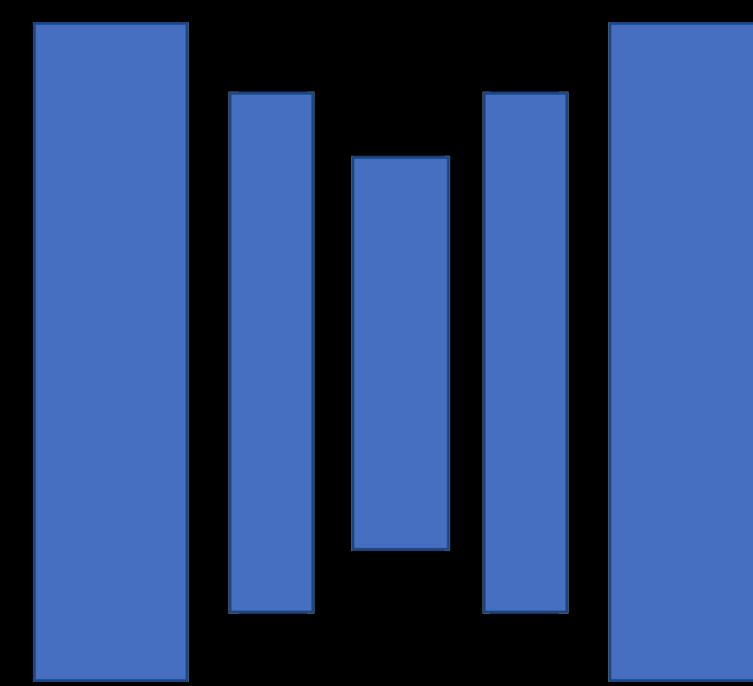
[Li et al. 2019]

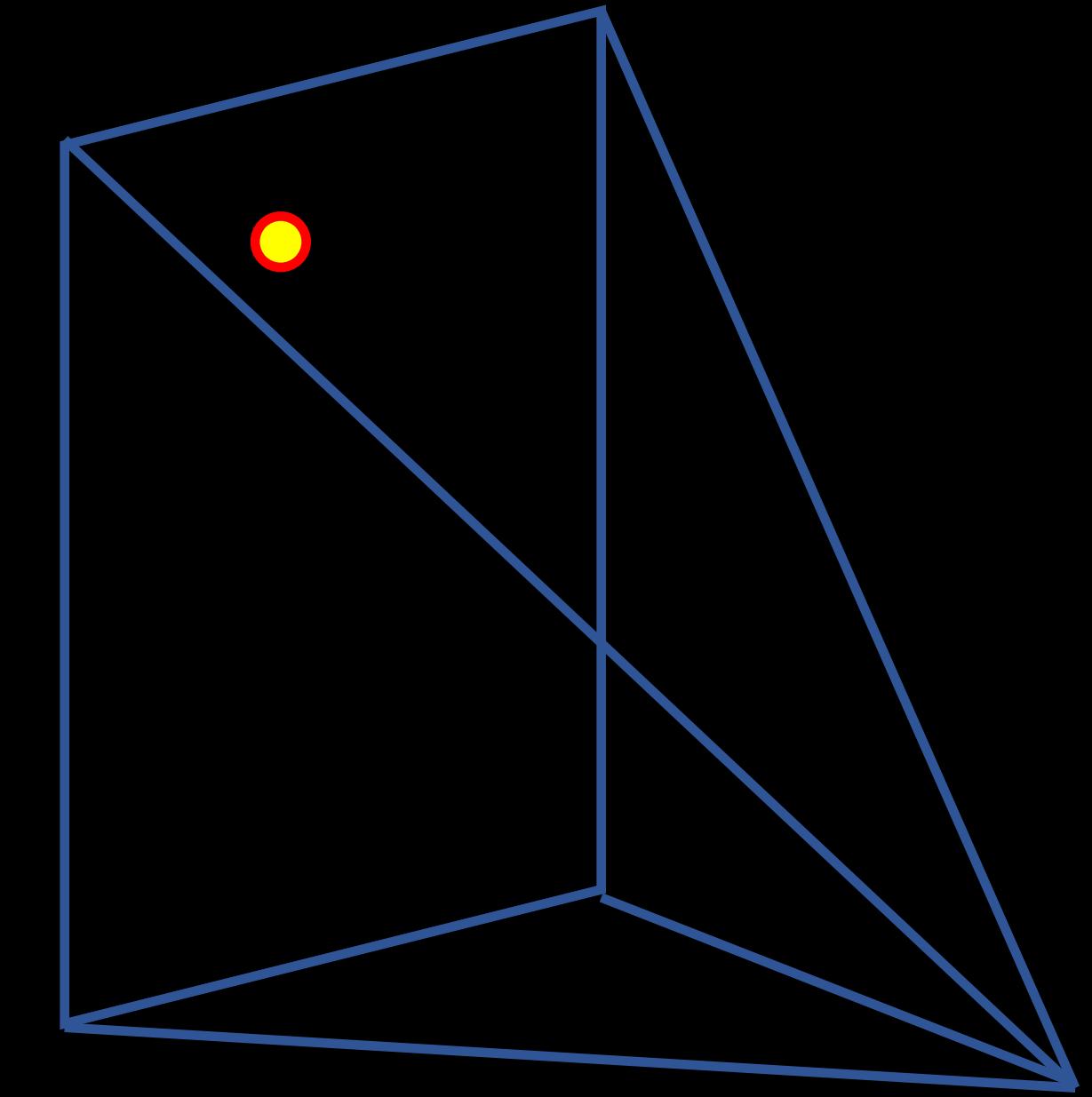
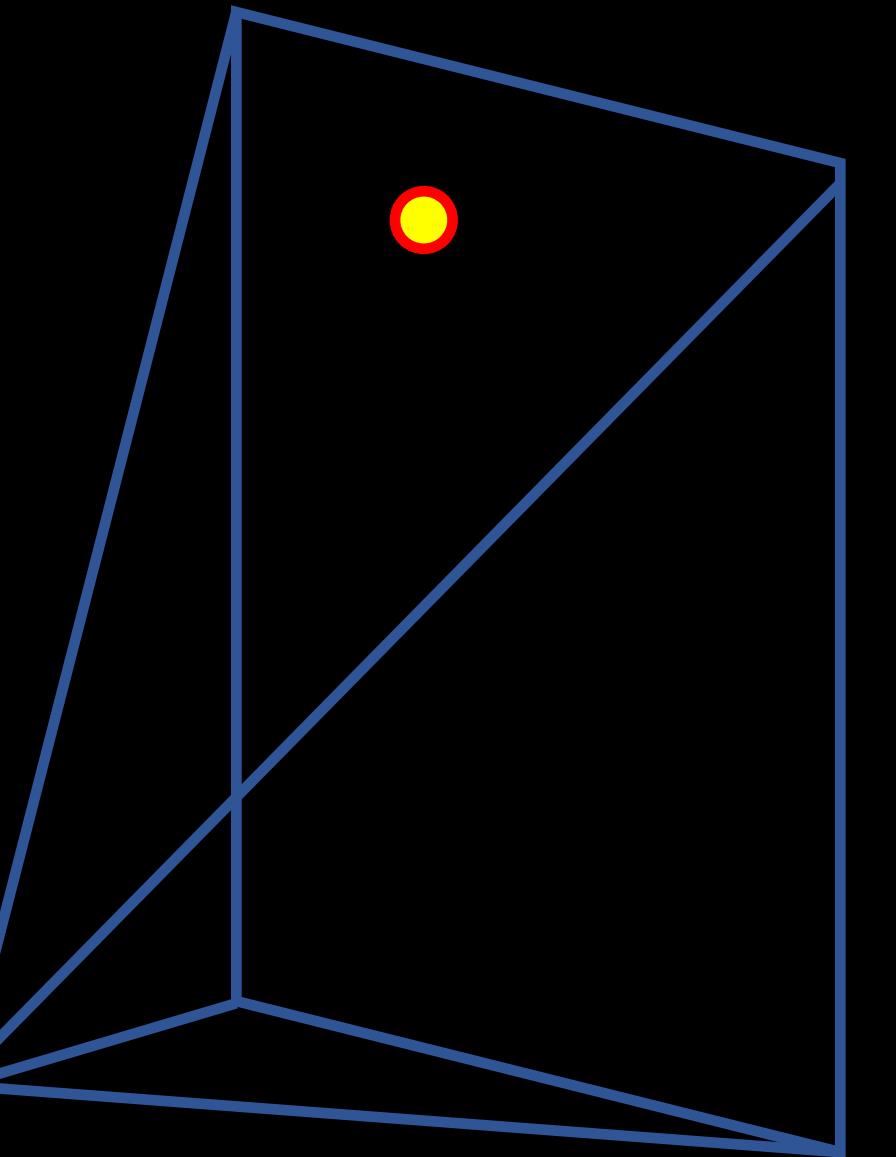
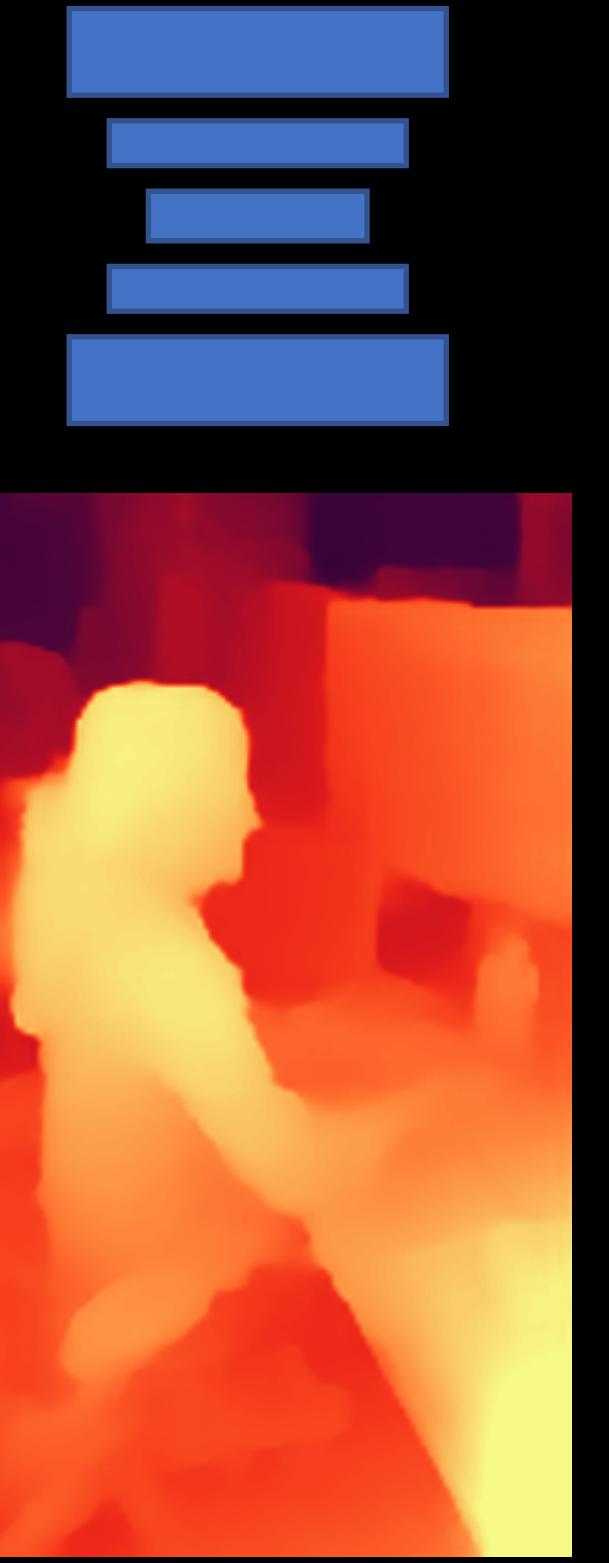
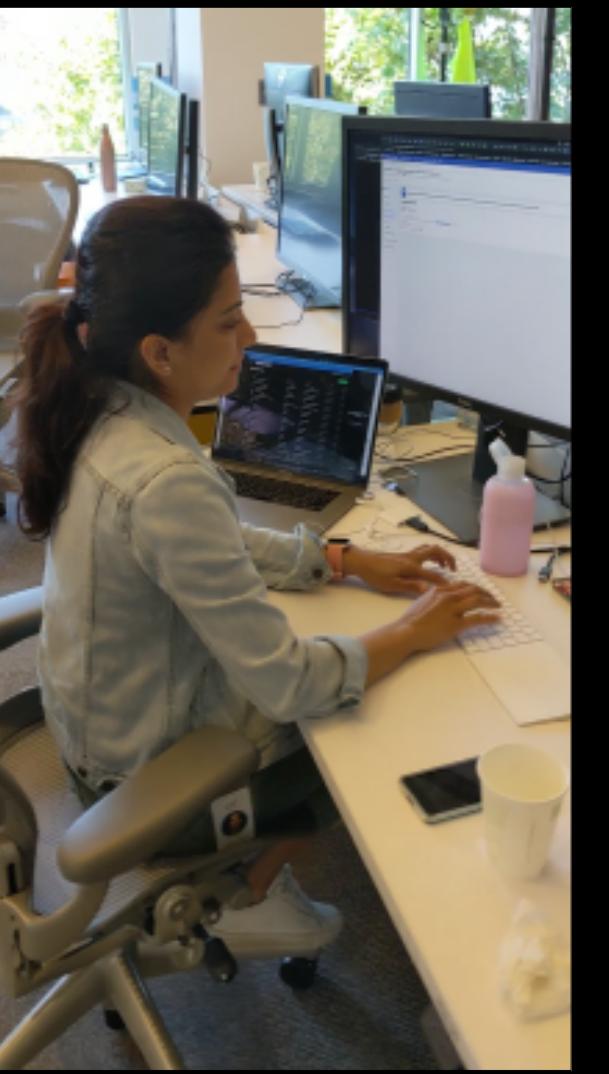
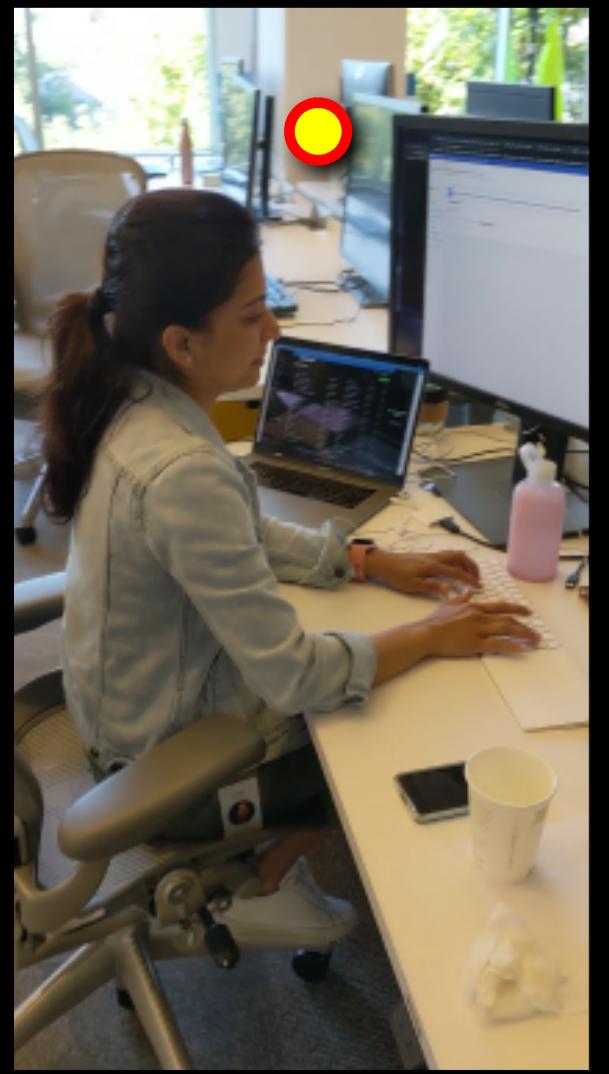
	Traditional Methods	Learning-based	Hybrid
Structure from Motion	Multi-View Stereo (COLMAP, Instant3D)	Single-frame Depth Estimation	Human Mannequin Challenge, WSVD, etc.
Dense	N	Semi	Y
Dynamics	N	N	Y
Geometric Consistency	Y	Y	3D points swim around

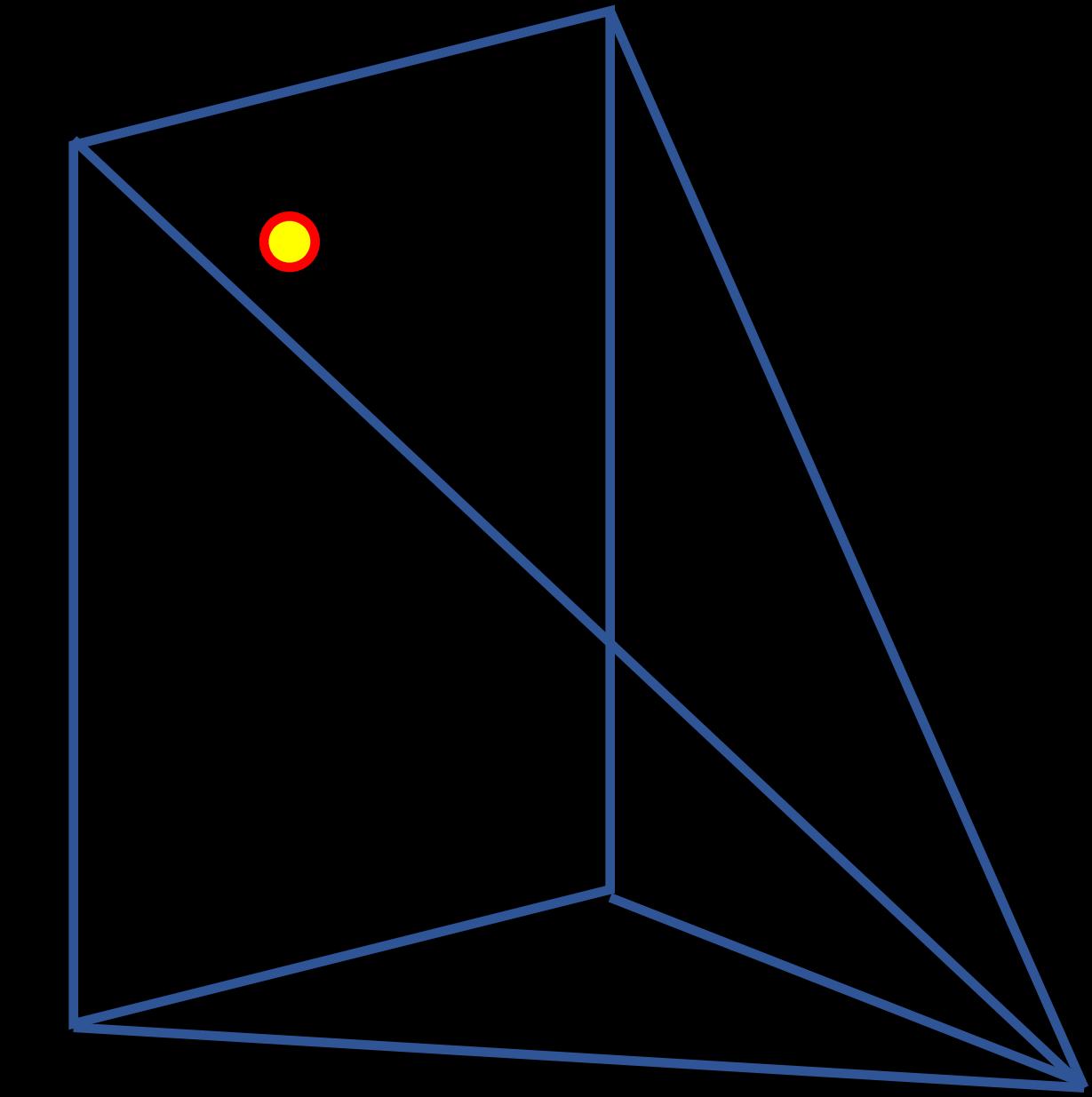
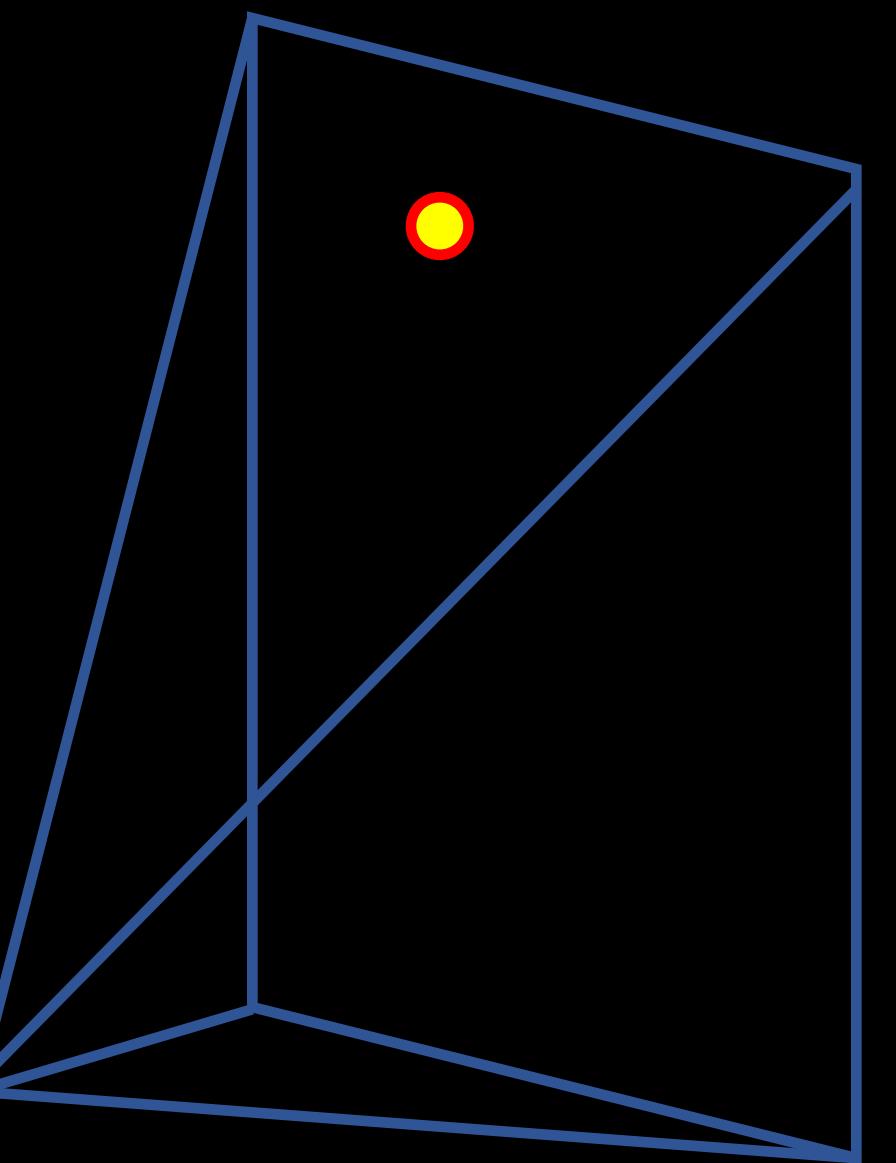
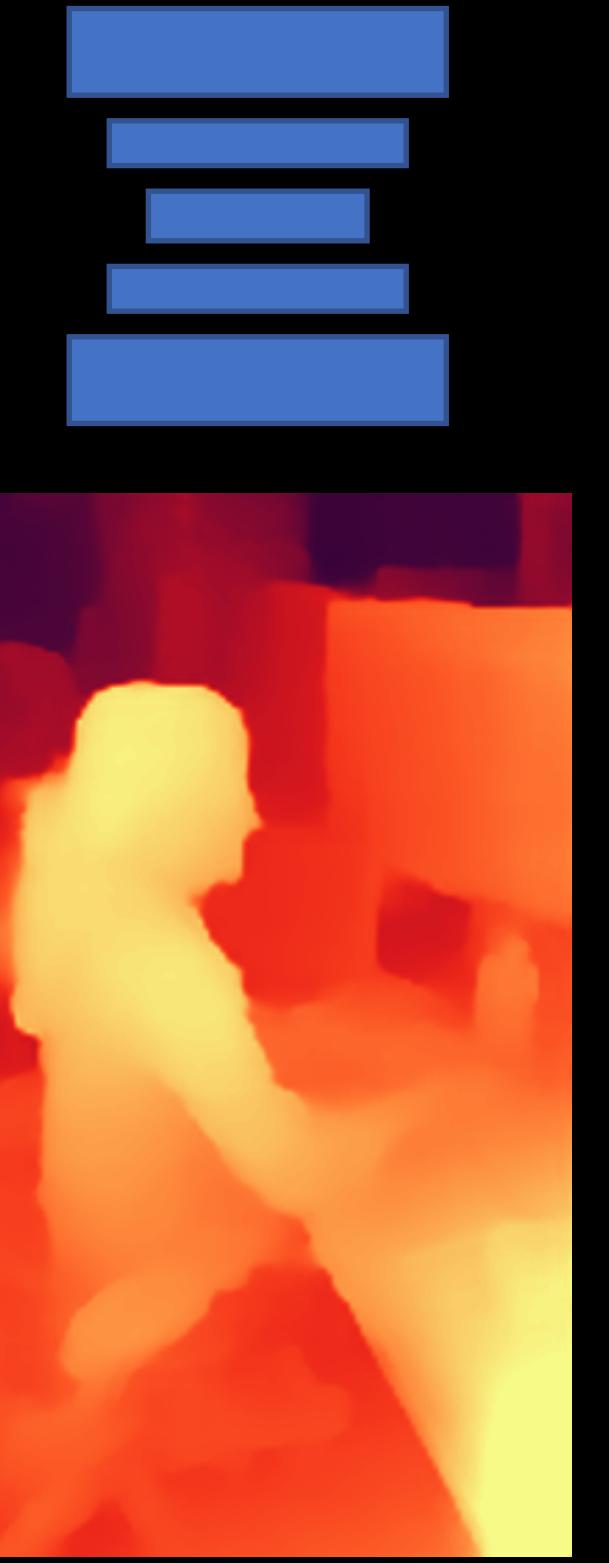
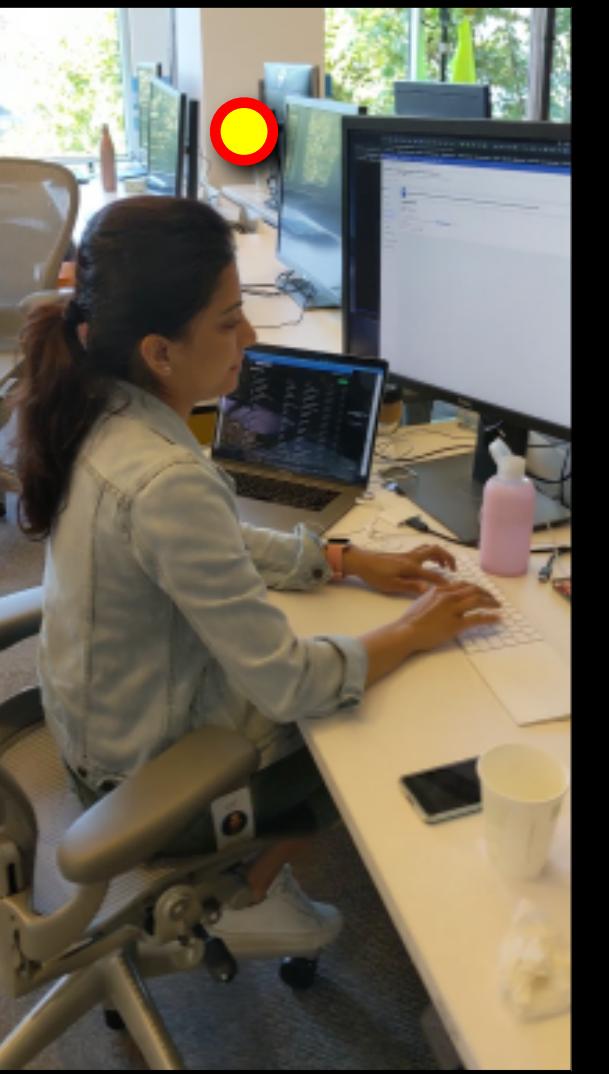
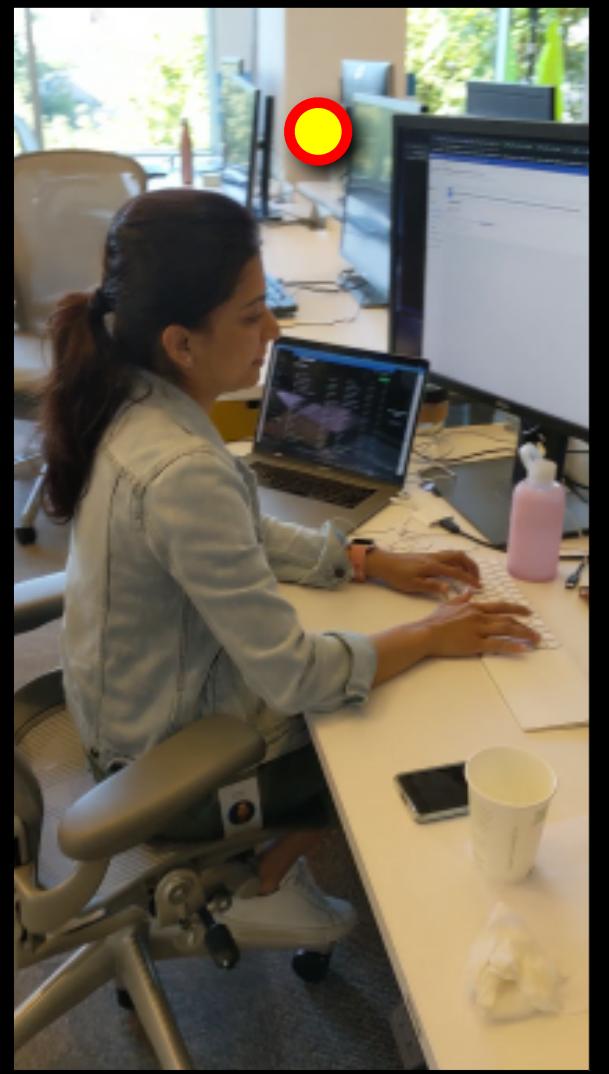
	Traditional Methods	Learning-based	Hybrid	Ours
Structure from Motion	Multi-View Stereo (COLMAP, Instant3D)	Single-frame Depth Estimation	Human Mannequin Challenge, WSVD, etc.	
Dense	N	Semi	Y	Y
Dynamics	N	N	Y	Moderate Motion
Geometric Consistency	Y	Y	N	3D points swim around

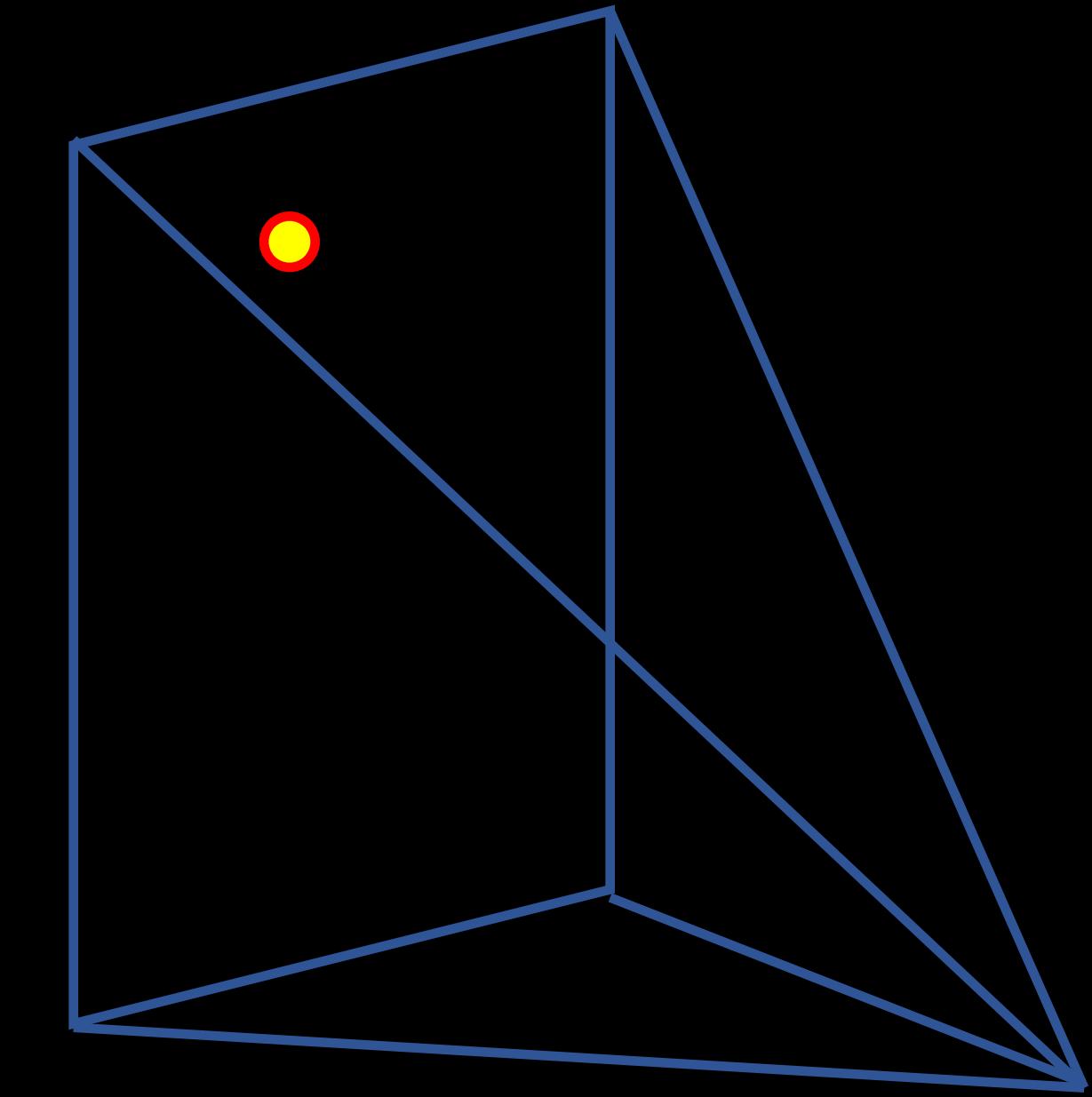
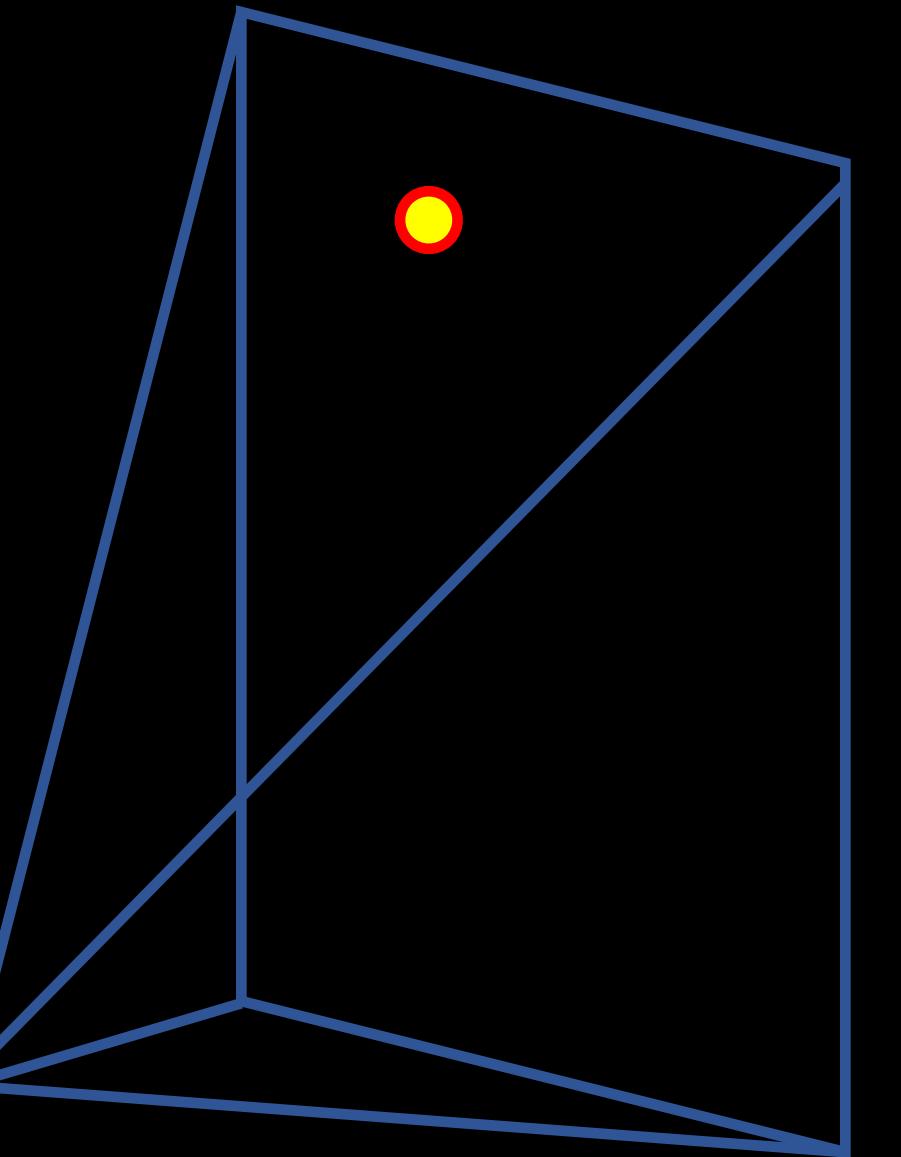
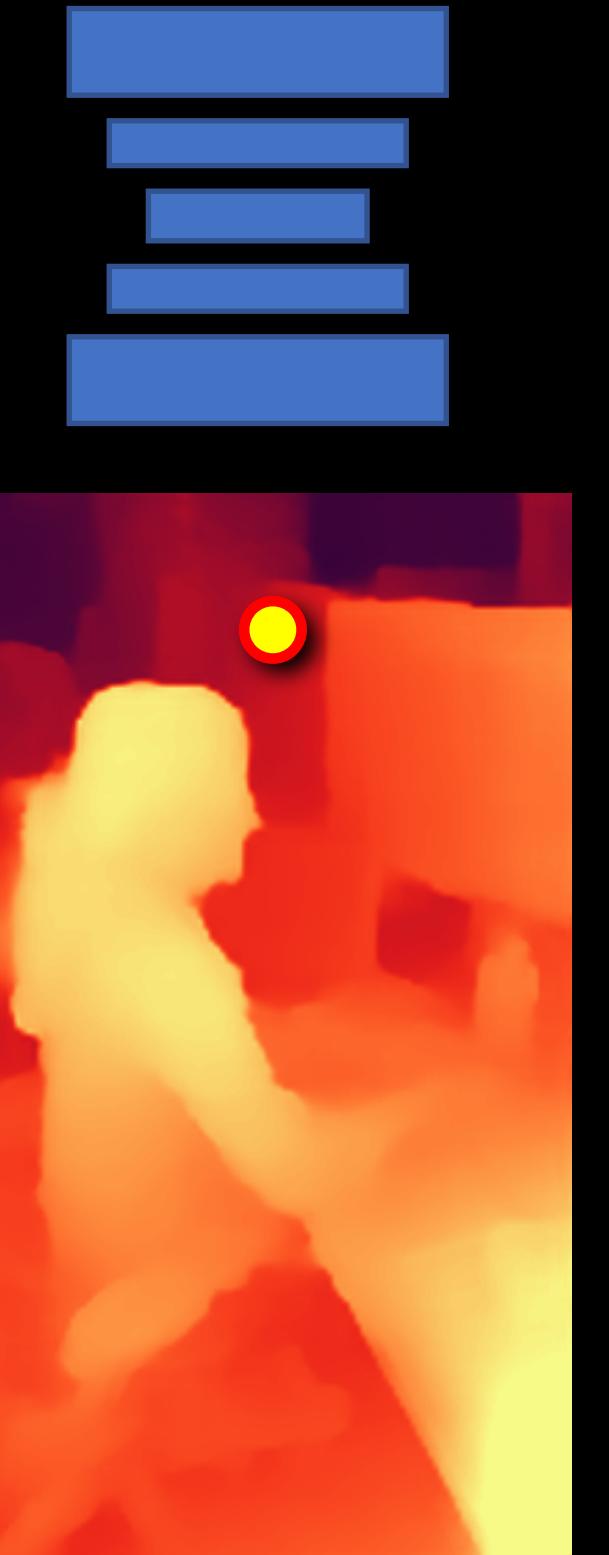
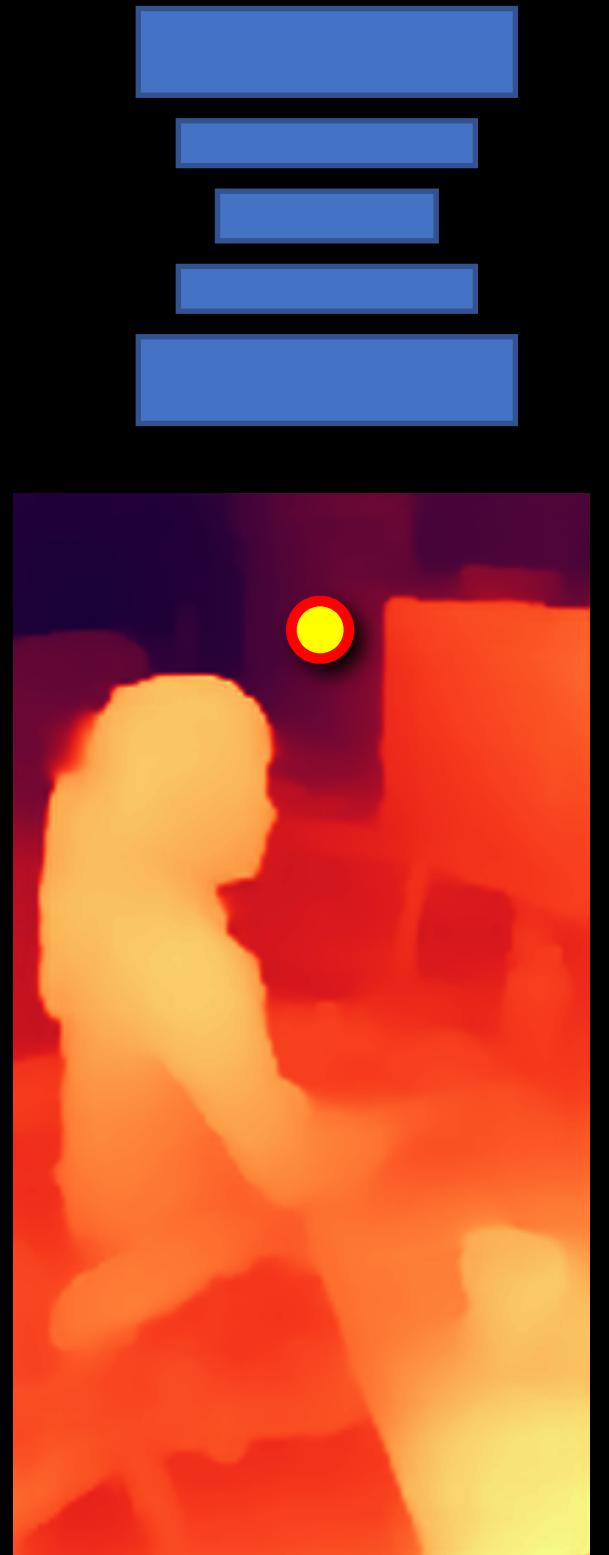
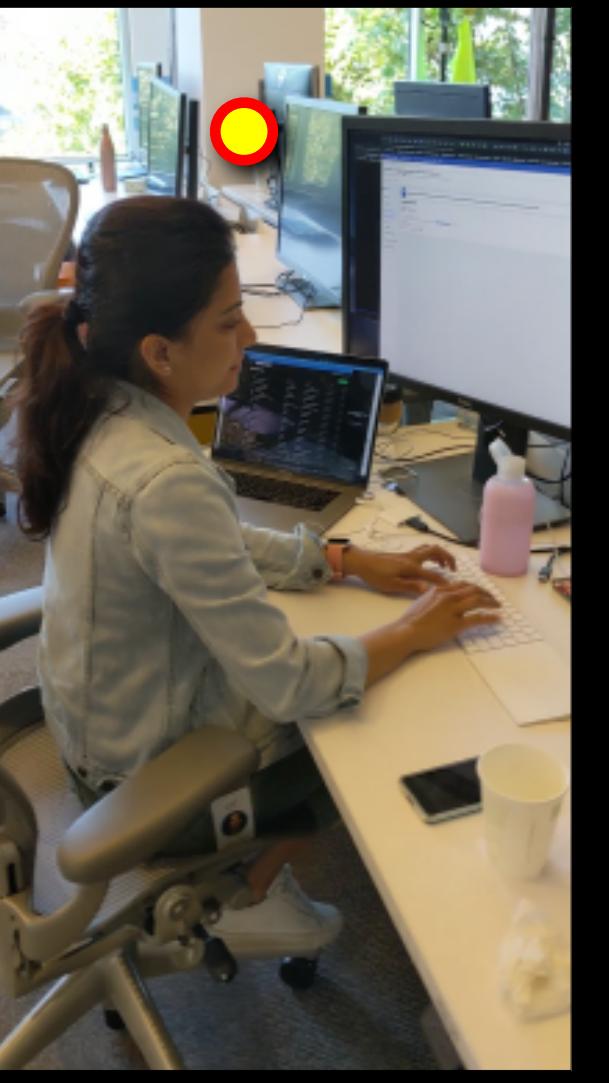
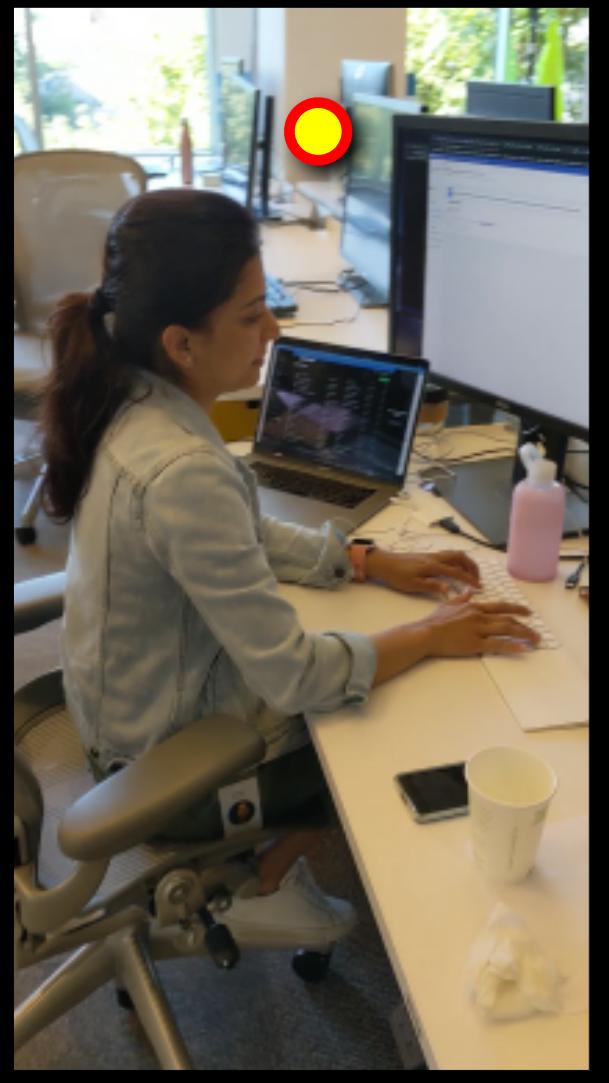
Method

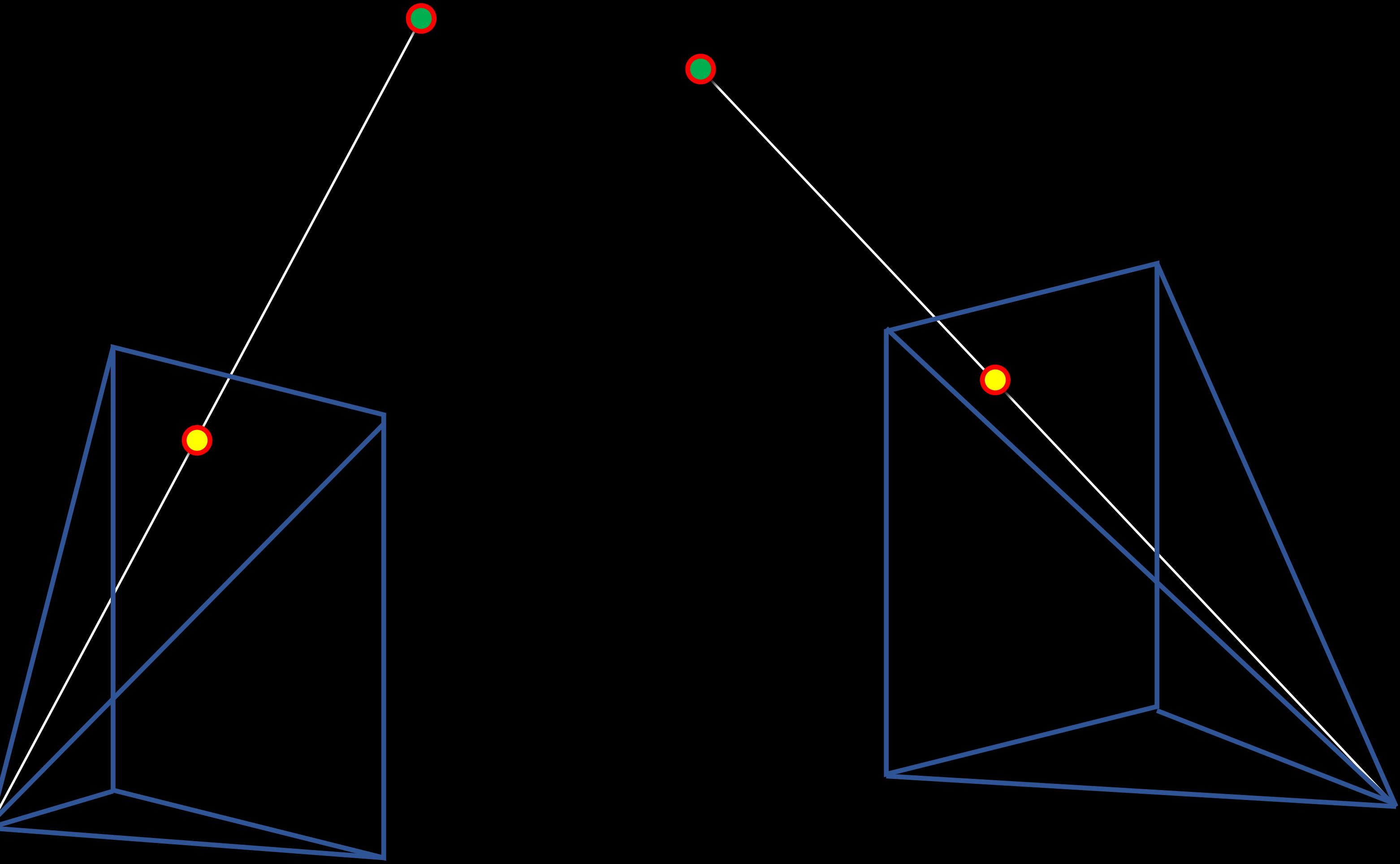
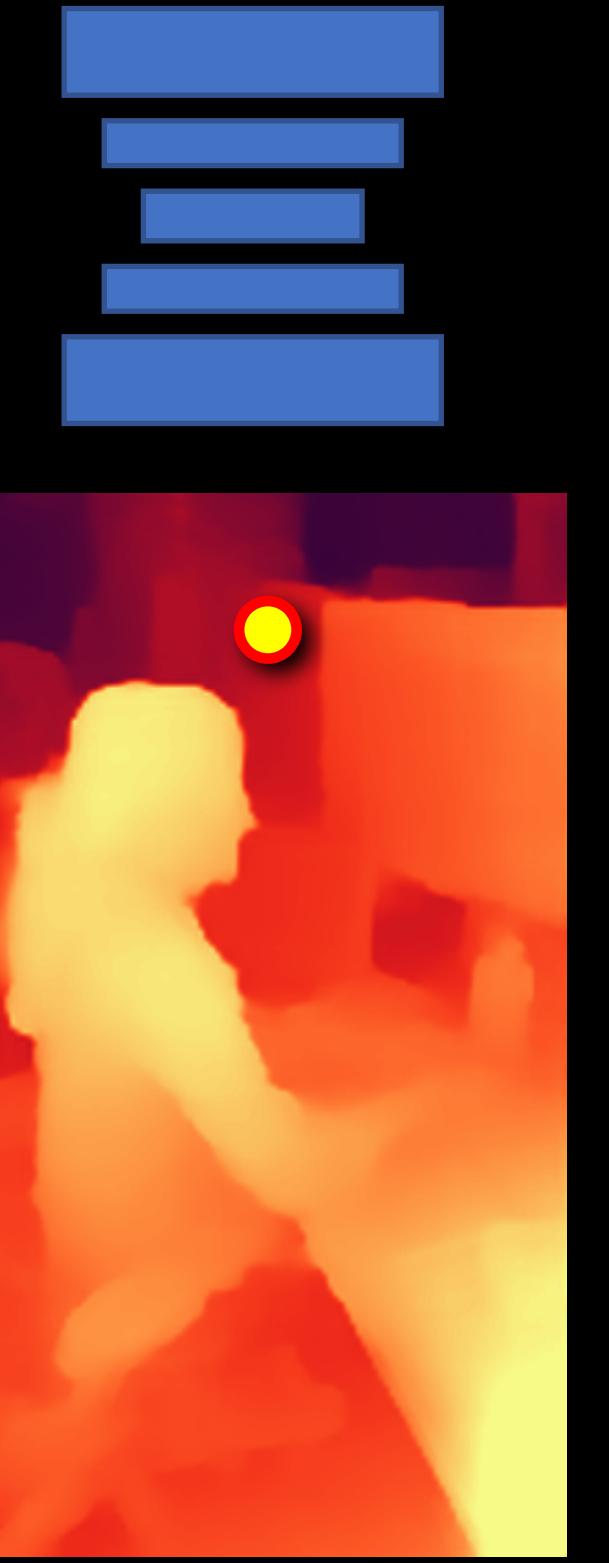
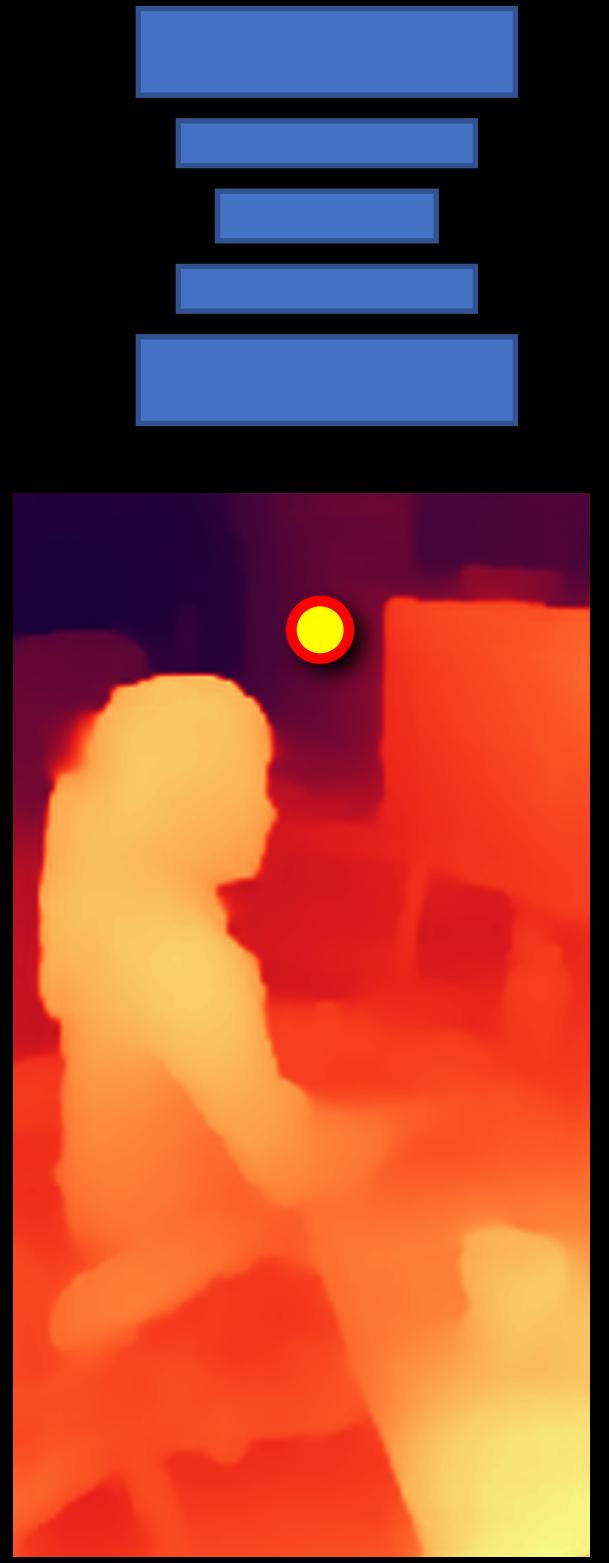
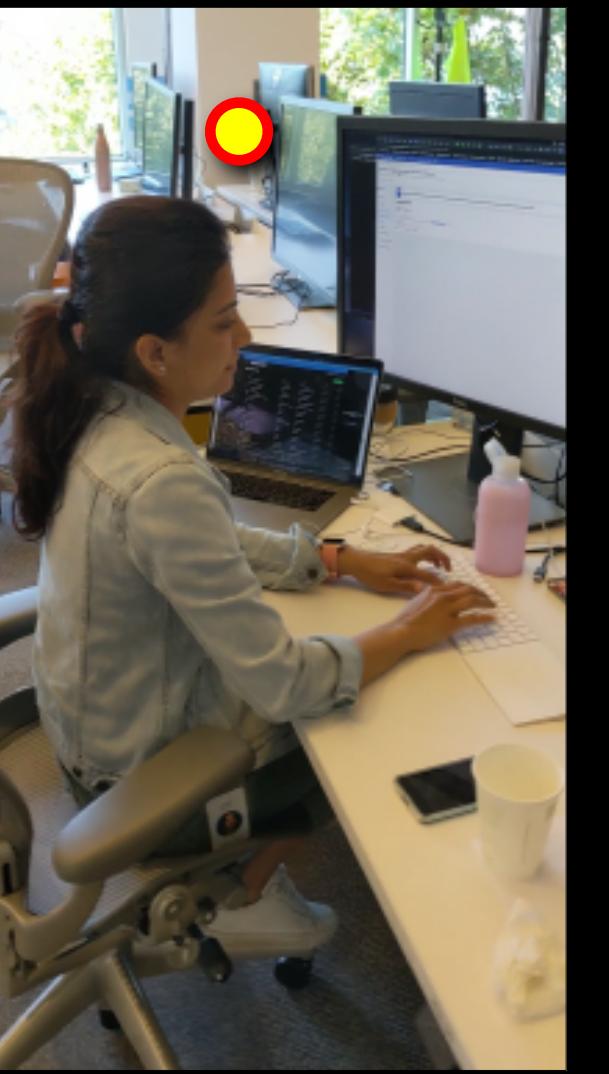
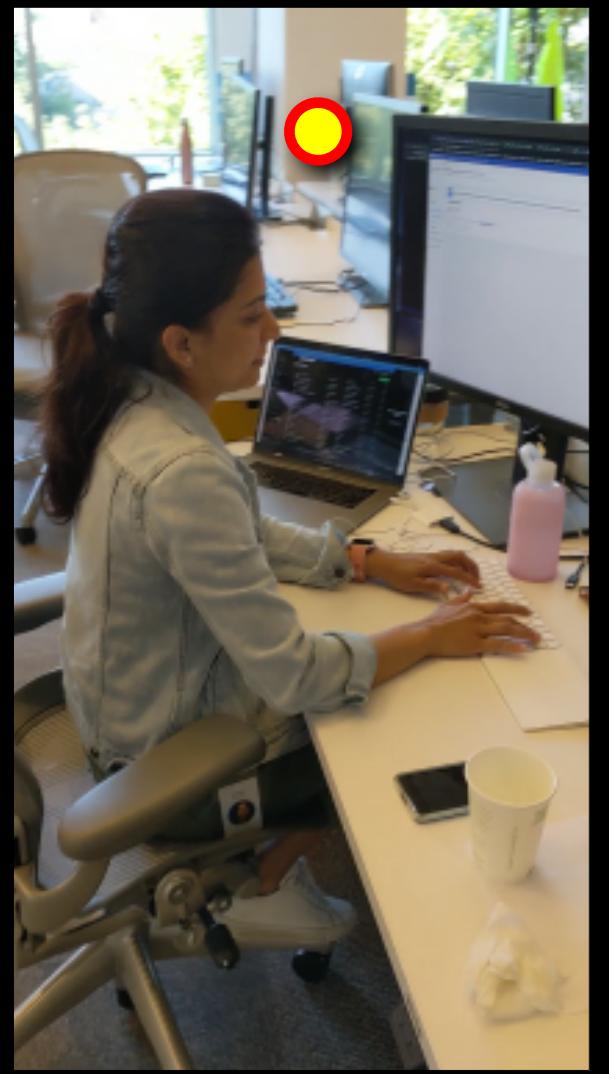


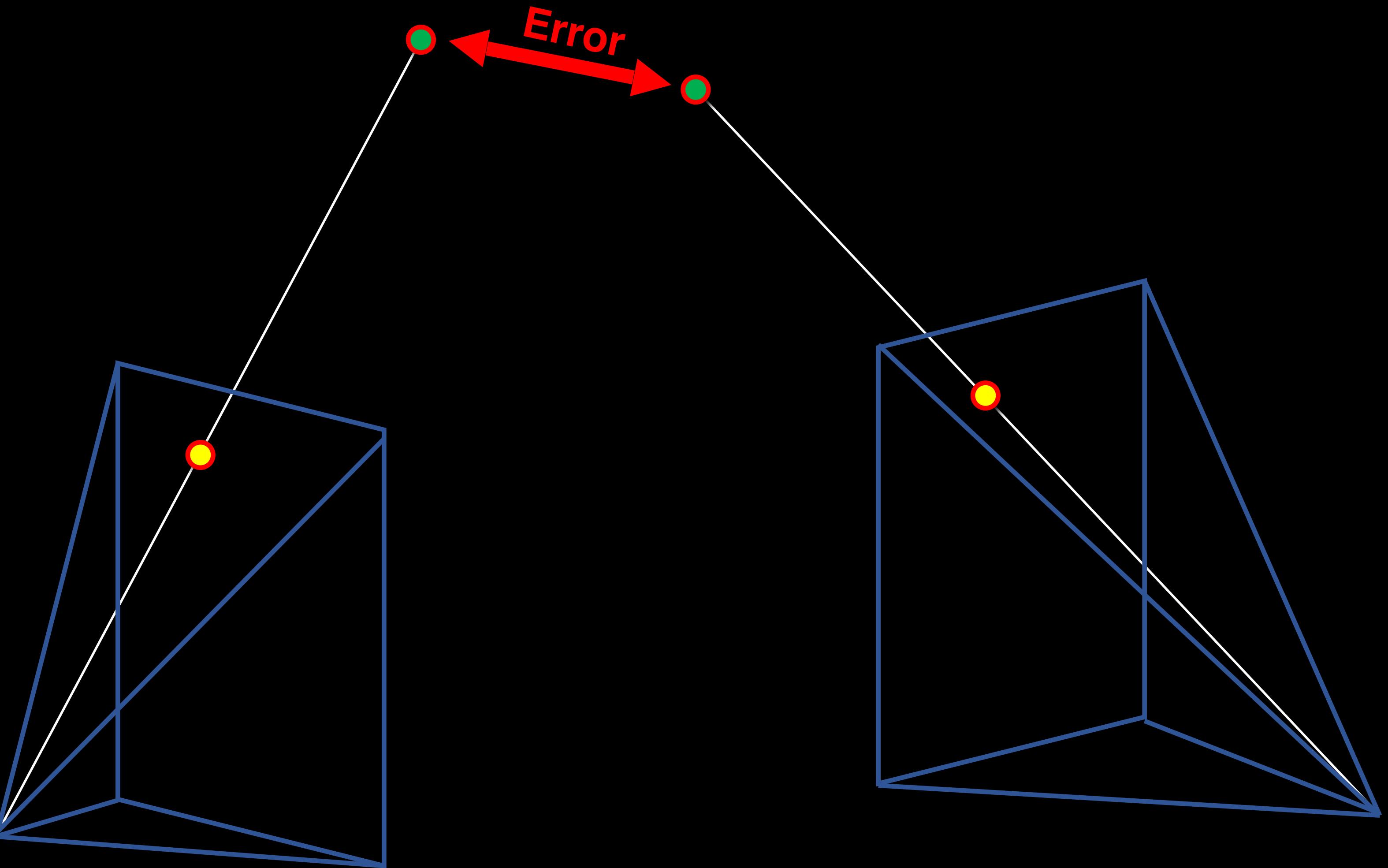
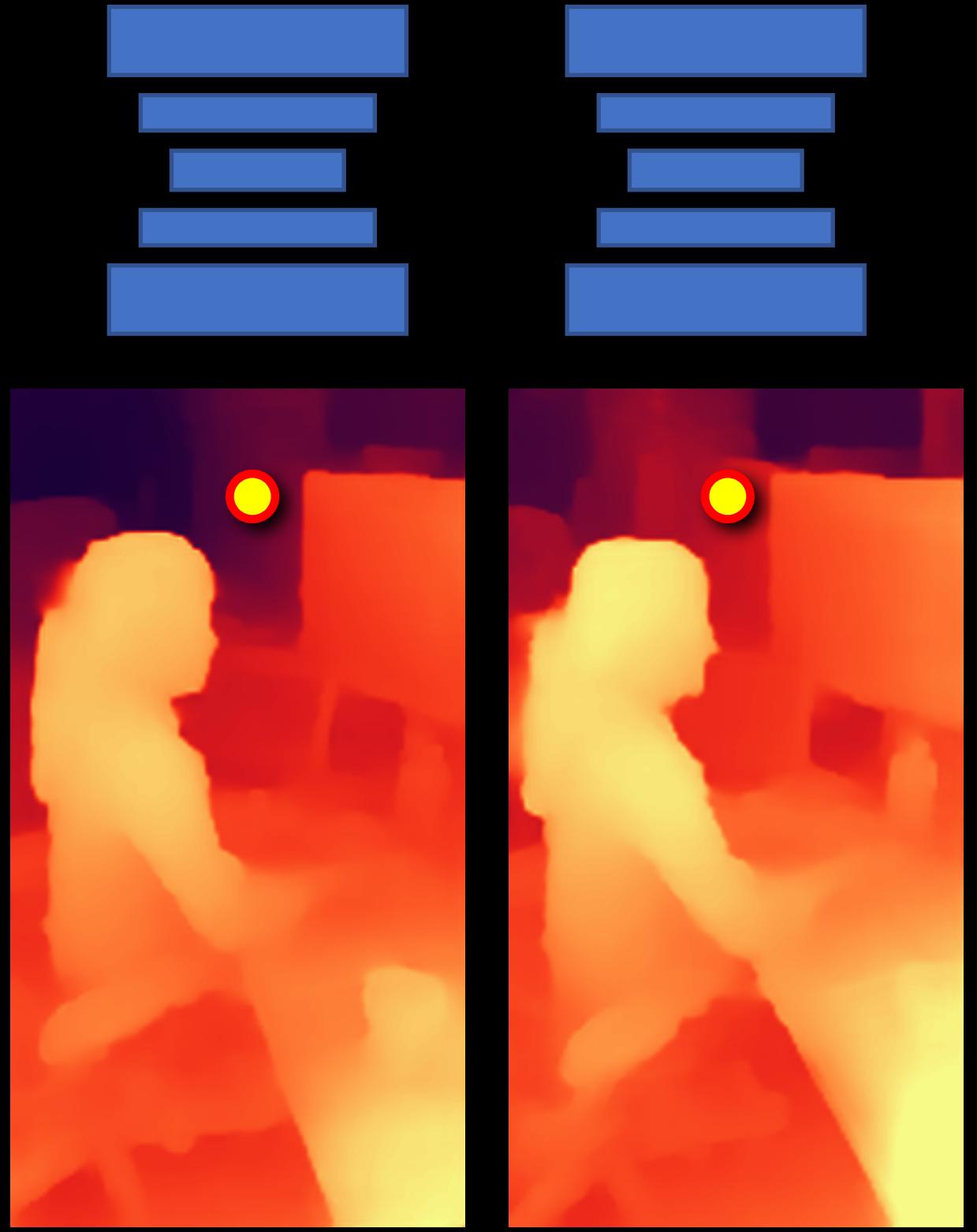
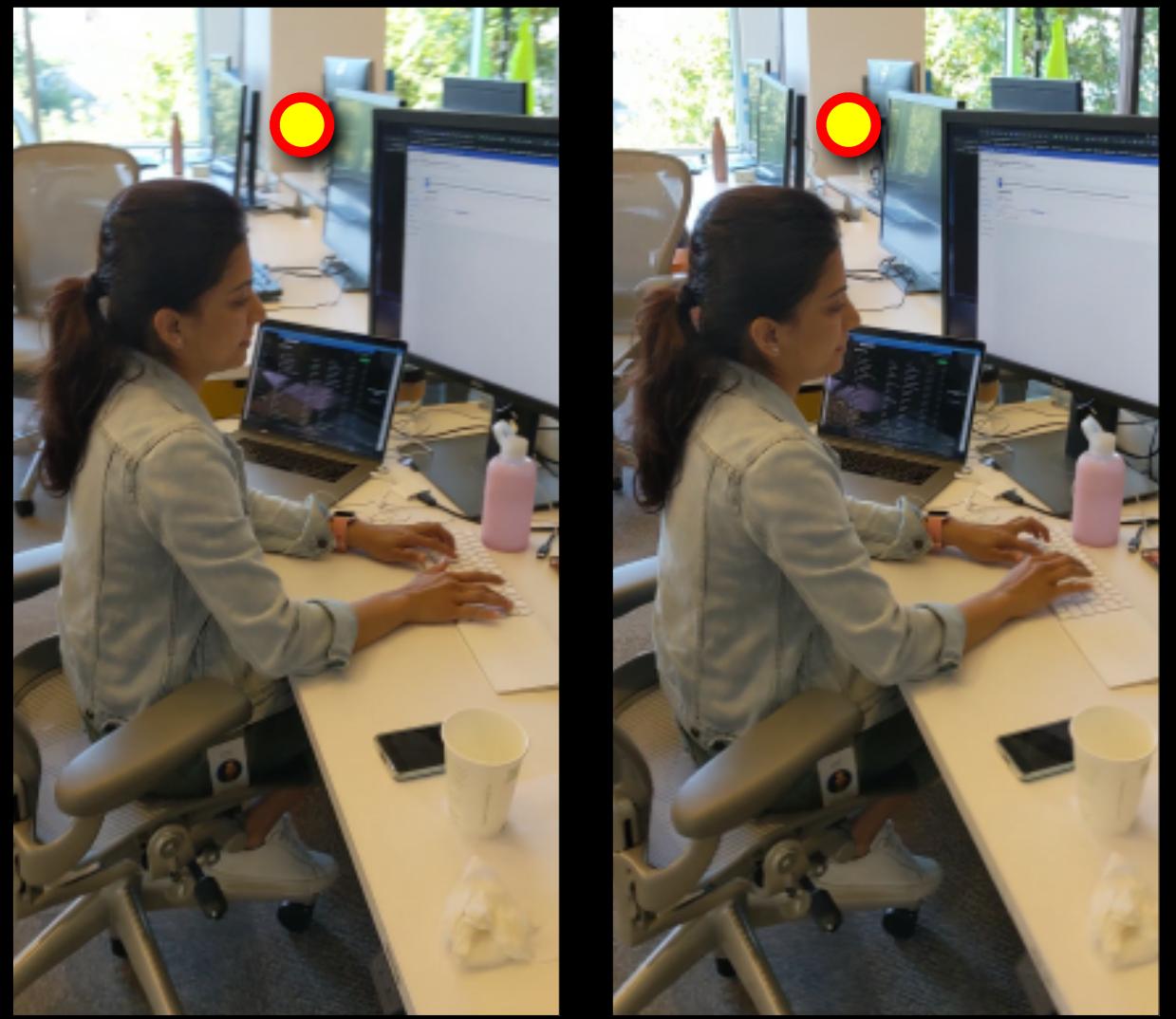


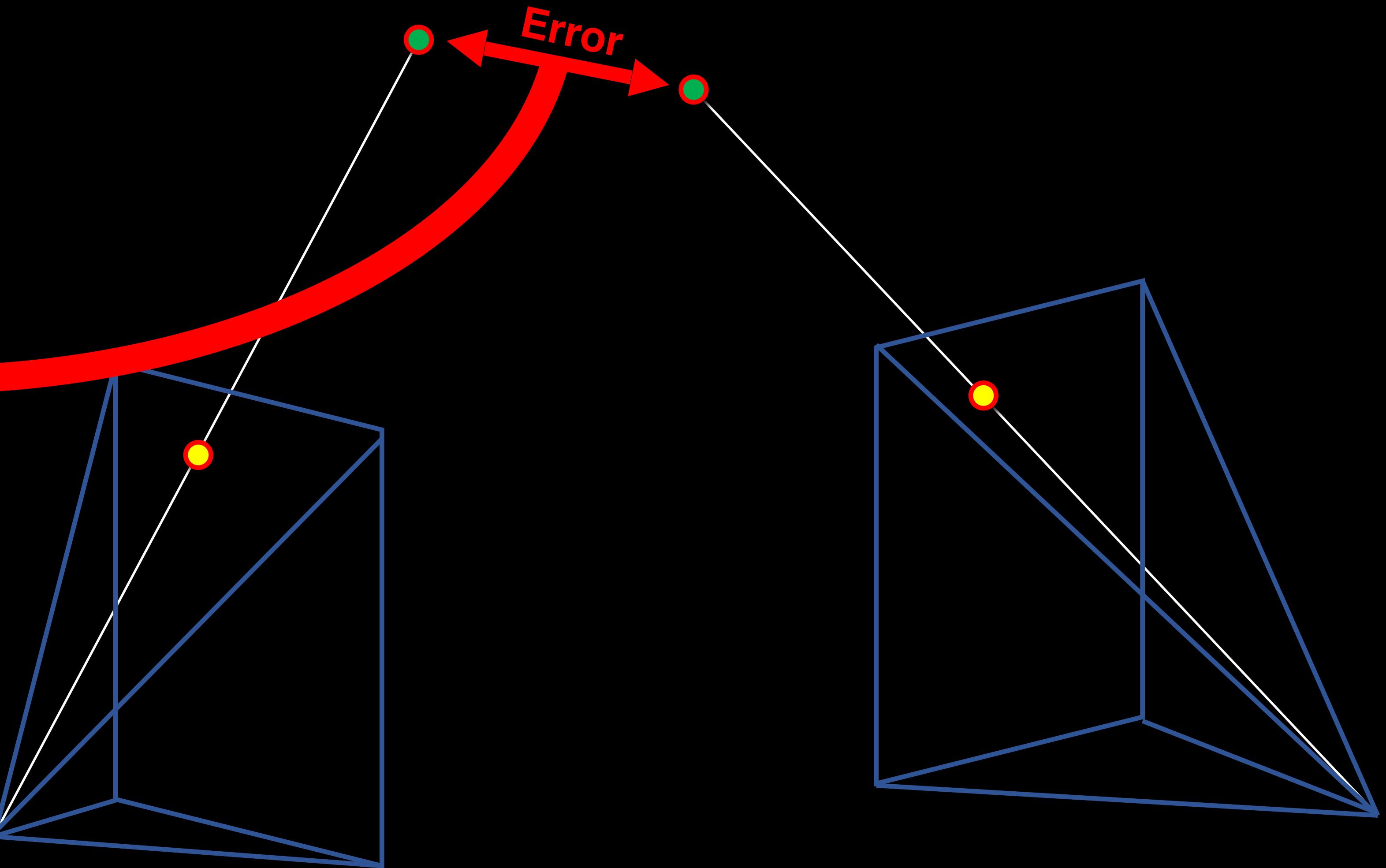
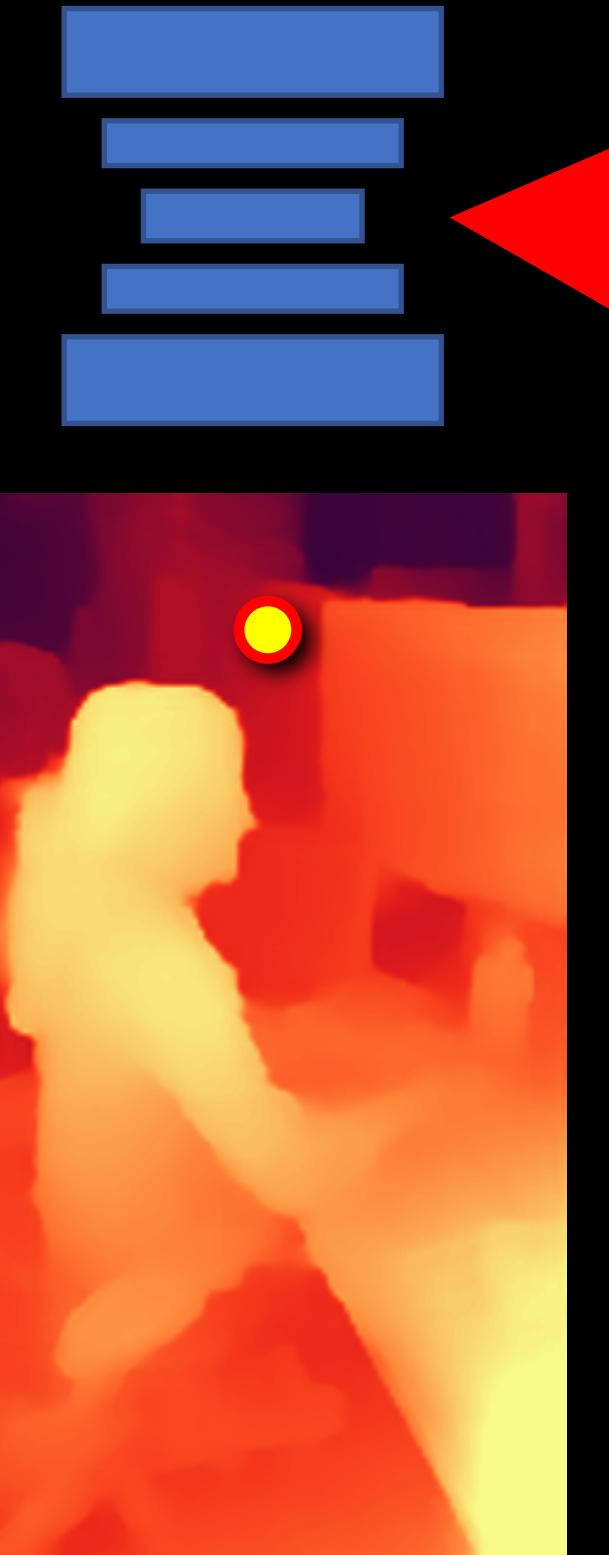
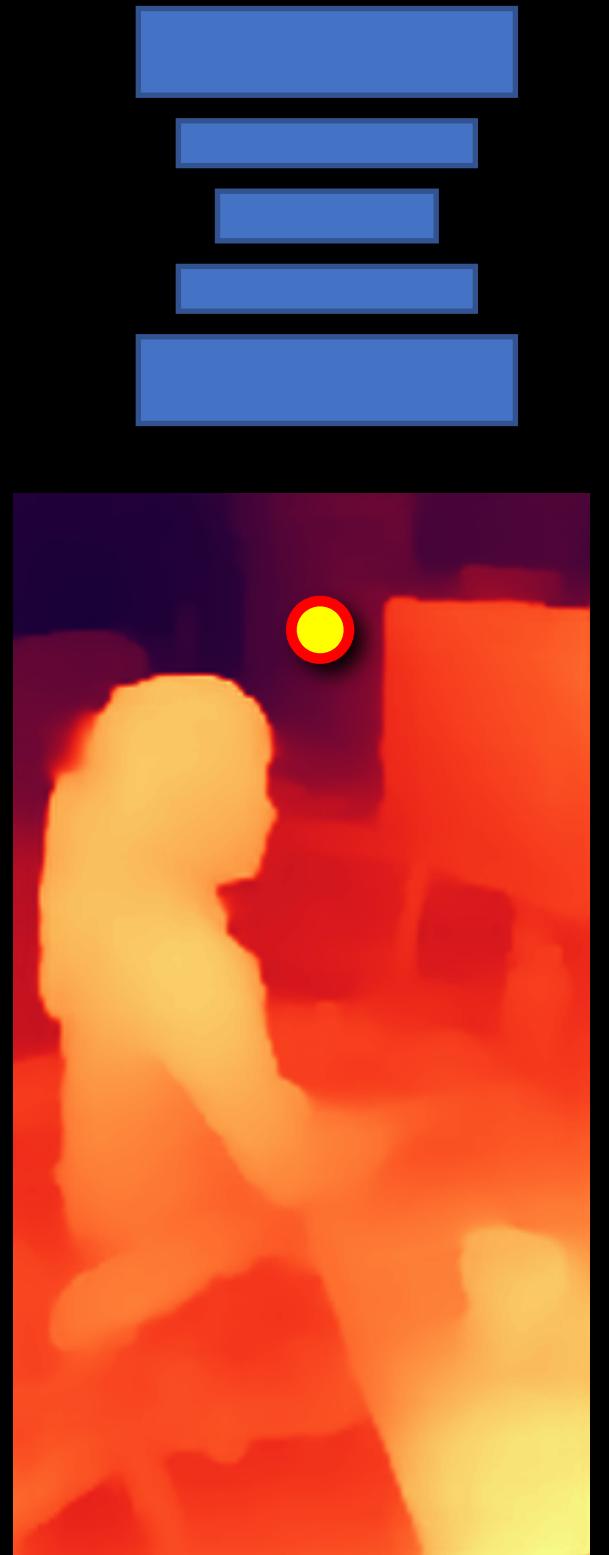
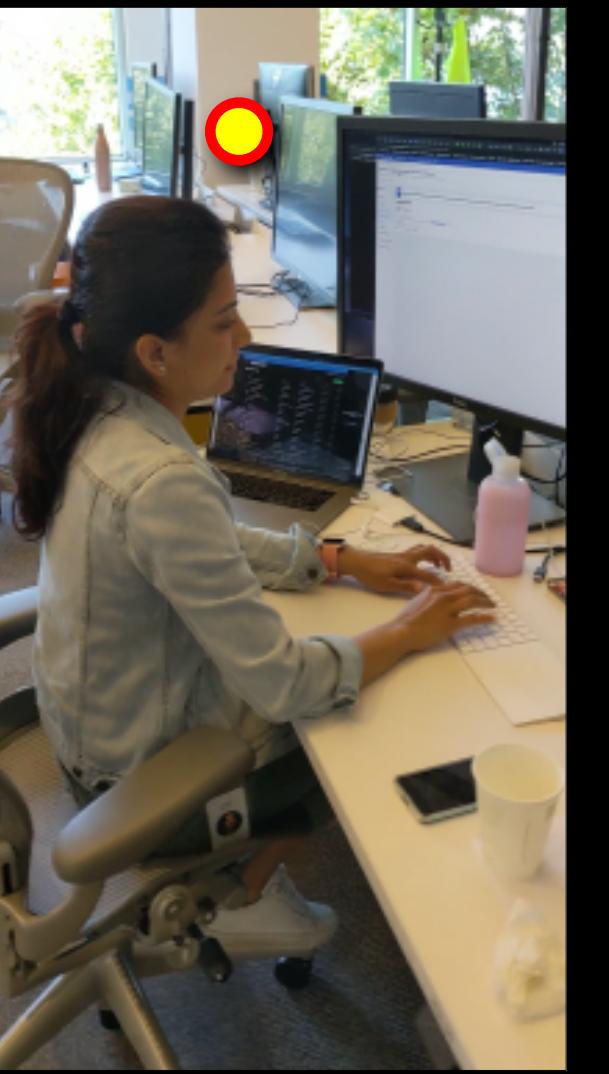
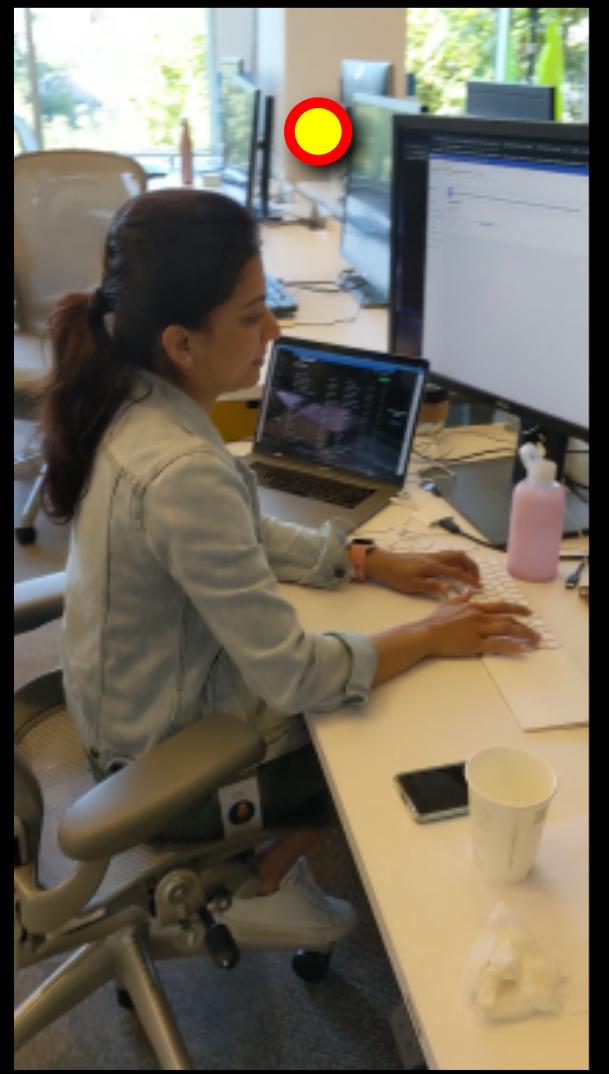


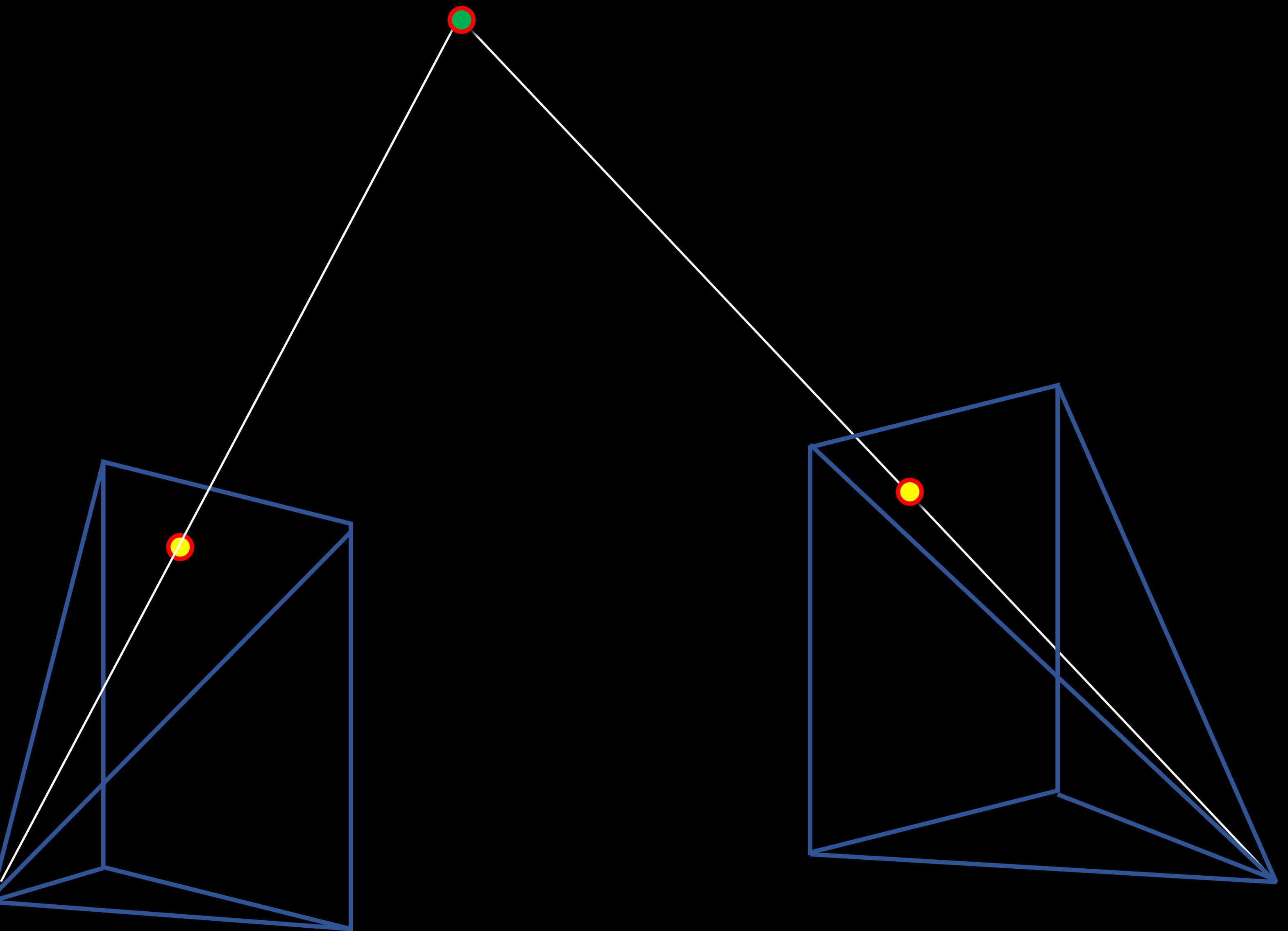
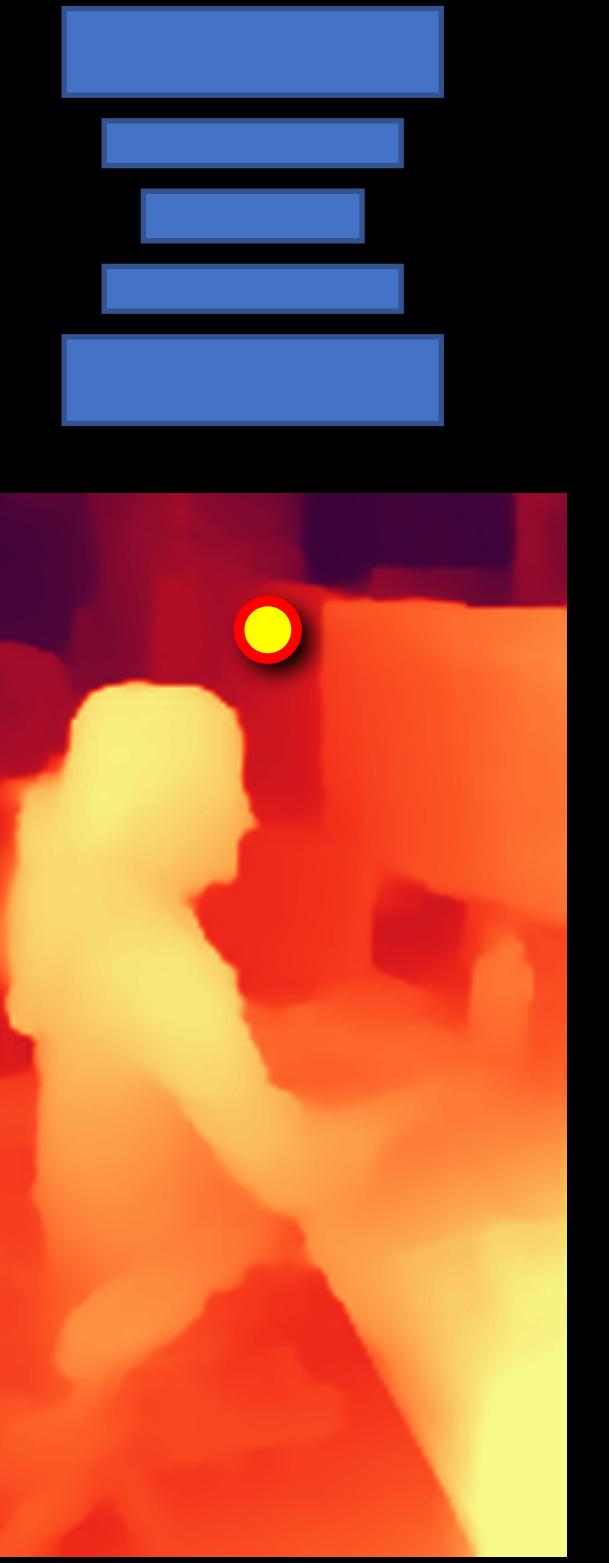
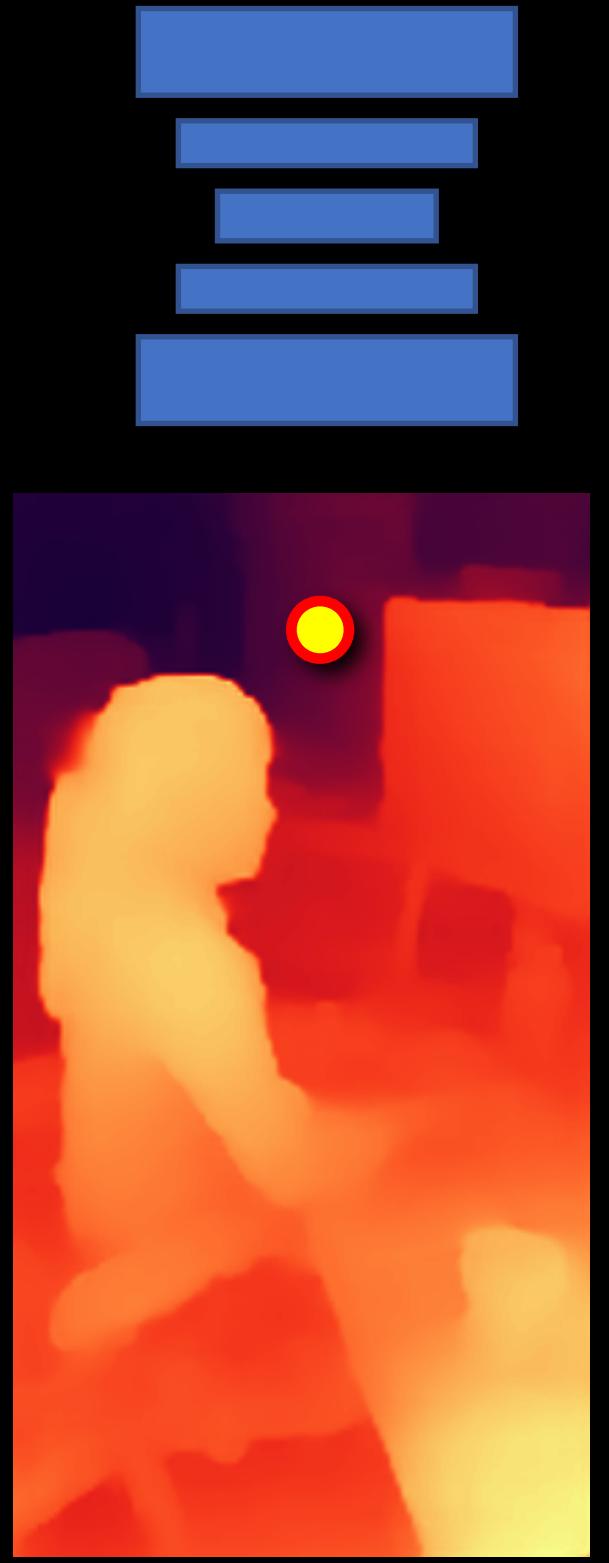
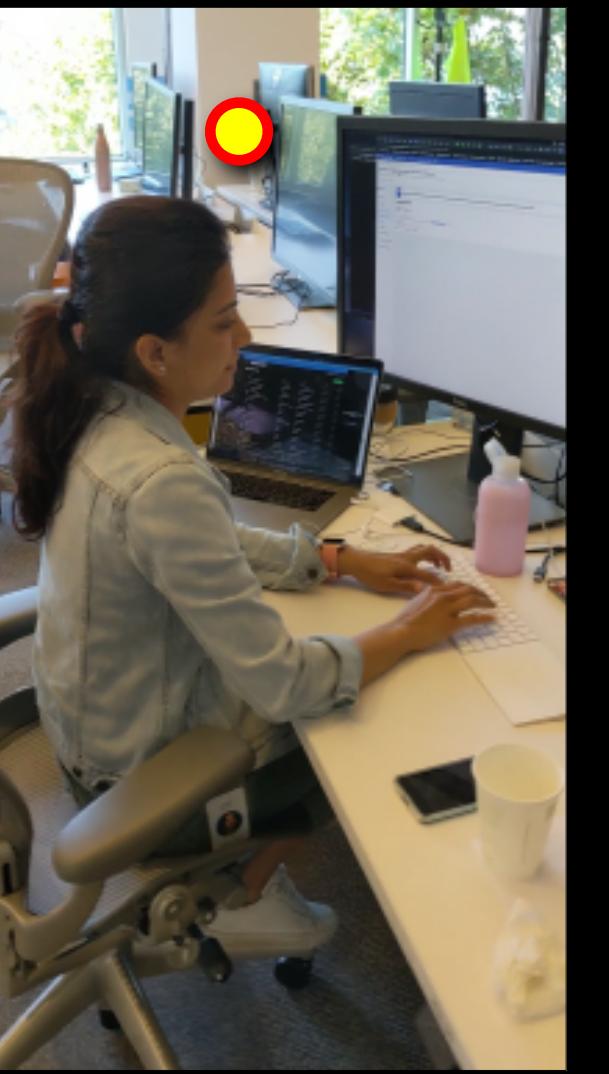
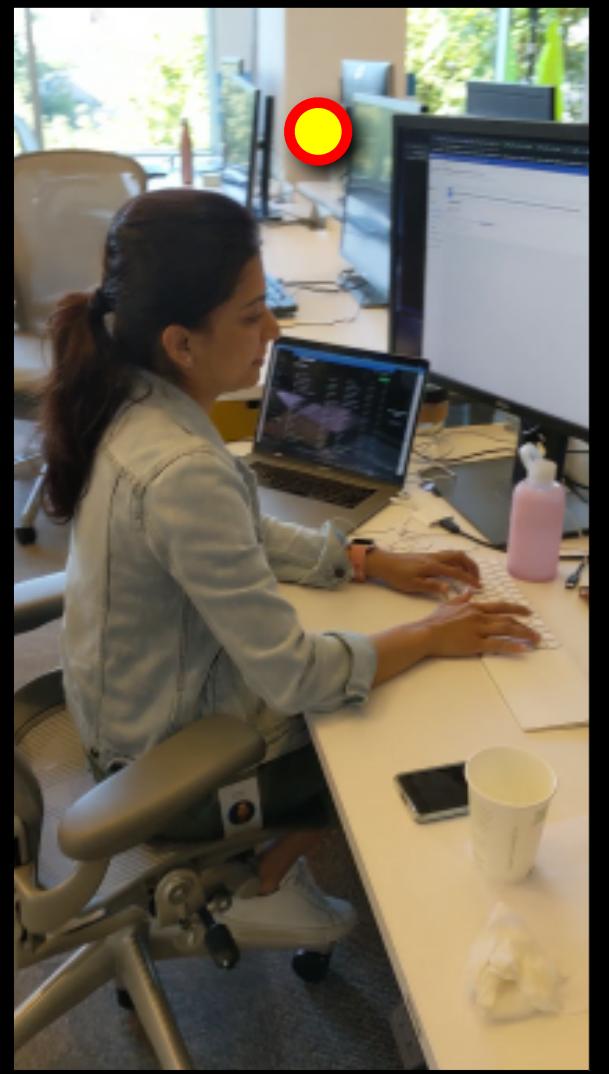


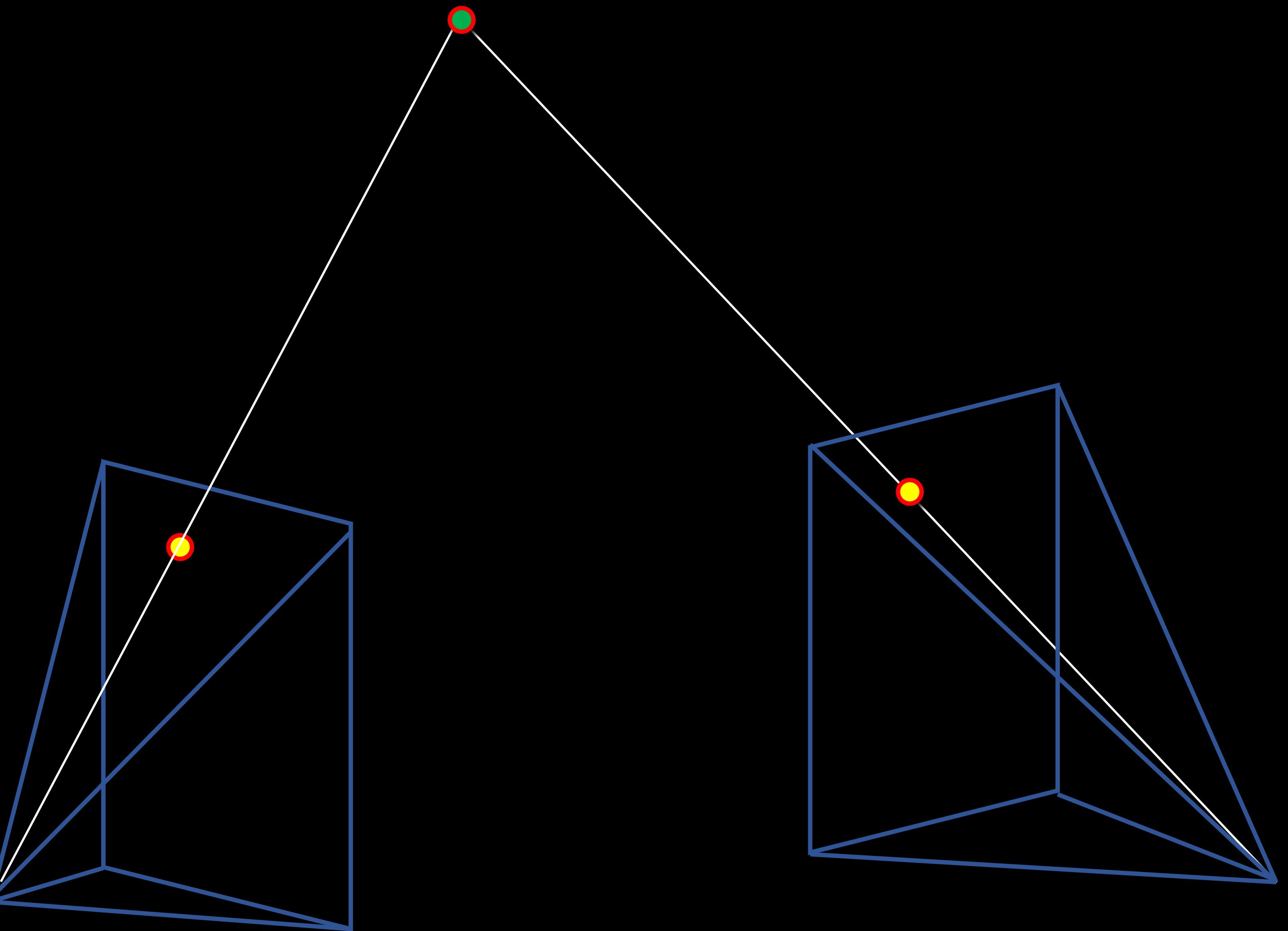
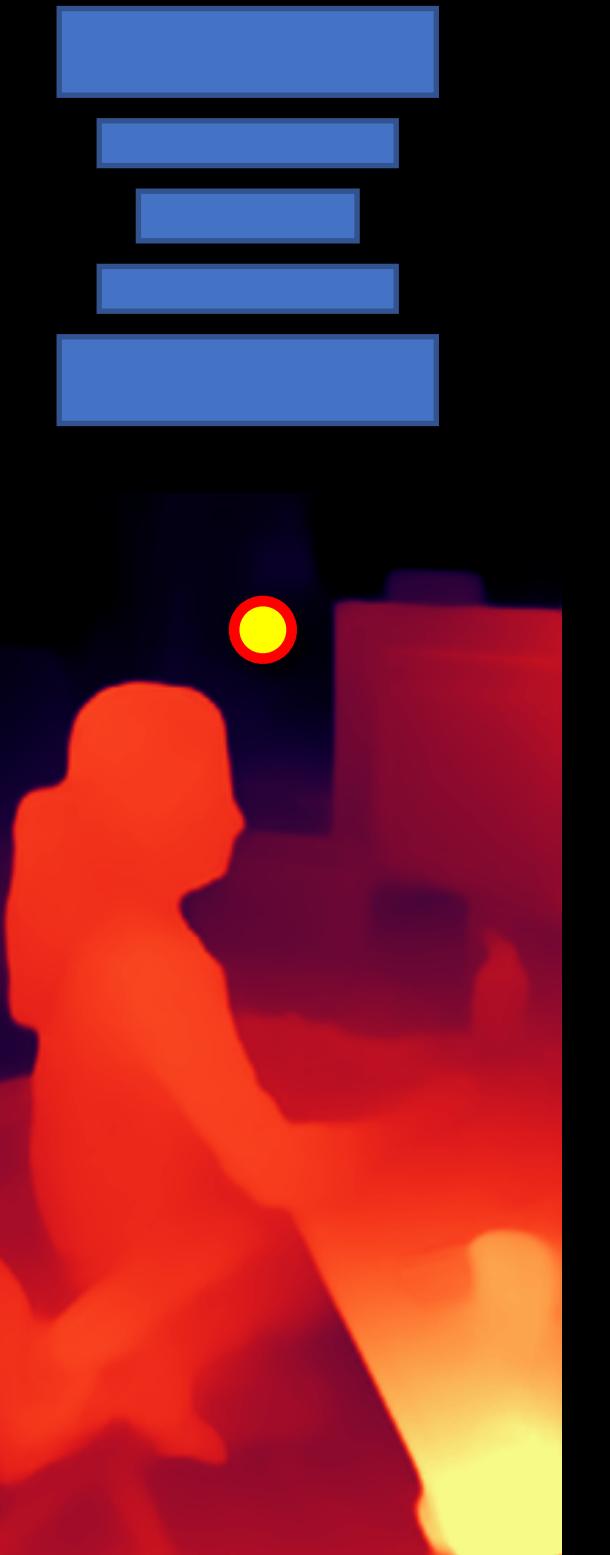
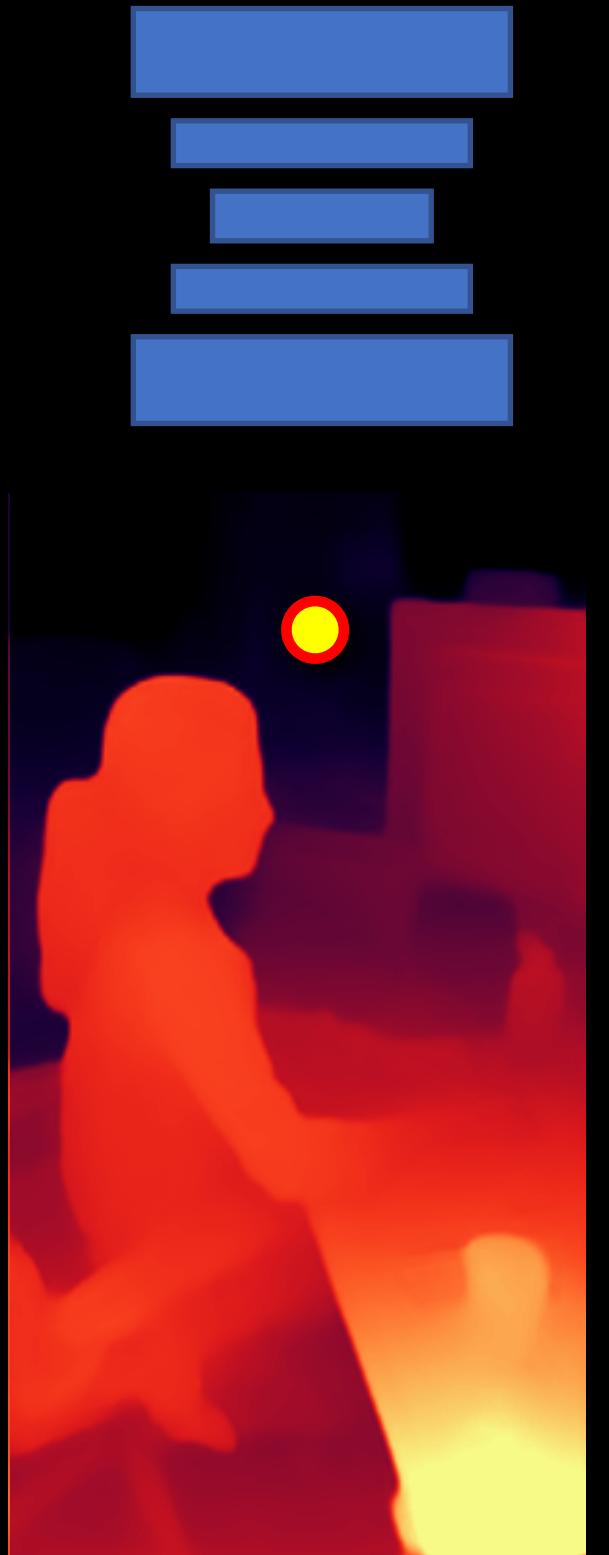
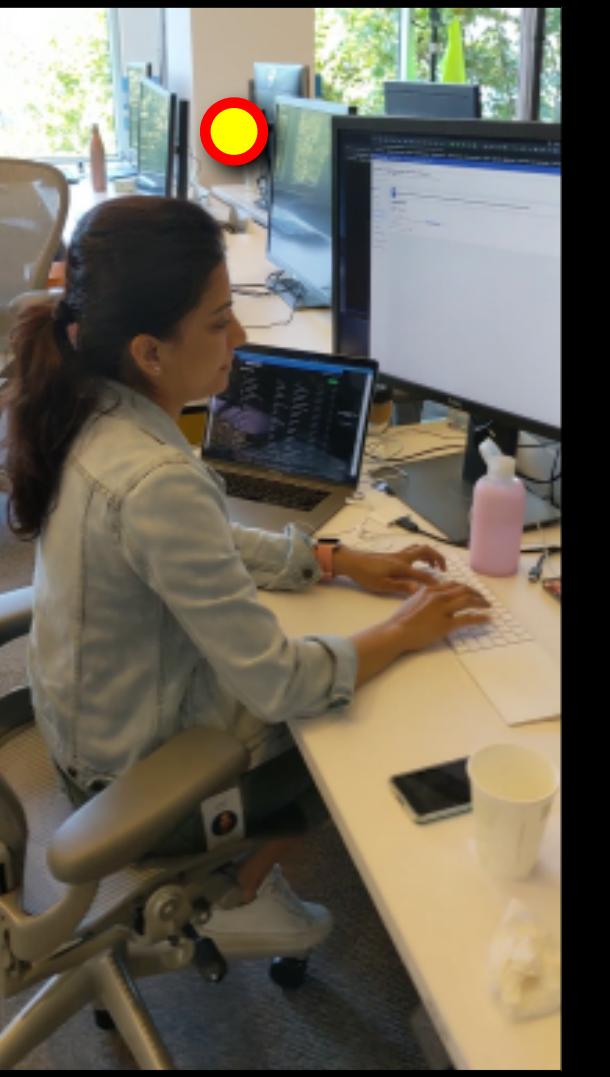
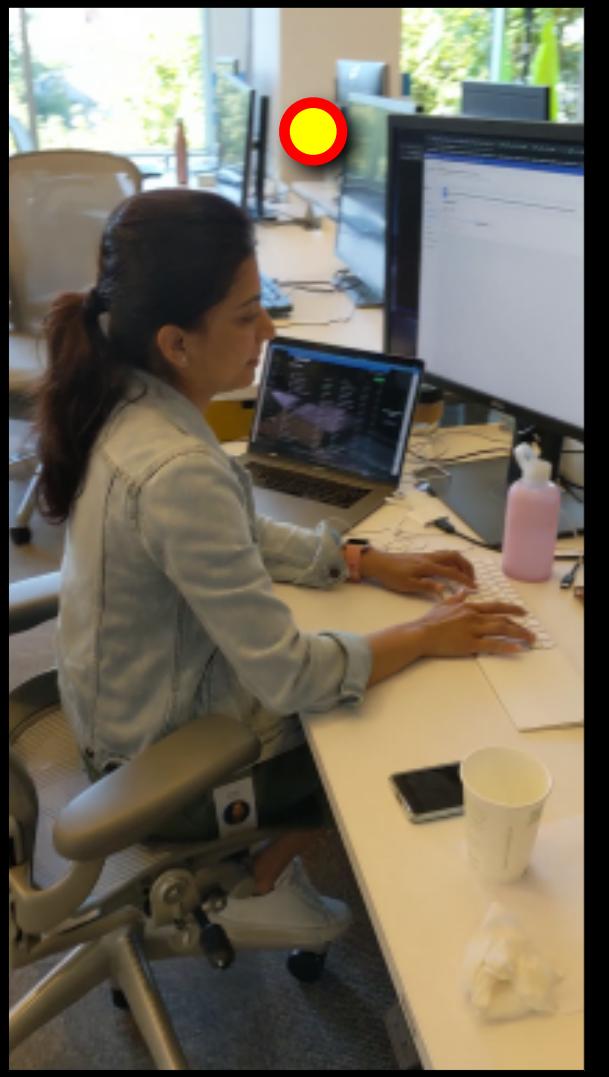




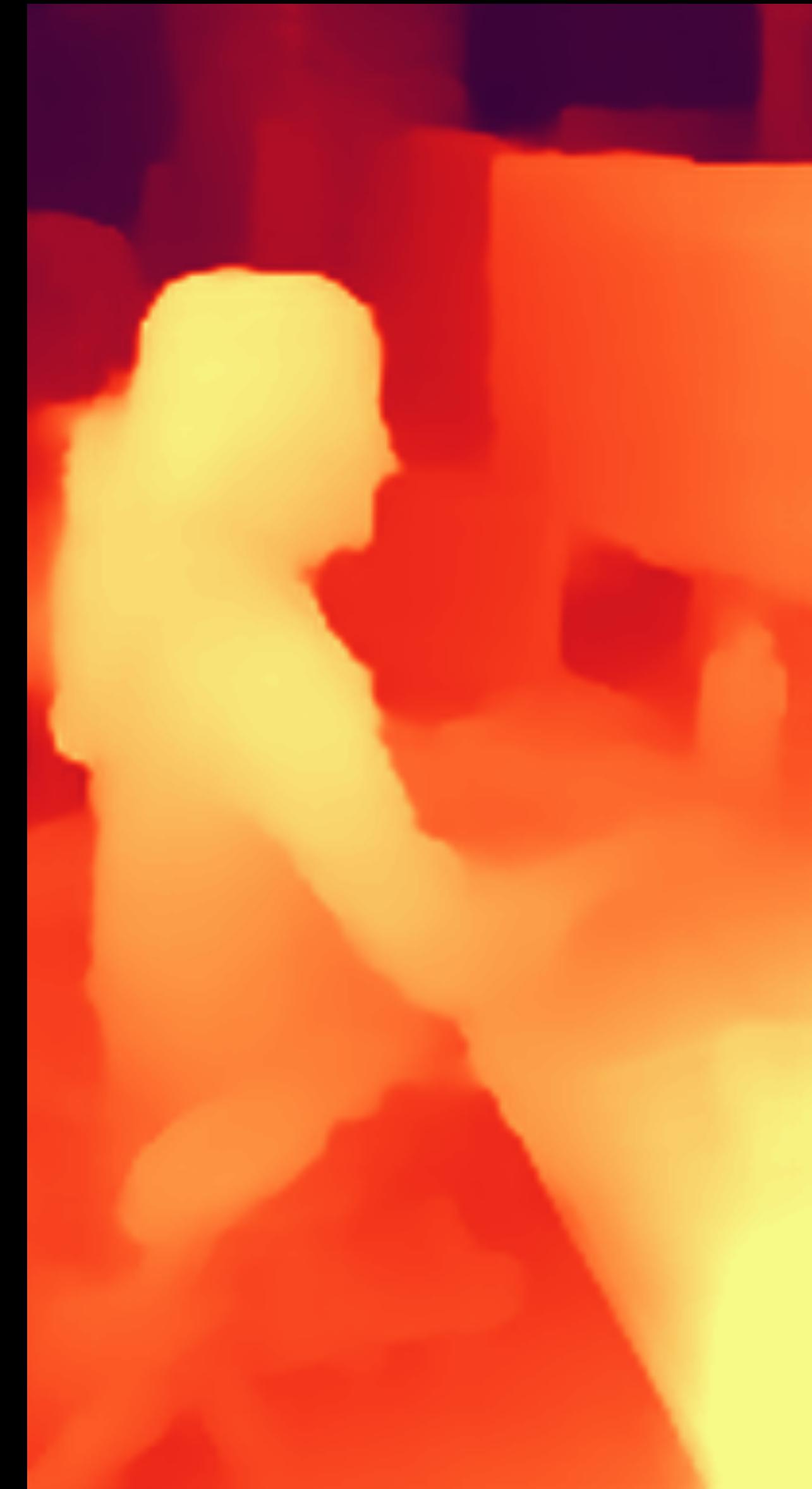


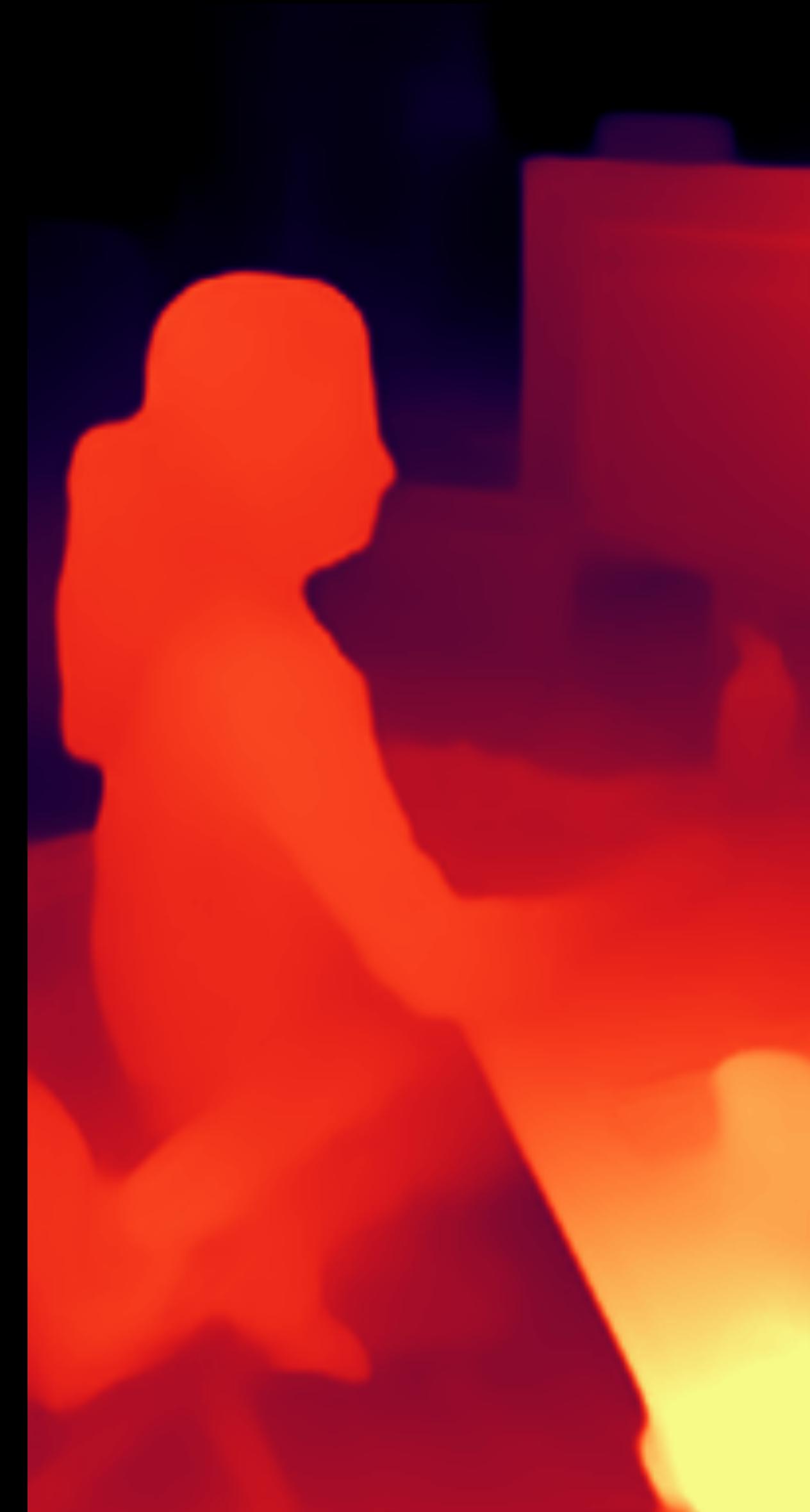


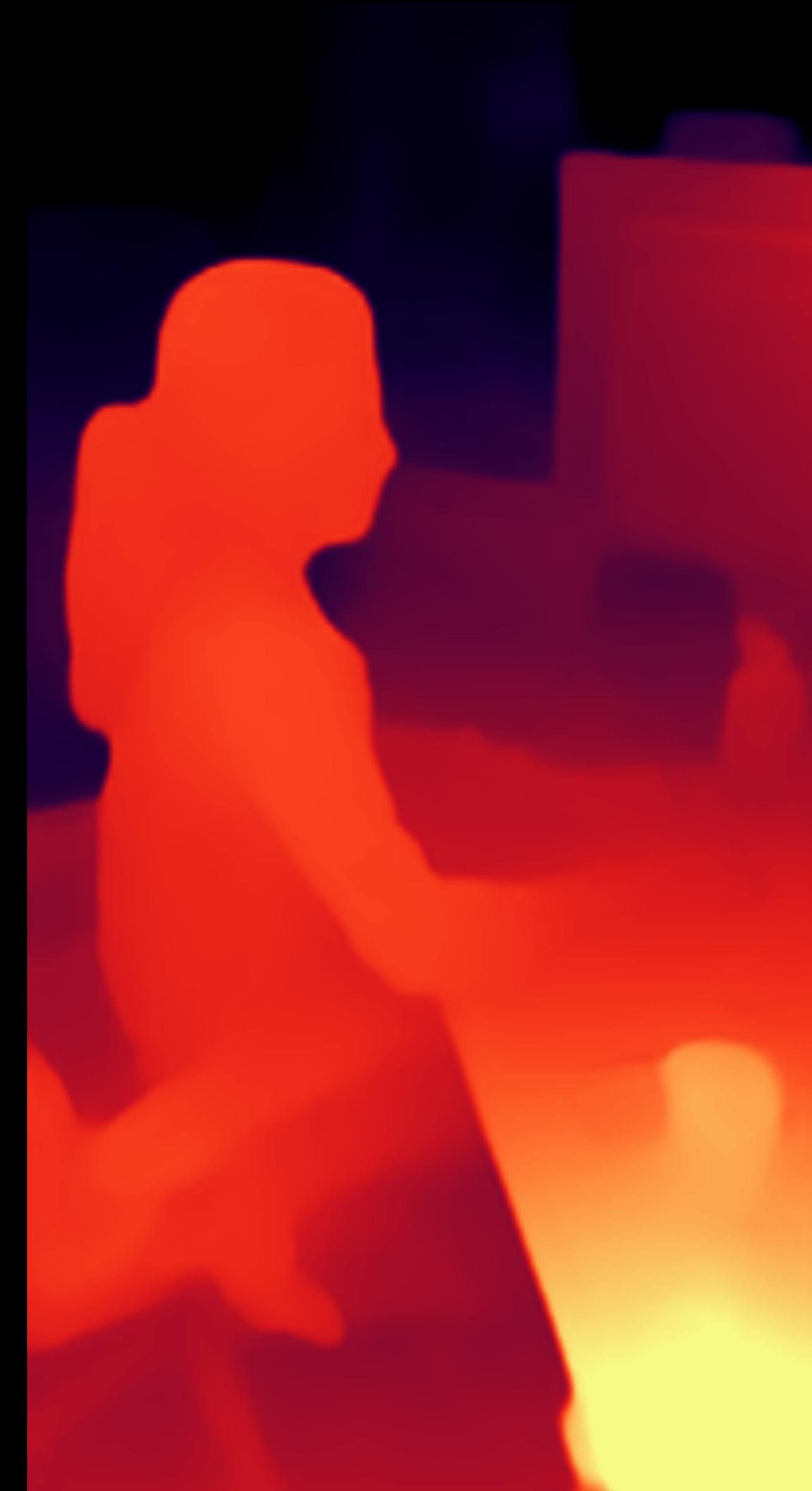


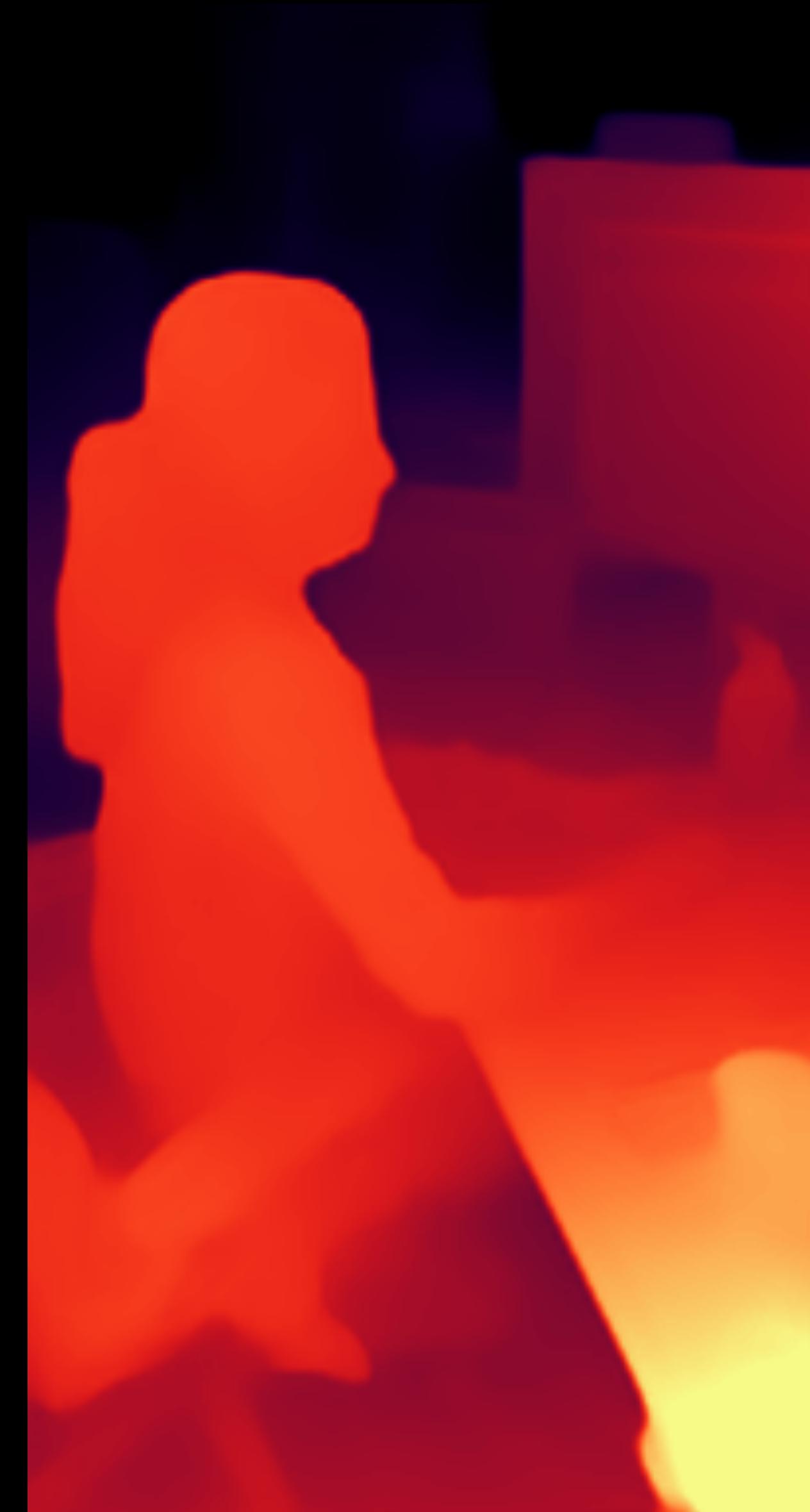


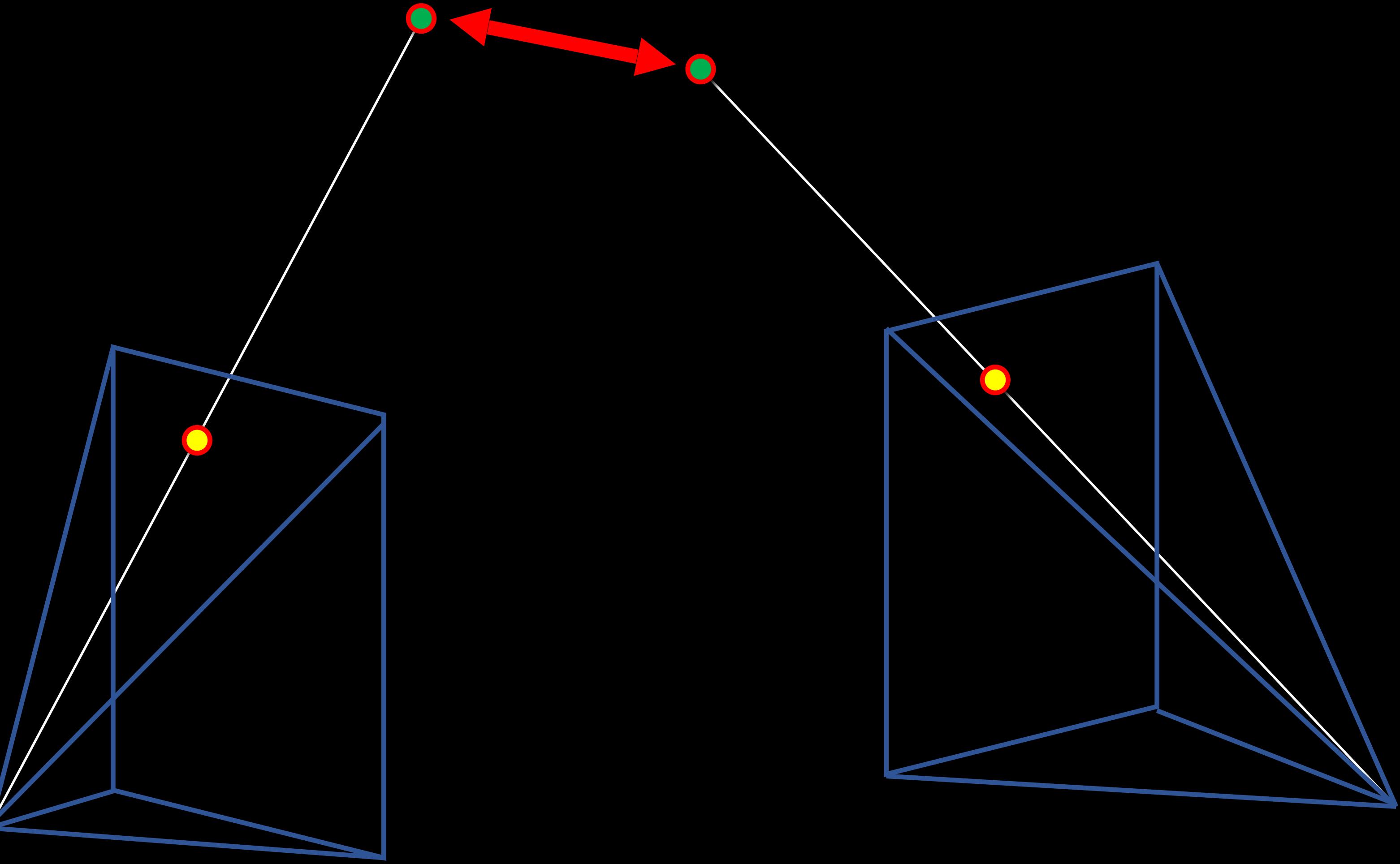
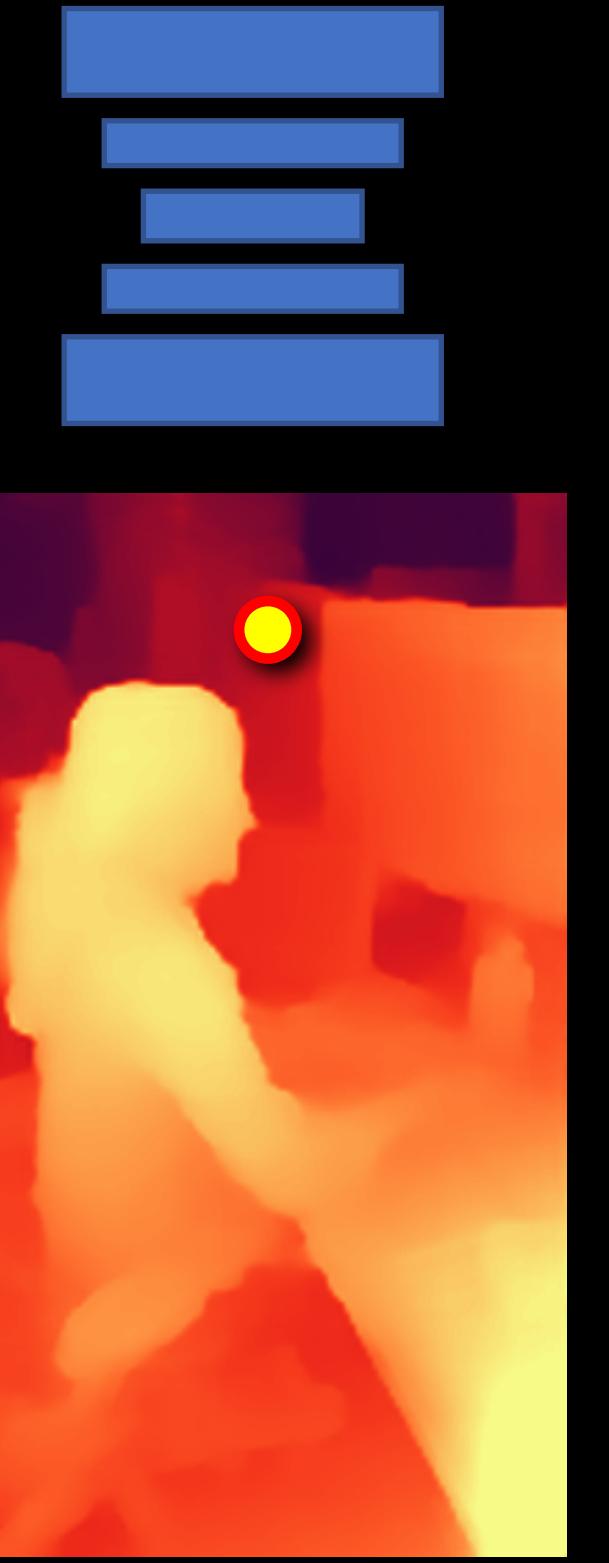
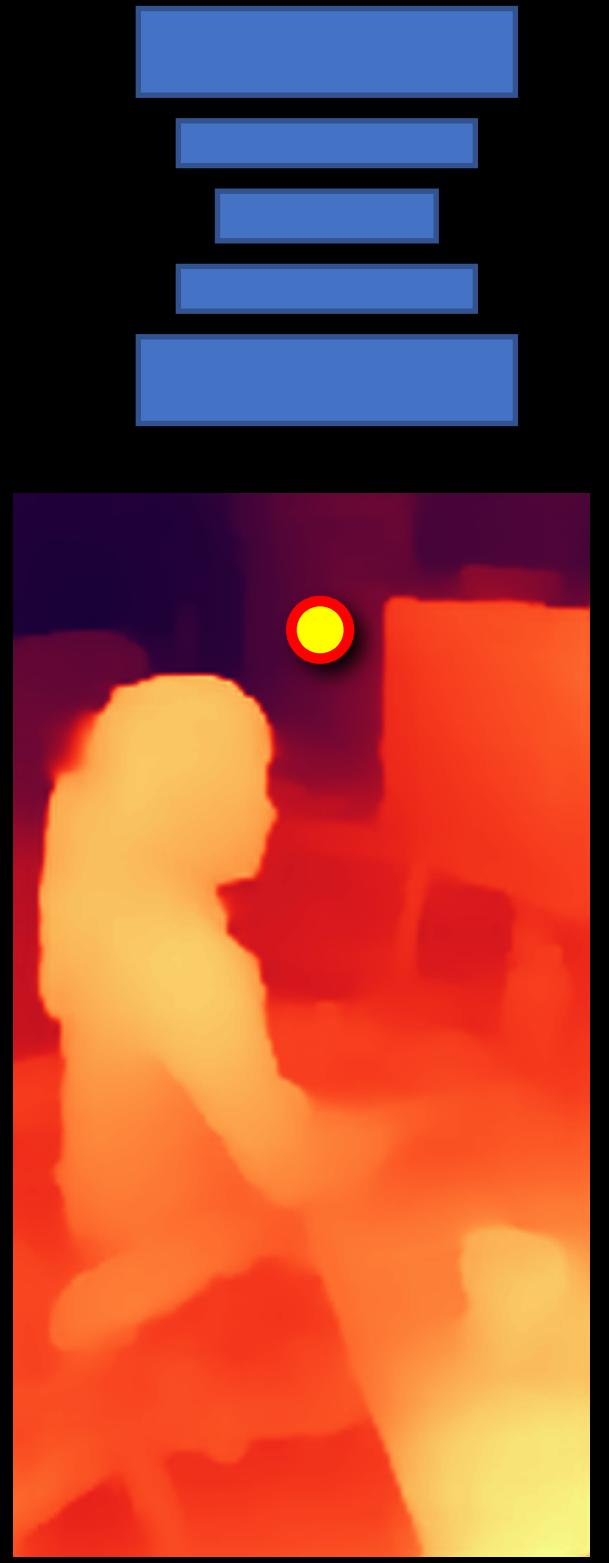
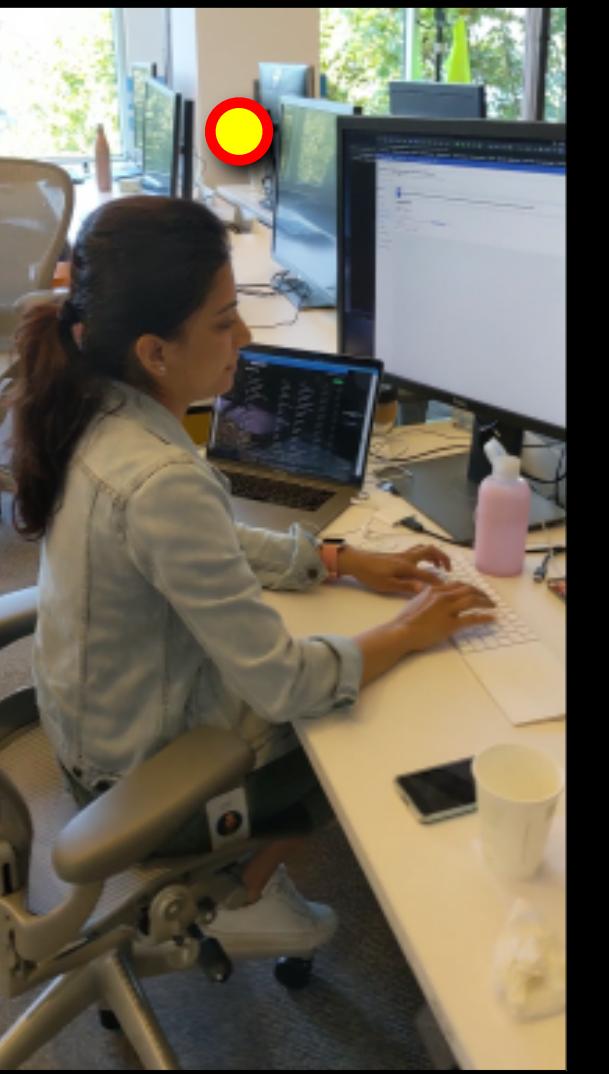
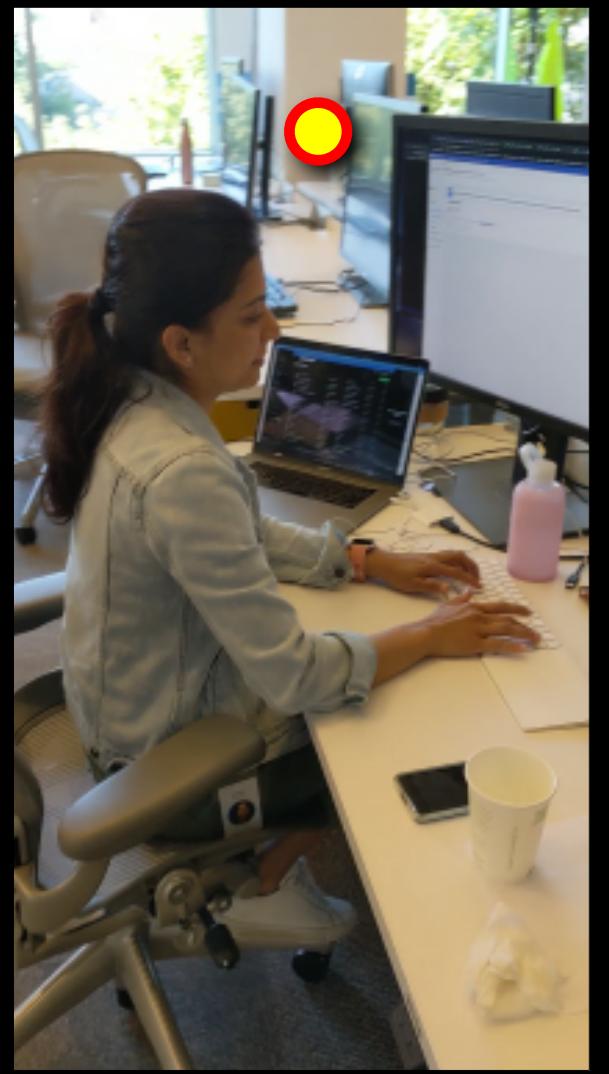


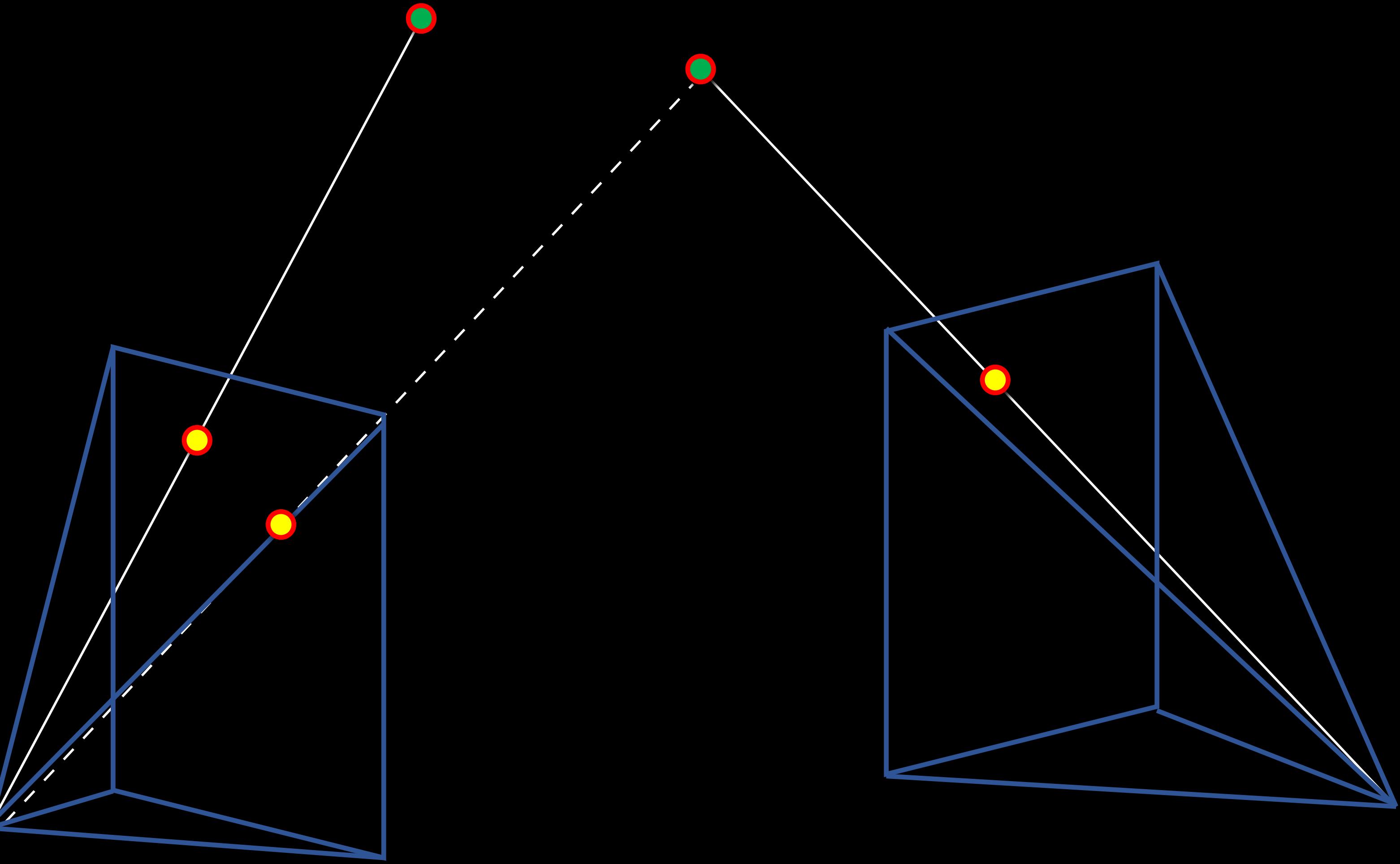
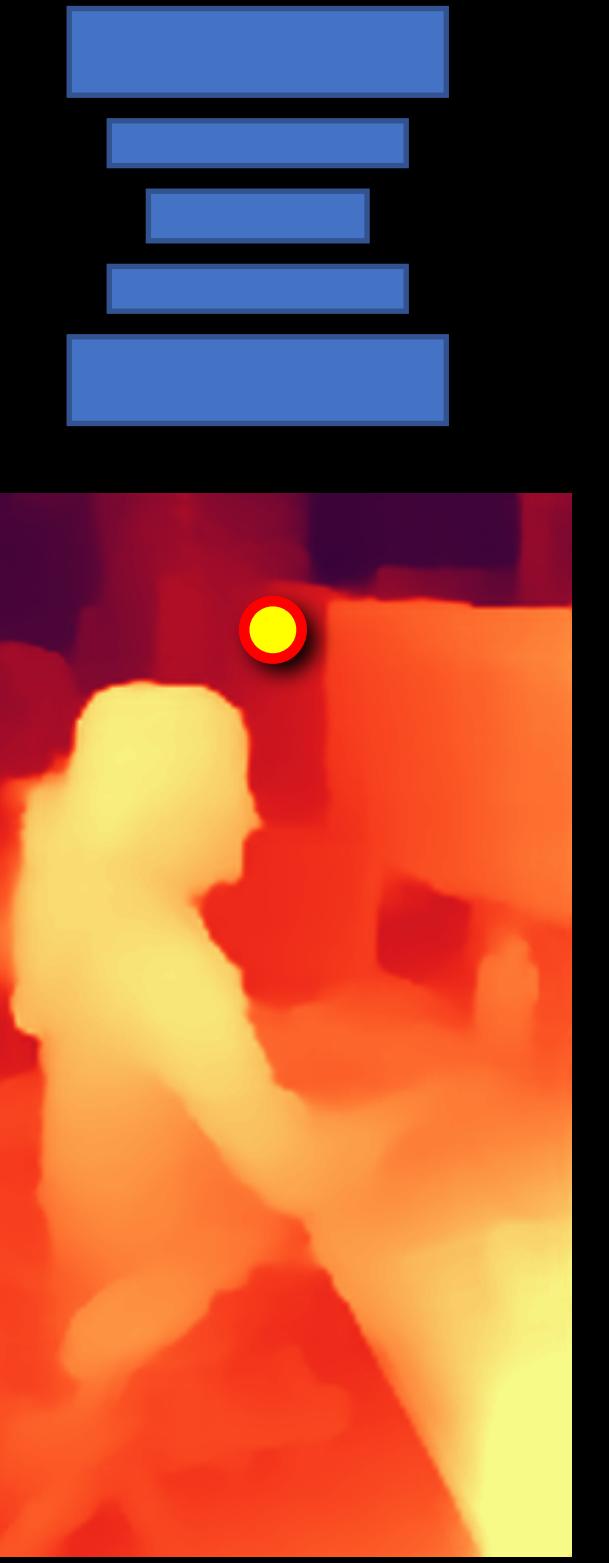
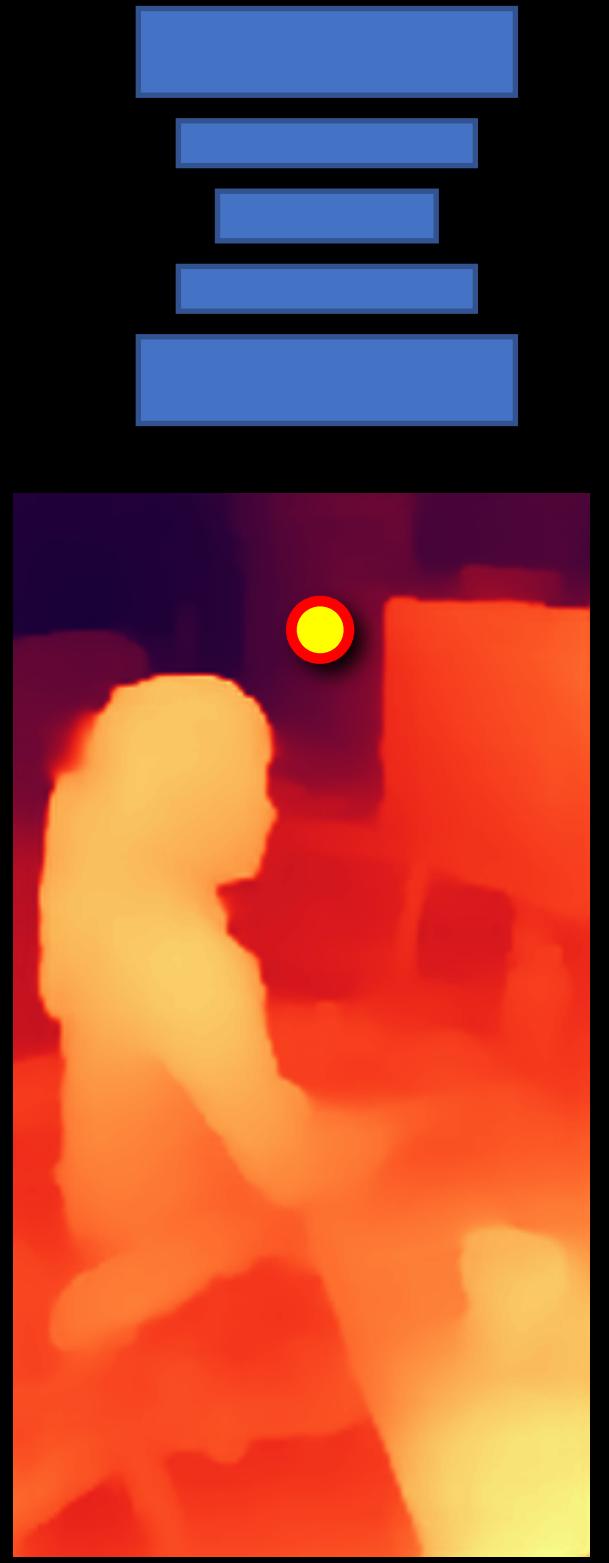
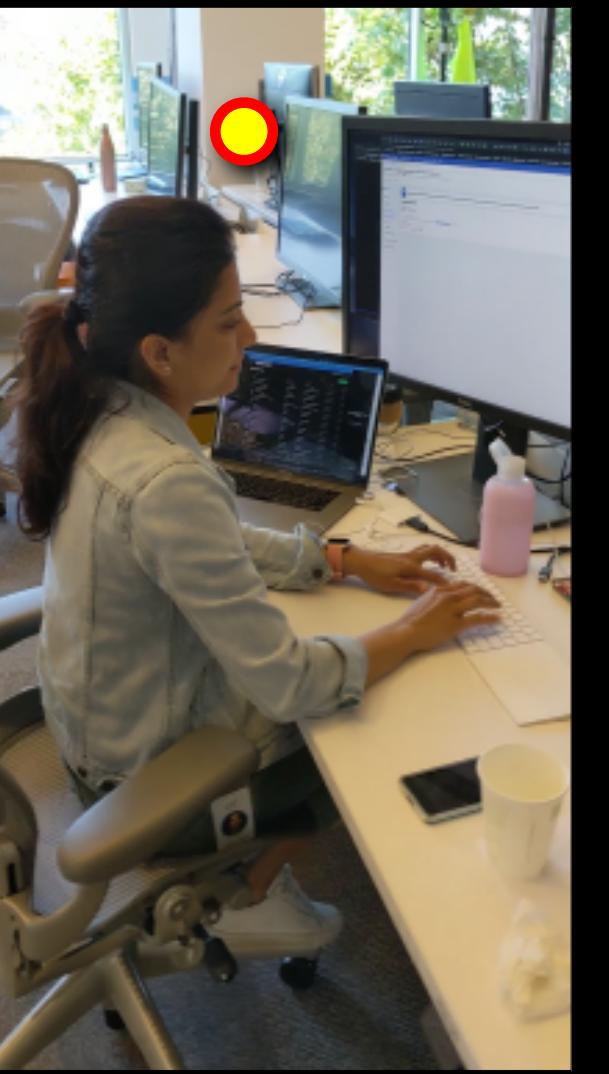
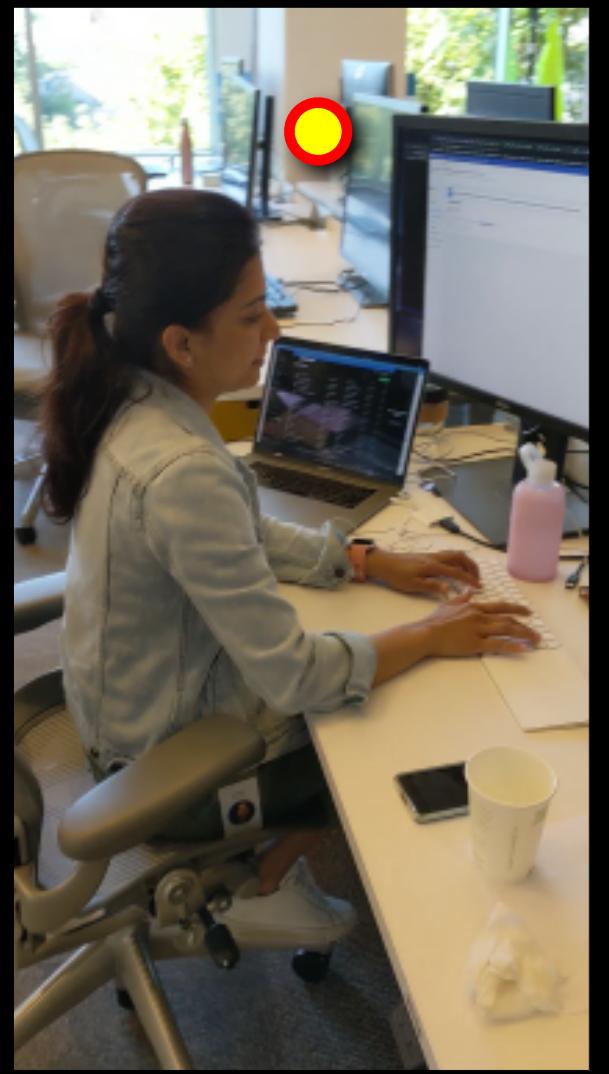


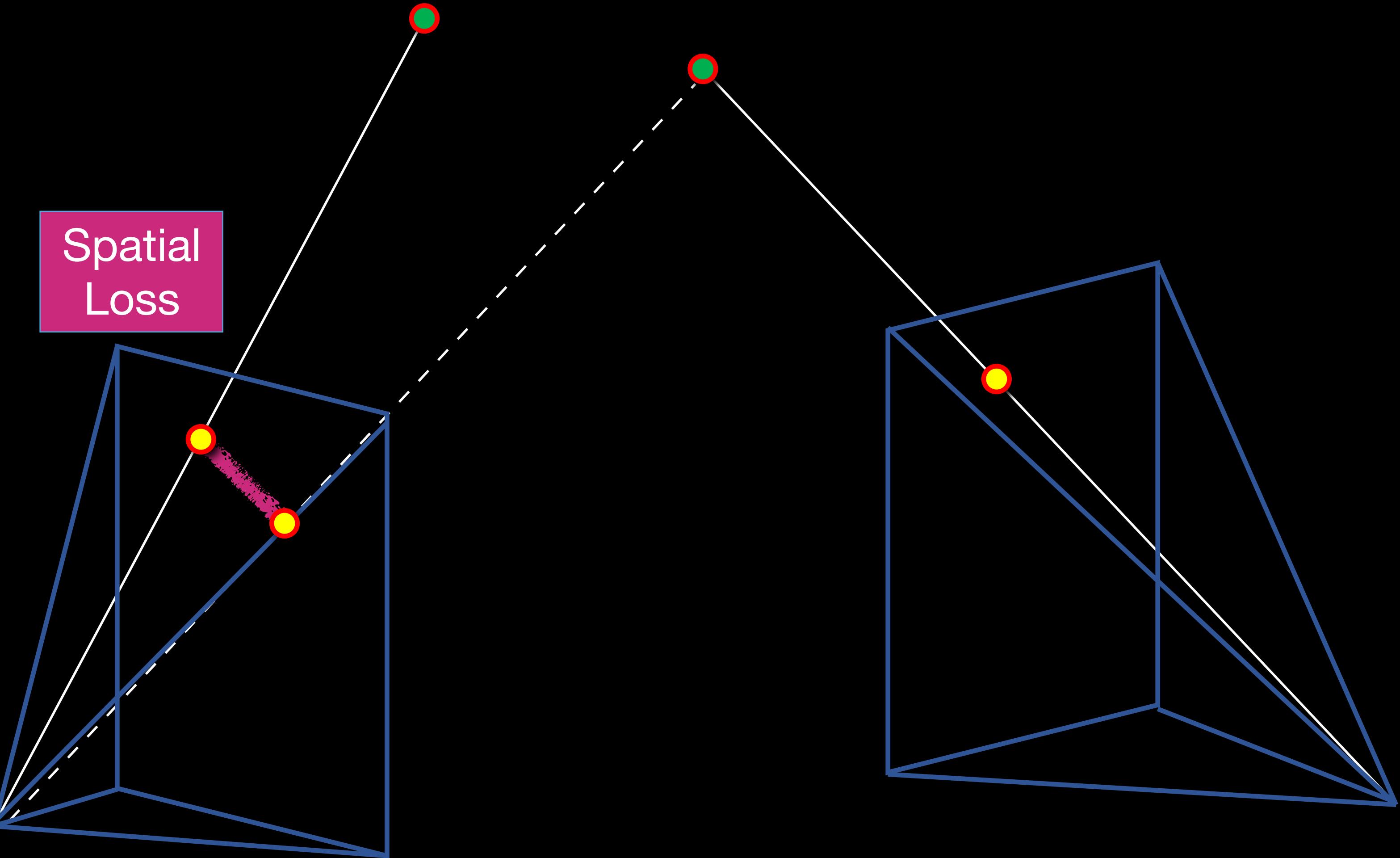
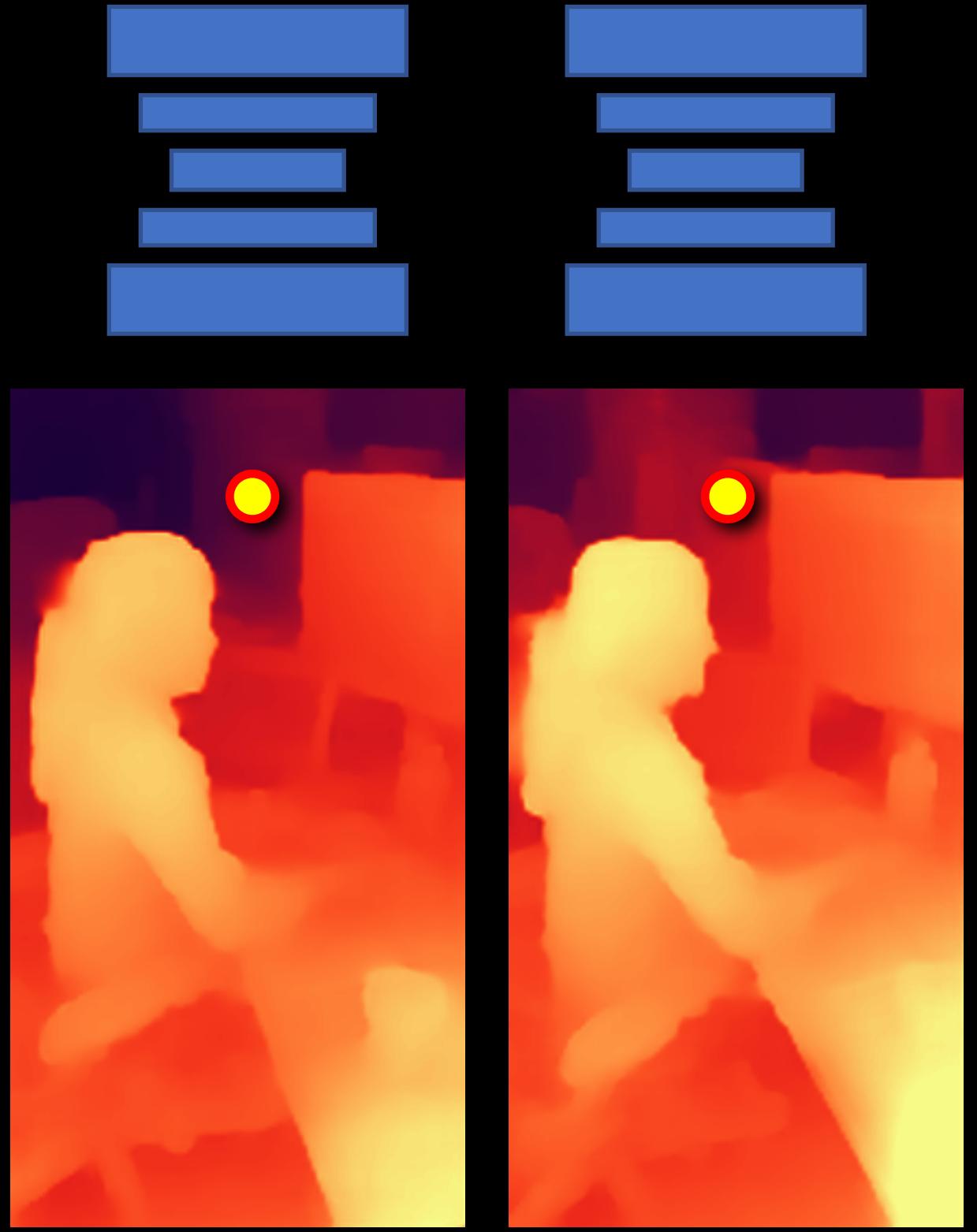
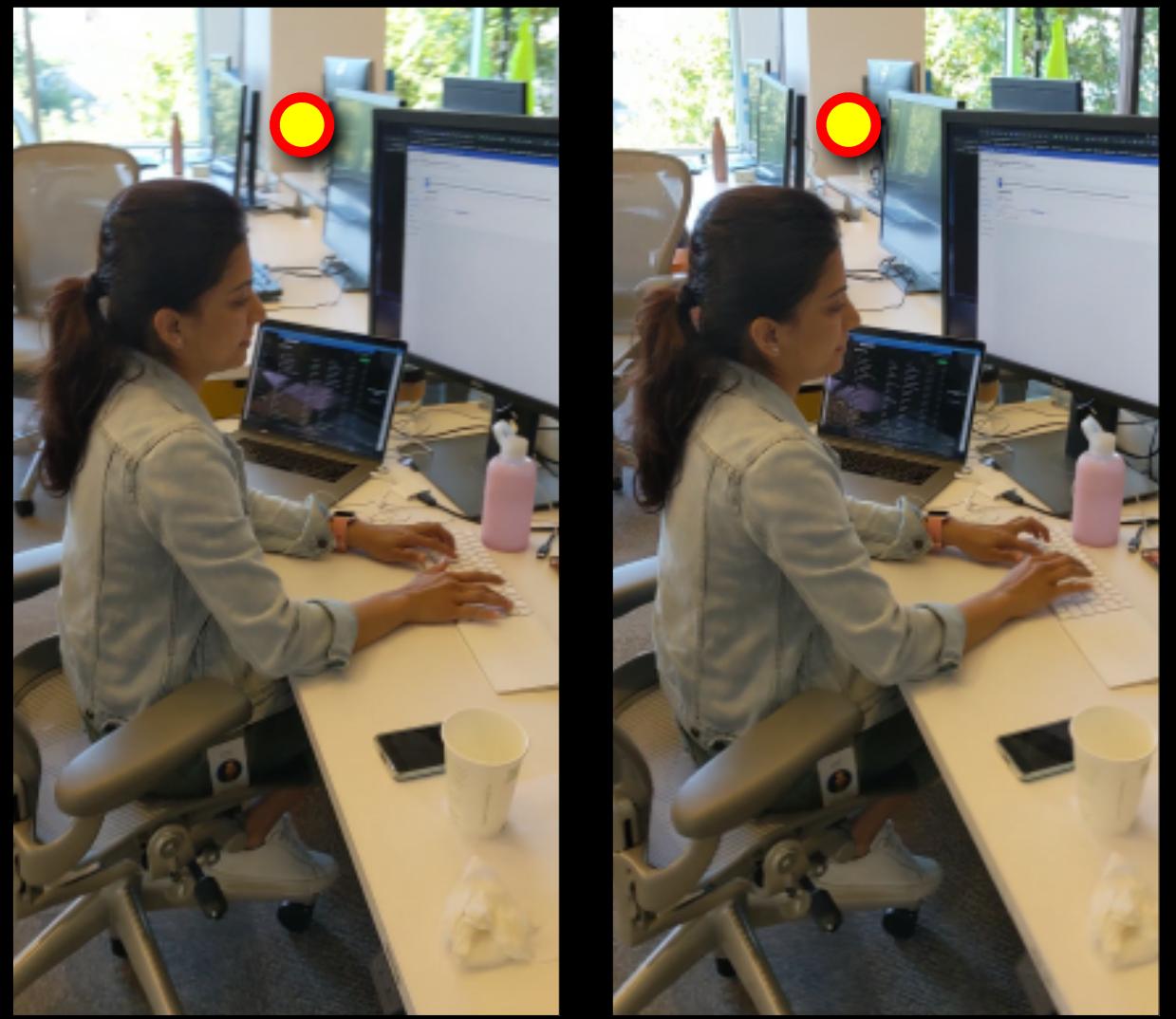


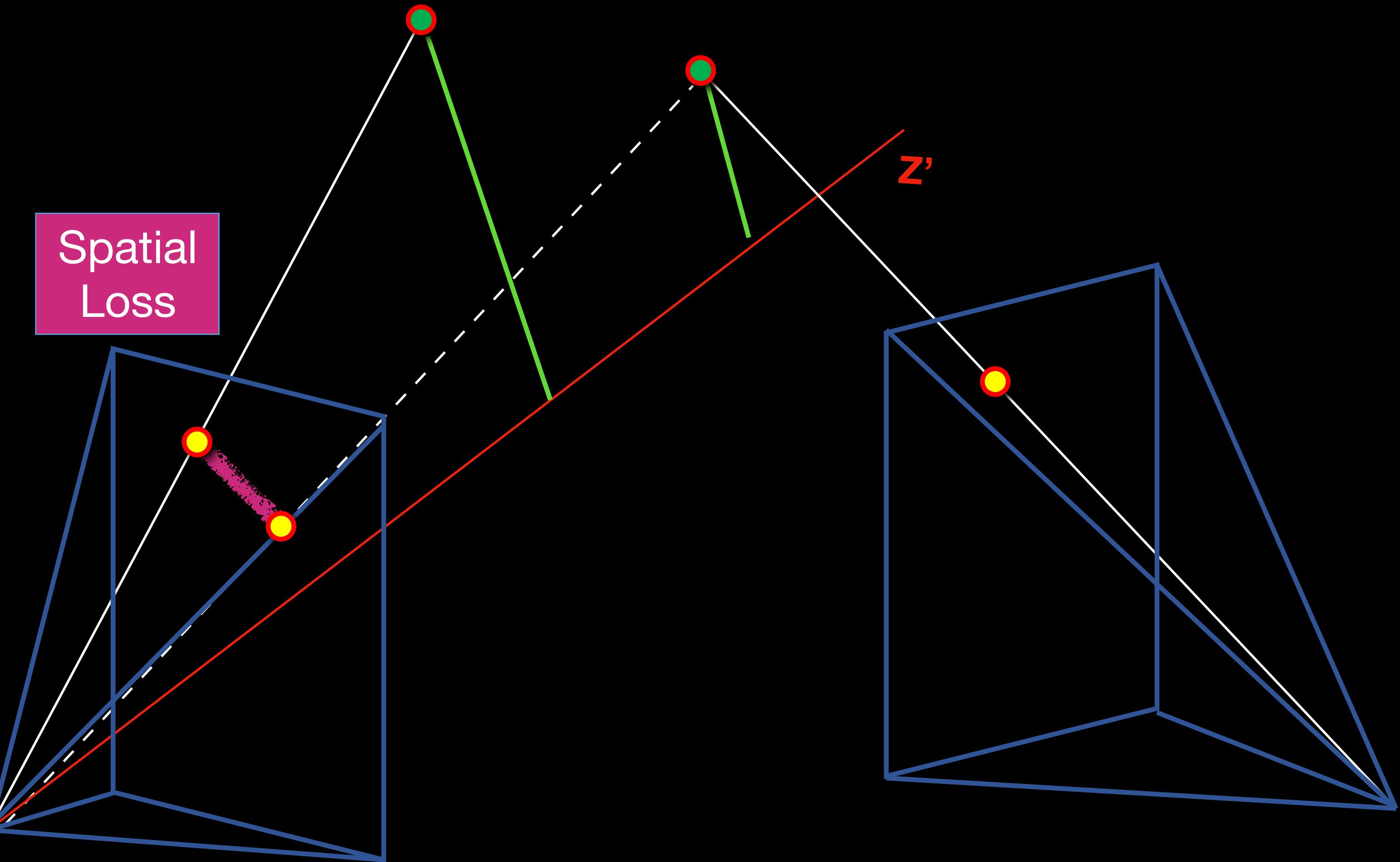
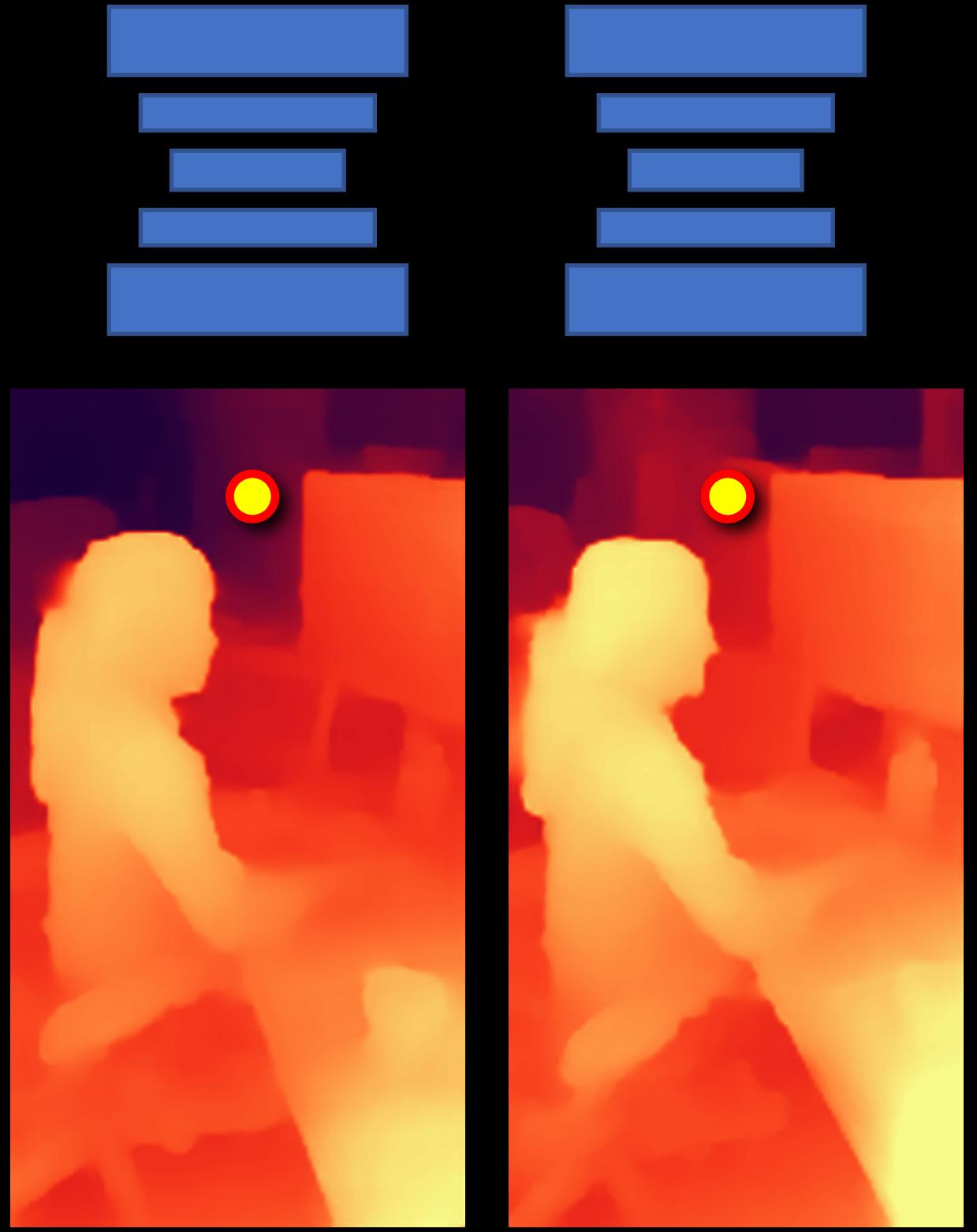
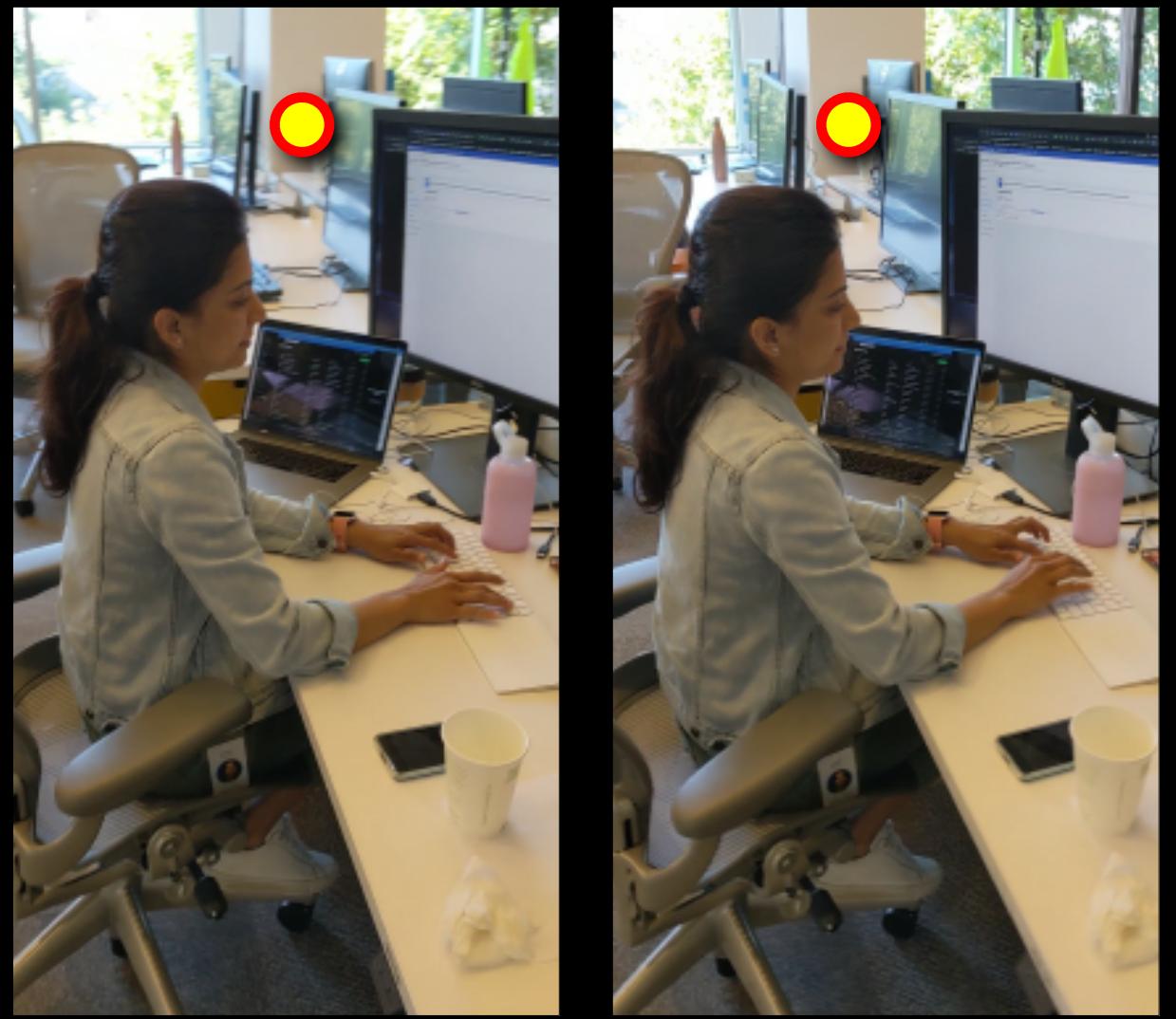


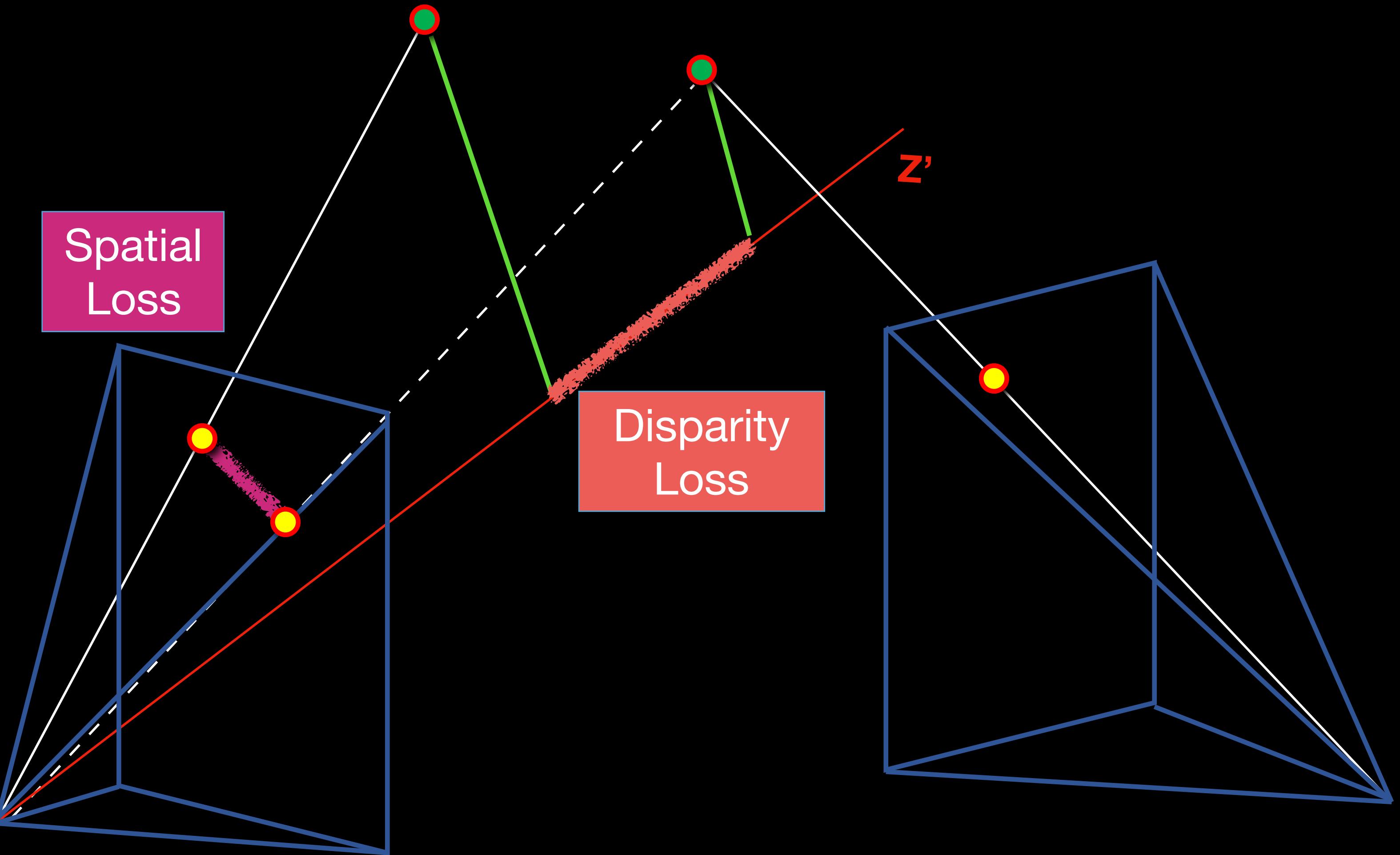
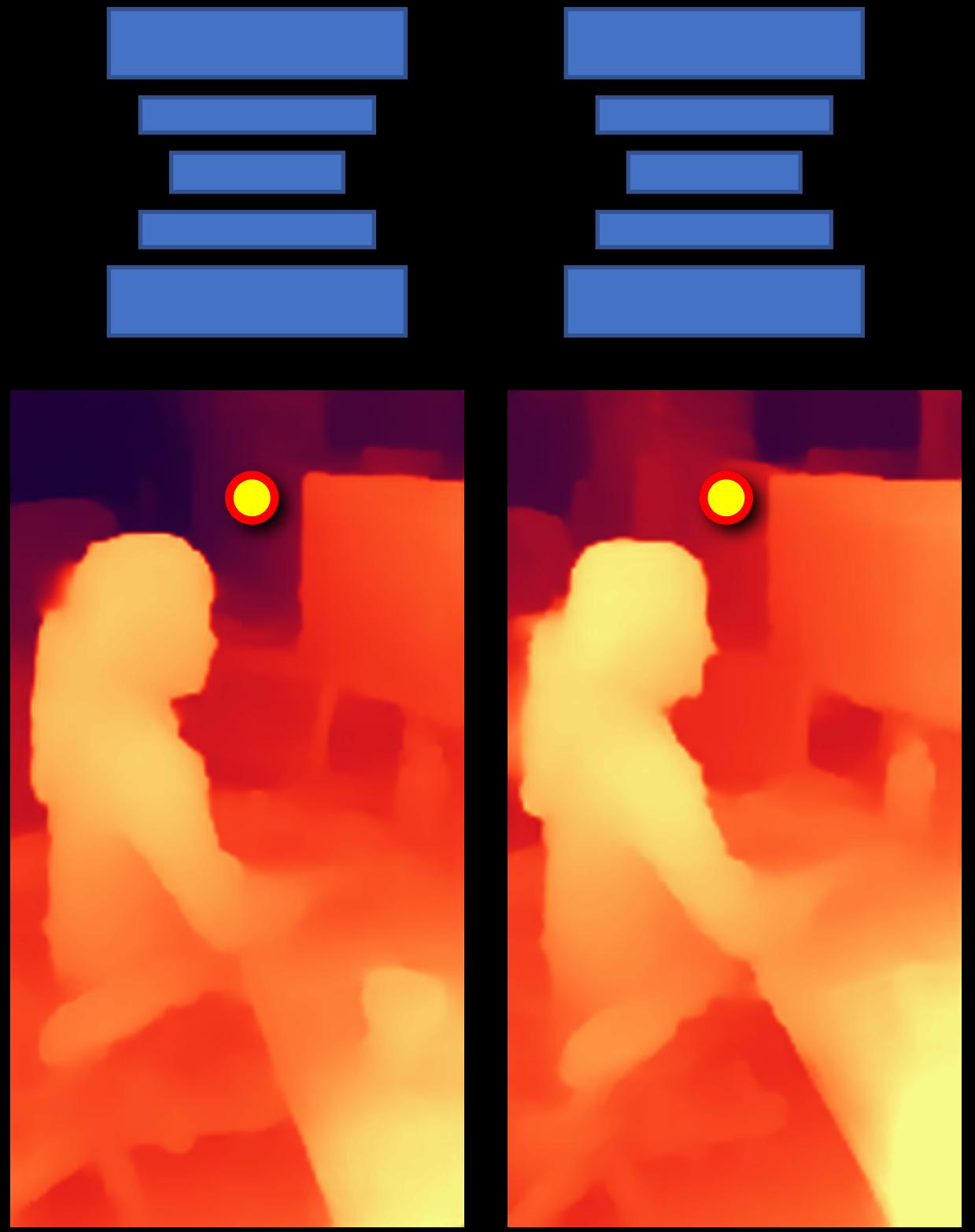
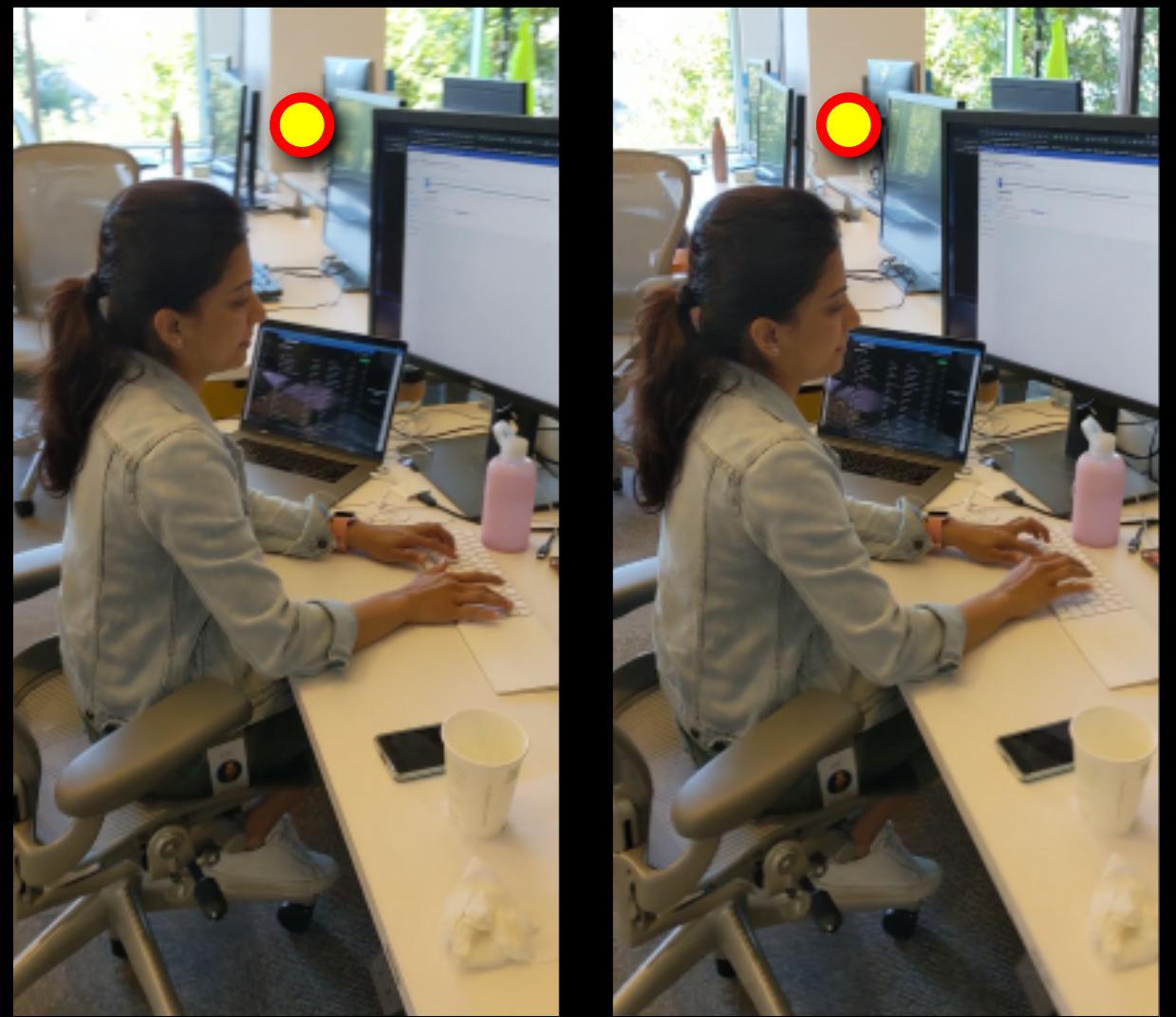


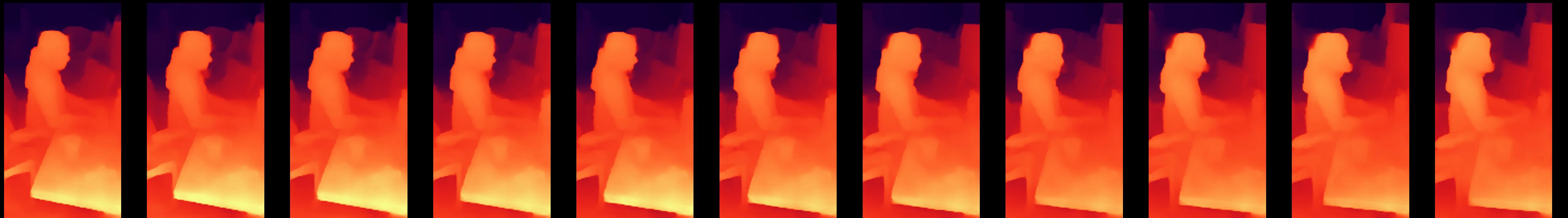
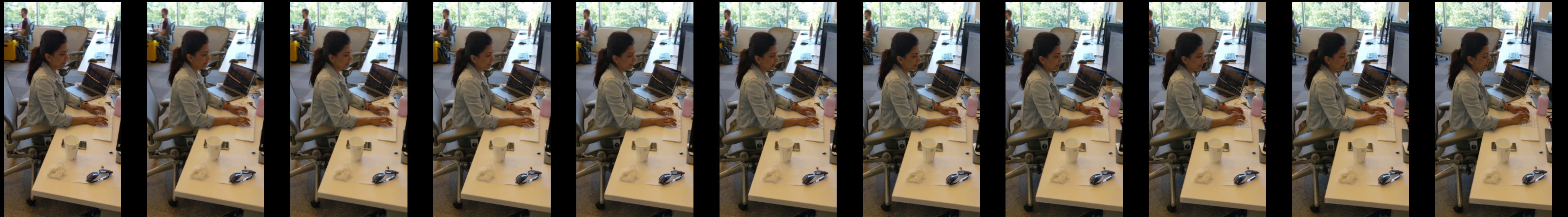


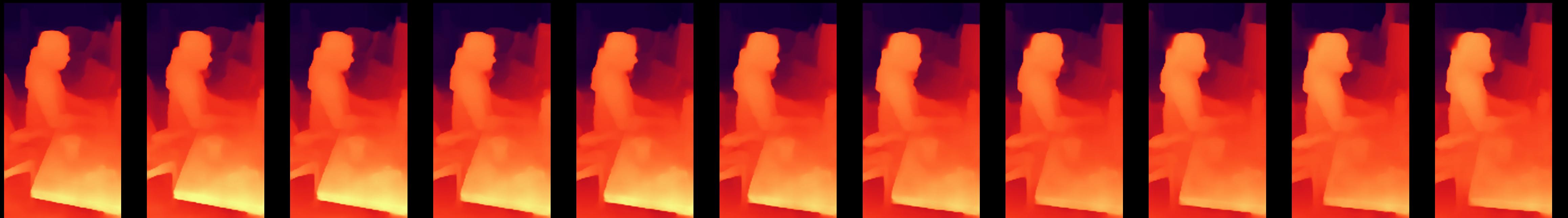
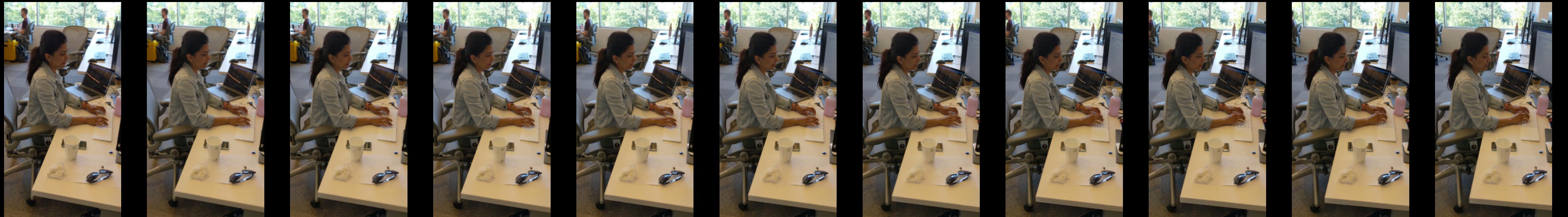


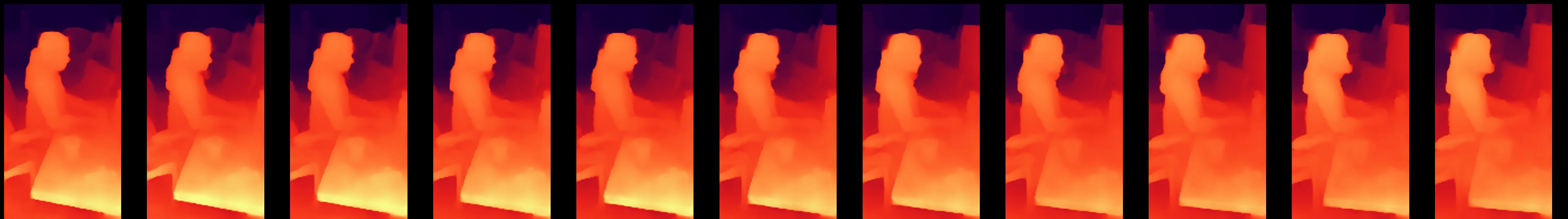
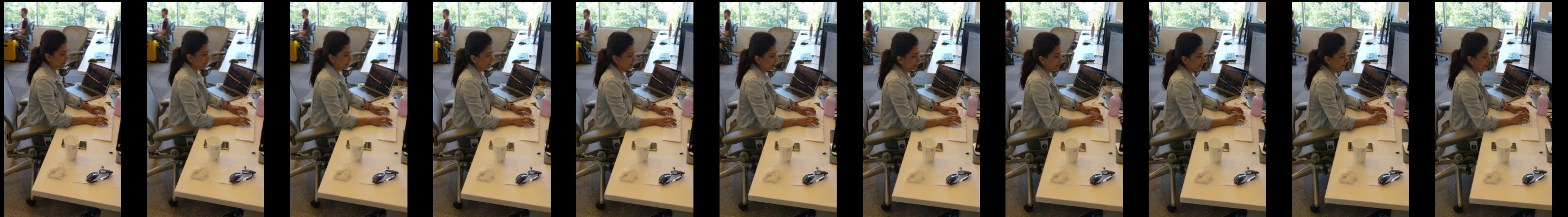


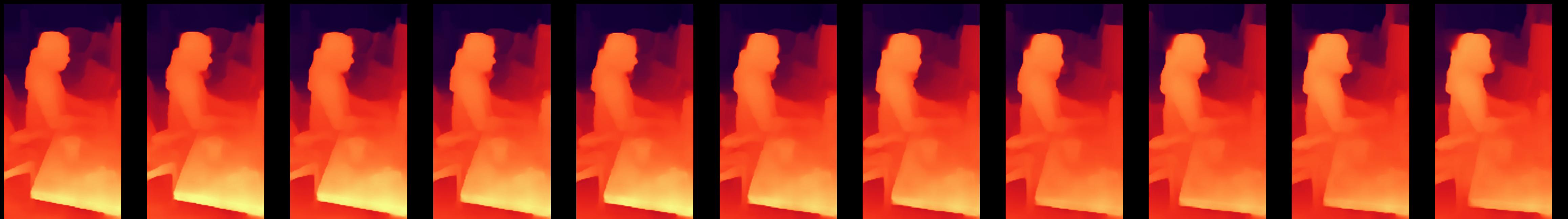
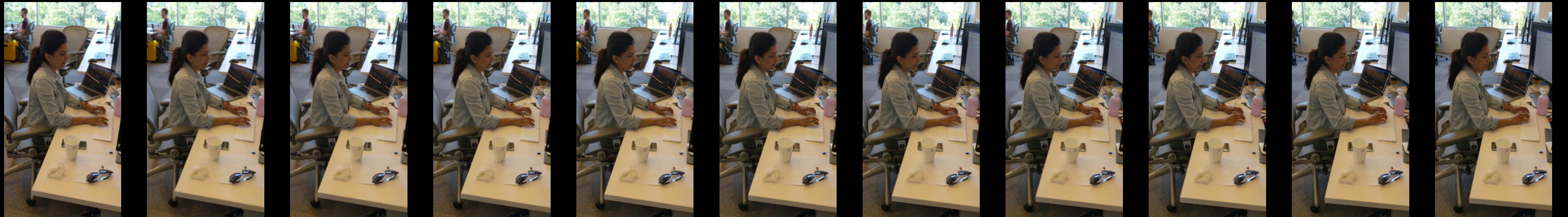


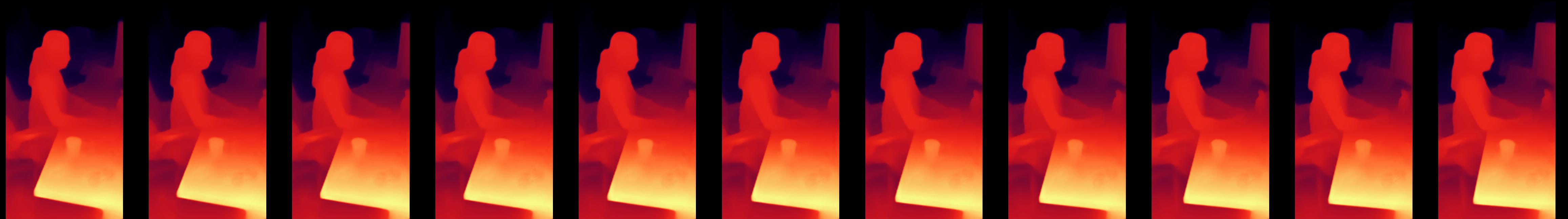
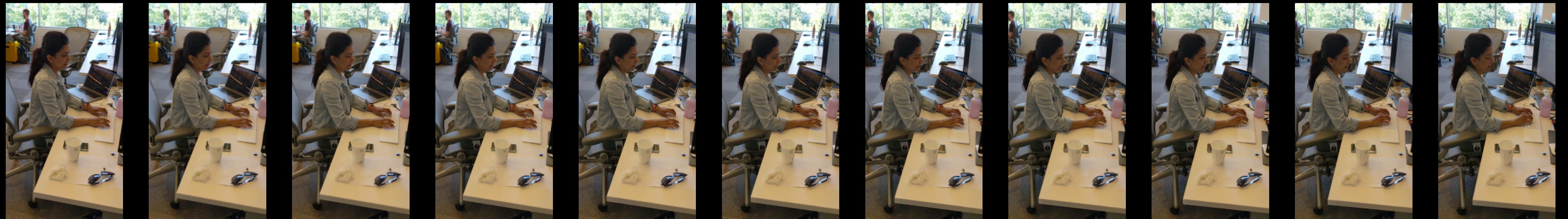


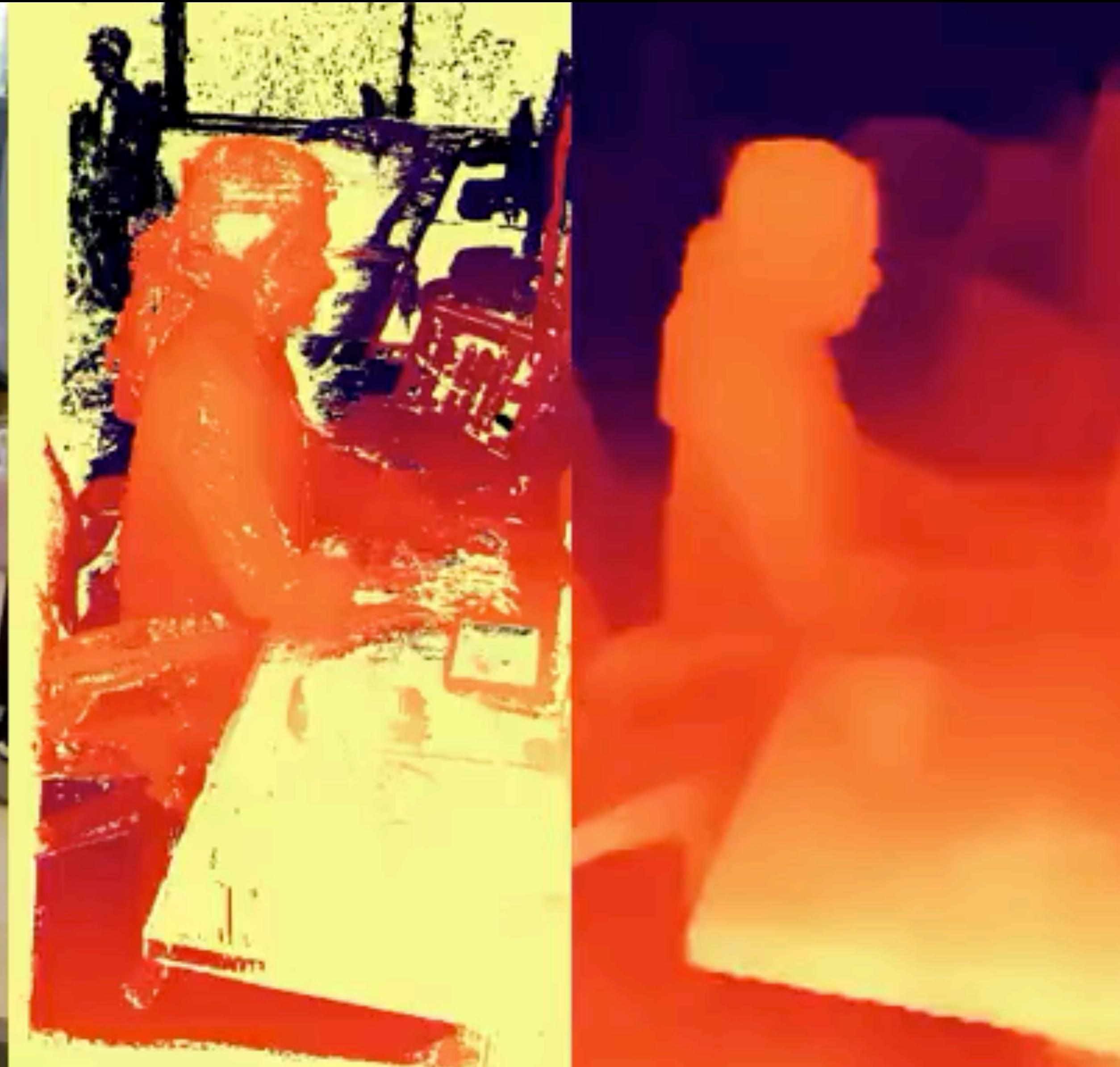








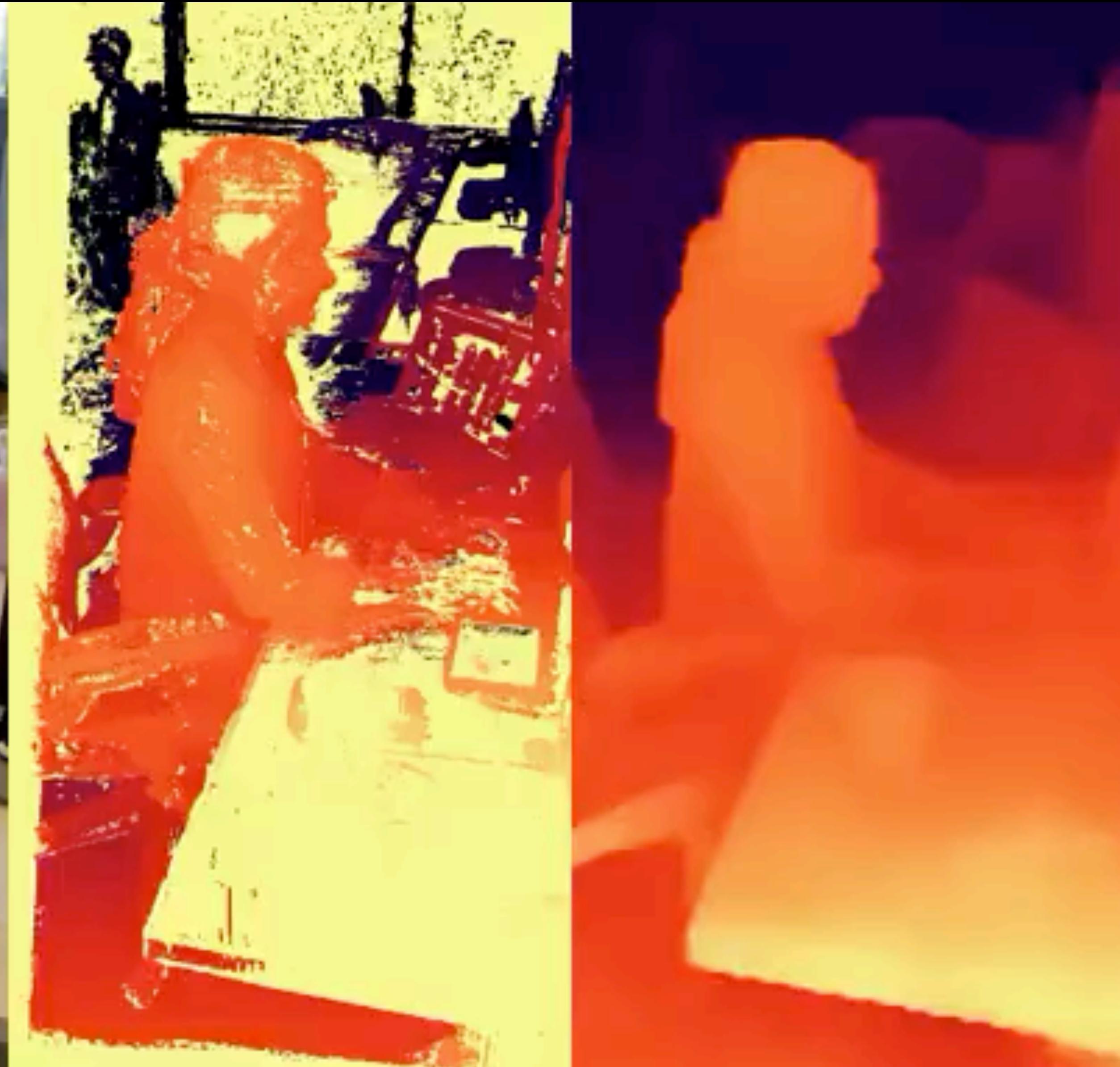




COLMAP
Schonberger et al.

Mannequin Challenge
Li et al. 2019

Ours



COLMAP
Schonberger et al.

Mannequin Challenge
Li et al. 2019

Ours

How to sample pairs?

How to sample pairs?

Exhaustive

- $O(N) \rightarrow O(N^2)$

How to sample pairs?

Exhaustive

- $O(N) \rightarrow O(N^2)$
- 100 frames (3.3s) \rightarrow 10K pairs

How to sample pairs?

Exhaustive

- $O(N) \rightarrow O(N^2)$
- 100 frames (3.3s) \rightarrow 10K pairs
- Expensive optical flow for 10K pairs

How to sample pairs?

Exhaustive

- $O(N) \rightarrow O(N^2)$

$O(N^2)$

- Expensive optical flow for 10K pairs

How to sample pairs?

Exhaustive

- $O(N) \rightarrow O(N^2)$

$O(N^2)$

- Expensive optical flow for 10K pairs

Hierarchical

How to sample pairs?

Exhaustive

- $O(N) \rightarrow O(N^2)$

$O(N^2)$

- Expensive optical flow for 10K pairs

Hierarchical



How to sample pairs?

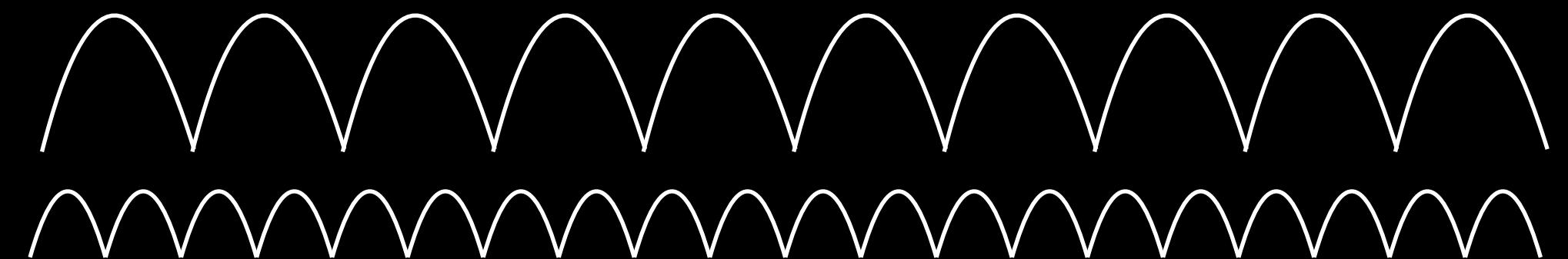
Exhaustive

- $O(N) \rightarrow O(N^2)$

$O(N^2)$

- Expensive optical flow for 10K pairs

Hierarchical



How to sample pairs?

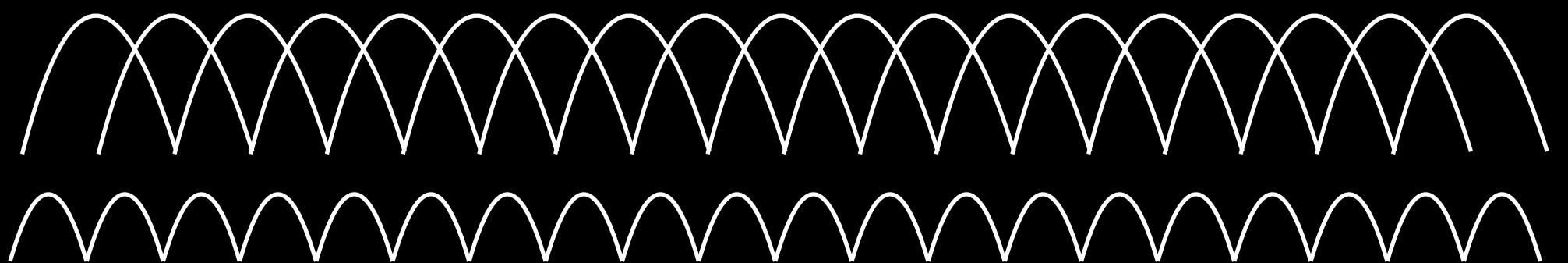
Exhaustive

- $O(N) \rightarrow O(N^2)$

$O(N^2)$

- Expensive optical flow for 10K pairs

Hierarchical



How to sample pairs?

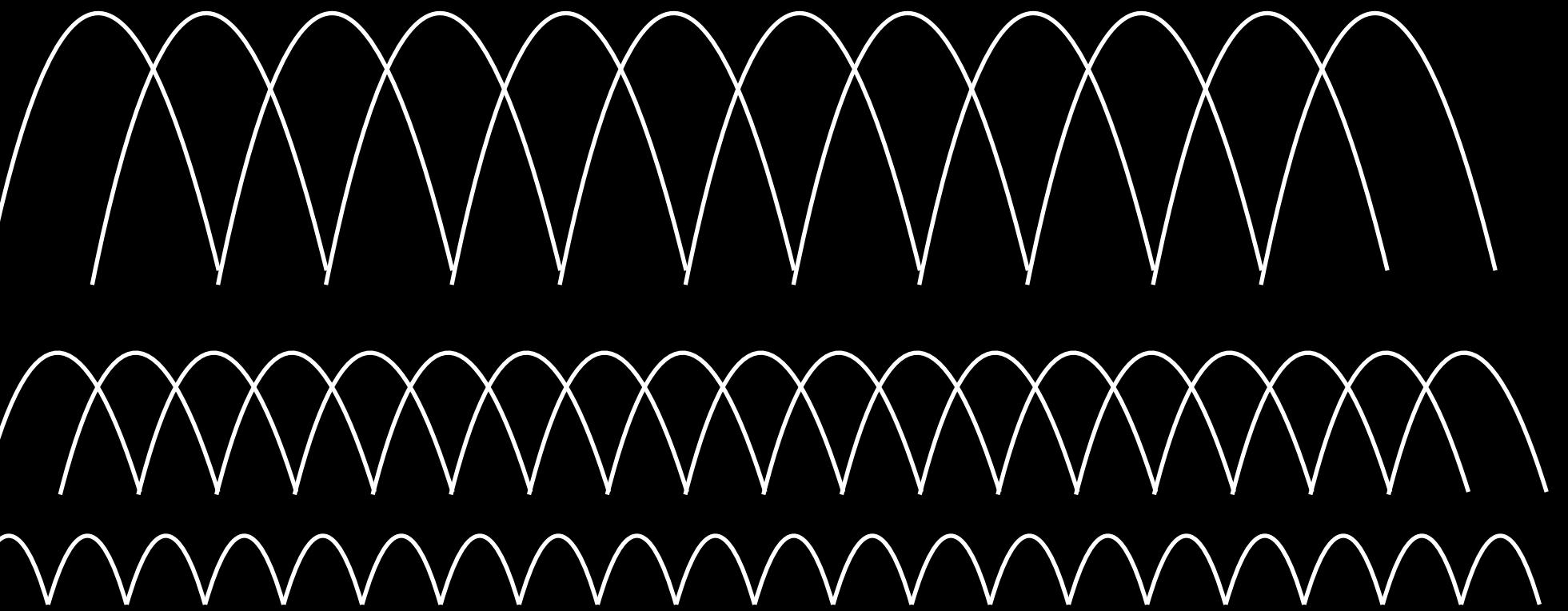
Exhaustive

- $O(N) \rightarrow O(N^2)$

$O(N^2)$

- Expensive optical flow for 10K pairs

Hierarchical



How to sample pairs?

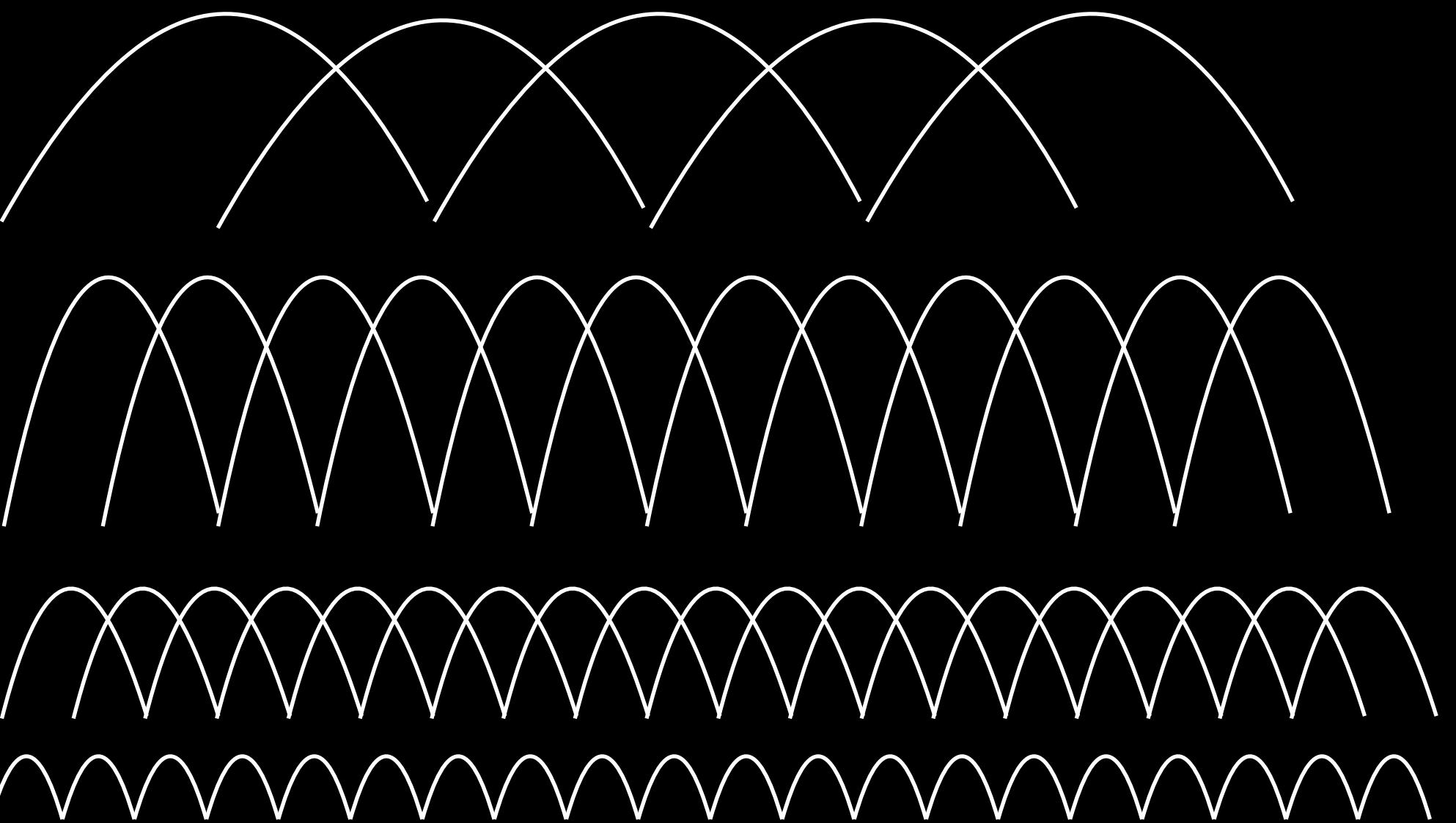
Exhaustive

- $O(N) \rightarrow O(N^2)$

$O(N^2)$

- Expensive optical flow for 10K pairs

Hierarchical



How to sample pairs?

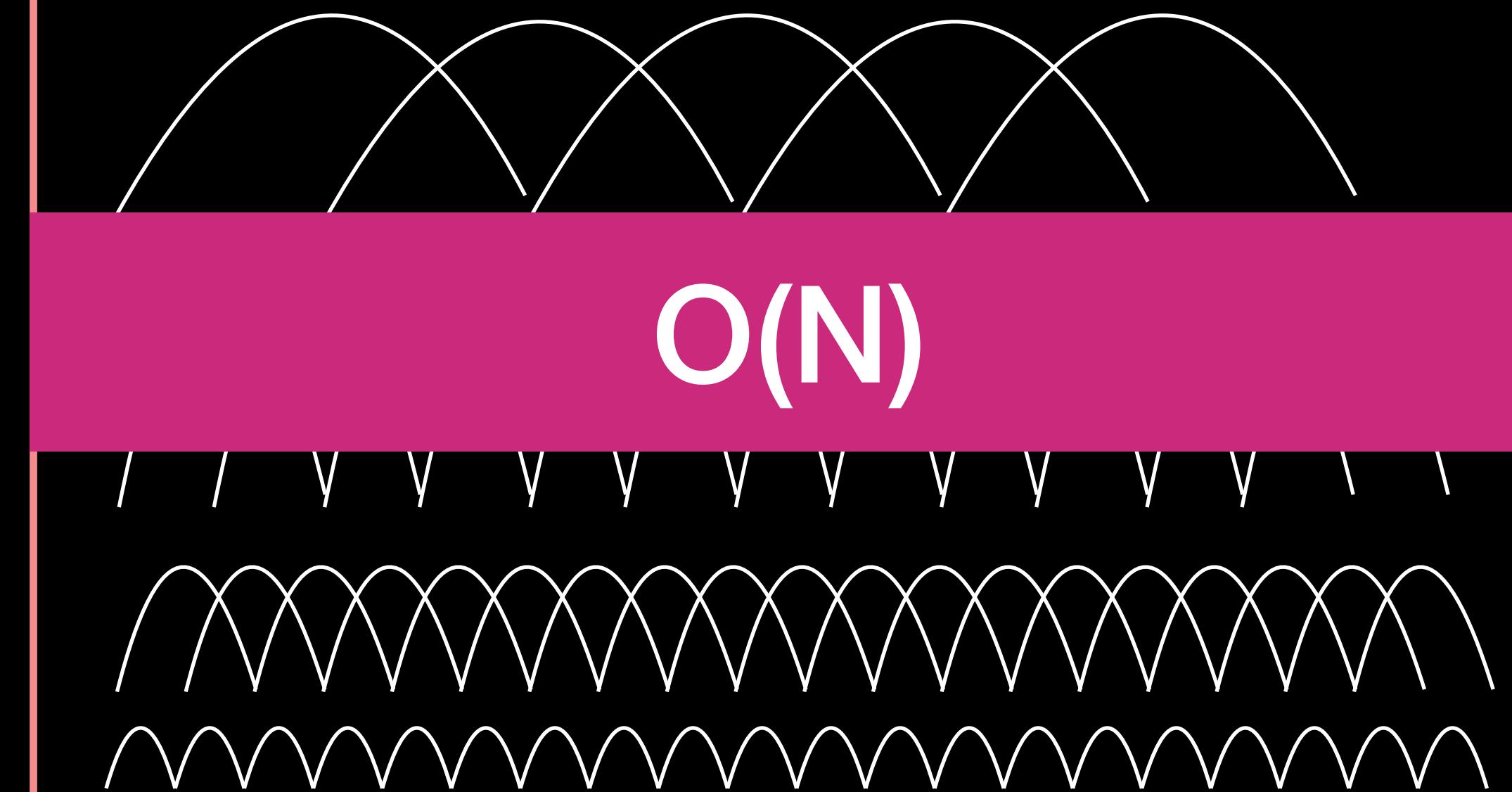
Exhaustive

- $O(N) \rightarrow O(N^2)$

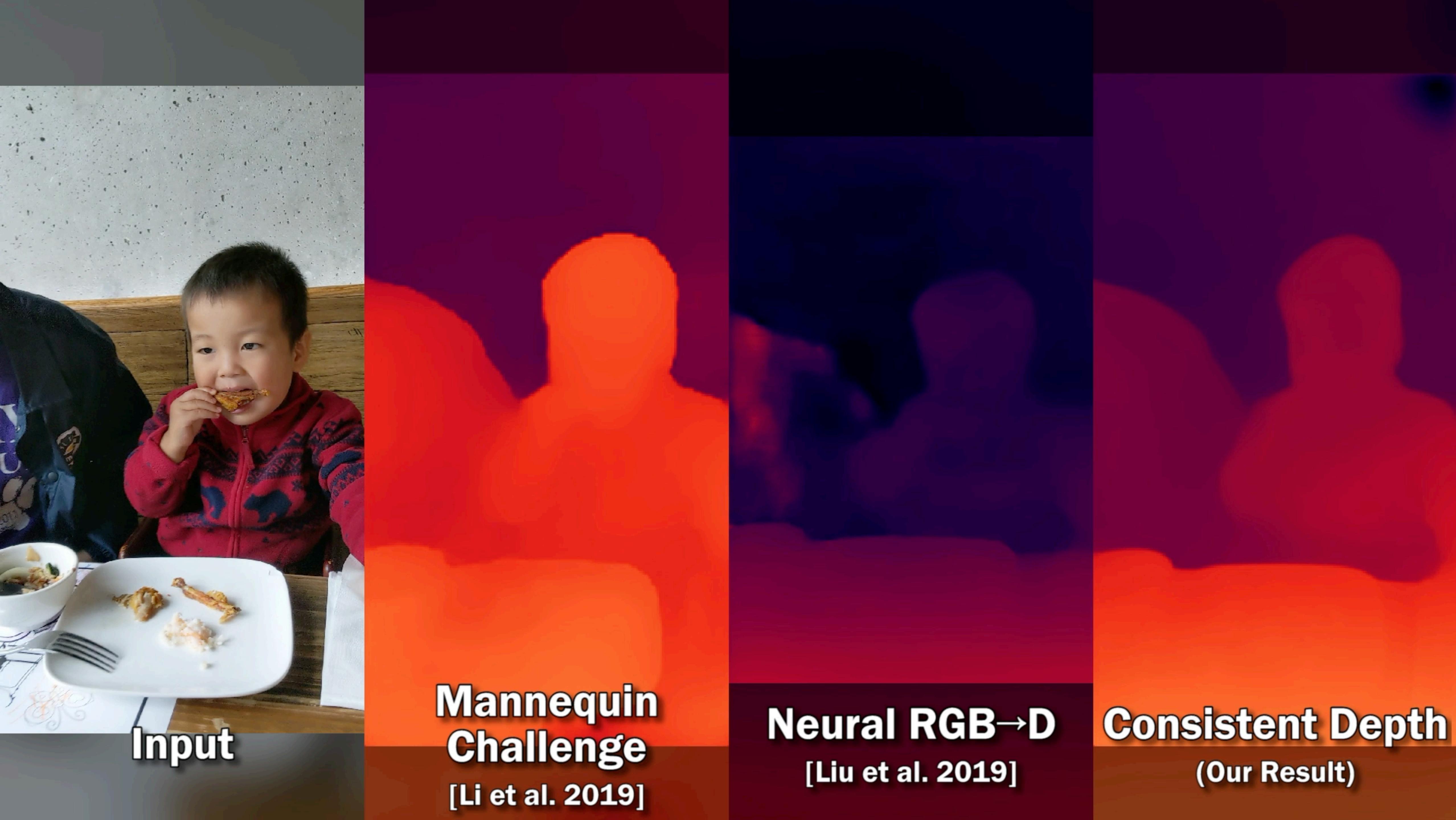
$O(N^2)$

- Expensive optical flow for 10K pairs

Hierarchical



Results

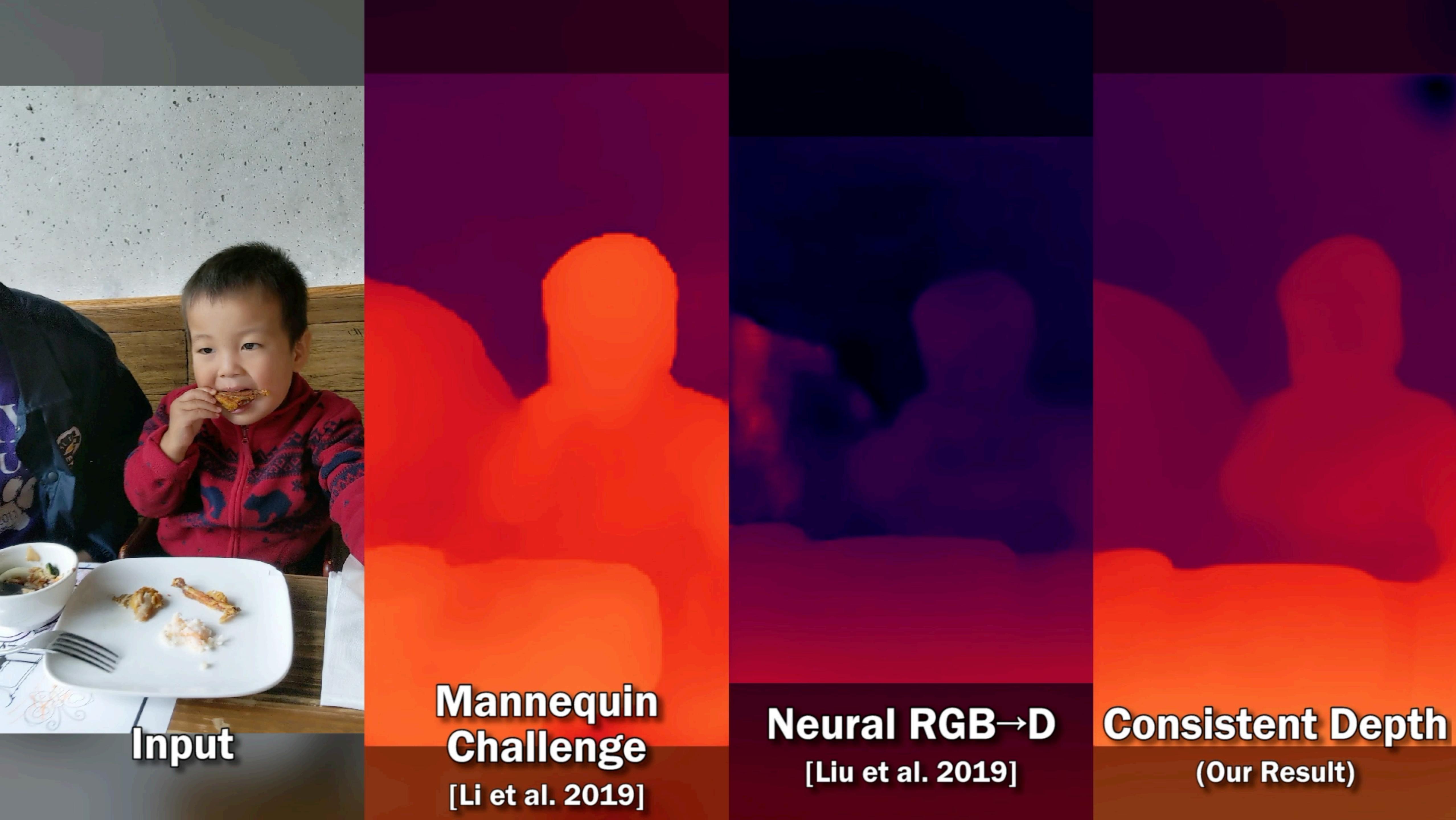


Input

**Mannequin
Challenge**
[Li et al. 2019]

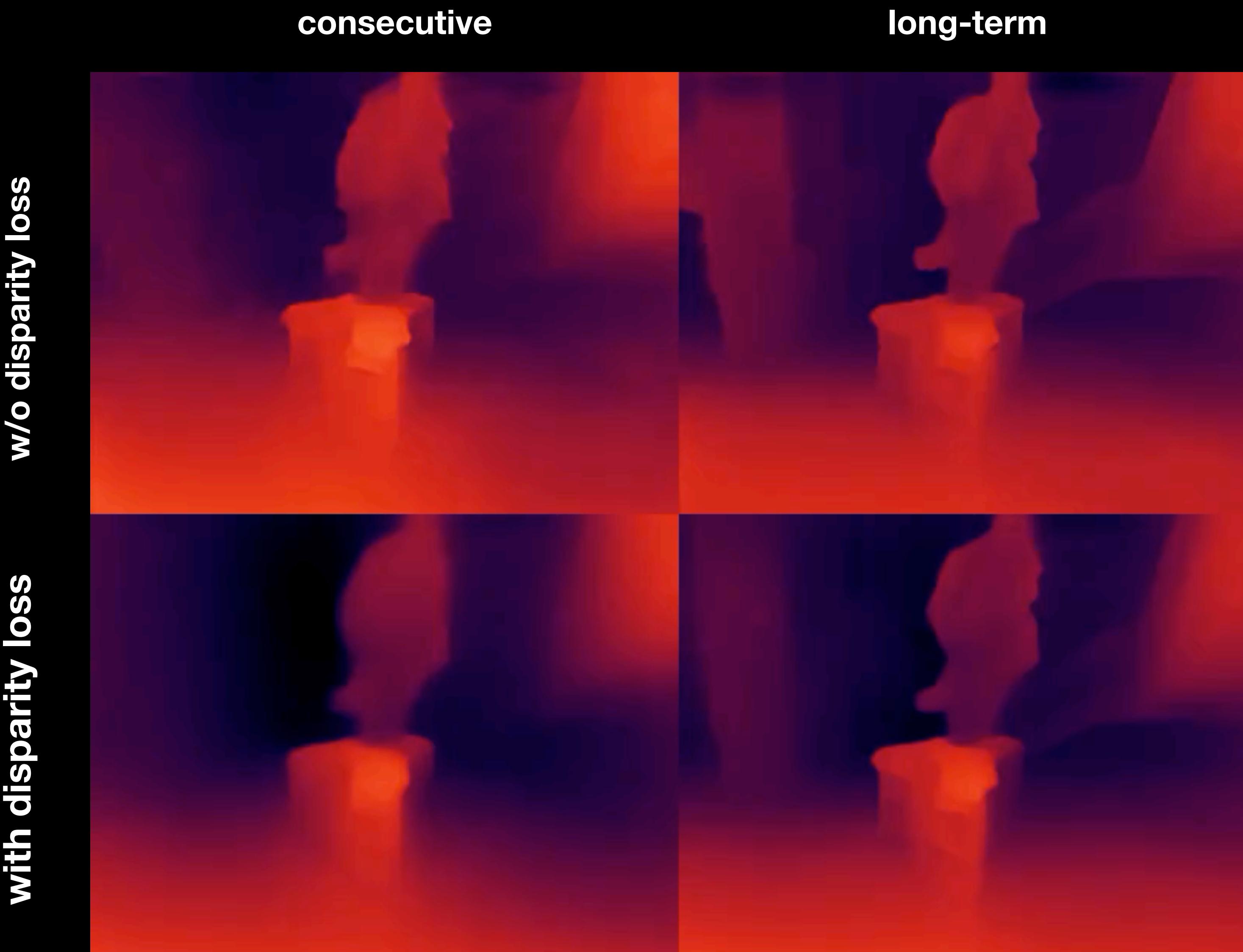
Neural RGB→D
[Liu et al. 2019]

Consistent Depth
(Our Result)

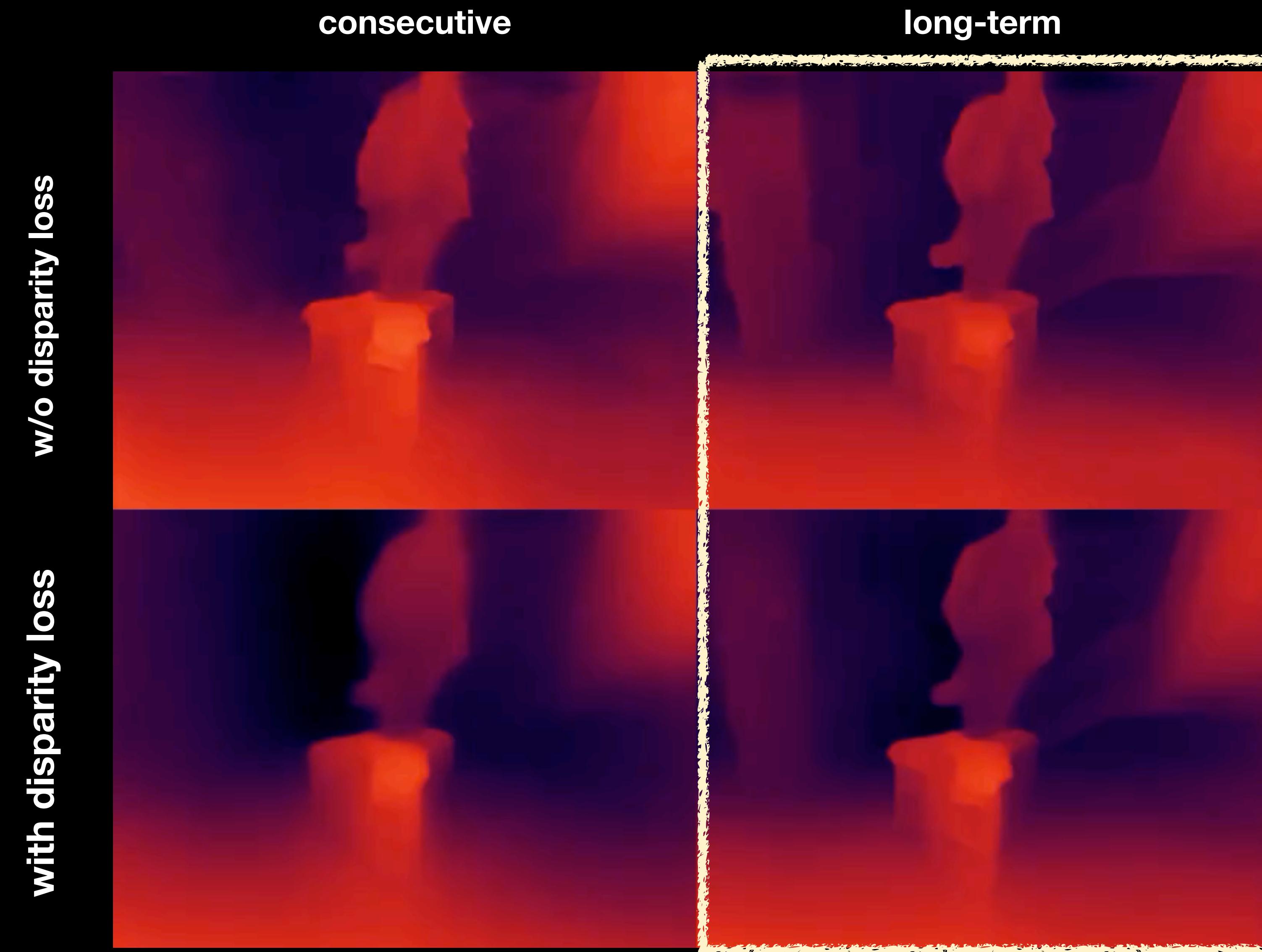


Long-term Constraint Matters

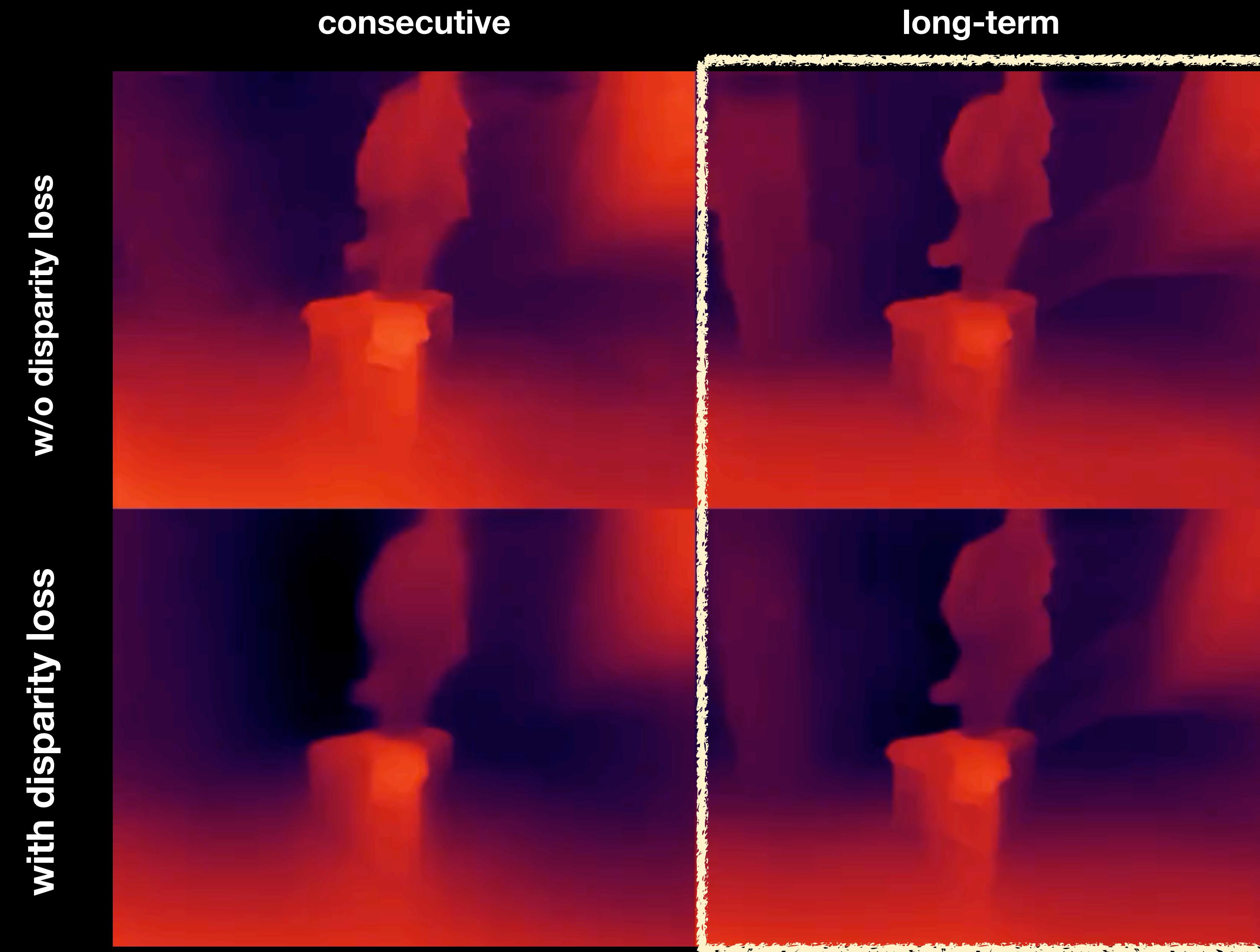
Long-term Constraint Matters



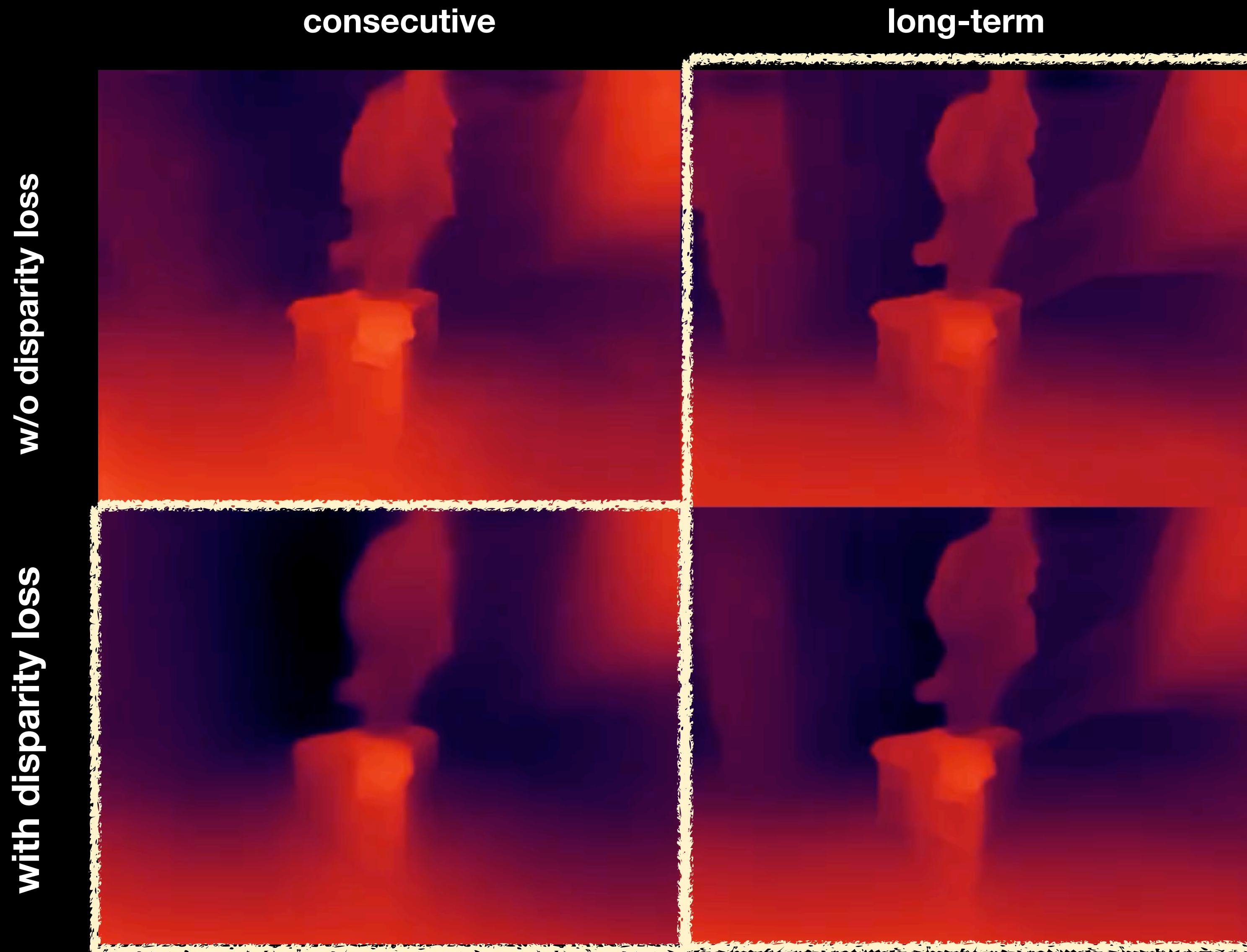
Long-term Constraint Matters



Long-term Constraint Matters



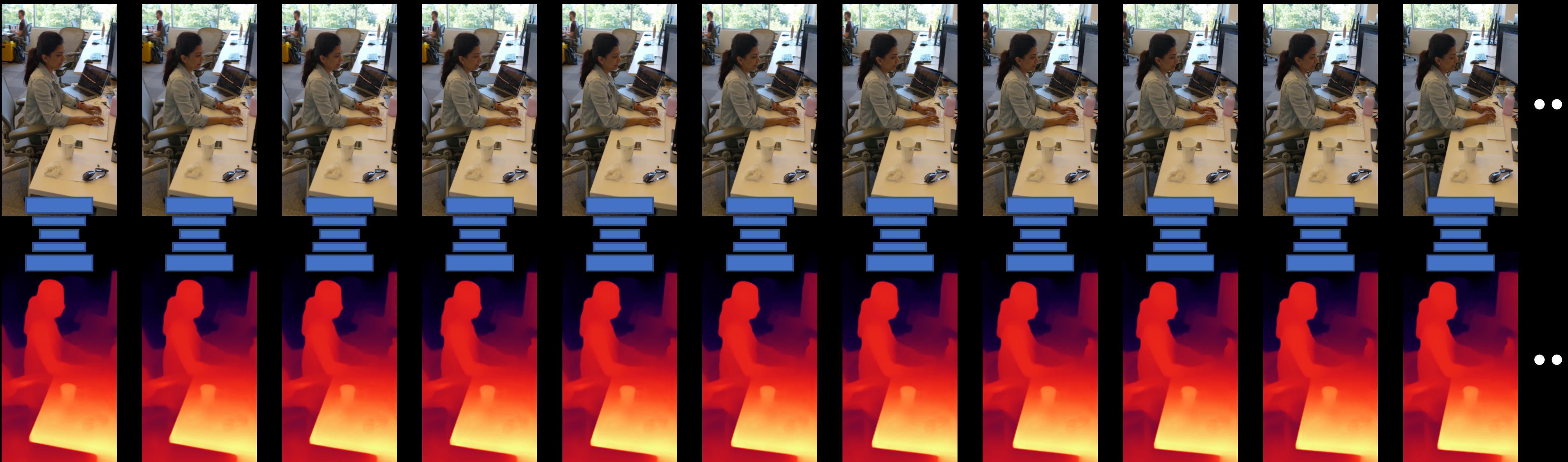
Long-term Constraint Matters



Why does it Work on Dynamic Videos?

Why does it Work on Dynamic Videos?

- Representation:



Why does it Work on Dynamic Videos?

- ~~Epipolar-aligned consistent motion~~
-

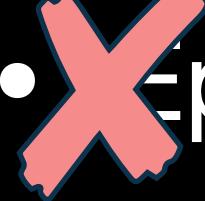


Why does it Work on Dynamic Videos?

- ~~Epipolar-aligned consistent motion~~
-

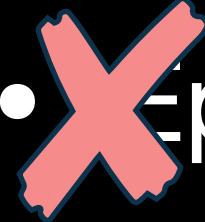


Why does it Work on Dynamic Videos?

-  Epipolar-aligned consistent motion
-  Non-epipolar-aligned consistent motion or inconsistent motion

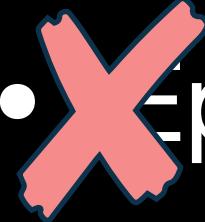


Why does it Work on Dynamic Videos?

-  Epipolar-aligned consistent motion
-  Non-epipolar-aligned consistent motion or inconsistent motion

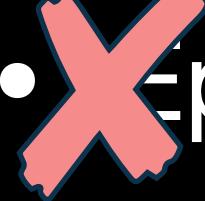


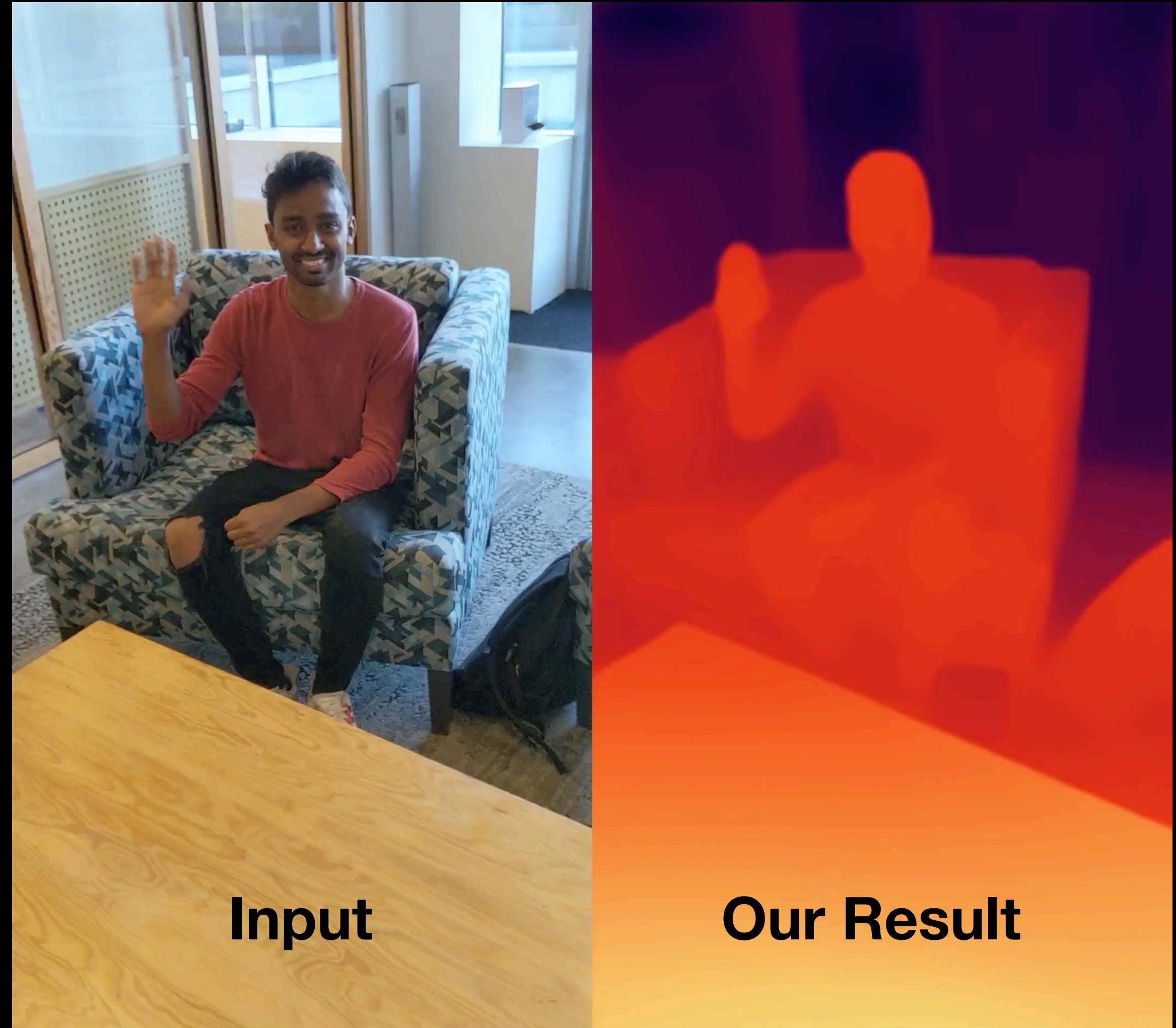
Why does it Work on Dynamic Videos?

-  Epipolar-aligned consistent motion
-  Non-epipolar-aligned consistent motion or inconsistent motion

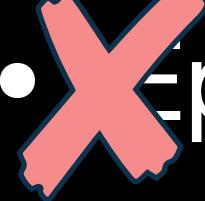


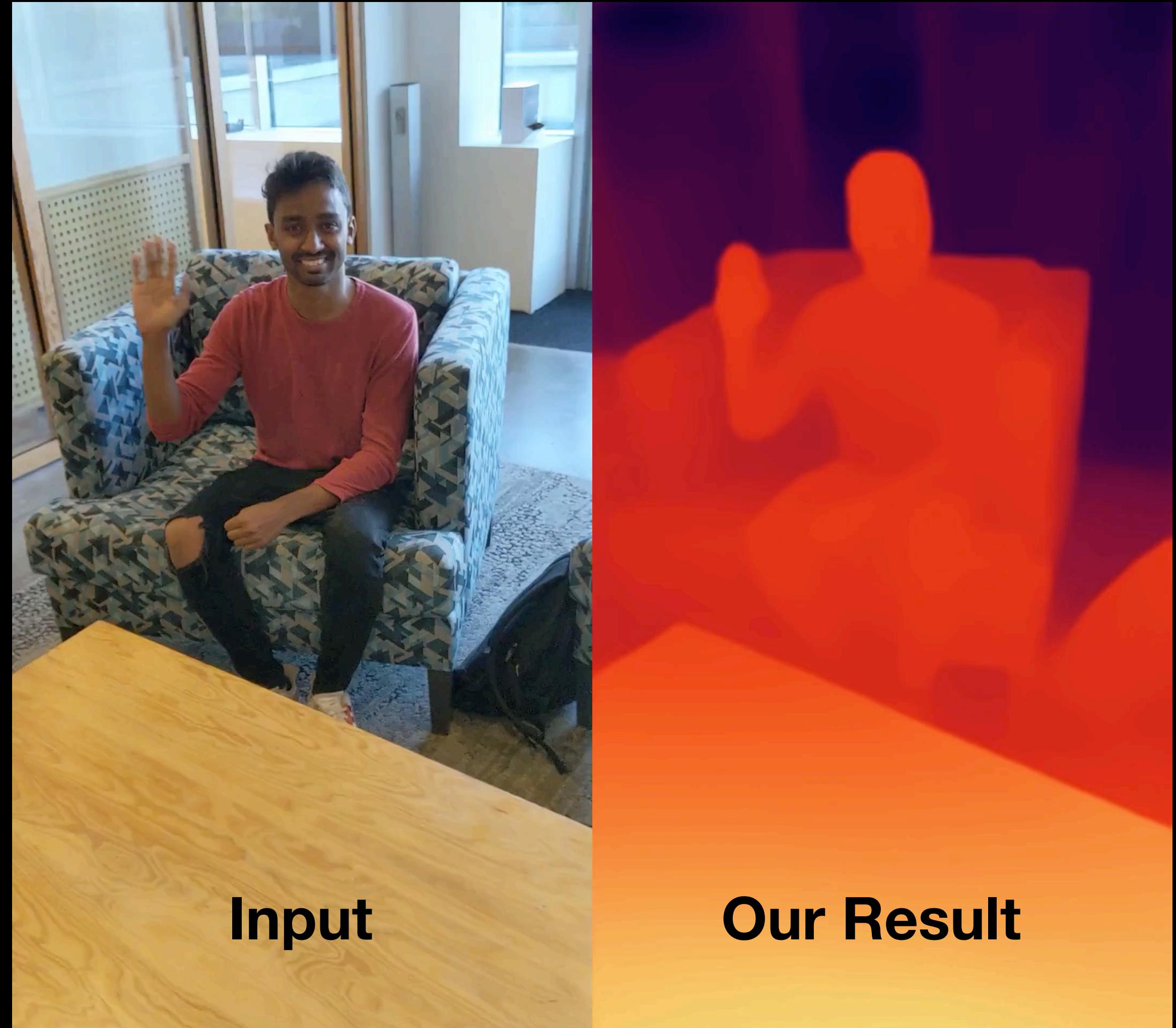
Why does it Work on Dynamic Videos?

-  Epipolar-aligned consistent motion
-  Non-epipolar-aligned consistent motion or inconsistent motion



Why does it Work on Dynamic Videos?

-  Epipolar-aligned consistent motion
-  Non-epipolar-aligned consistent motion or inconsistent motion



Project Webpage:
https://qrco.de/consistent_video_depth



Project Webpage

A Computational Time Machine





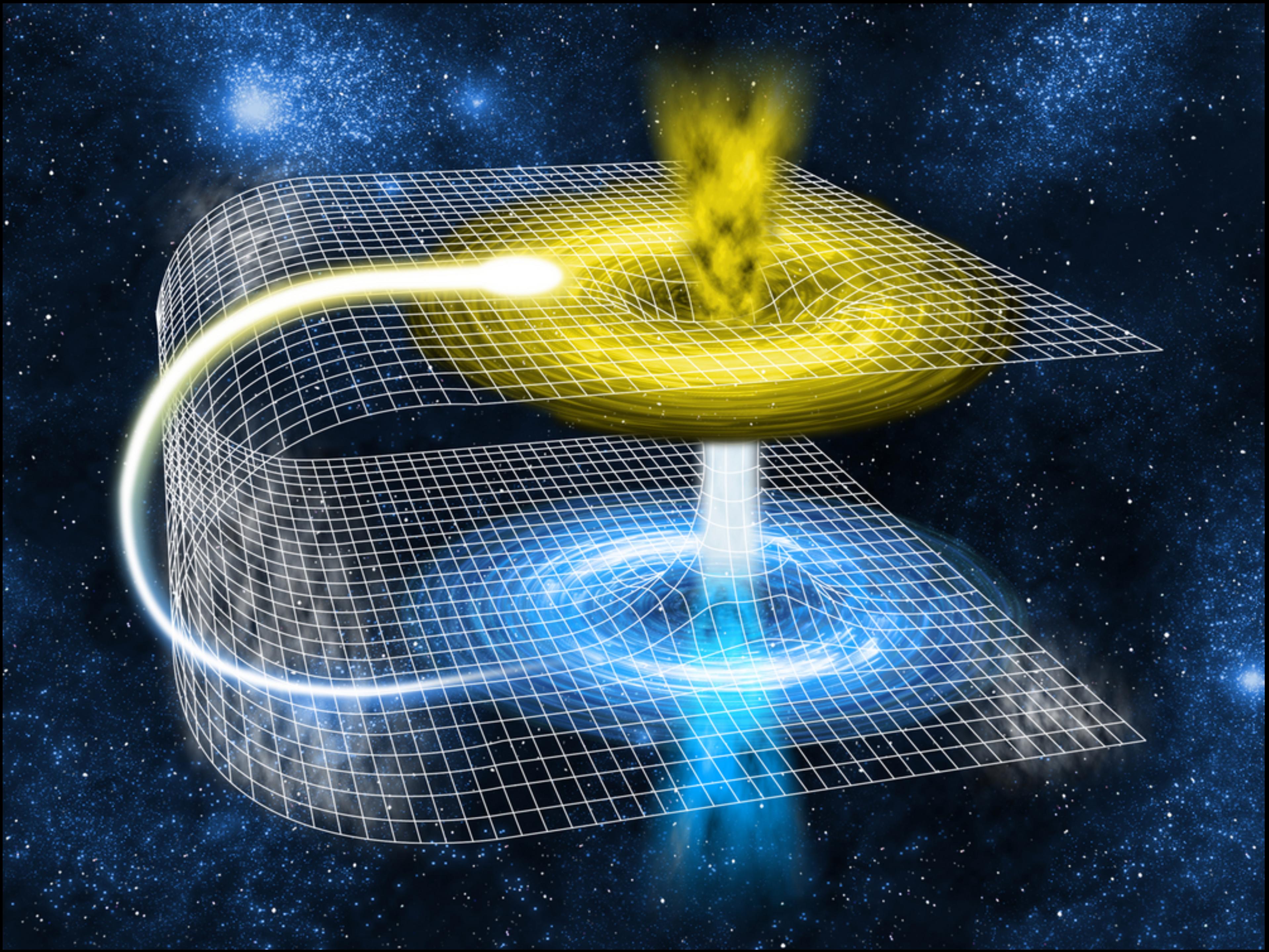








南京大屠杀





StyleGAN 2







KeystoneDepth: History in 3D



Xuan Luo
University of Washington



Yanmeng Kong
University of Washington



Jason Lawrence
Google Research



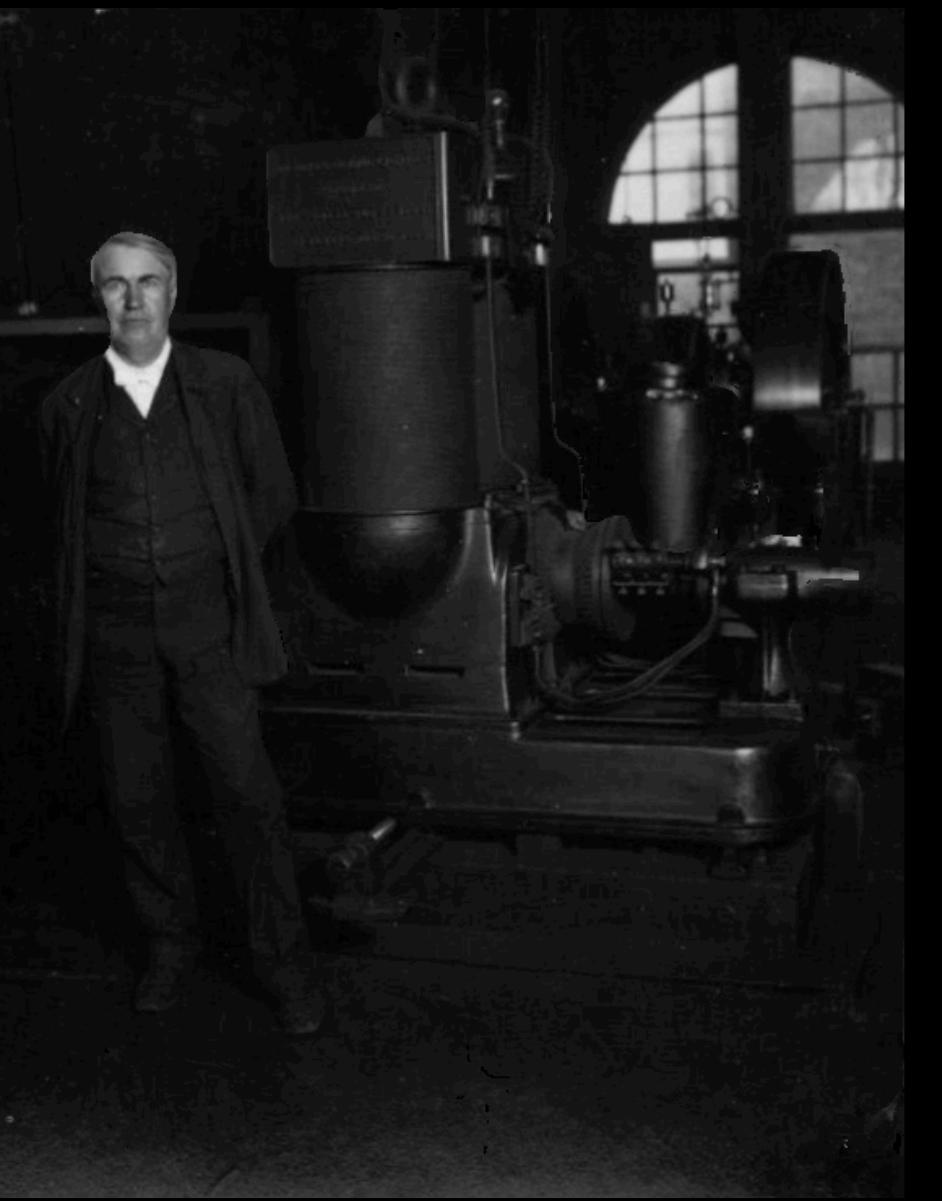
Ricardo Martin-Brualla
Google Research



Steven M. Seitz
University of Washington, Google Research



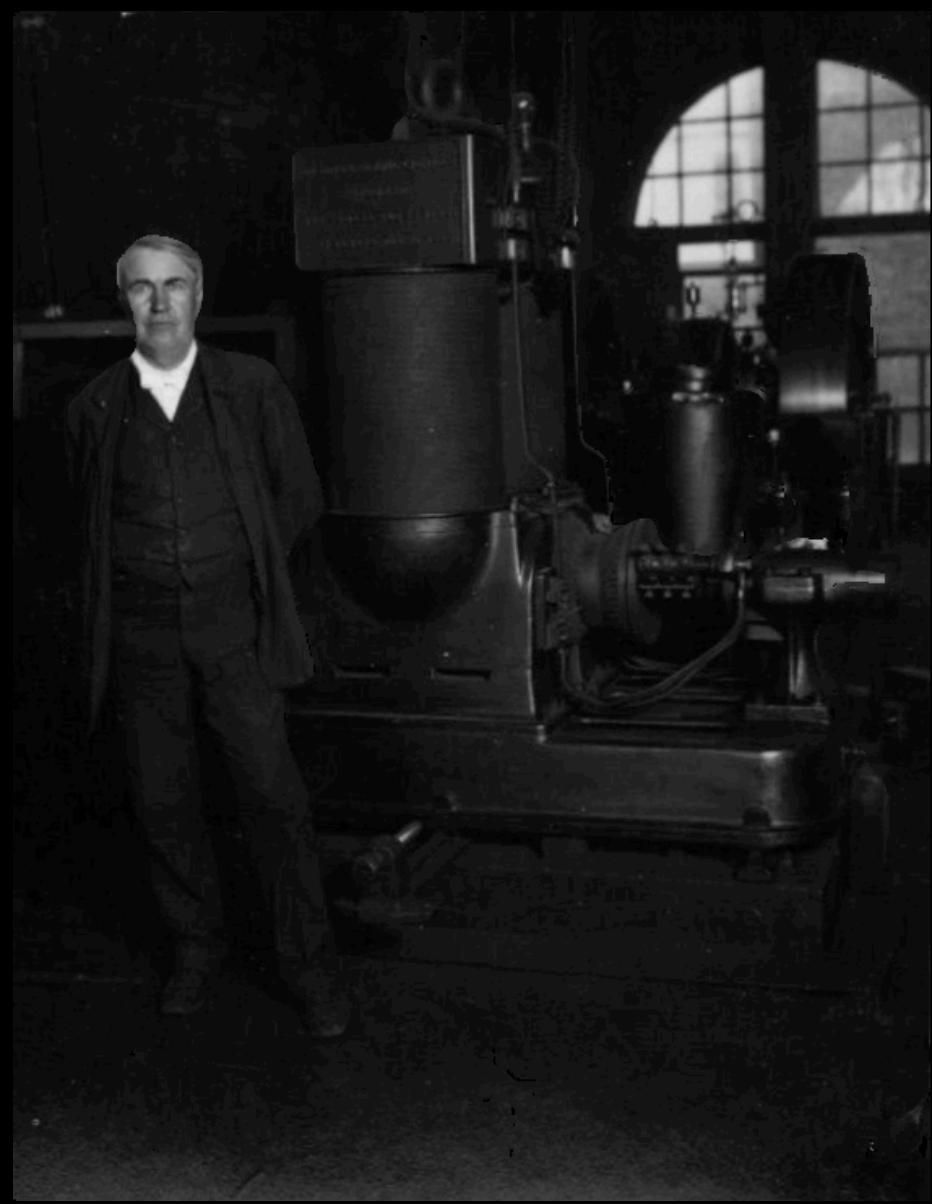
Thomas Edison



Thomas Edison's "dynamo", 1905



Thomas Edison



Thomas Edison's "dynamo", 1905



Fifth Avenue



Wall Street and Trinity Church



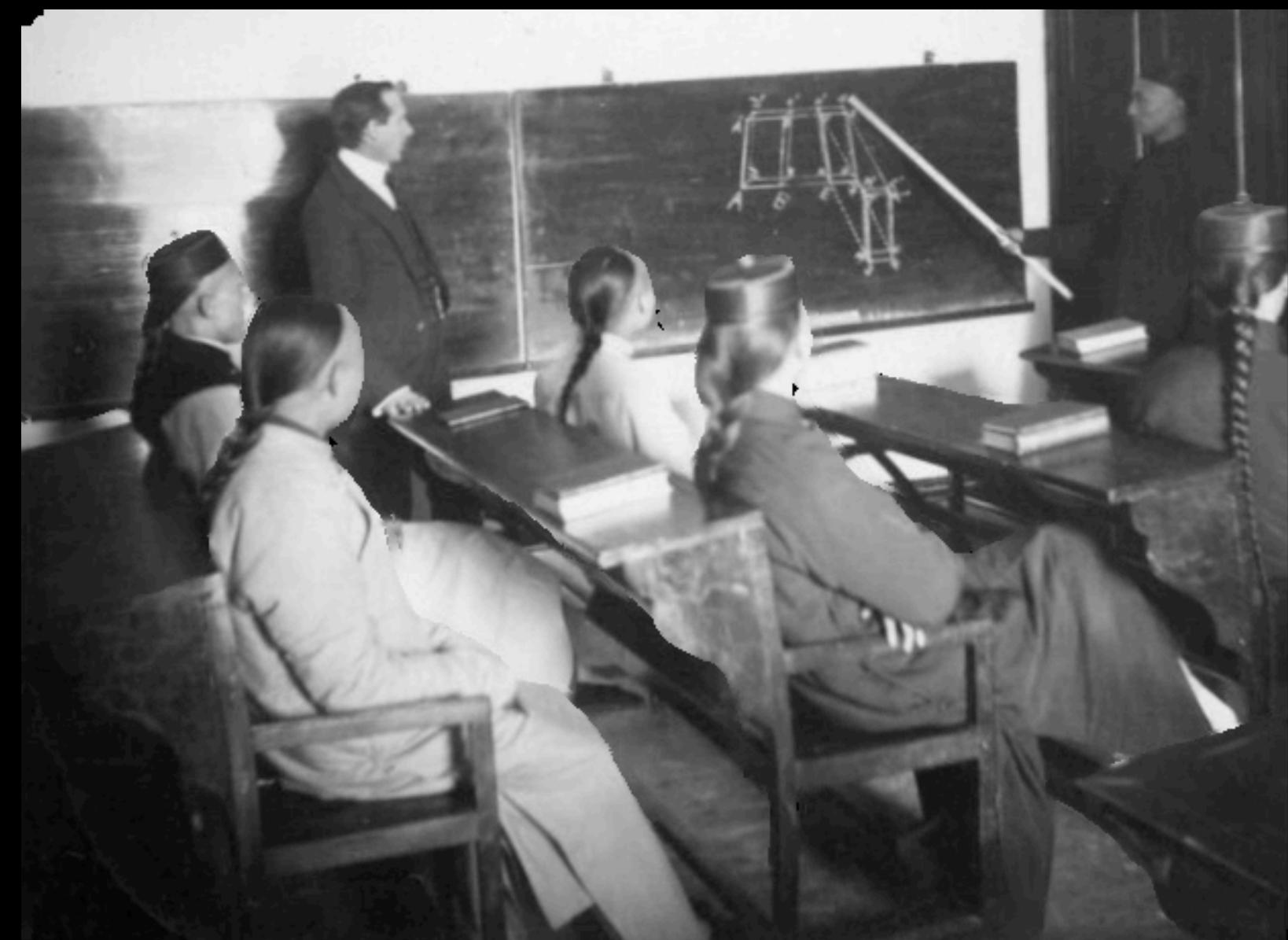
Hongzhang Li, Sept. 1900



Women in a Chinese home

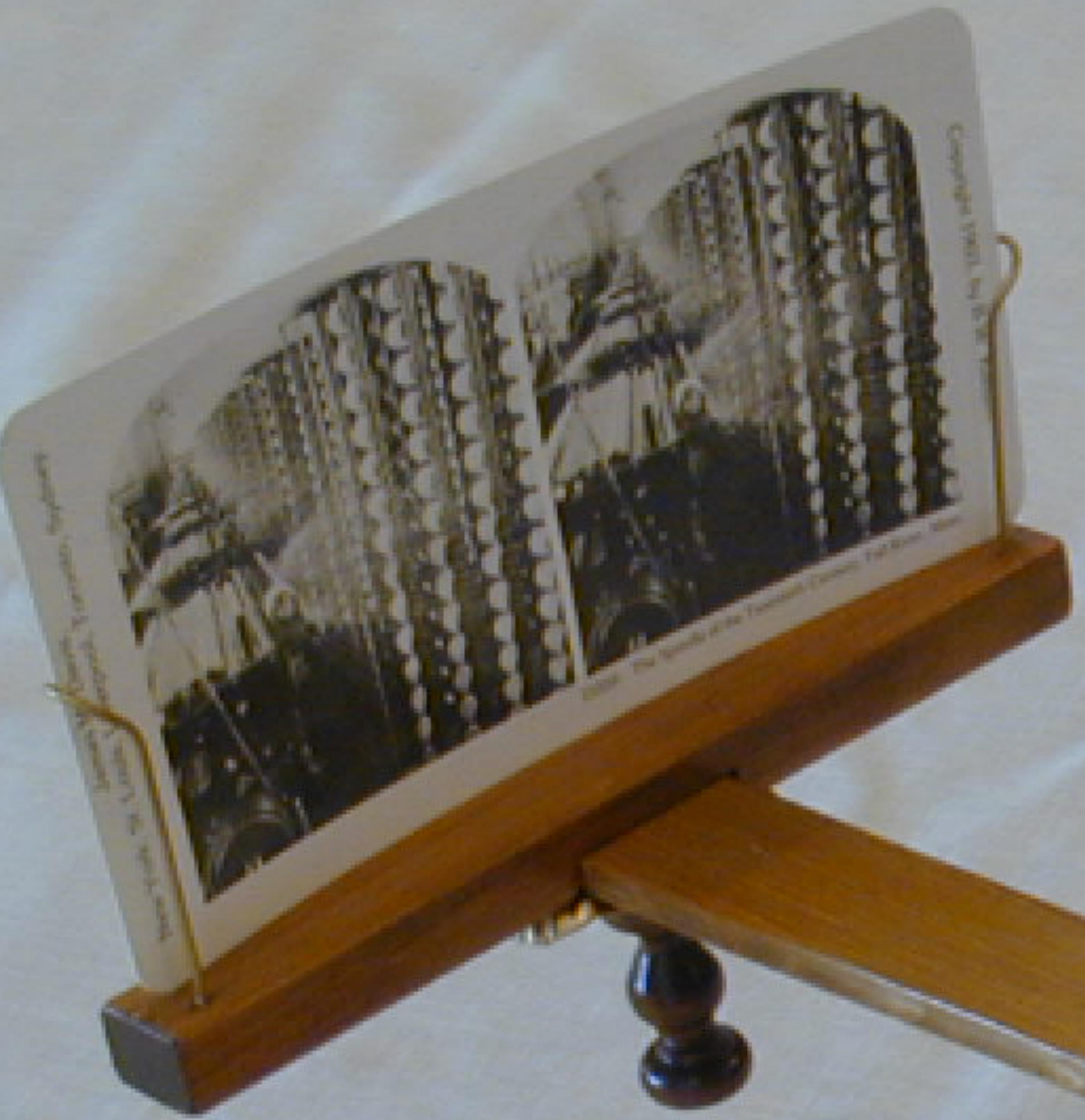


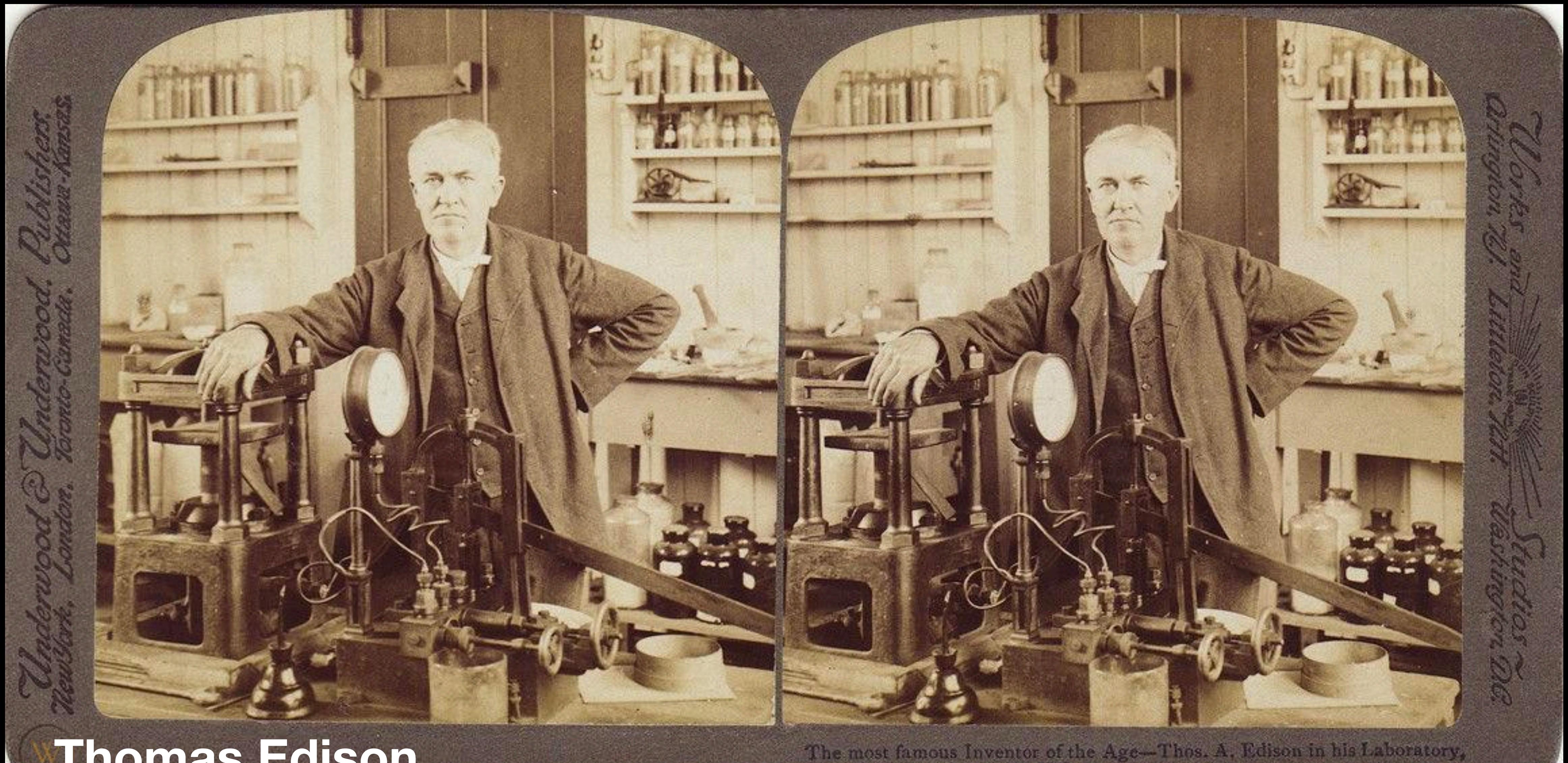
Large crafted dragons



Geometry class in Anglo Chinese College



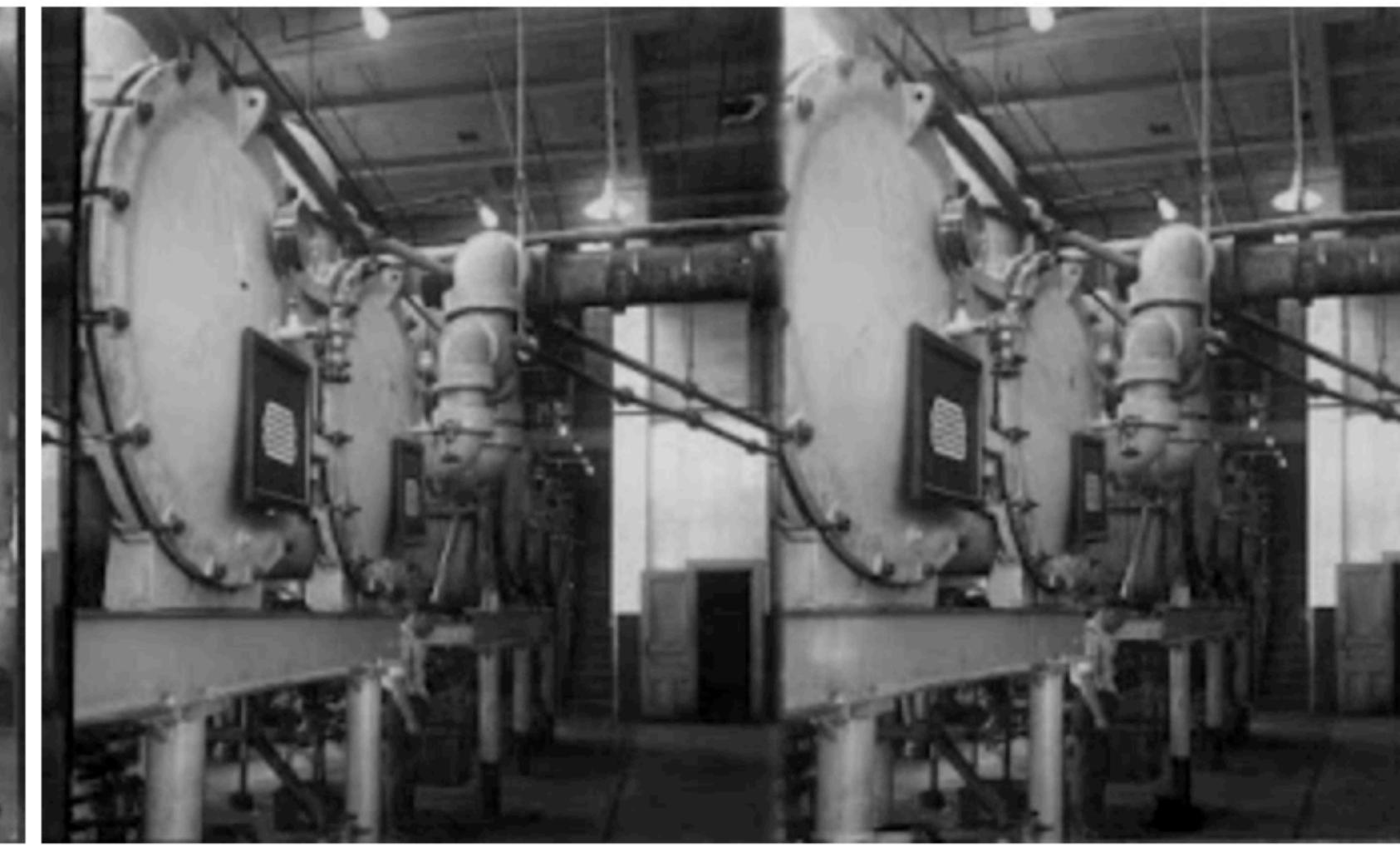
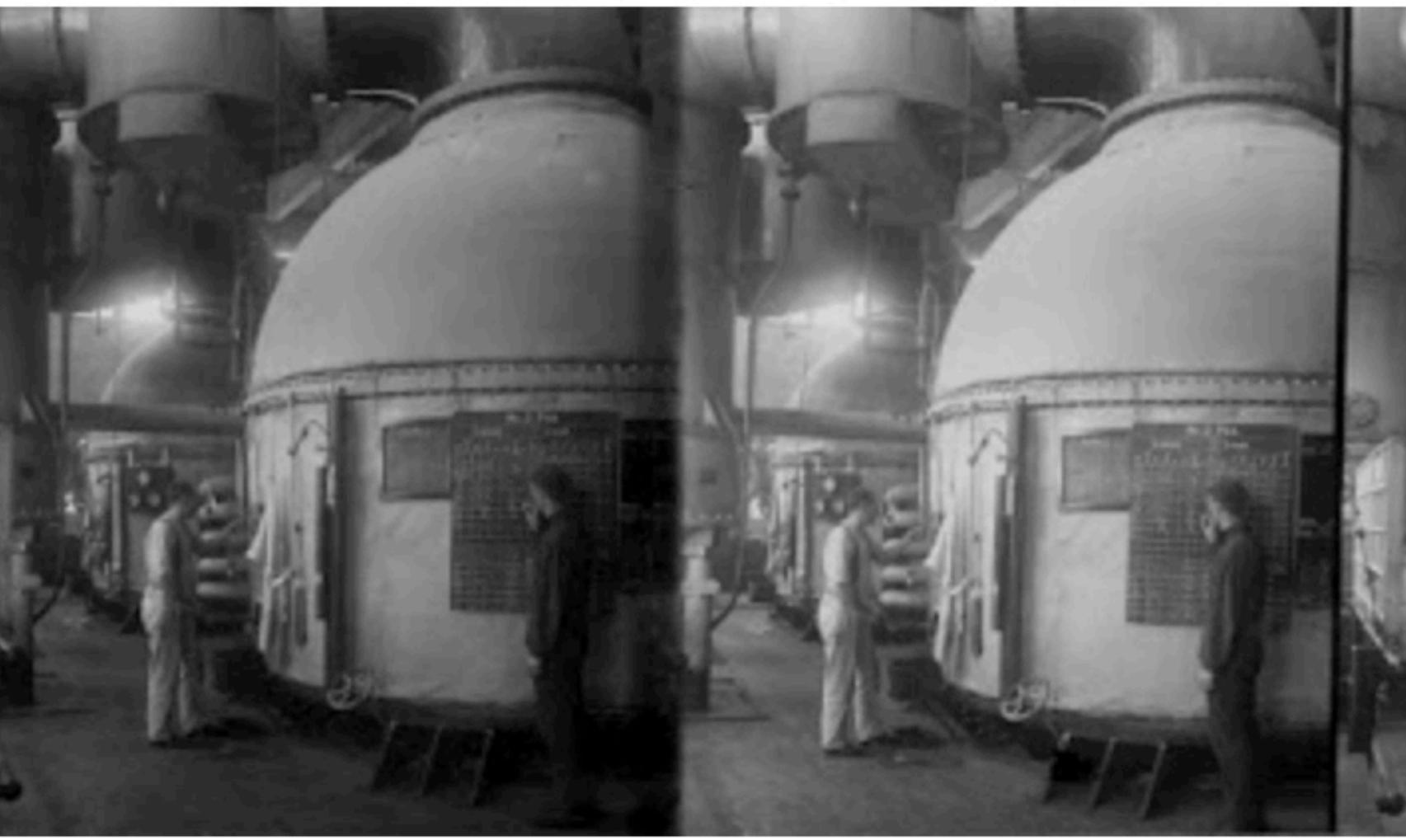
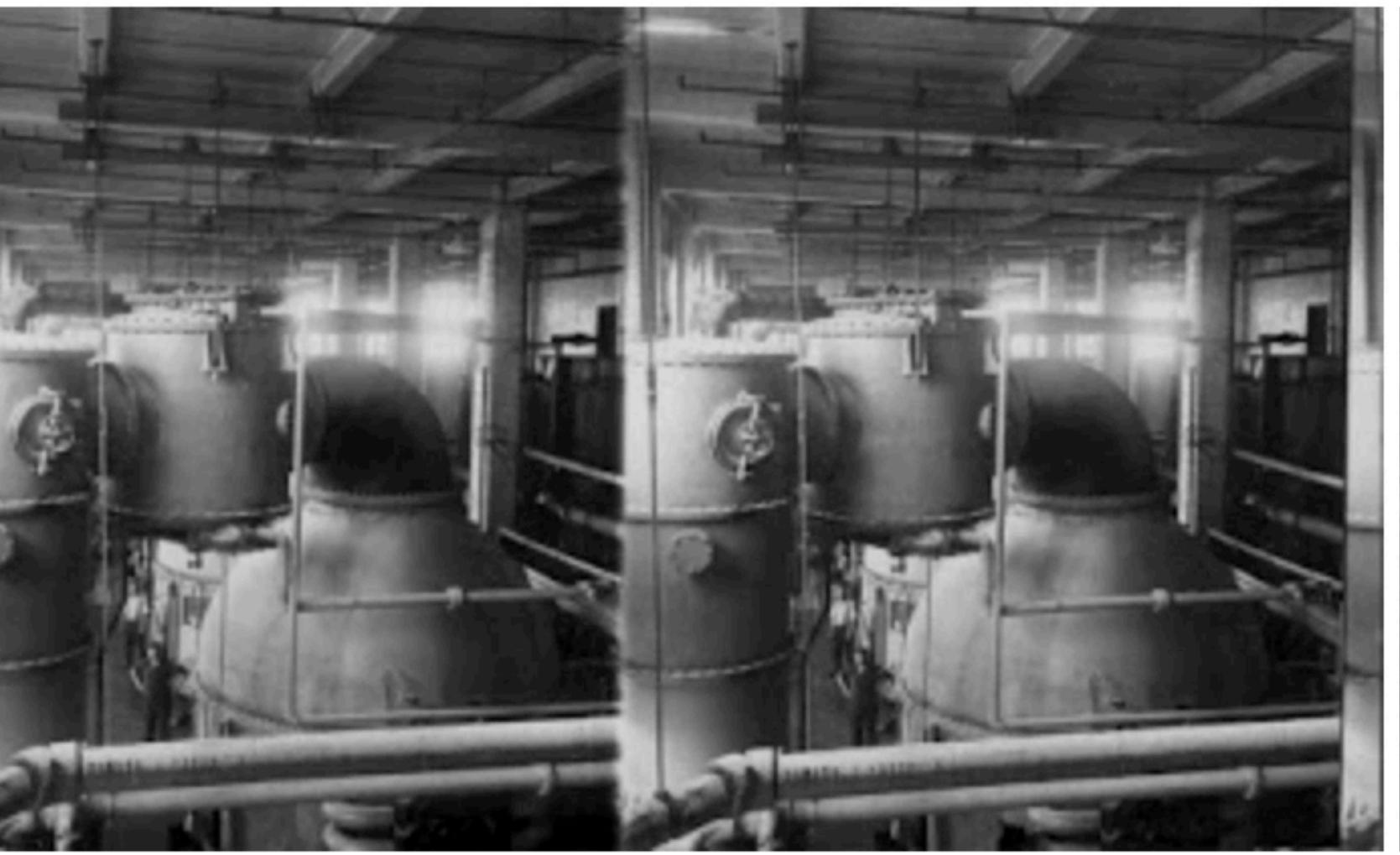
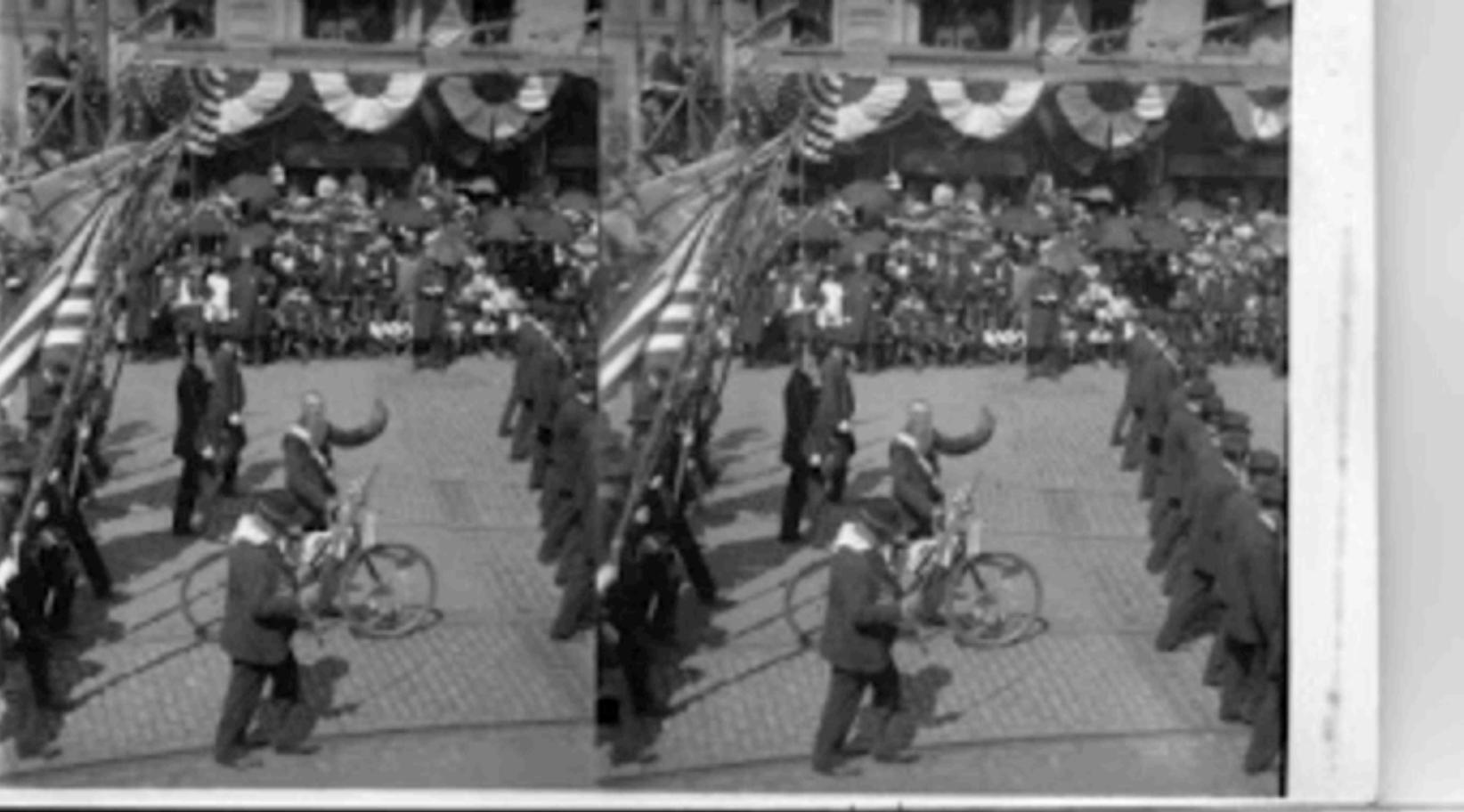
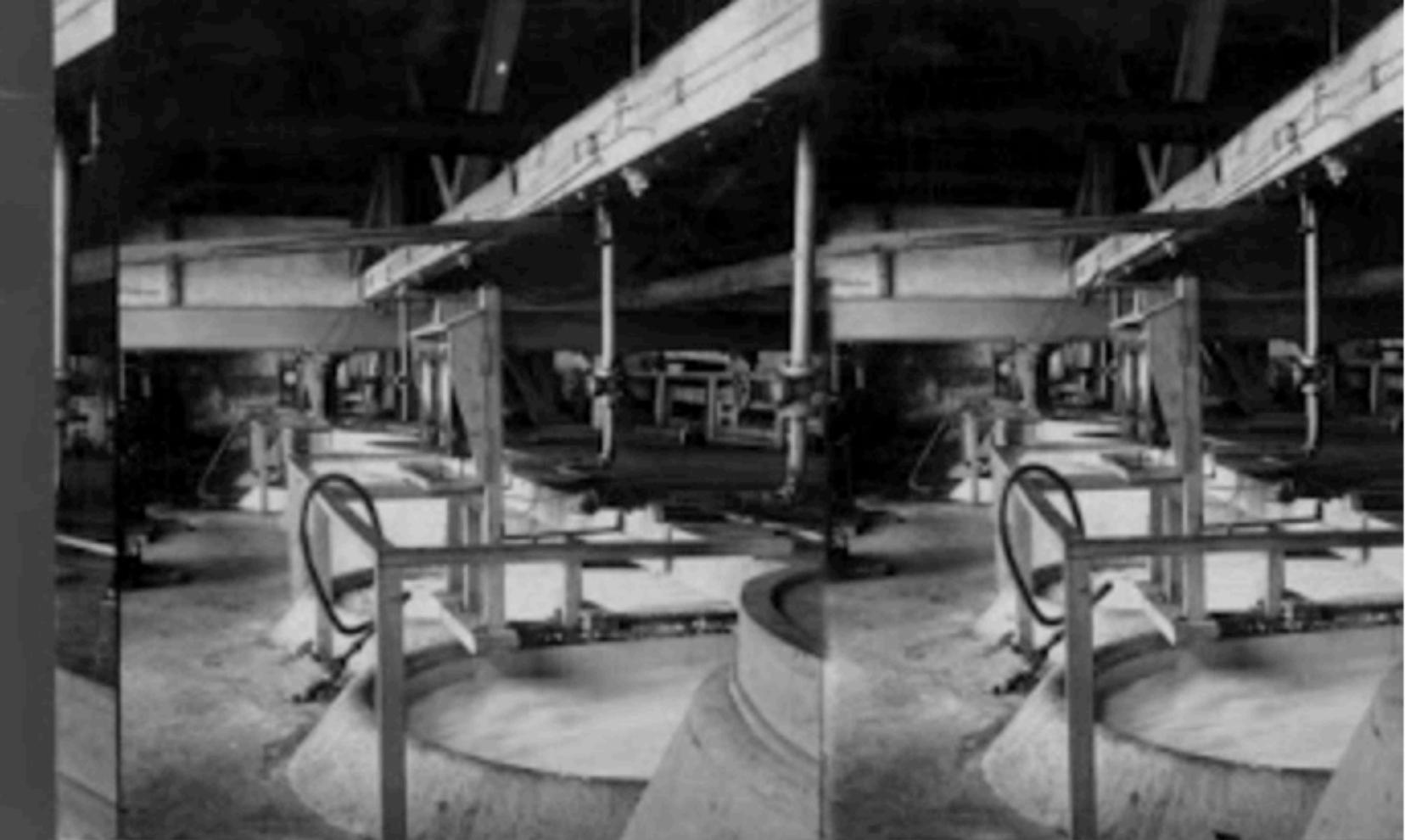


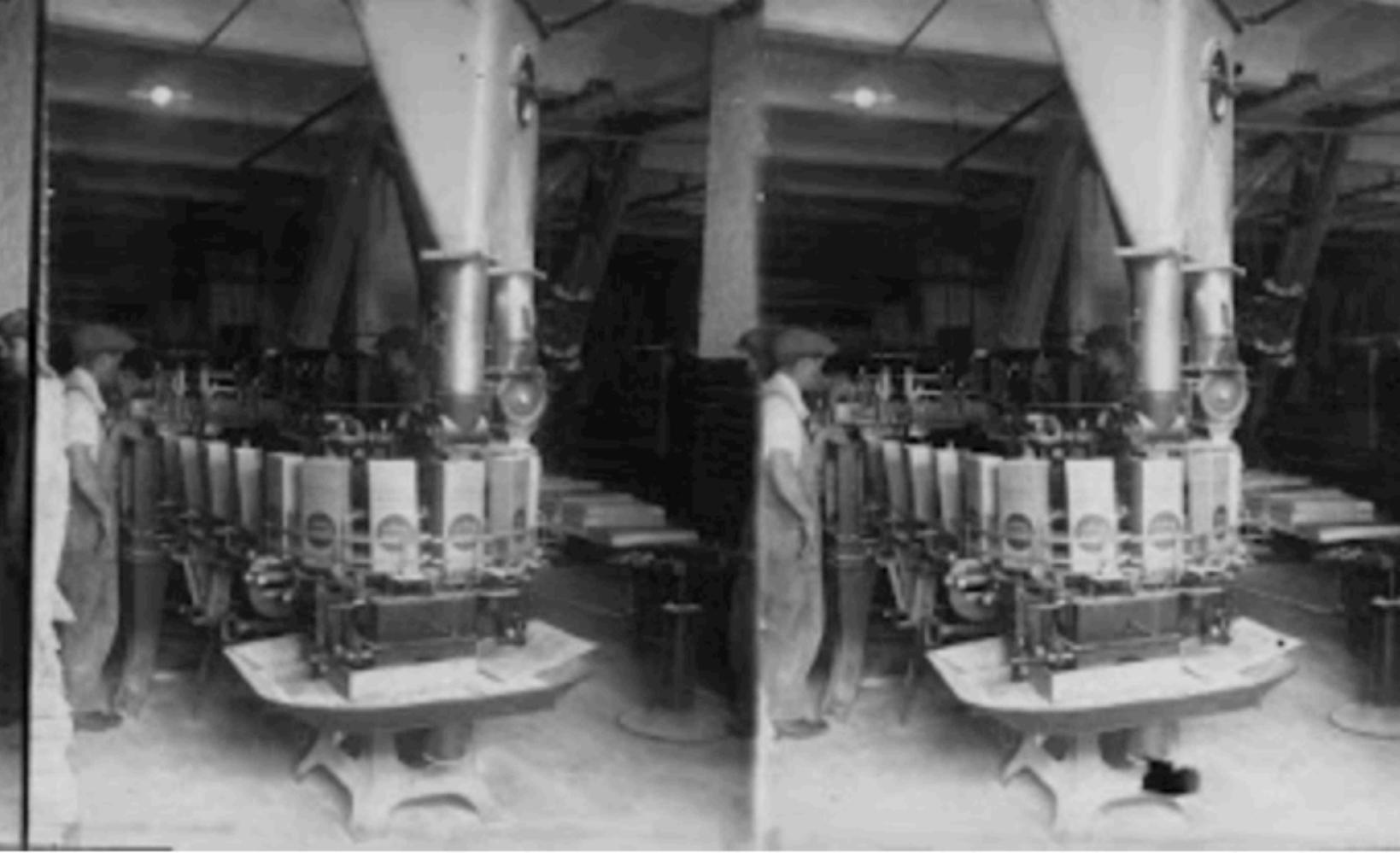
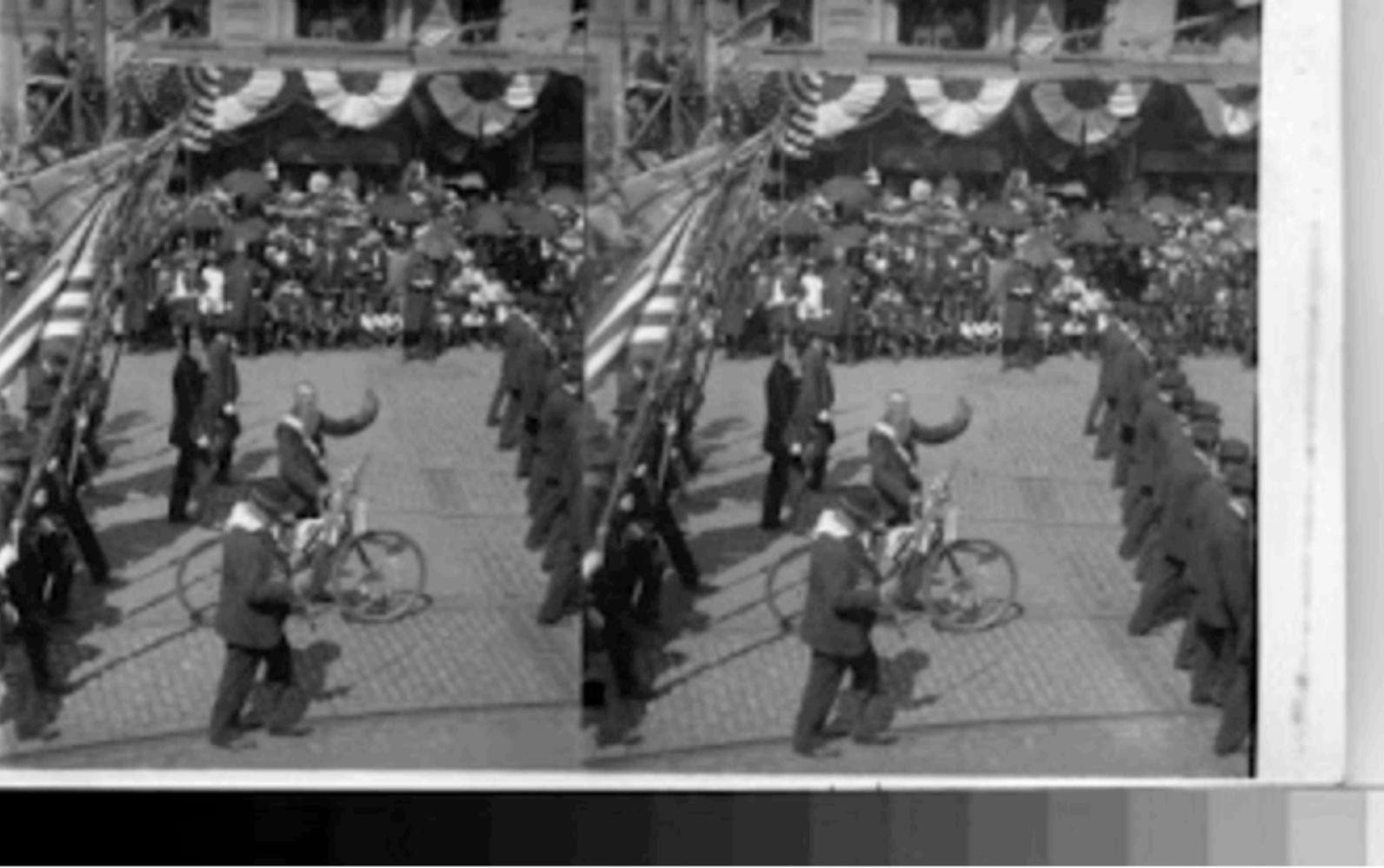
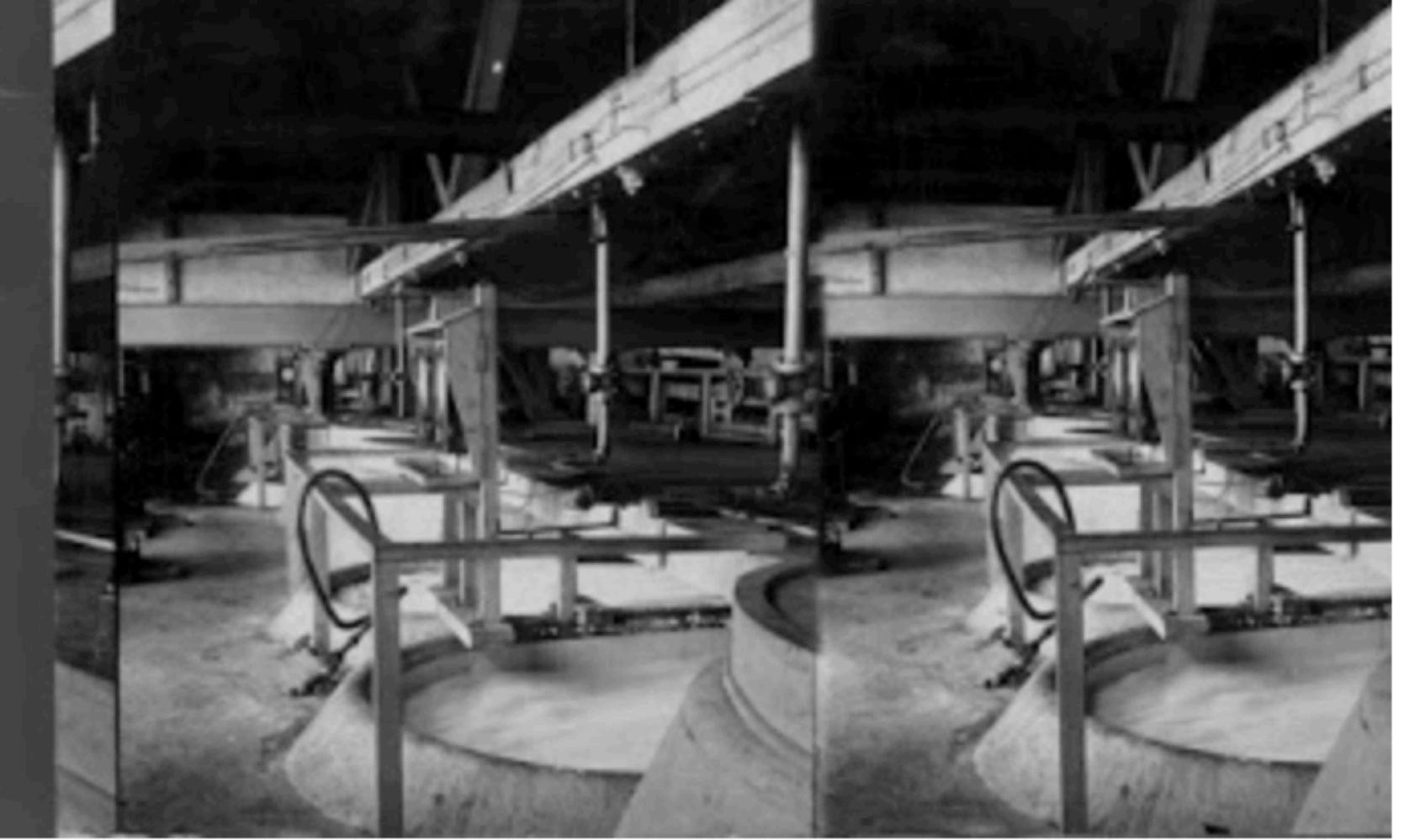


Thomas Edison
1847 - 1931

The most famous Inventor of the Age—Thos. A. Edison in his Laboratory,
East Orange, N. J., U.S.A.

Copyright 1901 by Underwood & Underwood





History in 3D

[Exhibitions](#)[Films](#)[Events](#)[Collections](#)

Collections



Keystone-Mast Collection

The Keystone-Mast Collection comprises over 350,000 stereoscopic photographs and negatives that depict the world between the late-nineteenth and mid-twentieth centuries. Firms represented in Keystone-Mast include Keystone View Company, Underwood and Underwood, B.W. Kilburn, H.C. White Company, Universal Photo Art Company (C.H. Graves), American Stereoscopic Company, W.H. Rau, Berry, Kelley and Chadwick, and George Rose. The photographs are generally not stereo cards, but contact prints from the original negatives, annotated by Keystone staff and photographers.

[Browse all 45350 Objects](#)

Collections

[Photographs](#)[Cameras](#)[Research Collections](#)[Sweeney Art Gallery](#)[Artists and Makers](#)[Advanced Search](#)[Create Account/Login](#)

[Exhibitions](#)[Films](#)[Events](#)[Collections](#)

Collections



Keystone-Mast Collection

The Keystone-Mast Collection comprises over 350,000 stereoscopic photographs and negatives that depict the world between the late-nineteenth and mid-twentieth centuries. Firms represented in Keystone-Mast include Keystone View Company, Underwood and Underwood, B.W. Kilburn, H.C. White Company, Universal Photo Art Company (C.H. Graves), American Stereoscopic Company, W.H. Rau, Berry, Kelley and Chadwick, and George Rose. The photographs are generally not stereo cards, but contact prints from the original negatives, annotated by Keystone staff and photographers.

[Browse all 45350 Objects](#)

Collections

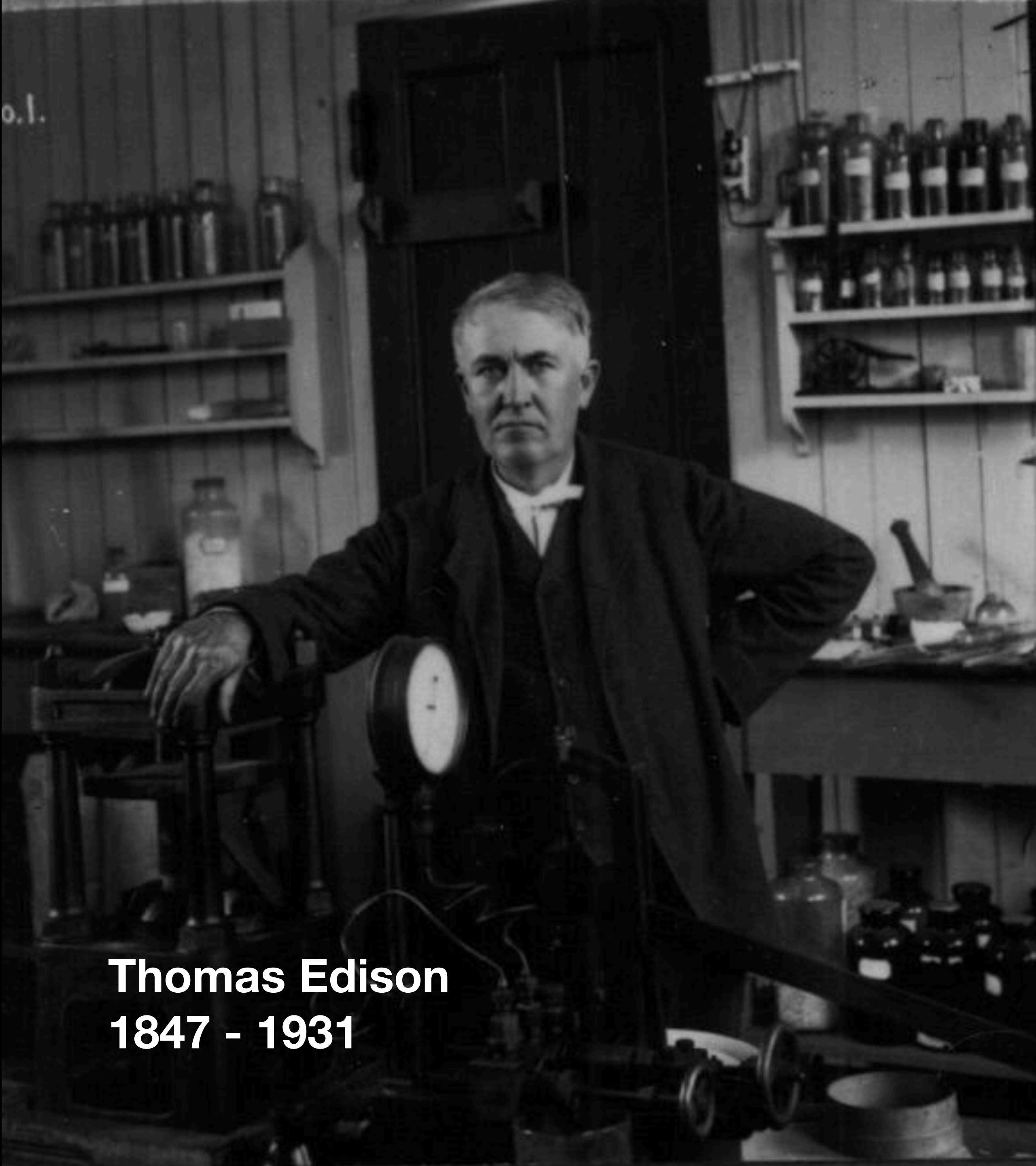
[Photographs](#)[Cameras](#)[Research Collections](#)[Sweeney Art Gallery](#)[Artists and Makers](#)[Advanced Search](#)[Create Account/Login](#)



Thomas Edison
1847 - 1931

VII962

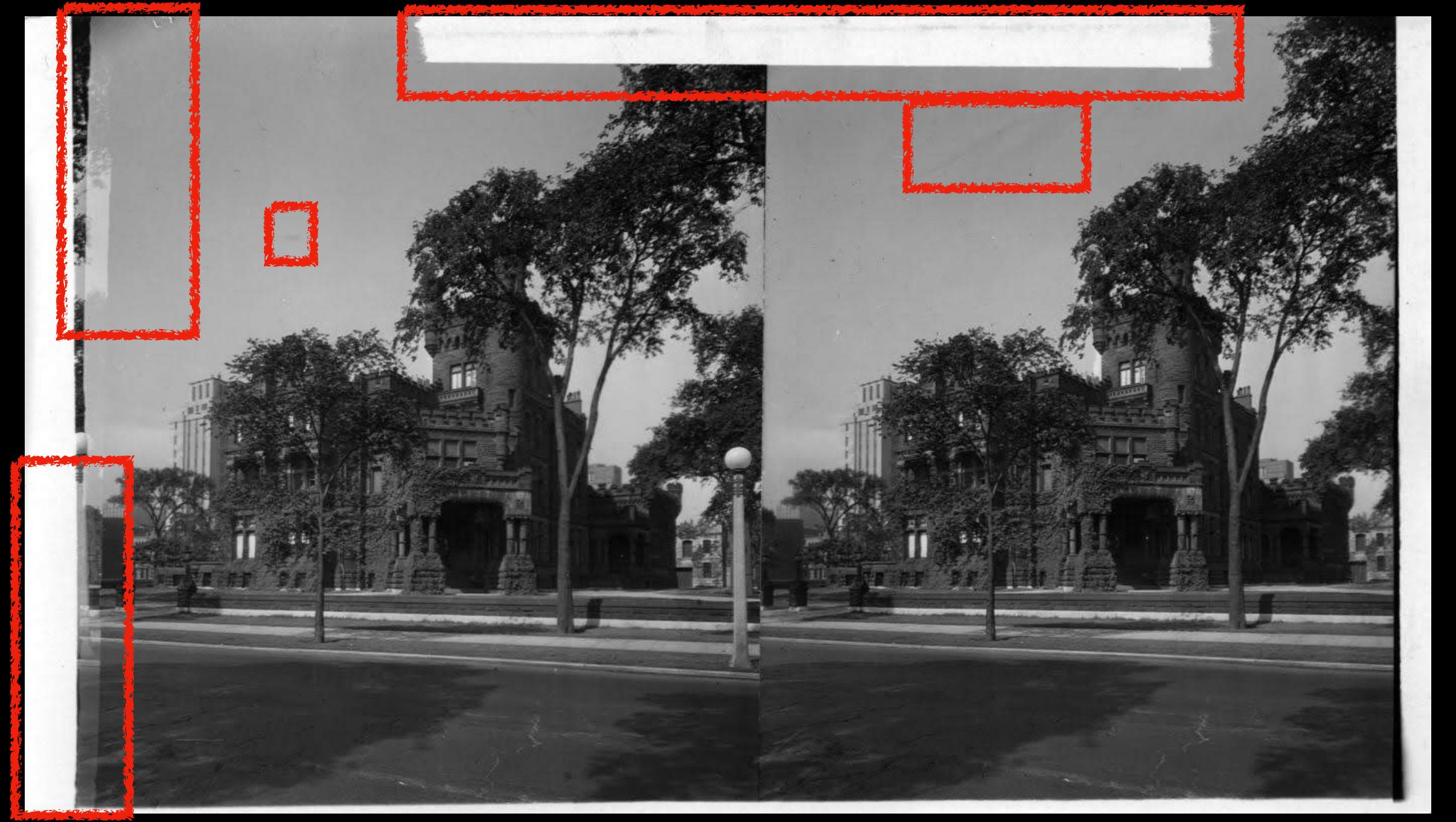
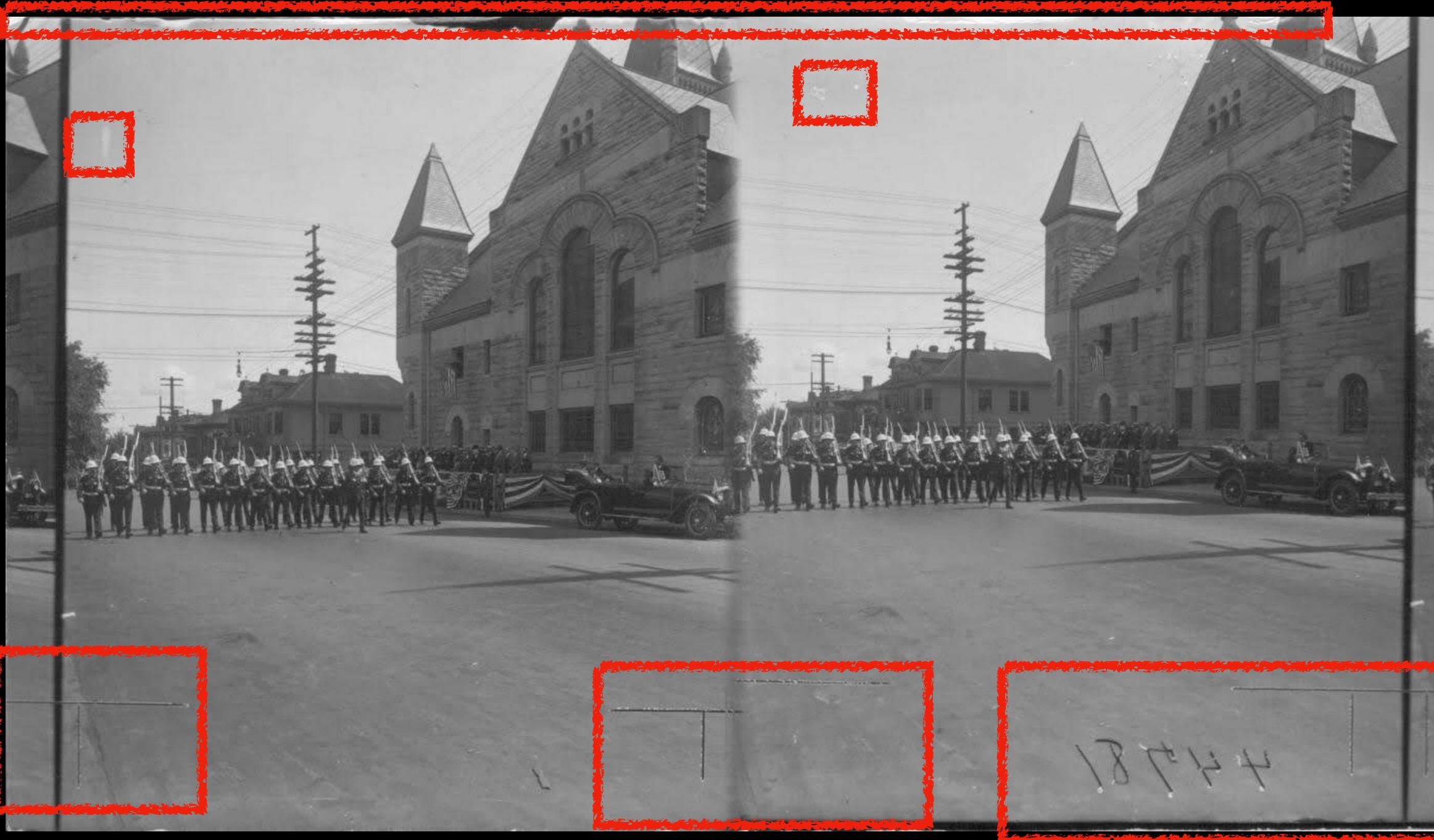
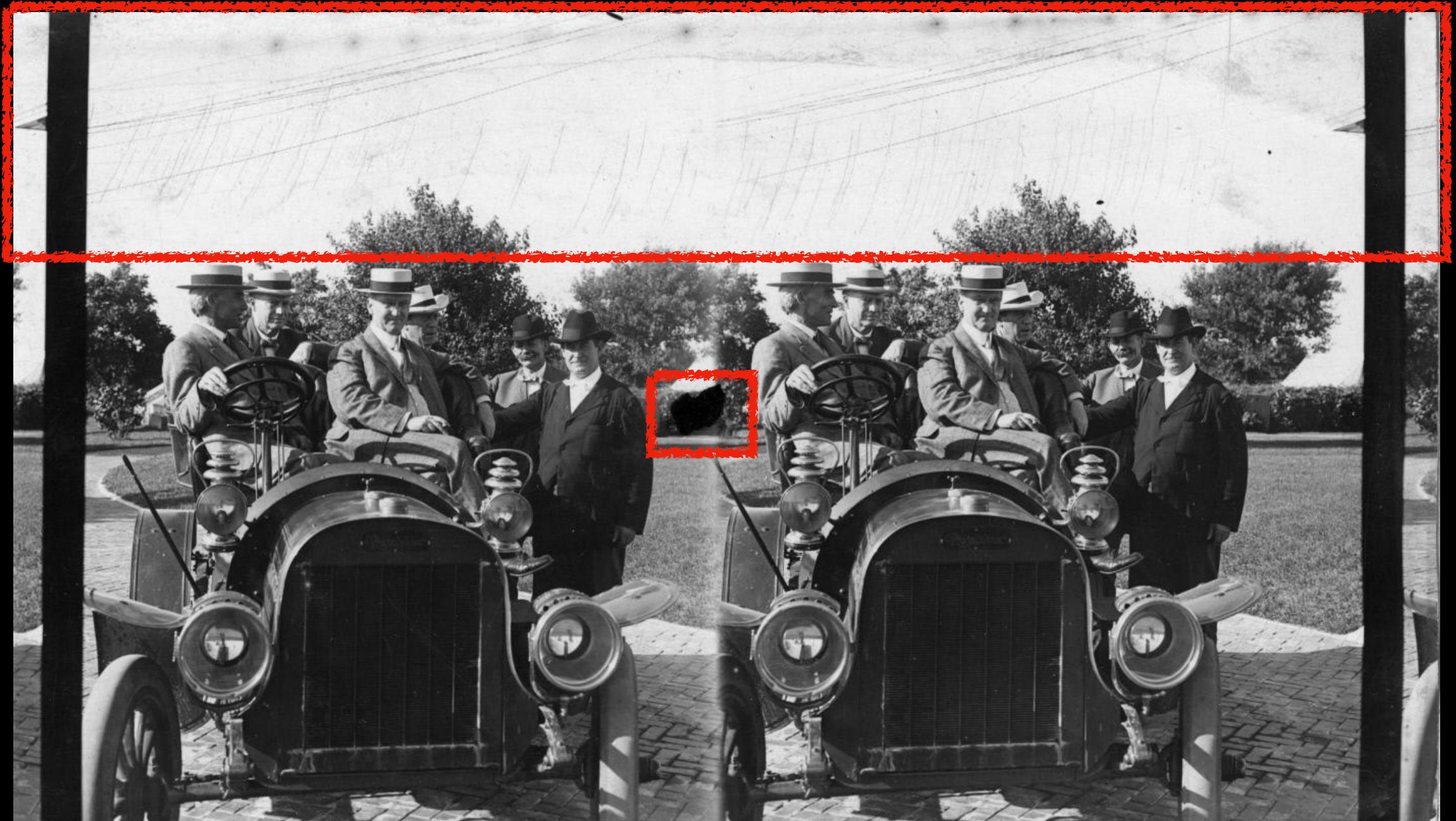
o.l.

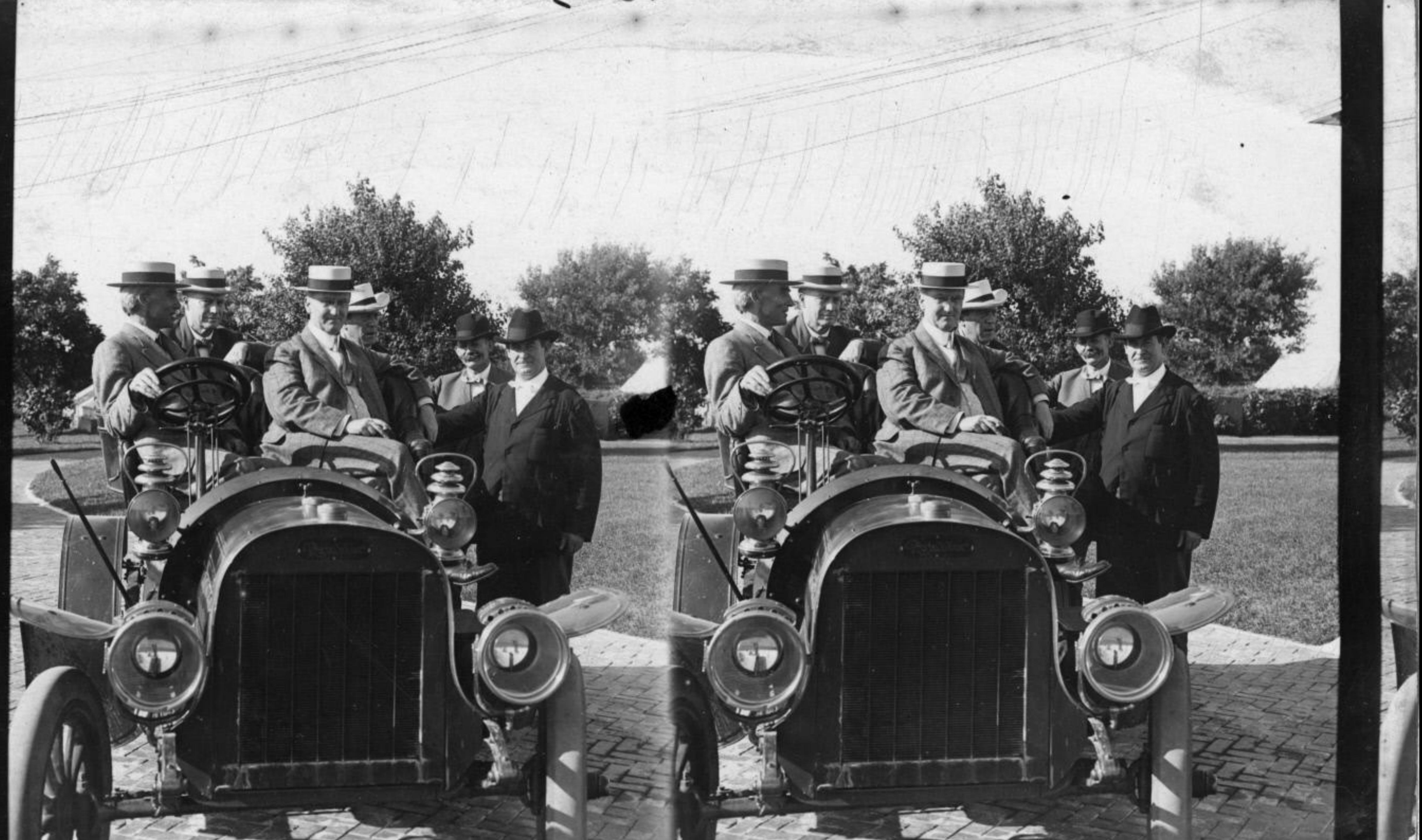


Thomas Edison
1847 - 1931



Remove Bad Images







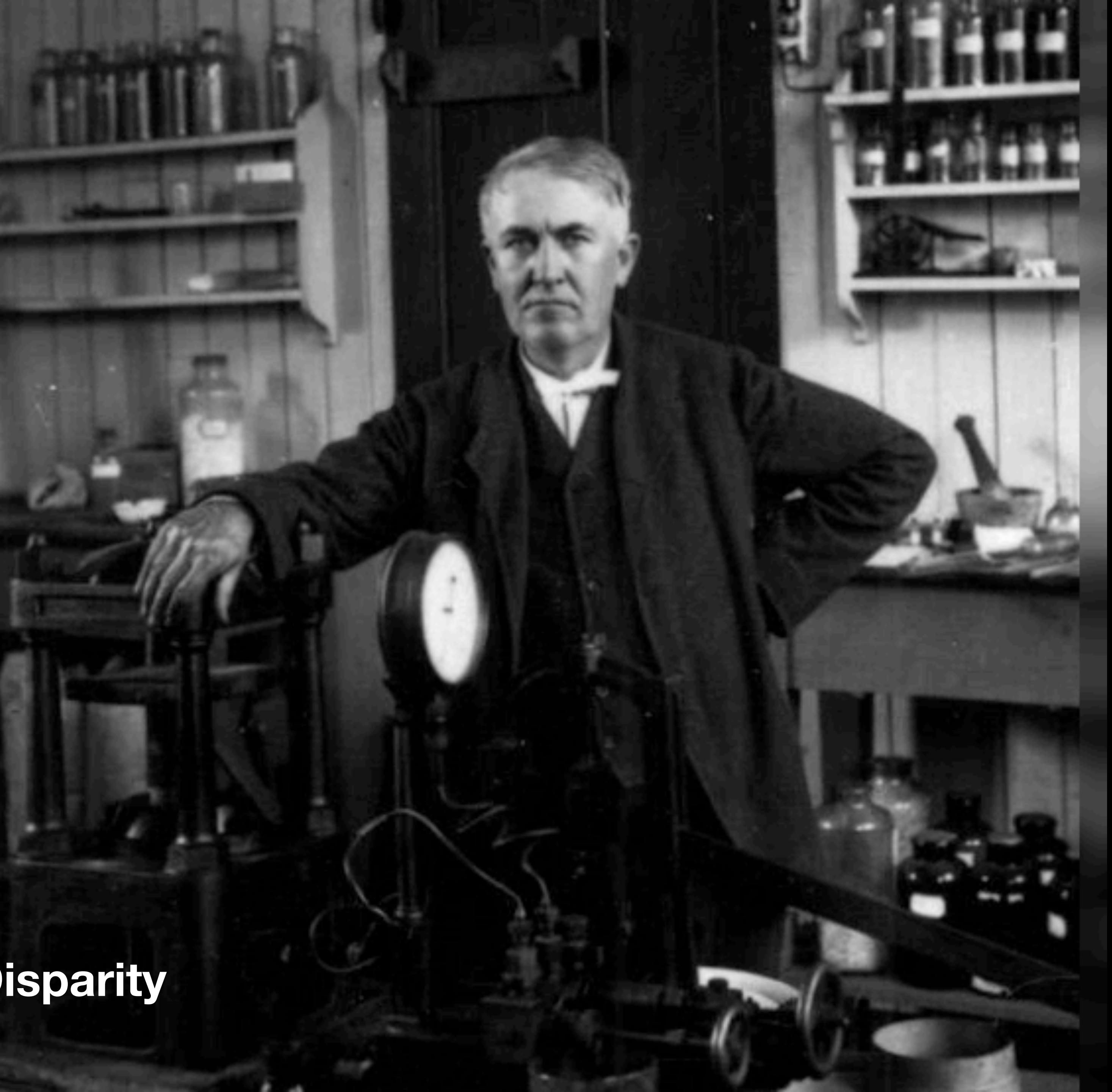




Cropped and Rectified Images



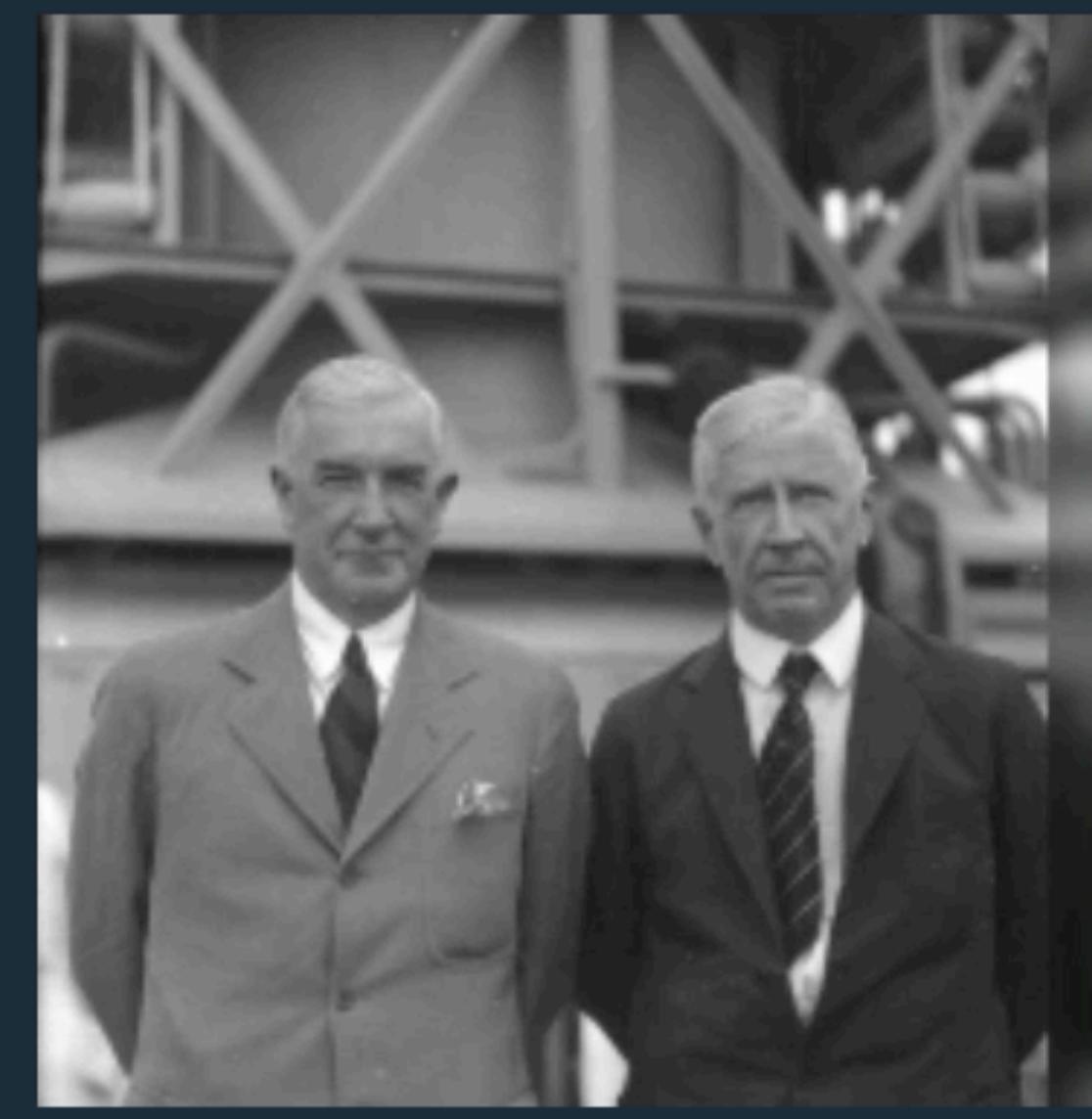
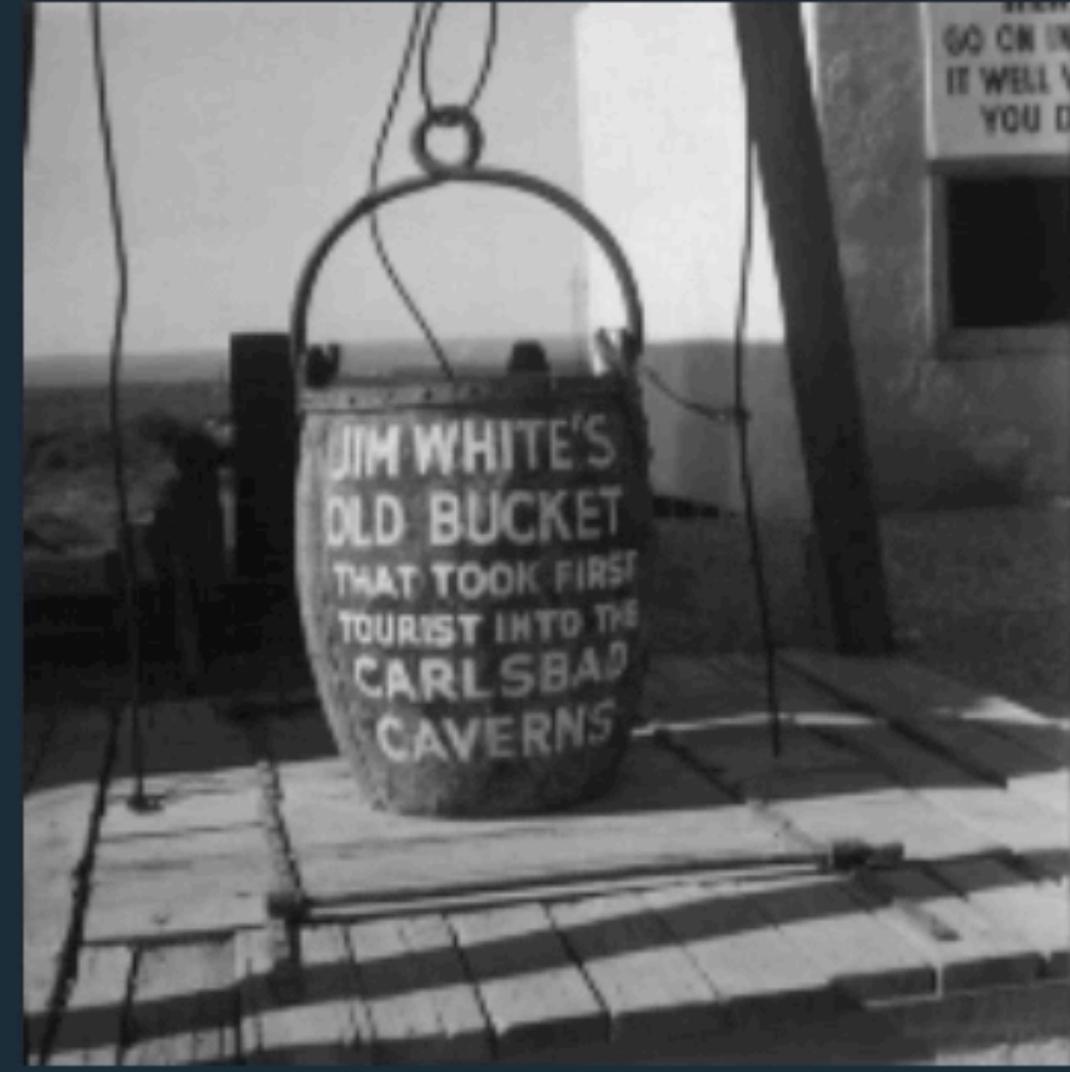
Cropped and Rectified Images



Estimated Disparity

Estimated Disparity





KeystoneDepth

- 37,239 entries
- Rectified antique stereo images
 - + Camera calibration
- Disparity maps
- Metadata (description, date, places, subjects)



KeystoneDepth





keystonedepth.cs.washington.edu



KeystoneDepth



<https://time-travel-rephotography.github.io/>

Time-Travel Rephotography



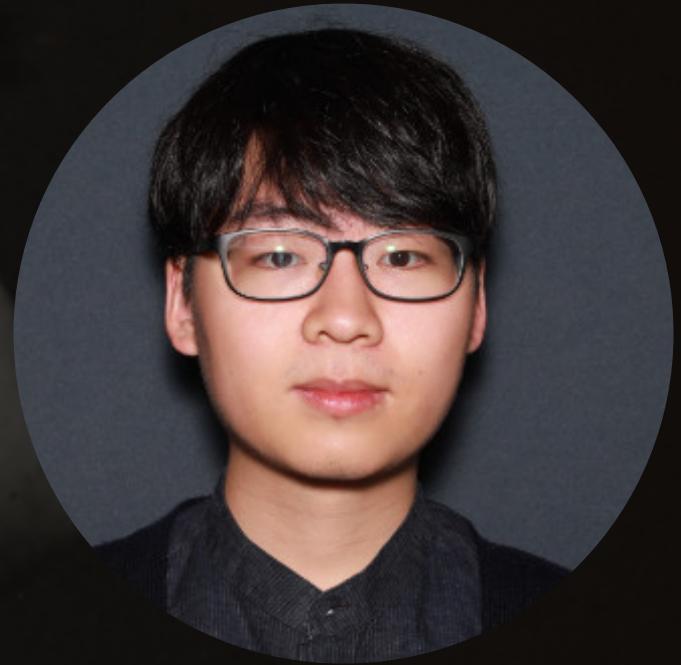
Xuan Luo

University of Washington



Cecilia Zhang

UC Berkeley



Paul Yoo

University of Washington



Ricardo Martin-Brualla

Google Research



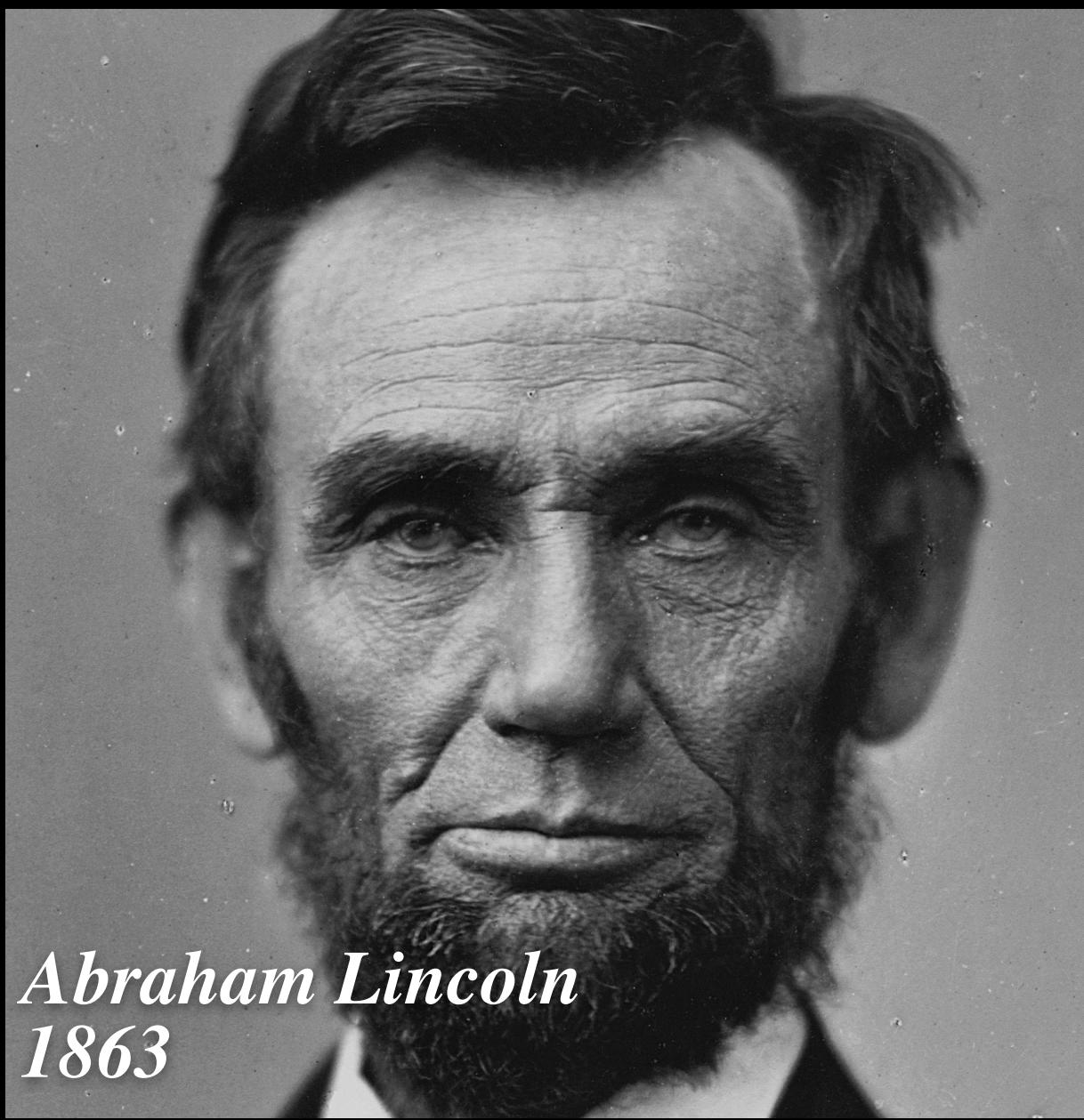
Jason Lawrence

Google Research

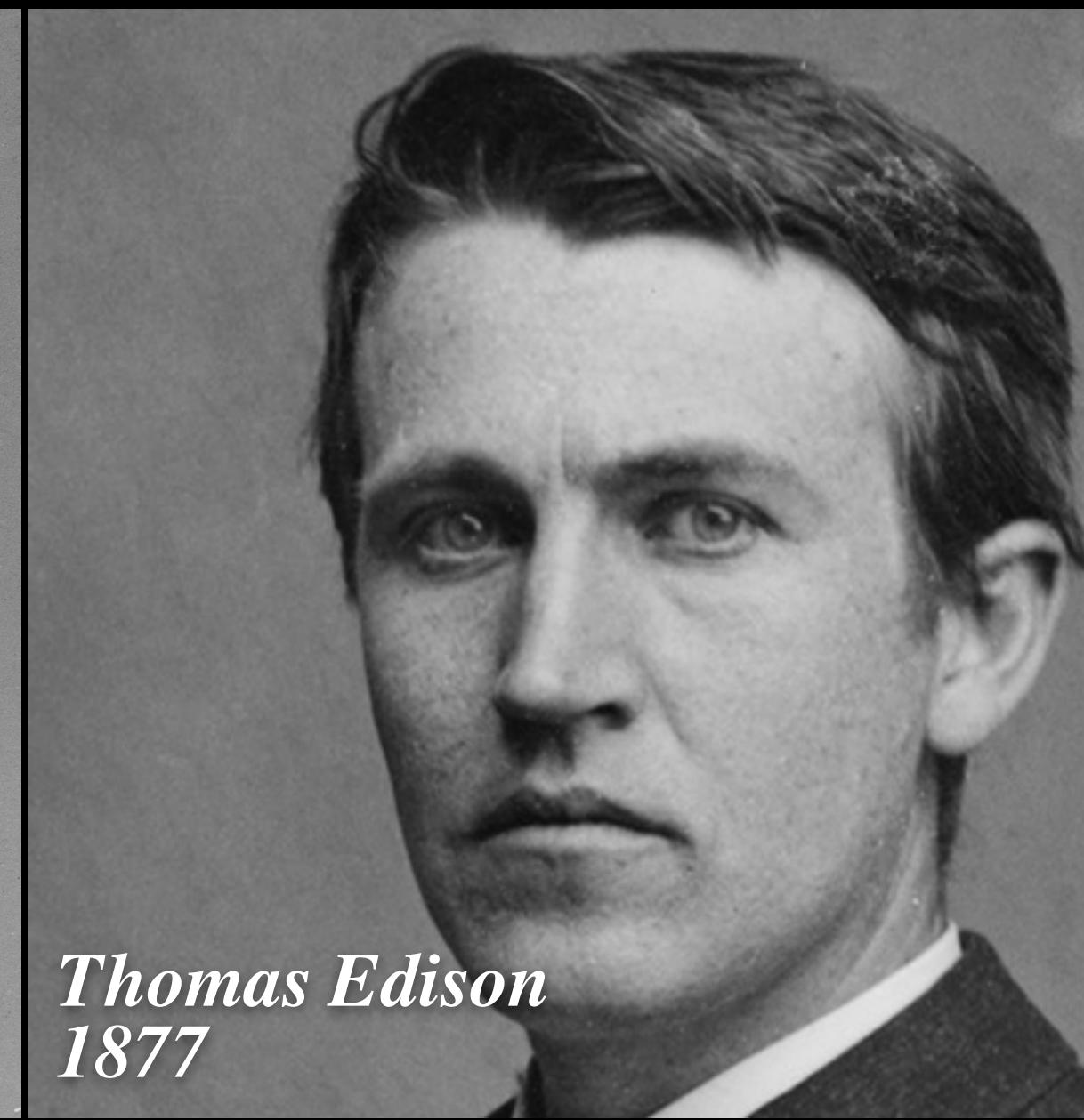


Steven M. Seitz

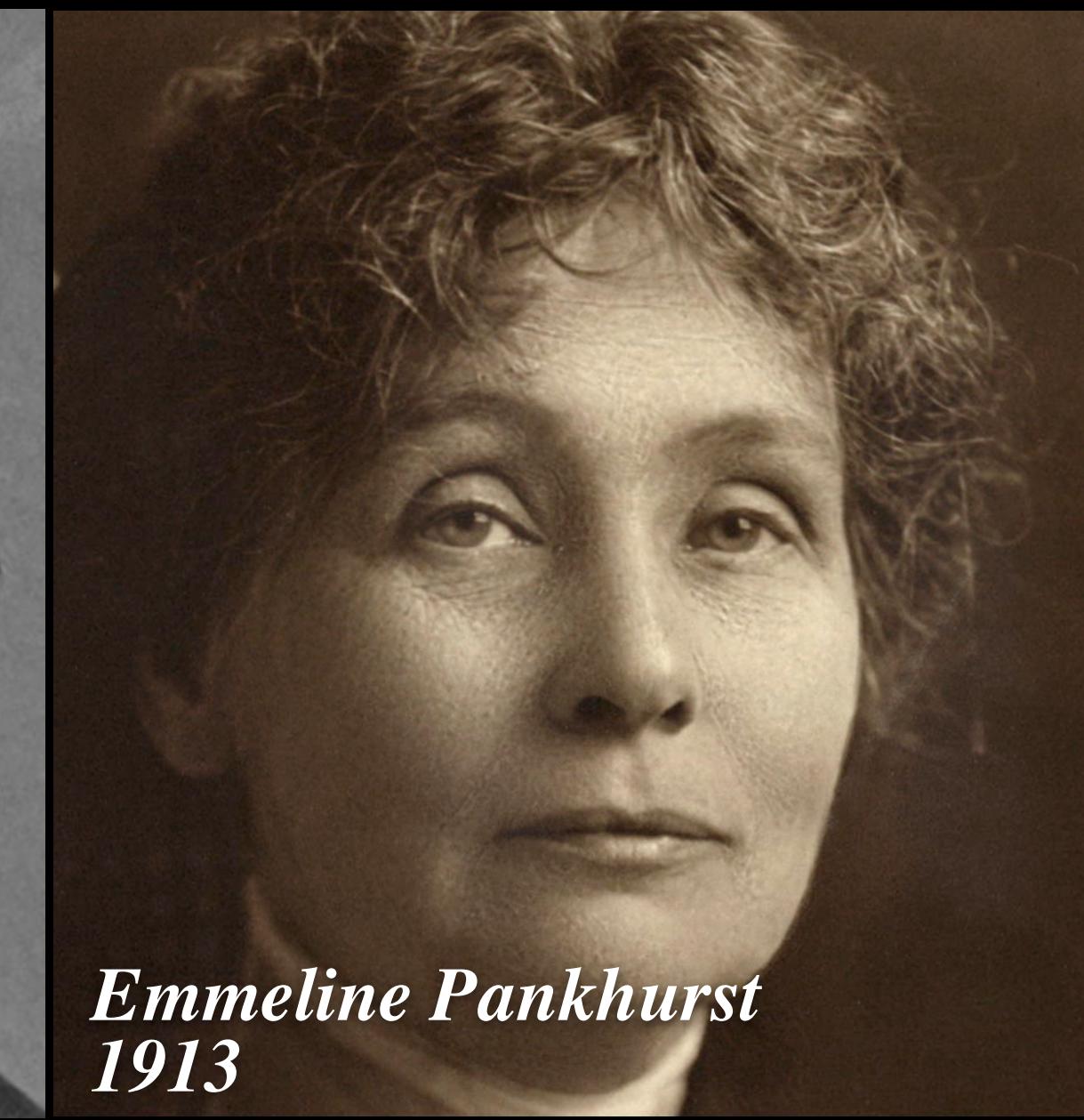
University of Washington
Google Research



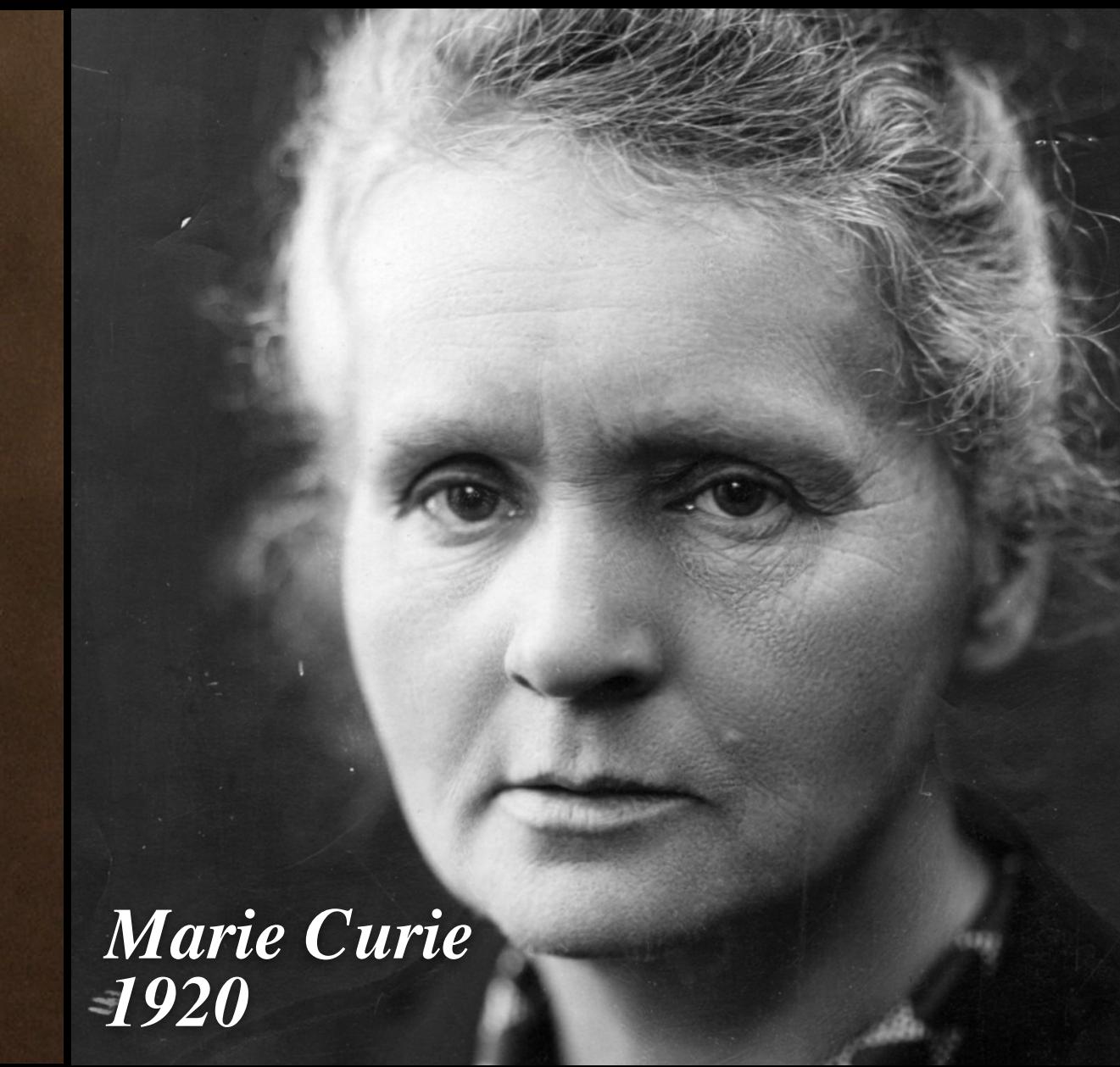
Abraham Lincoln
1863



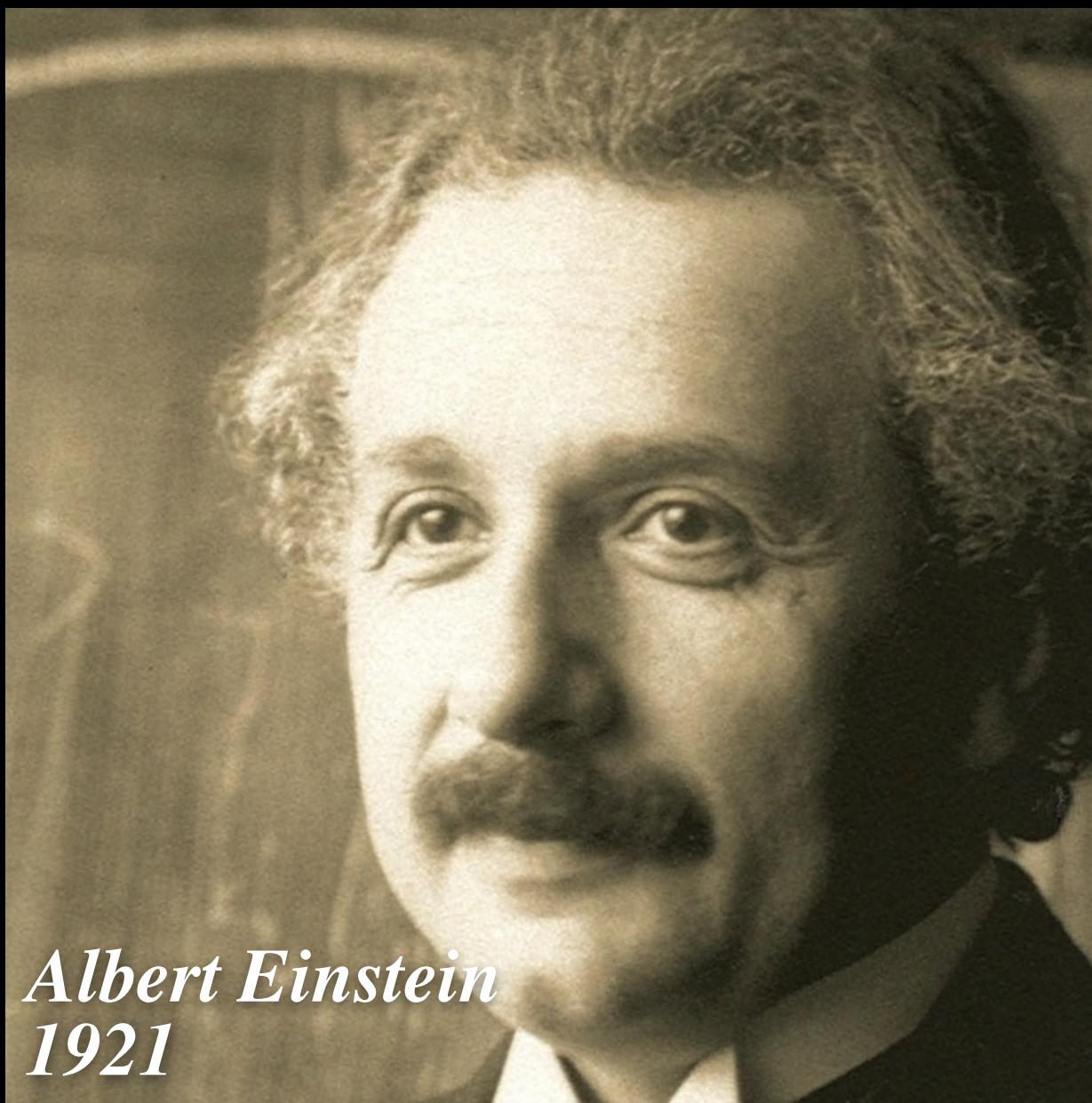
Thomas Edison
1877



Emmeline Pankhurst
1913



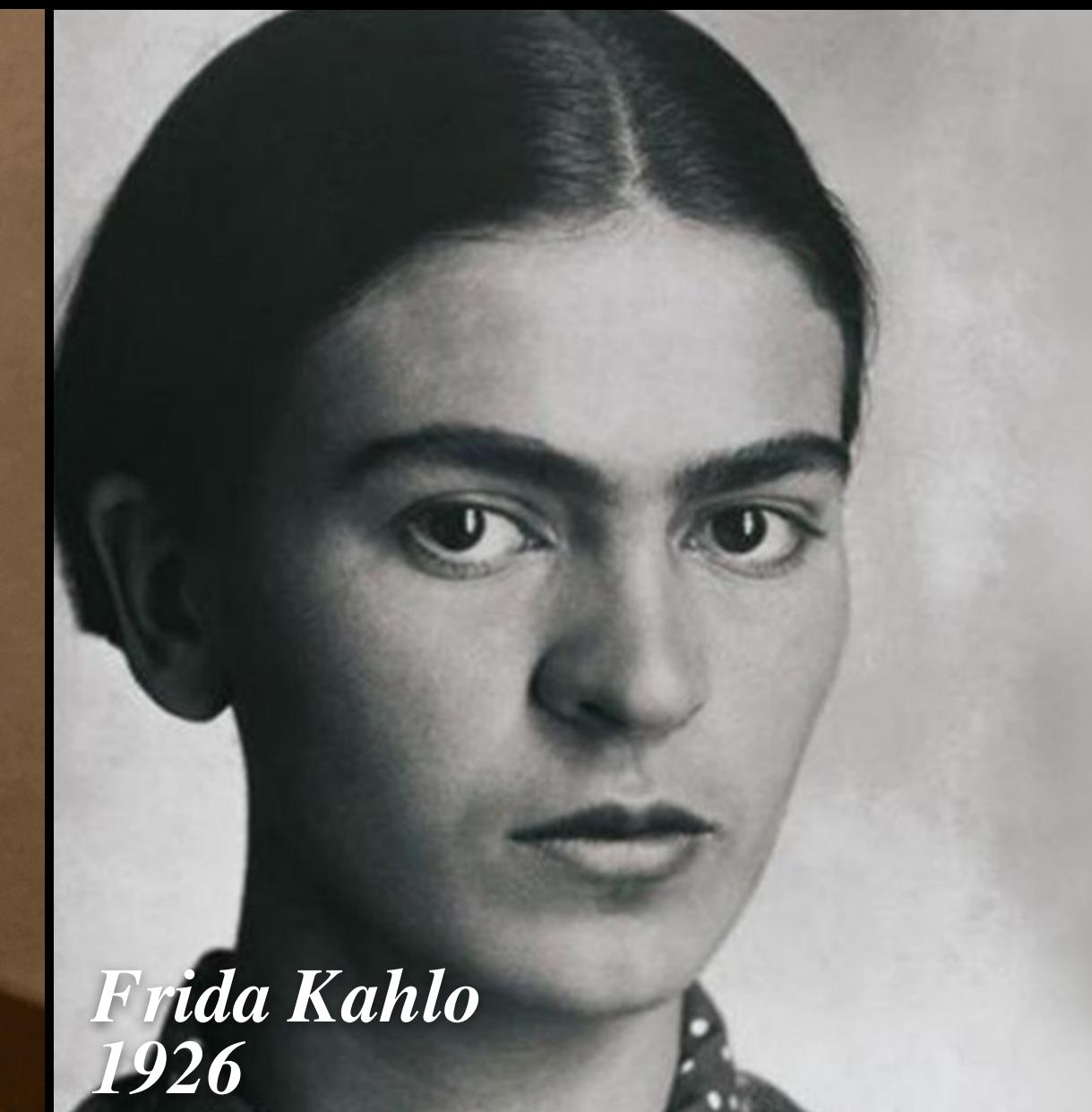
Marie Curie
1920



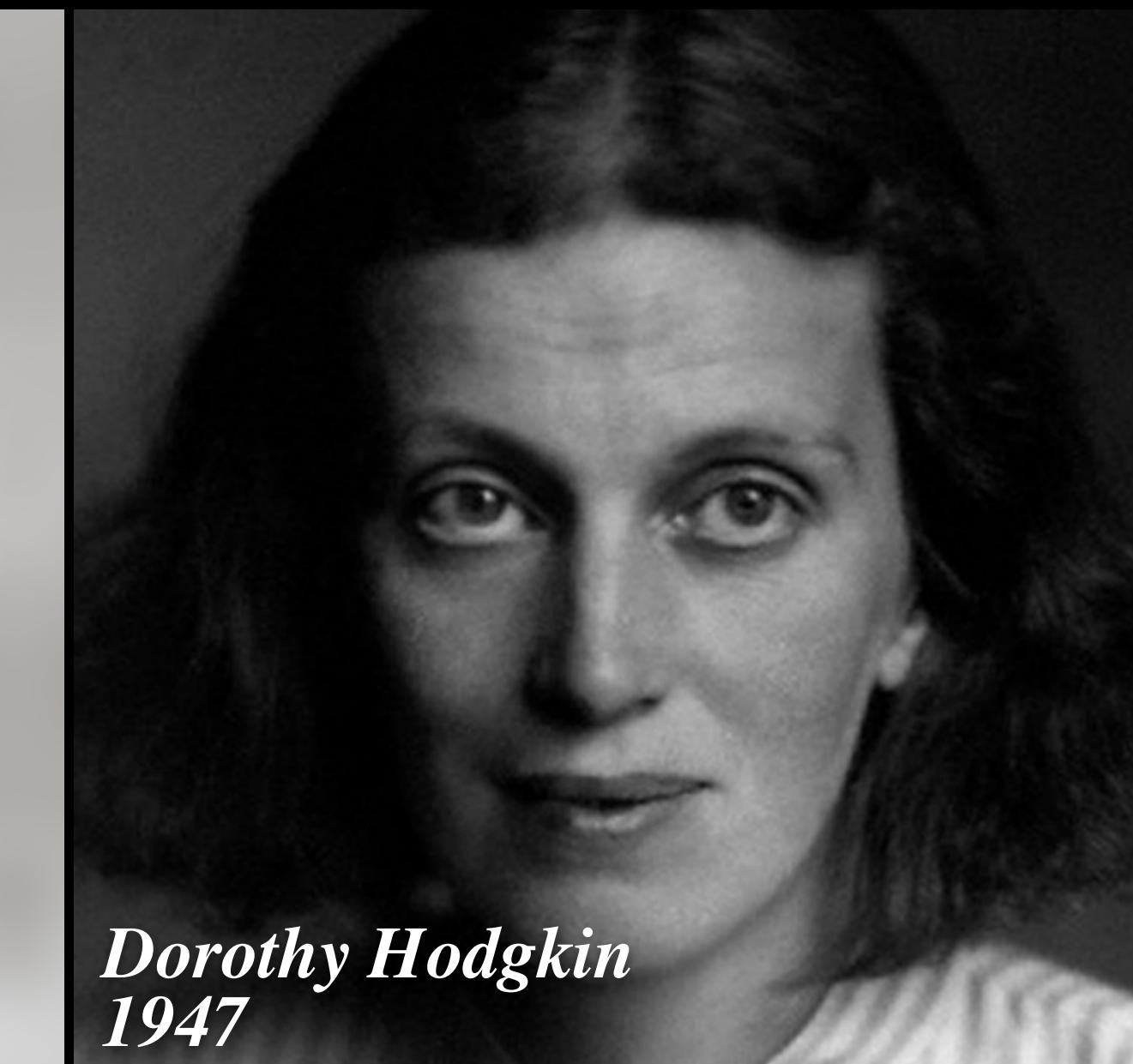
Albert Einstein
1921



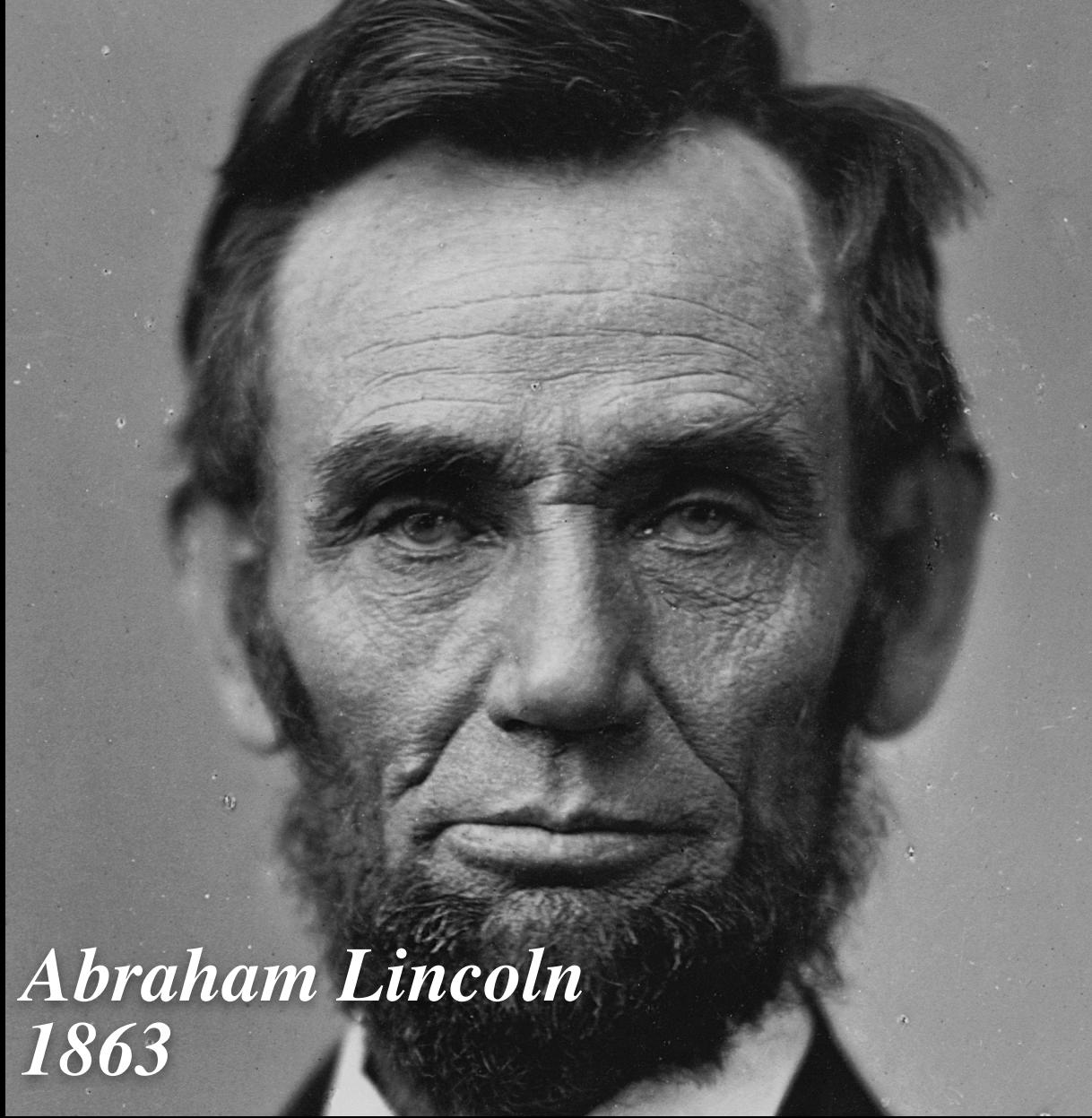
Franz Kafka
1923



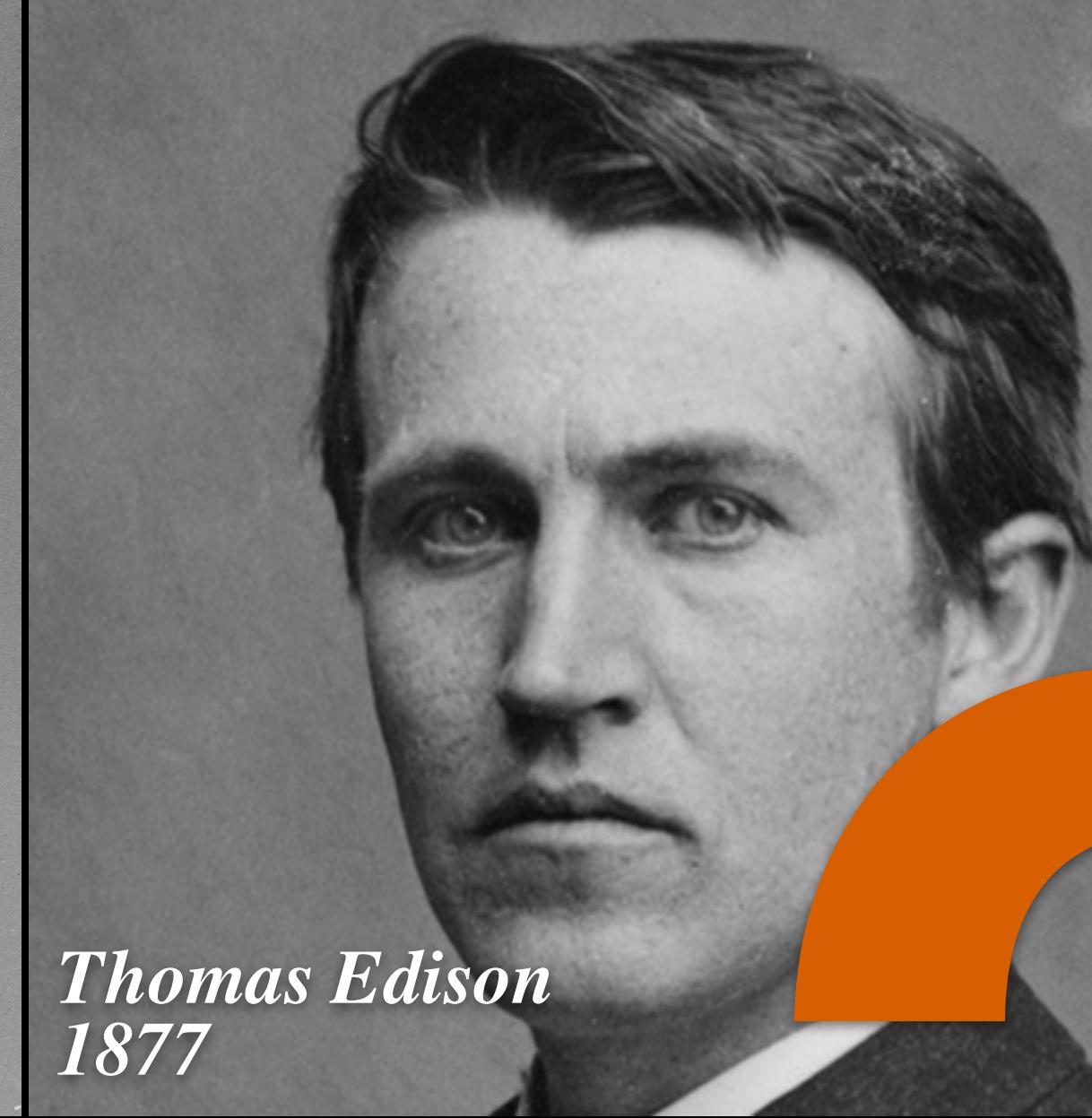
Frida Kahlo
1926



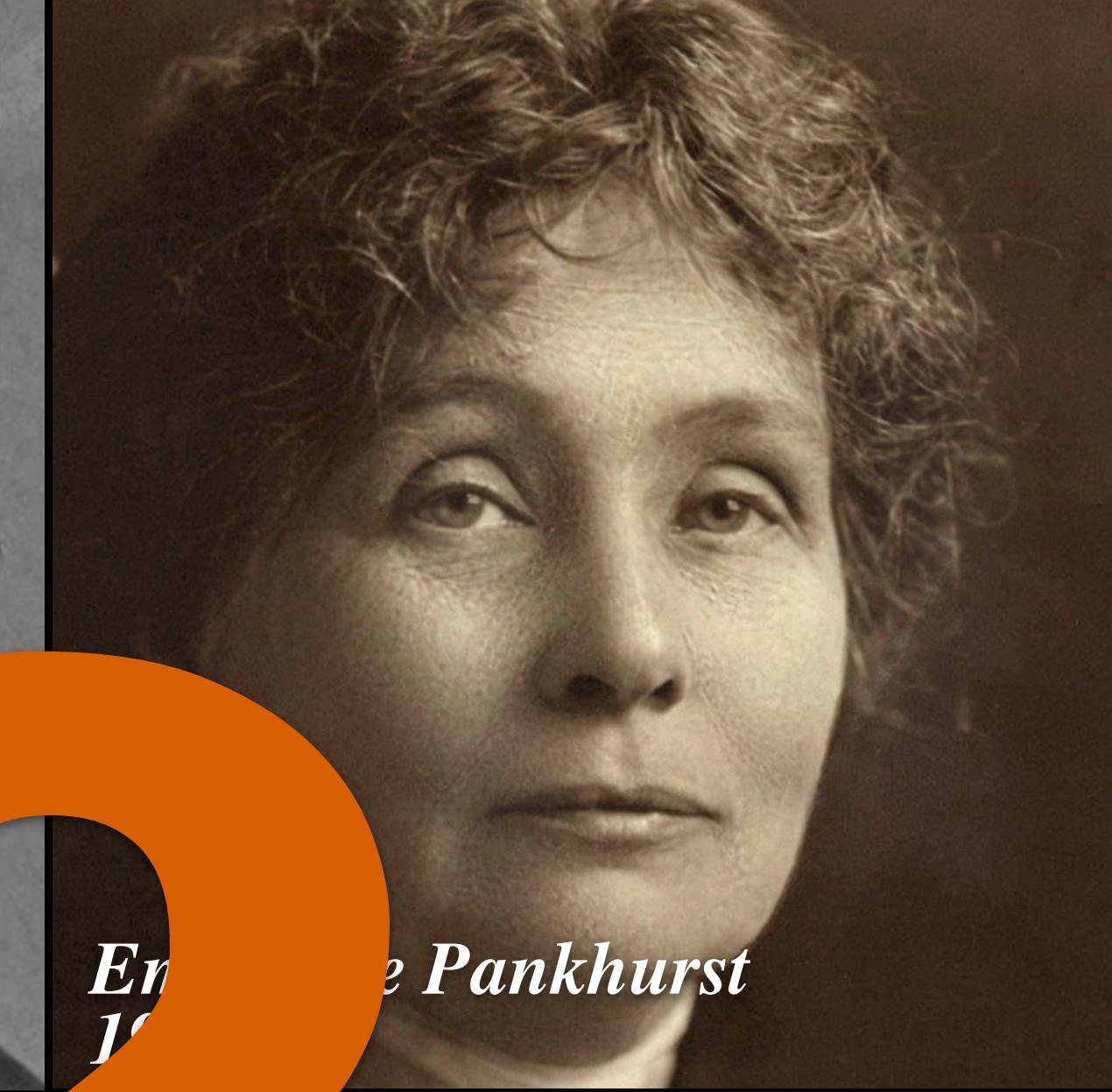
Dorothy Hodgkin
1947



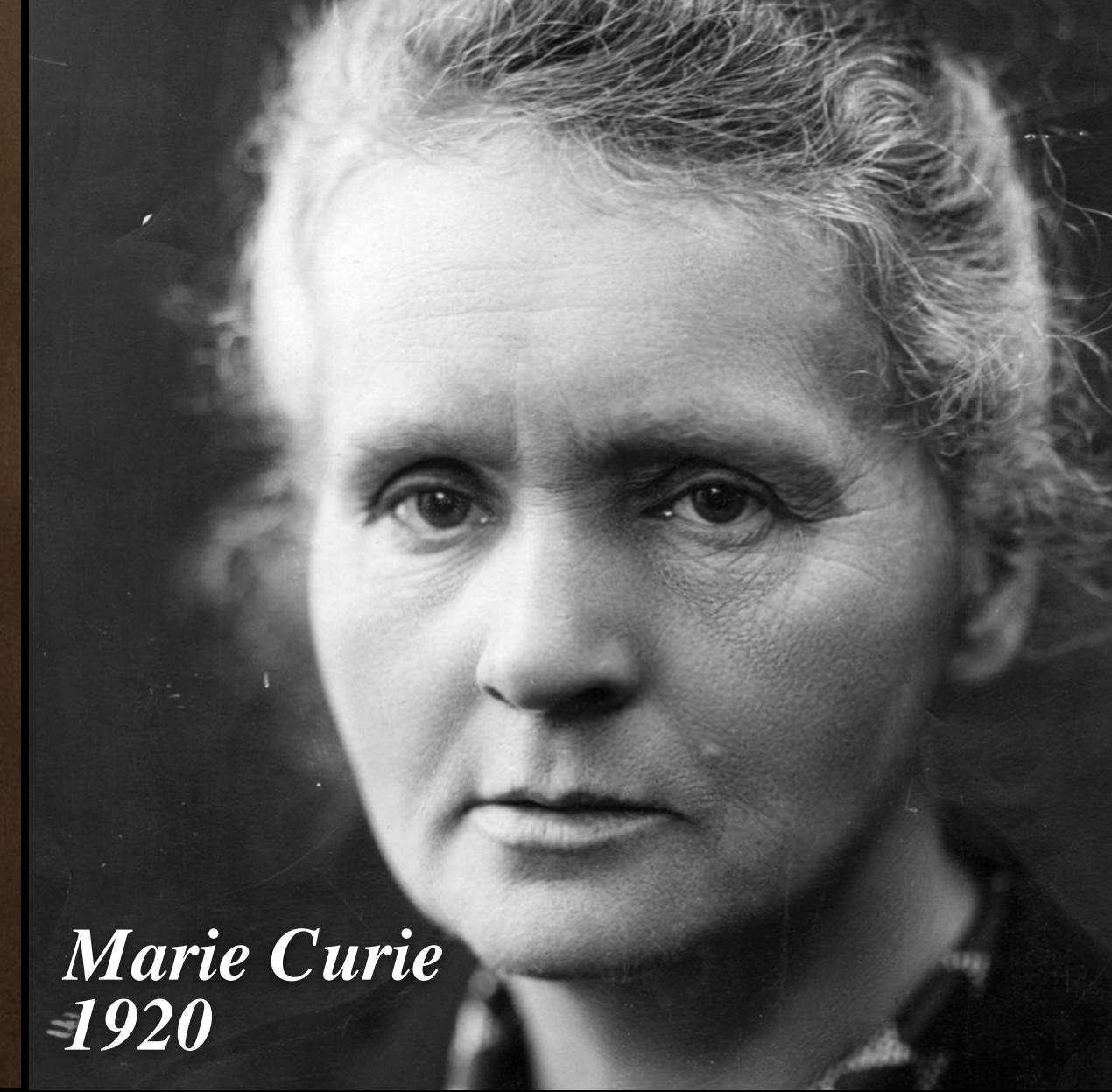
Abraham Lincoln
1863



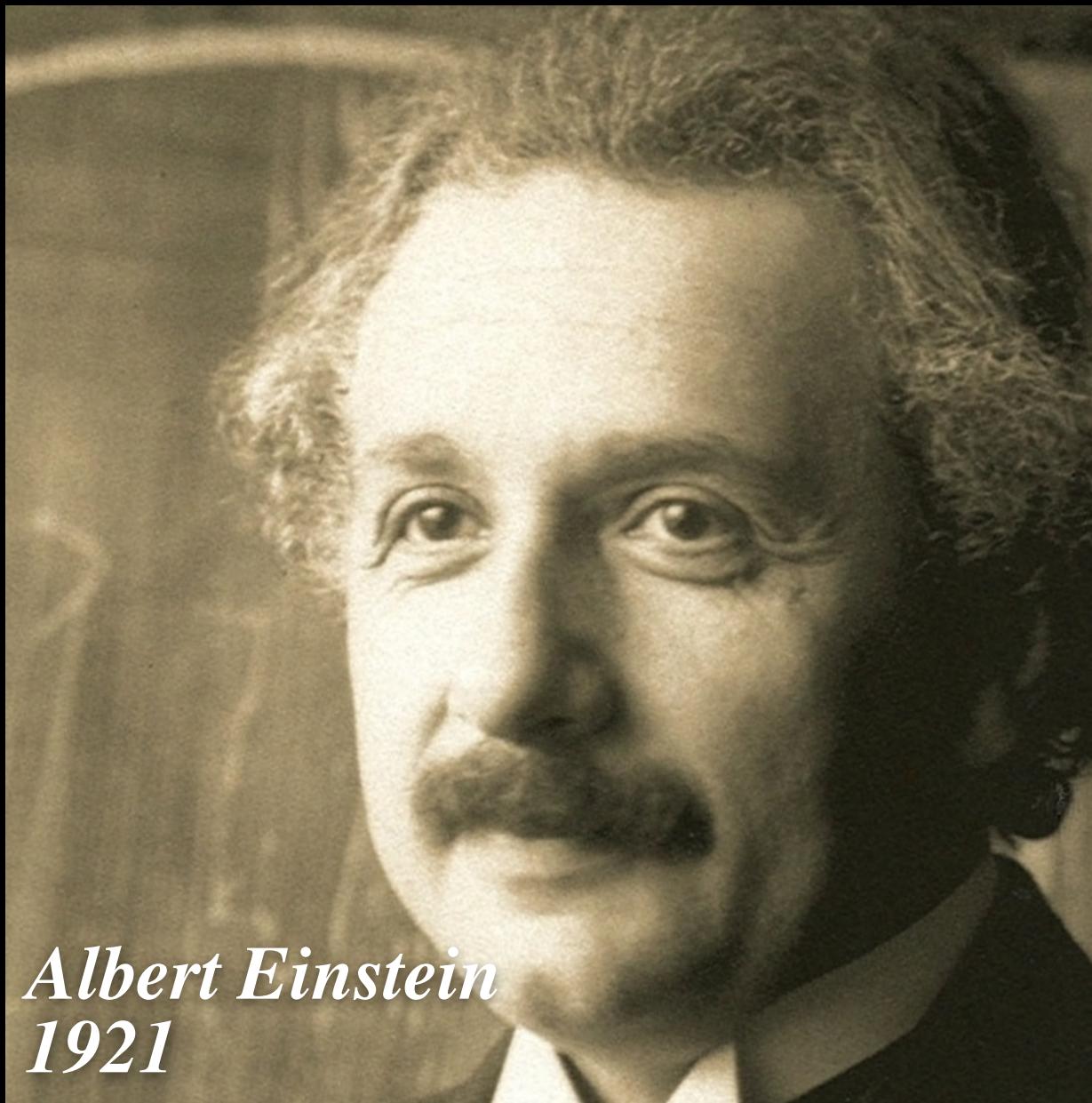
Thomas Edison
1877



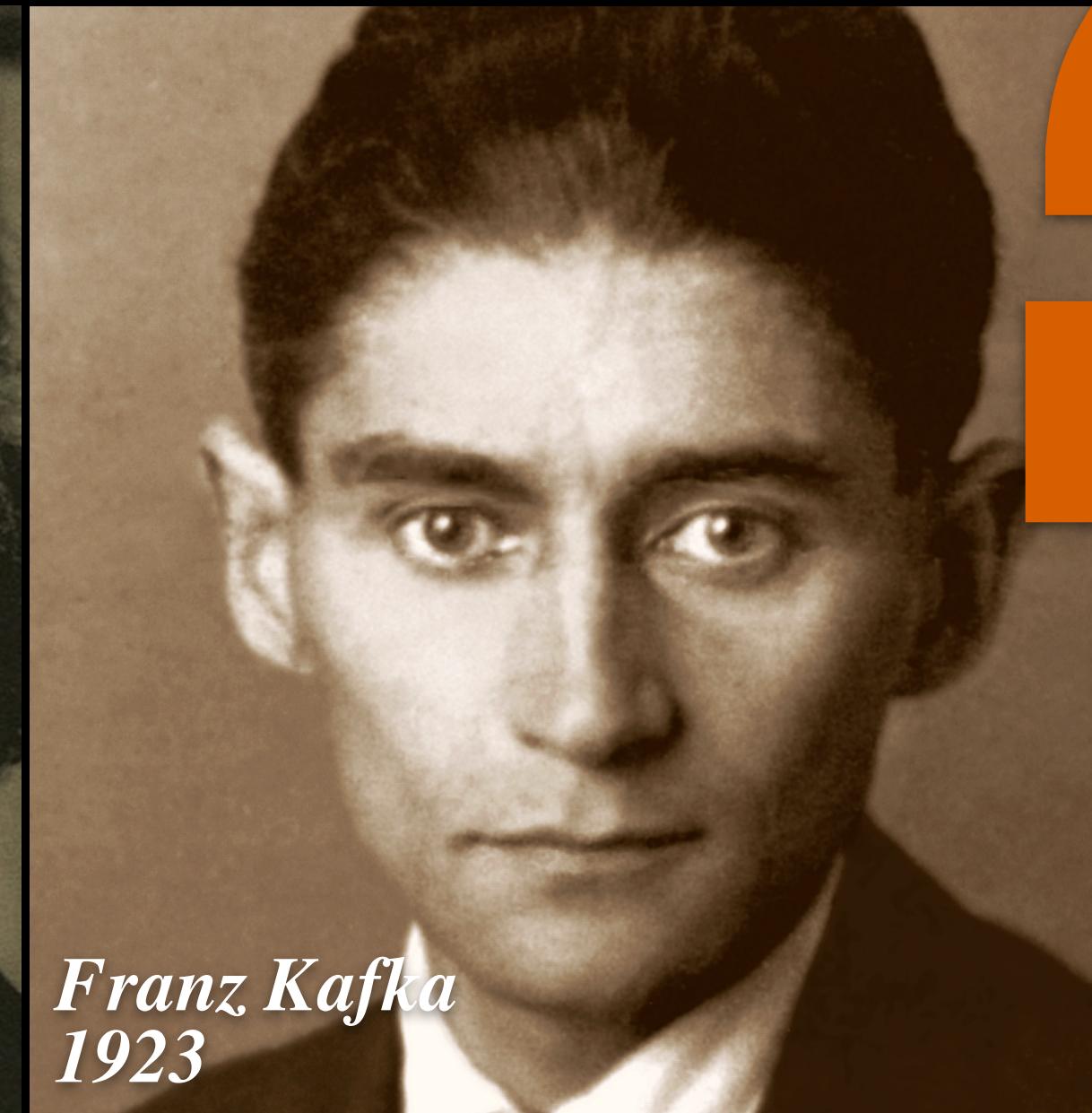
Emily Pankhurst
1908



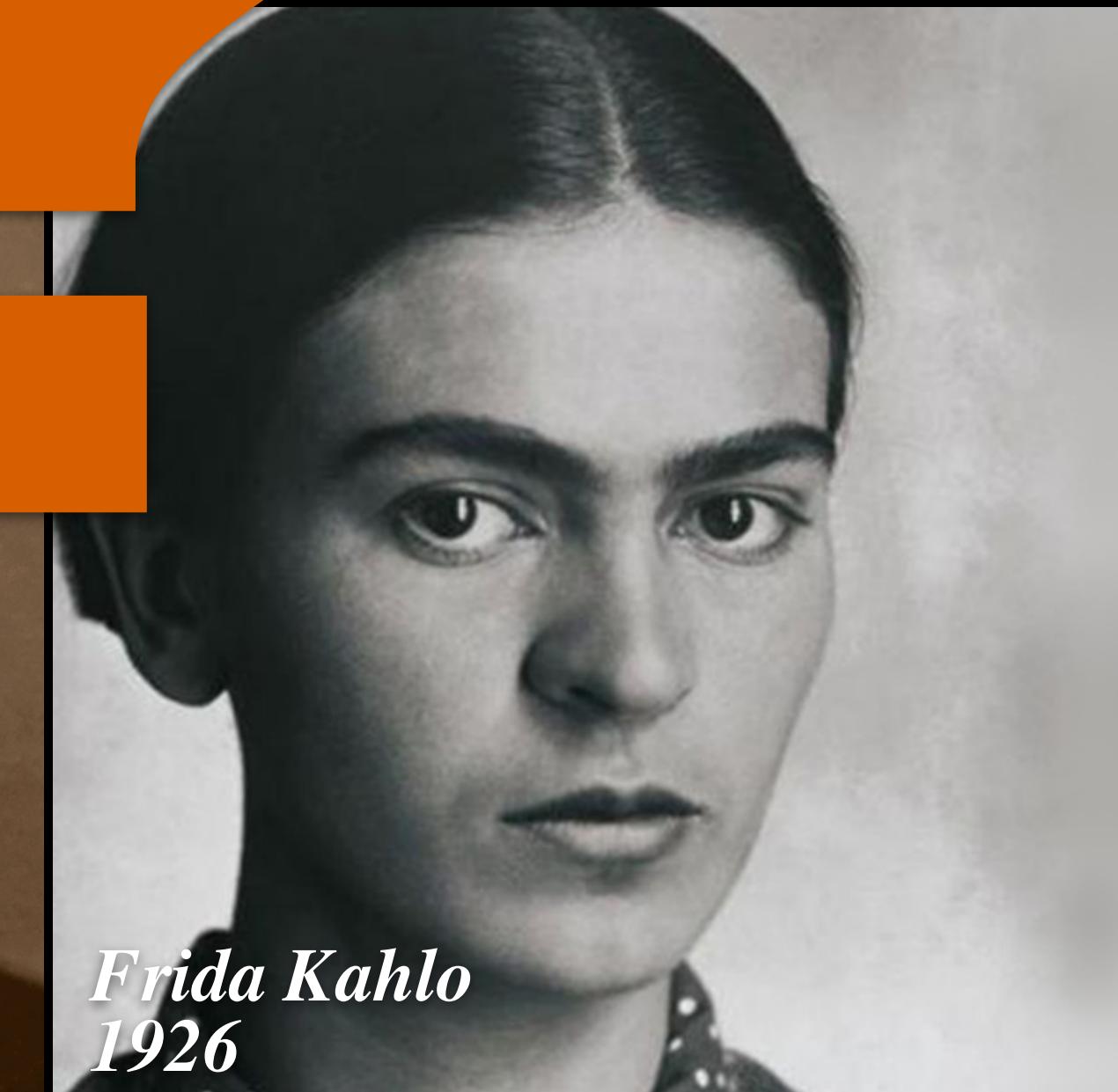
Marie Curie
1920



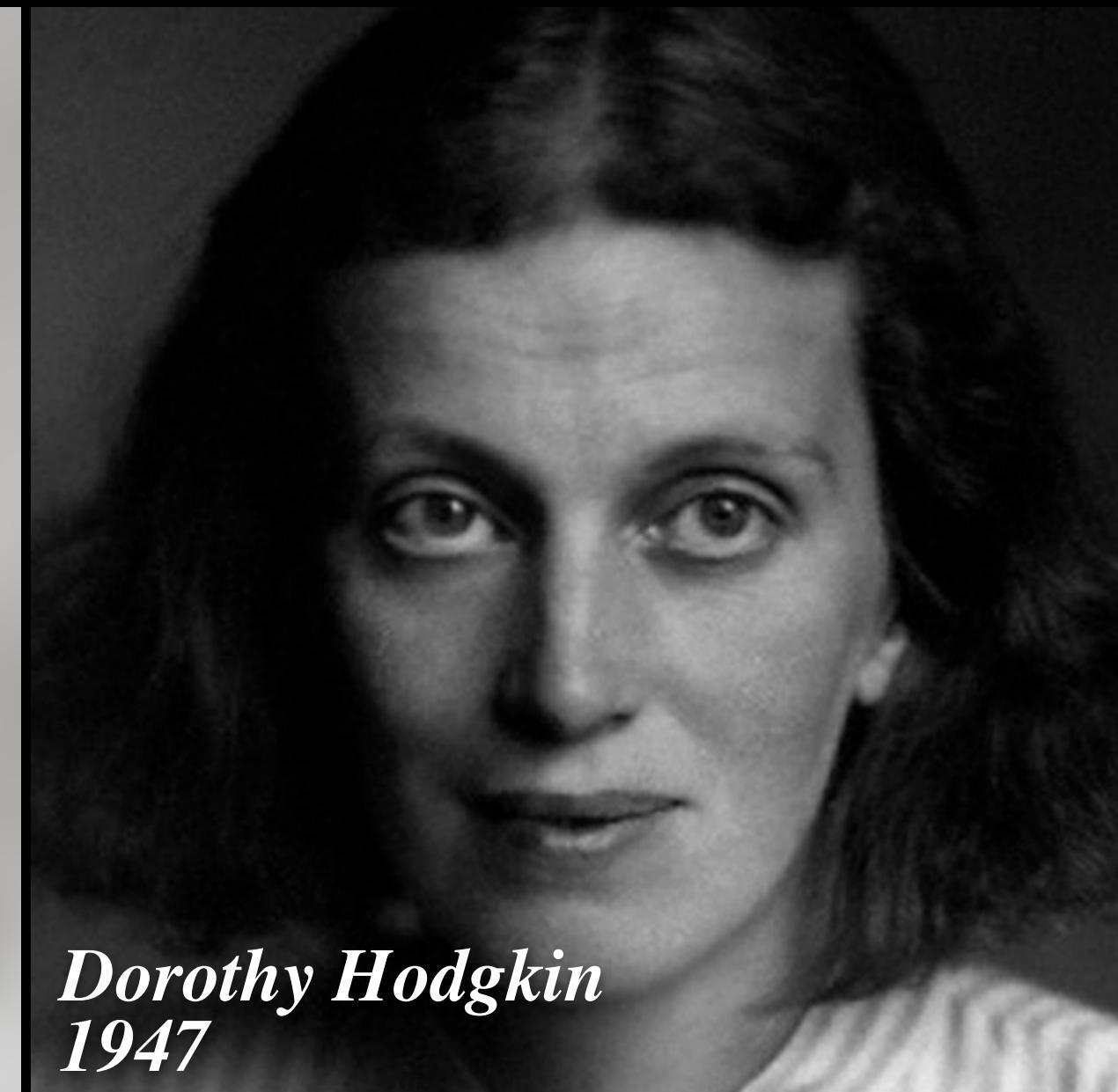
Albert Einstein
1921



Franz Kafka
1923

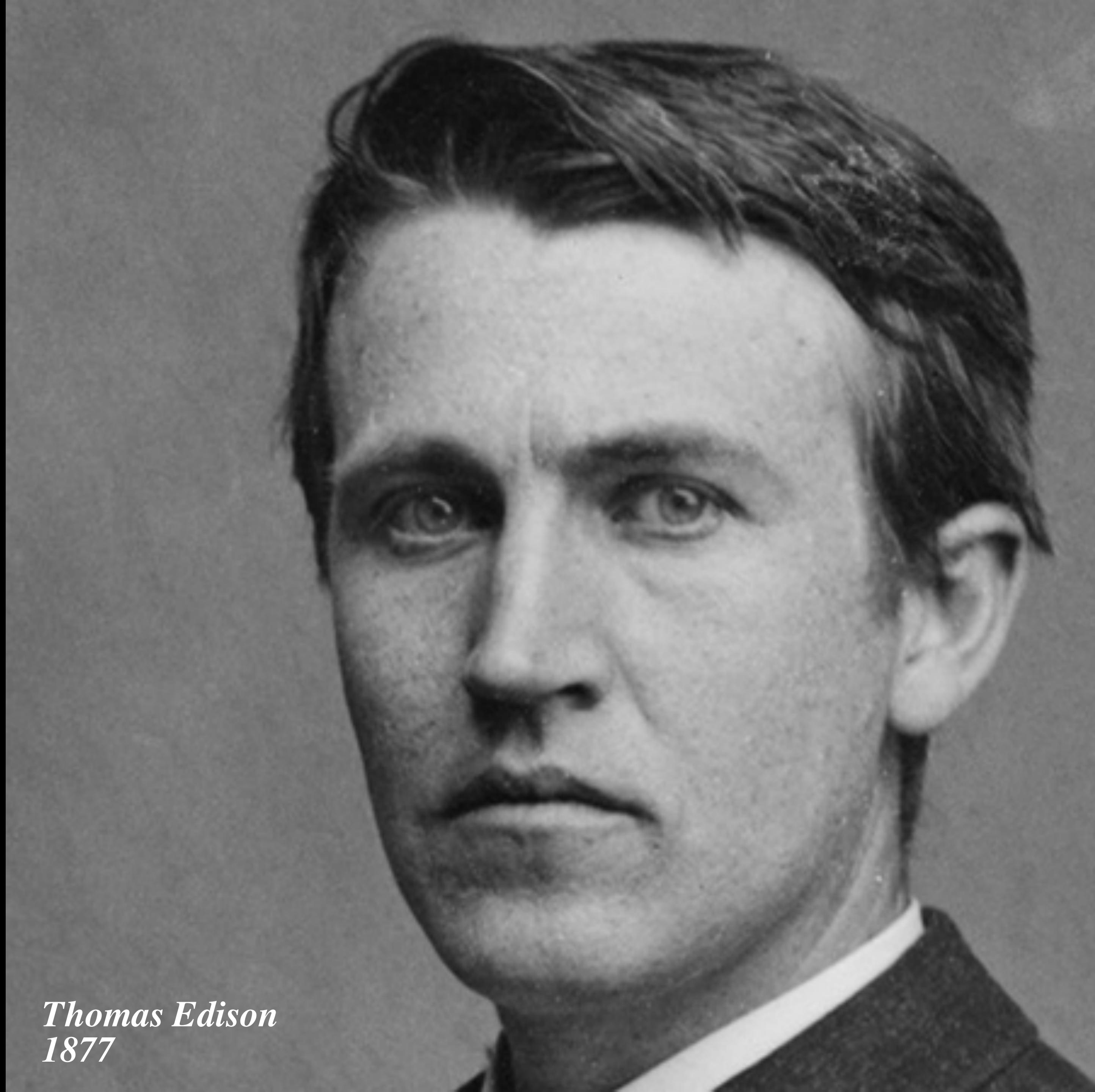


Frida Kahlo
1926

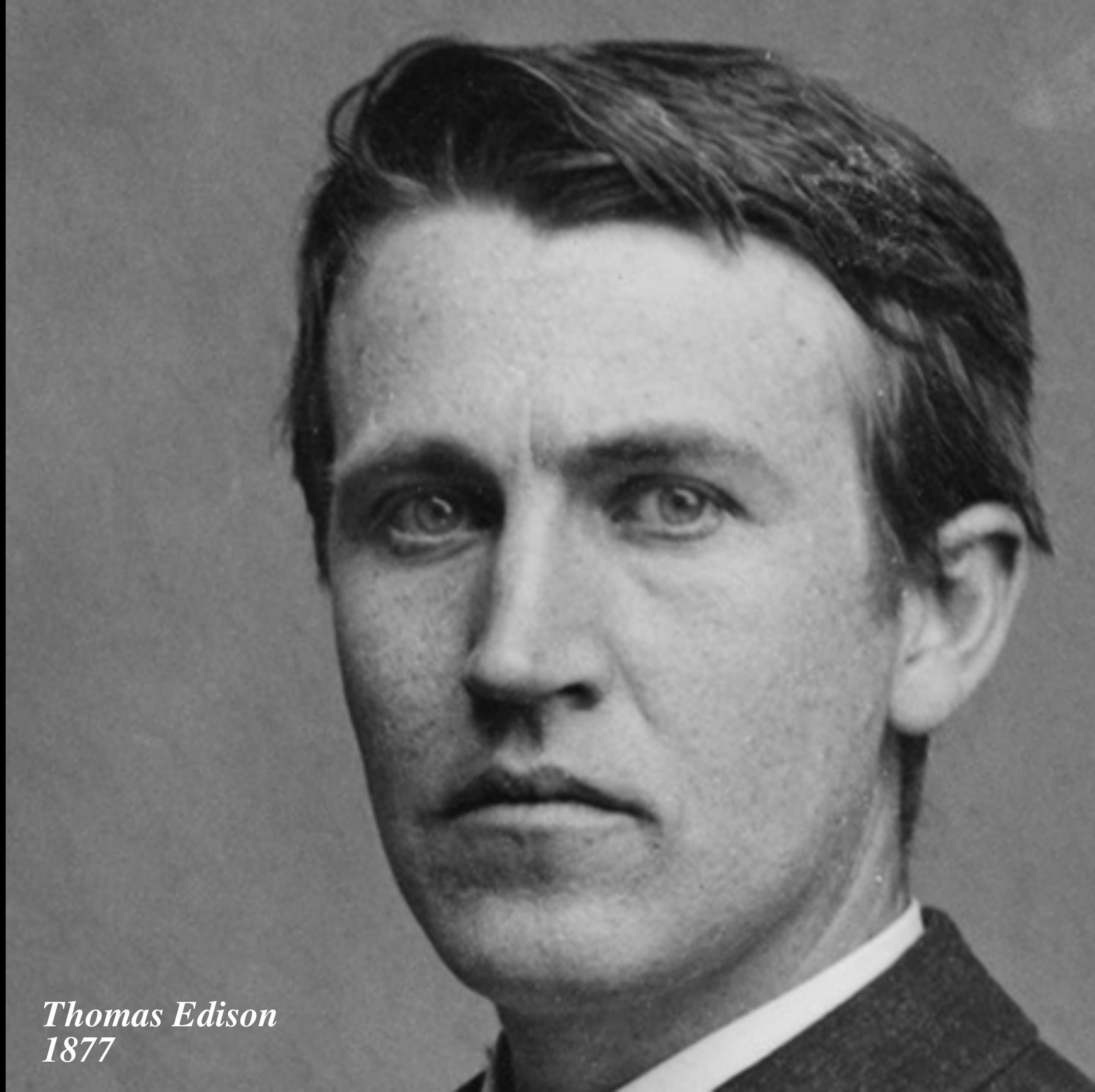


Dorothy Hodgkin
1947

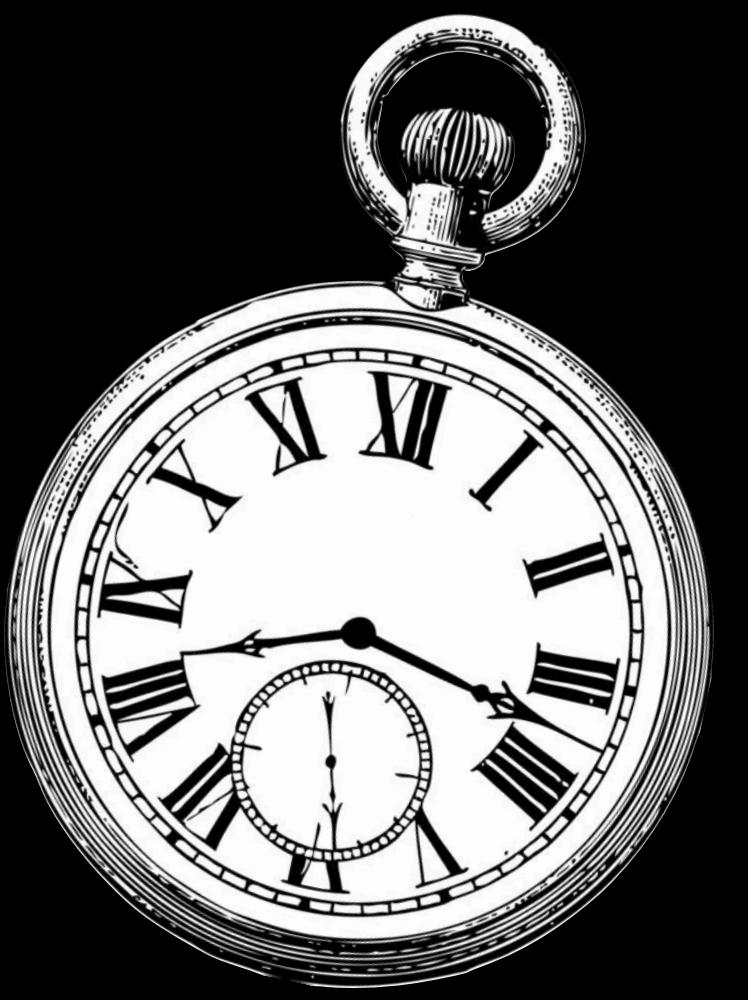




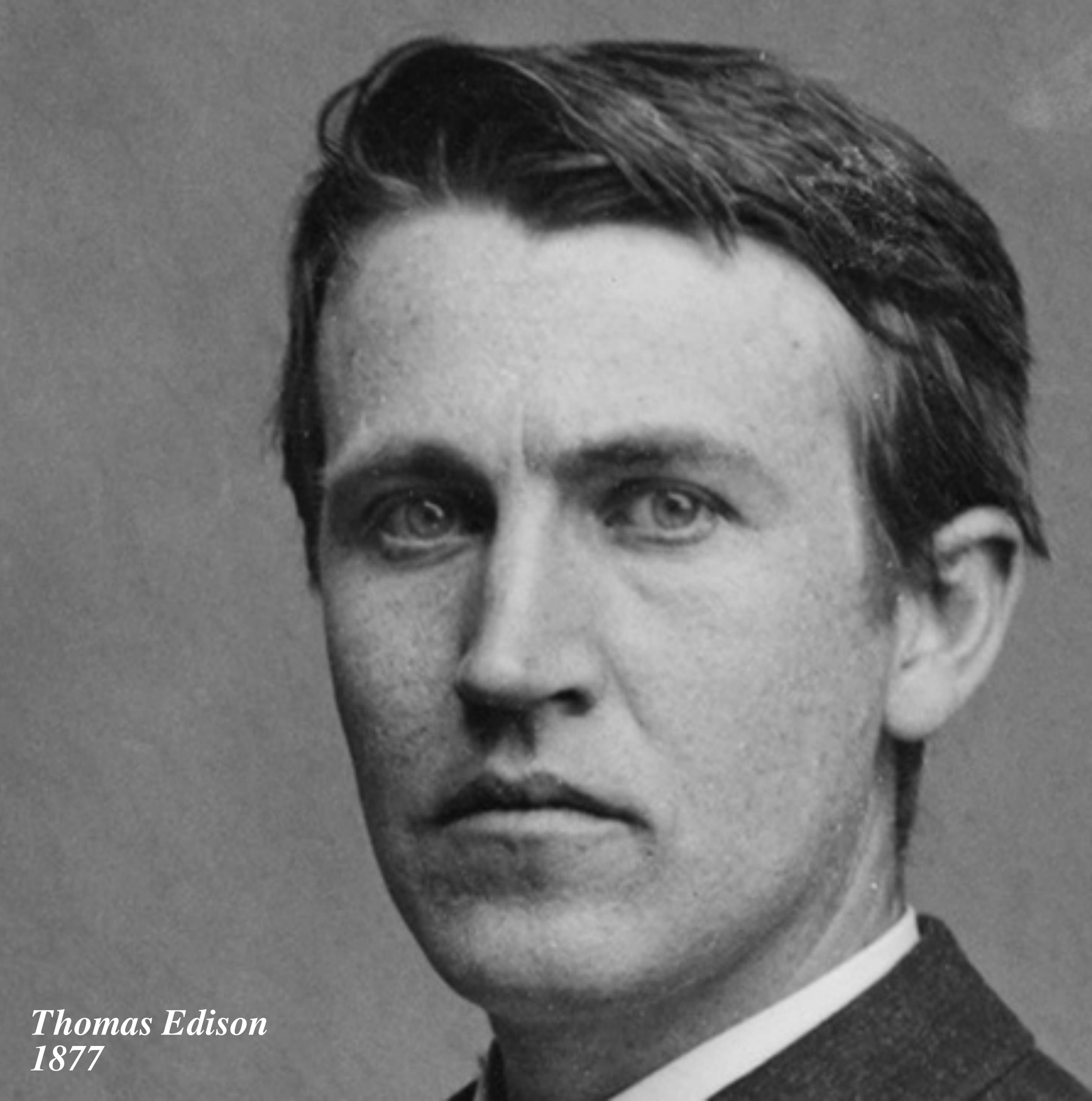
Thomas Edison
1877

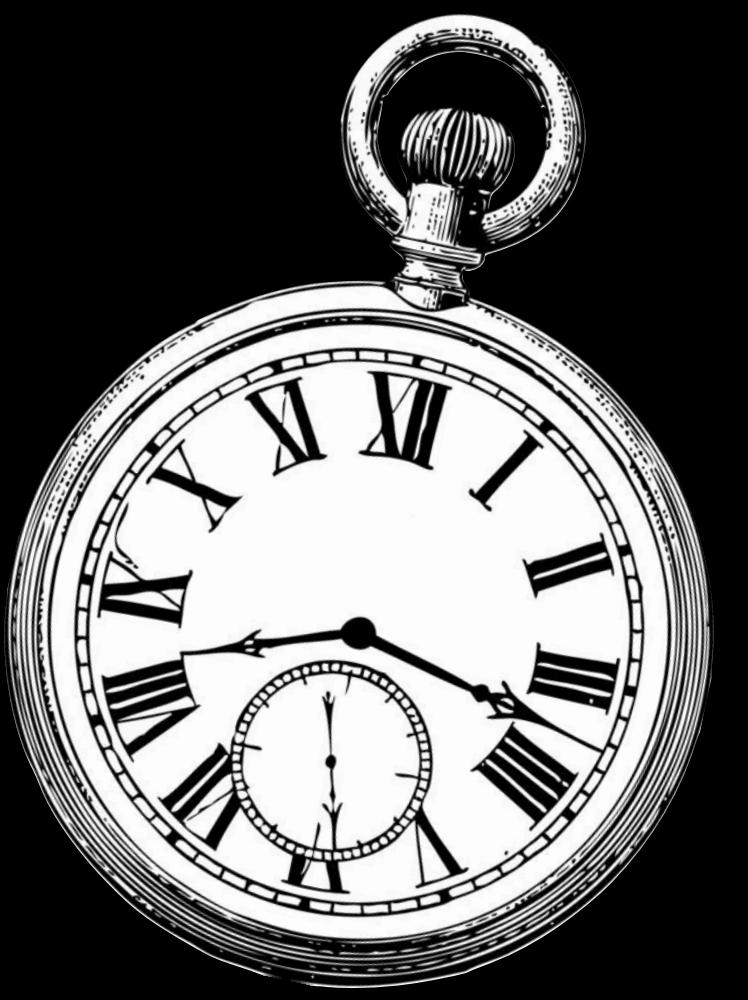


Thomas Edison
1877

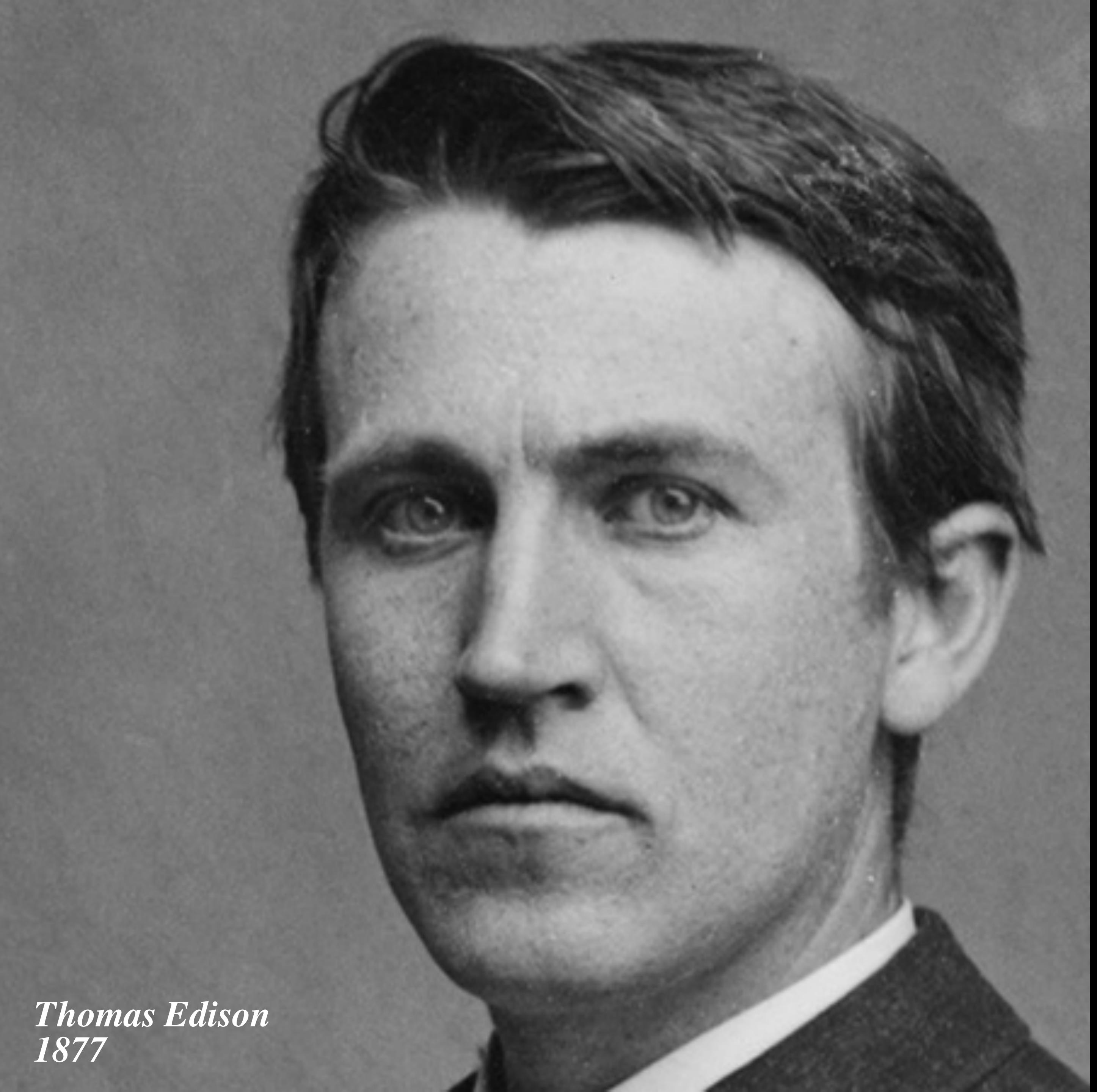


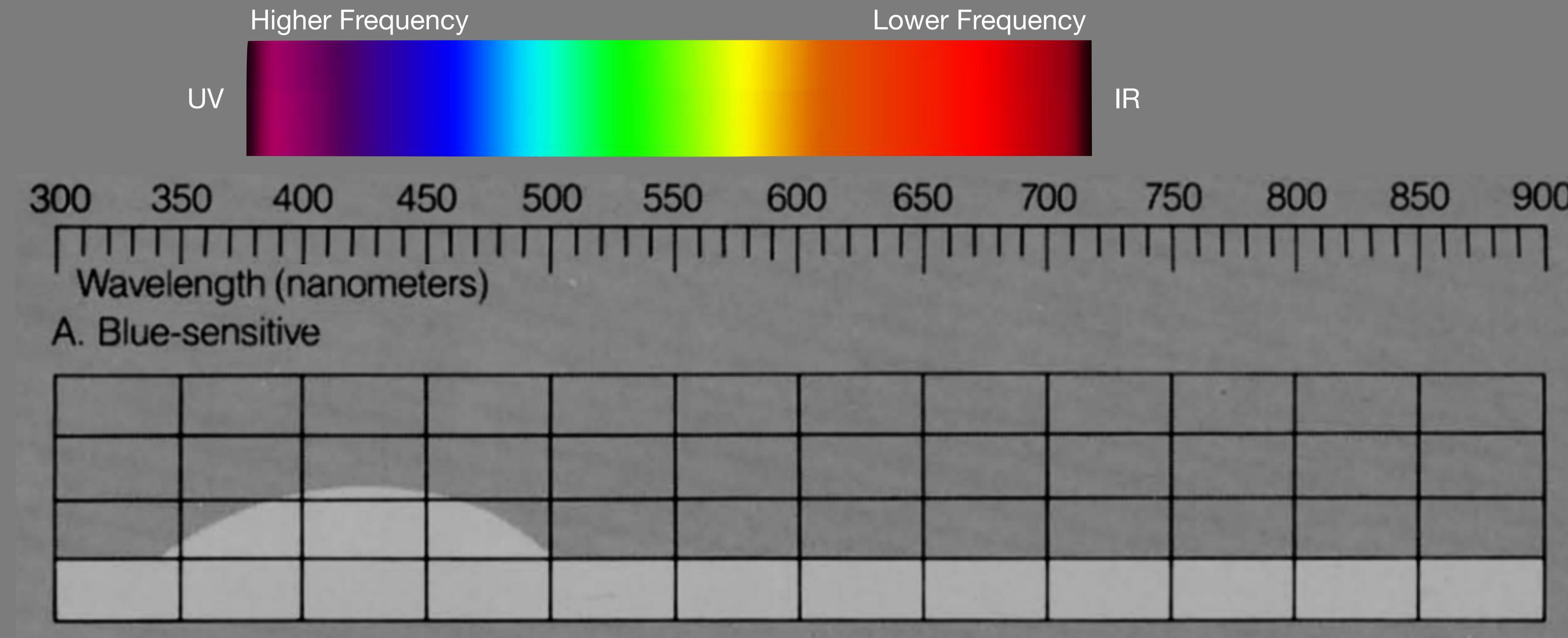
Thomas Edison
1877

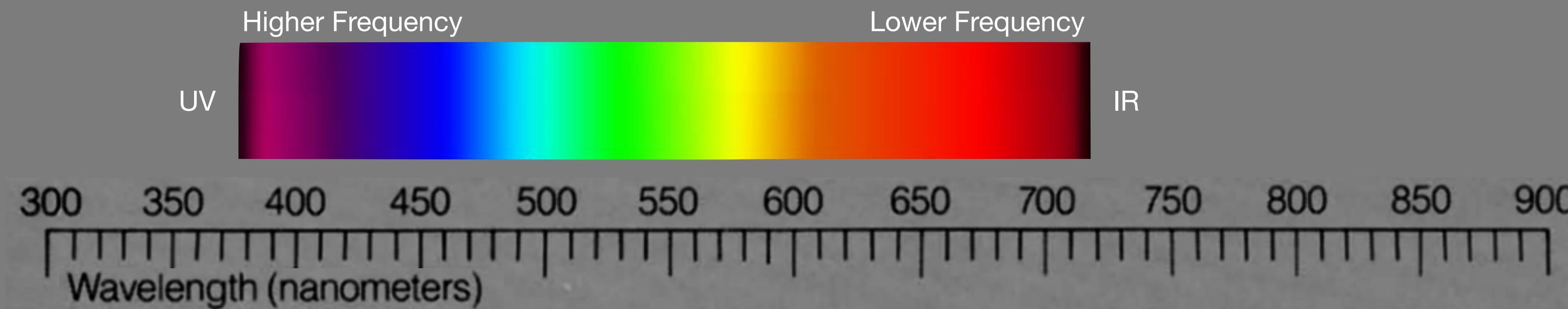




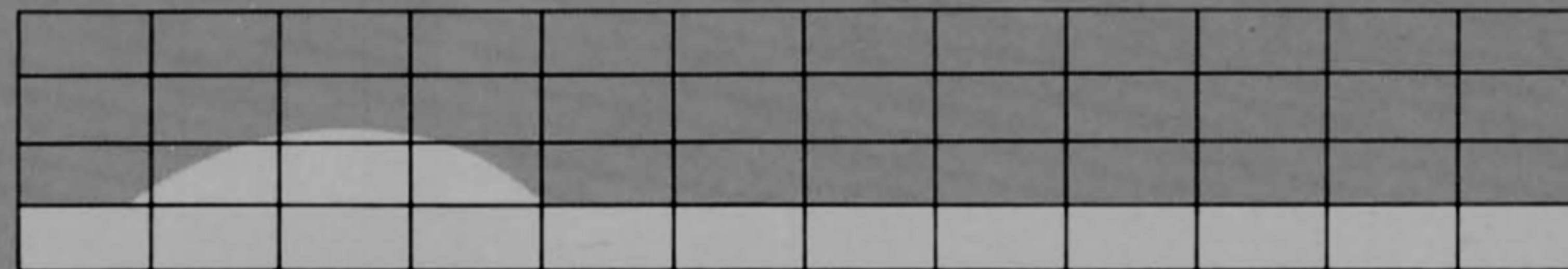
Thomas Edison
1877



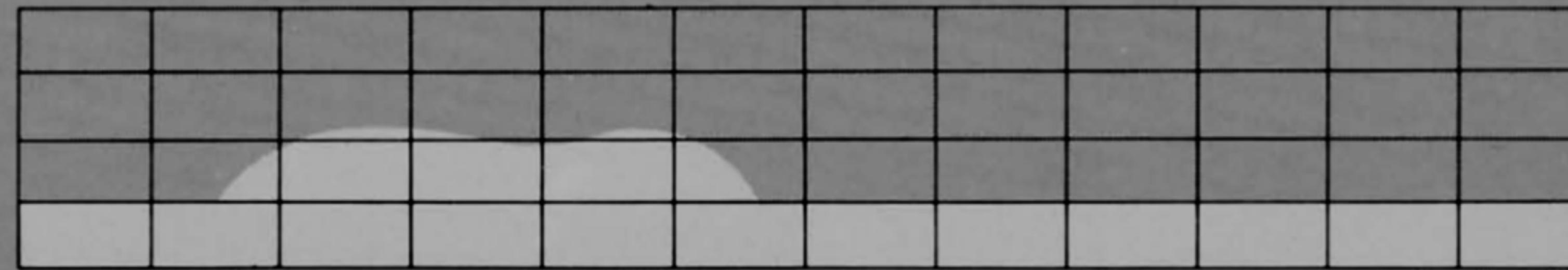




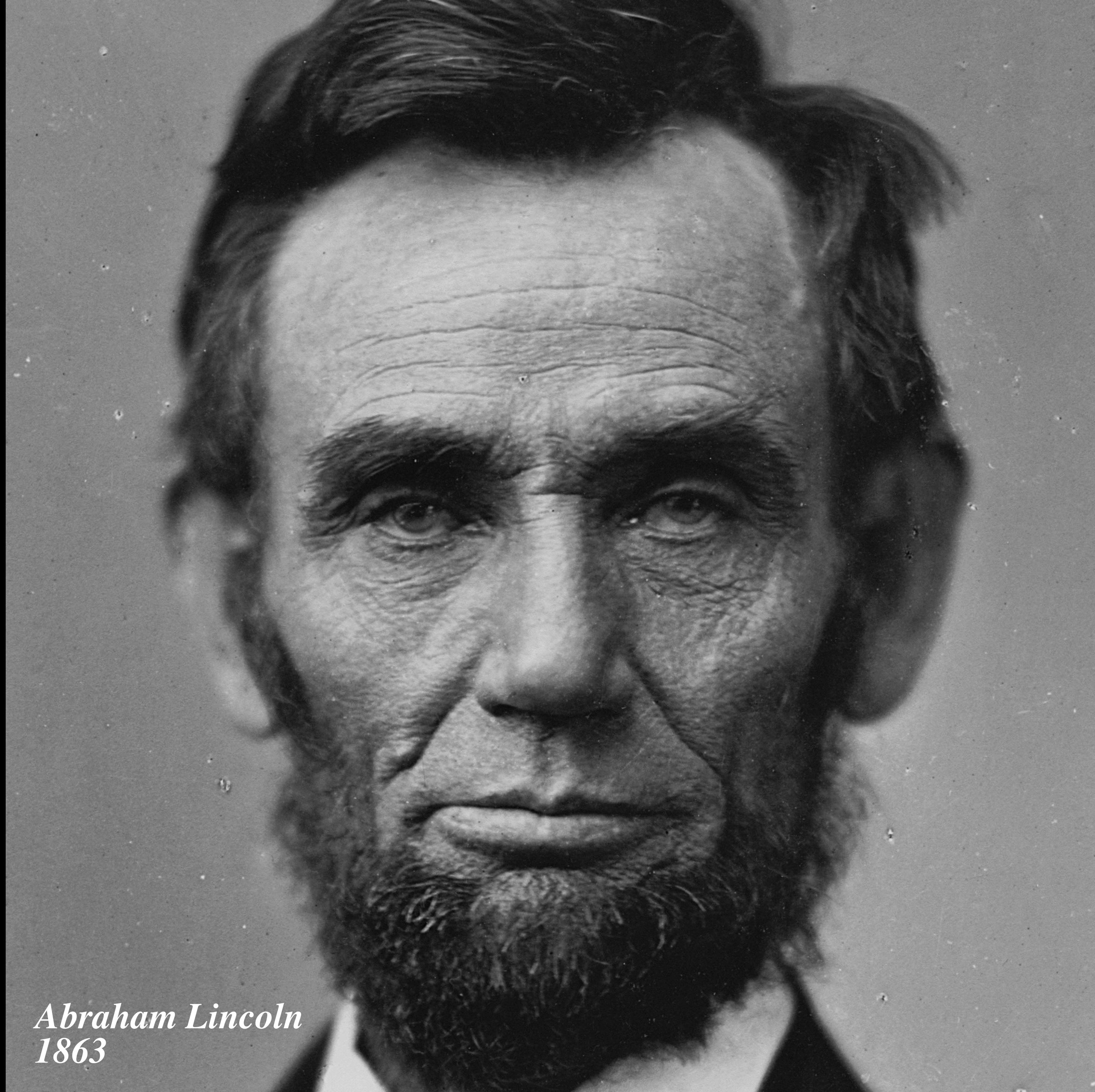
A. Blue-sensitive



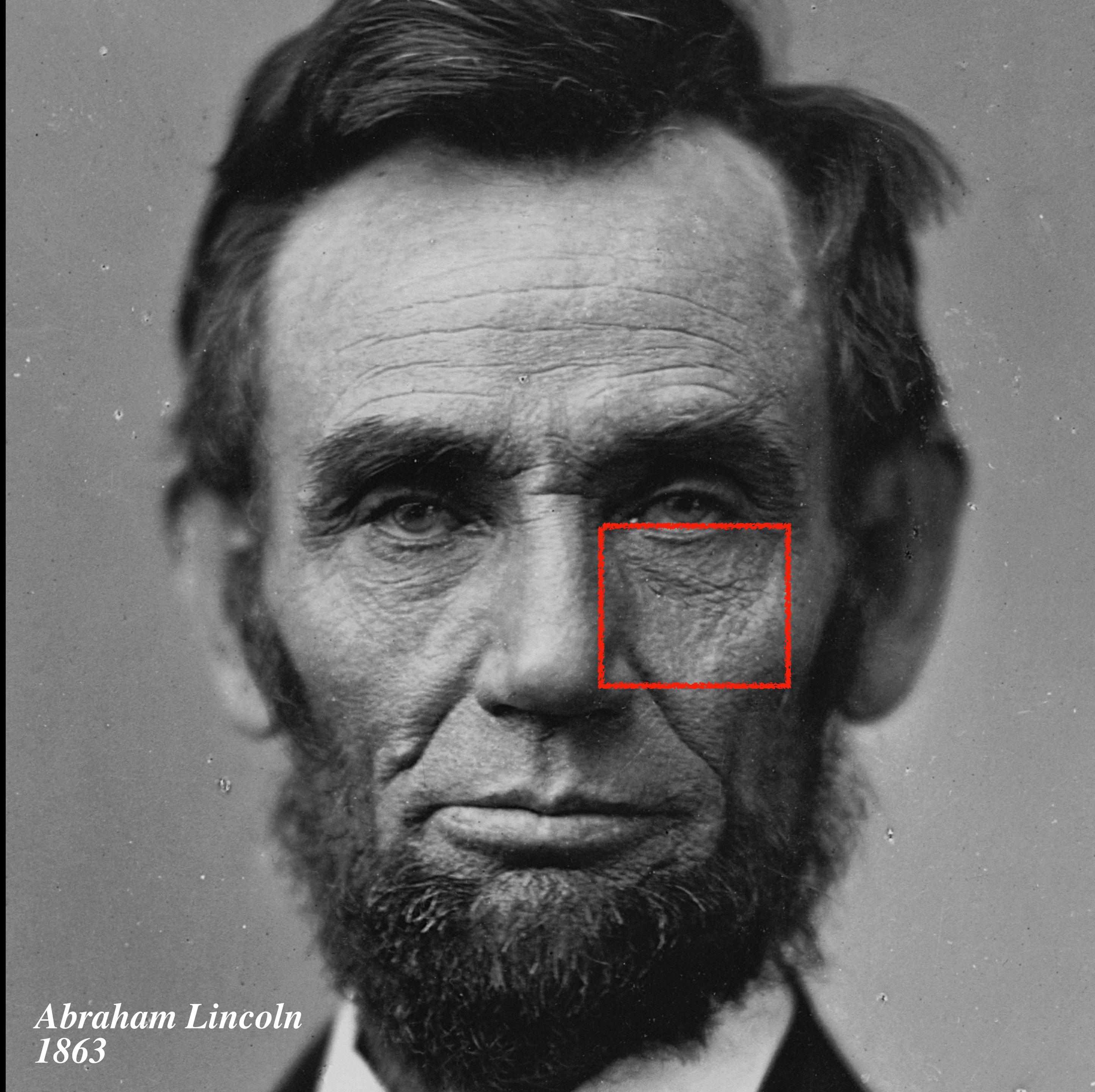
B. Orthochromatic



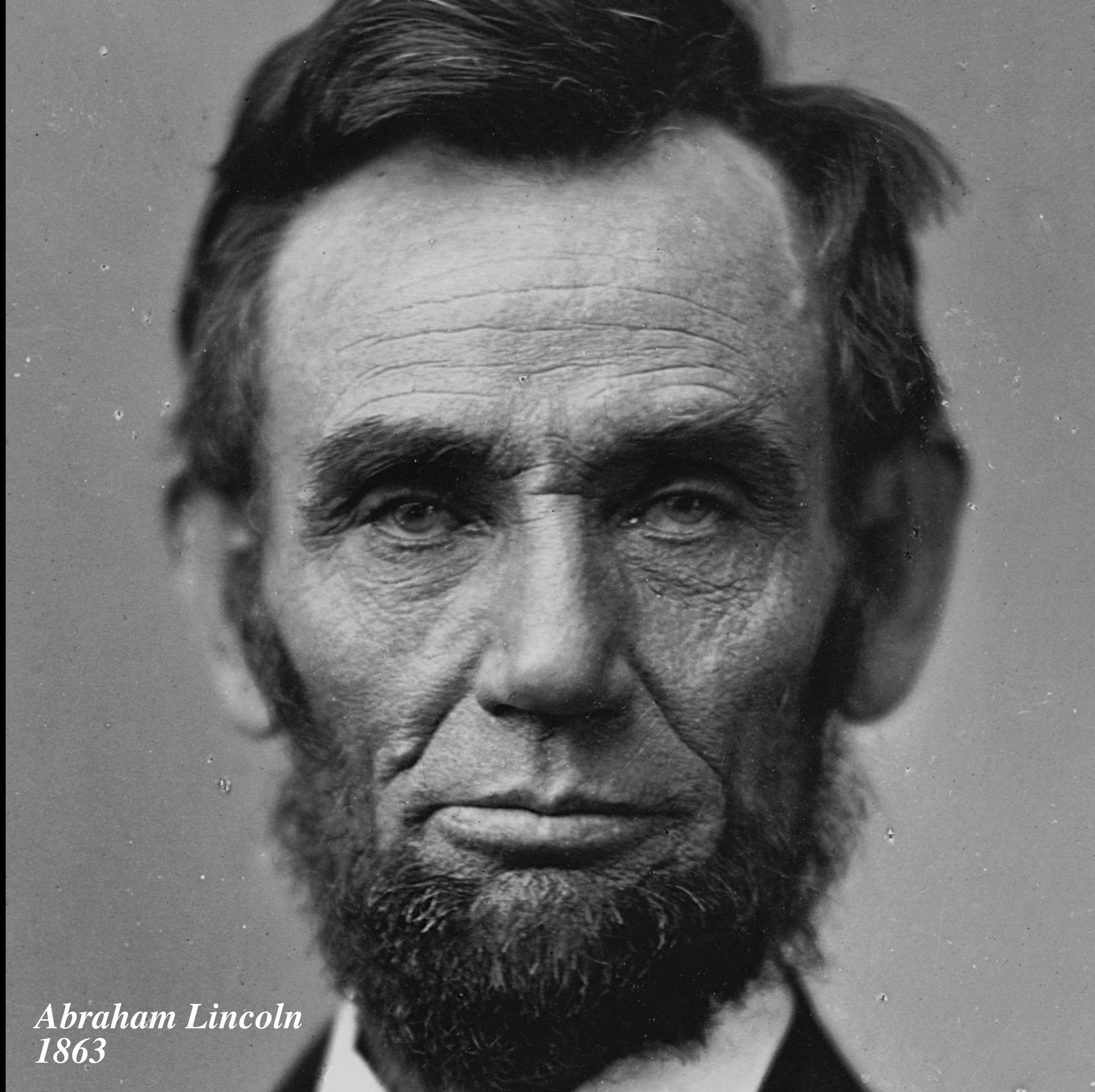
1873 -



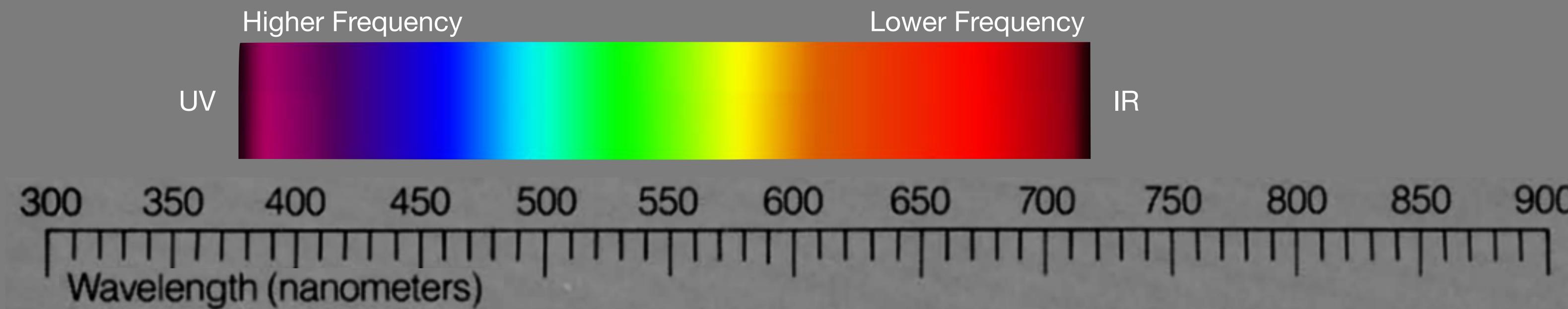
Abraham Lincoln
1863



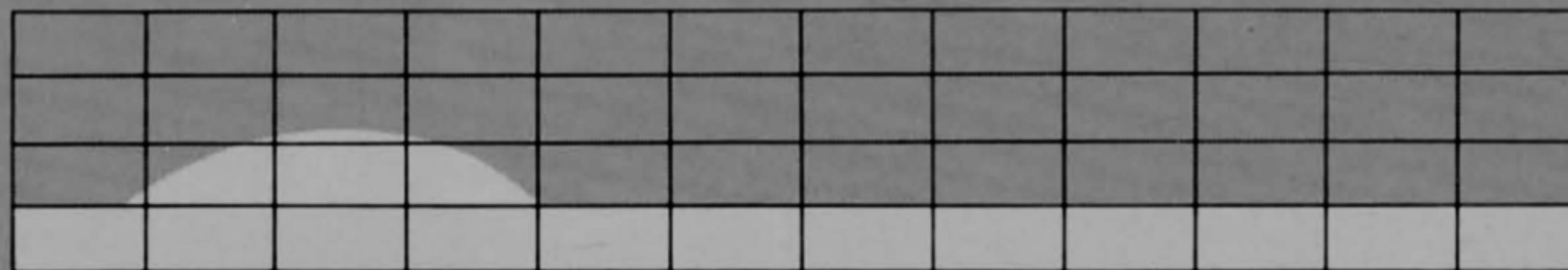
Abraham Lincoln
1863



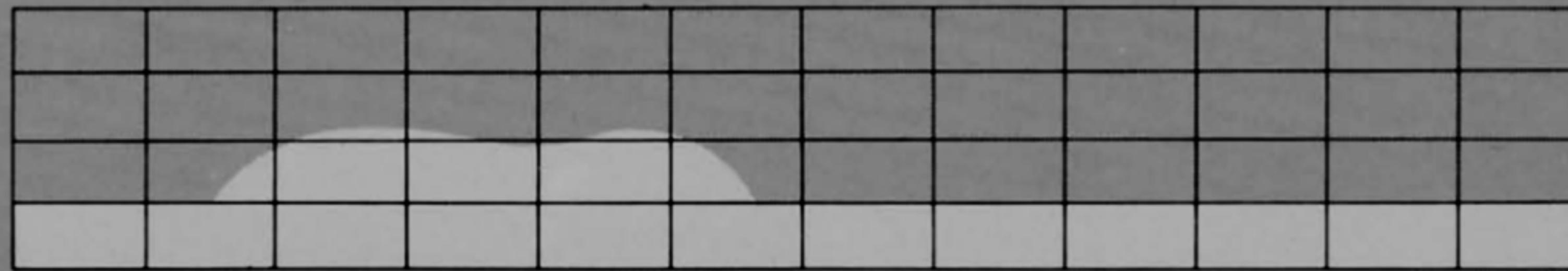
Abraham Lincoln
1863



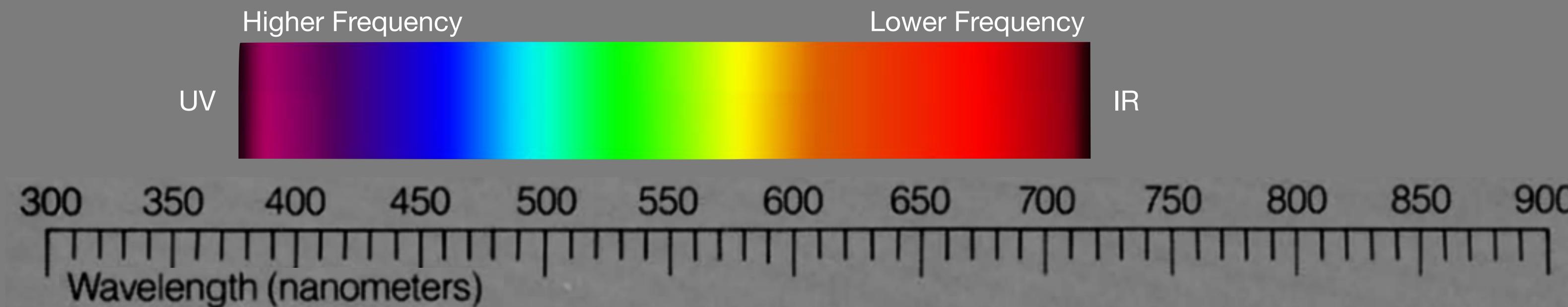
A. Blue-sensitive



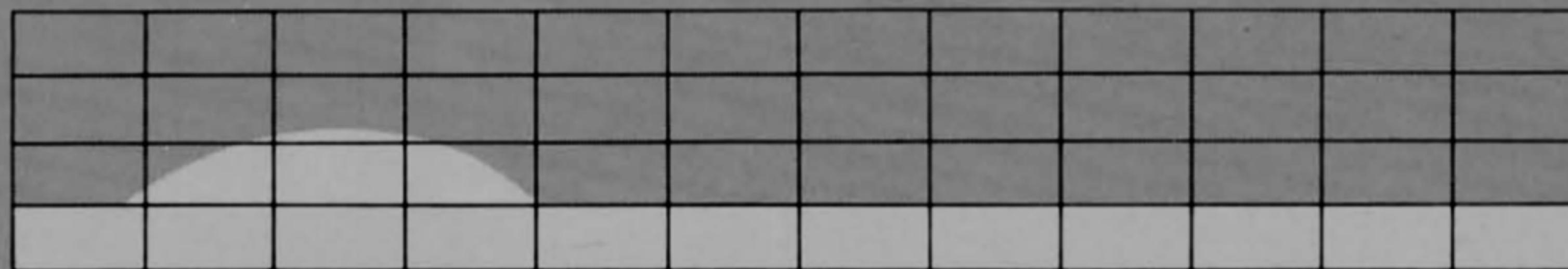
B. Orthochromatic



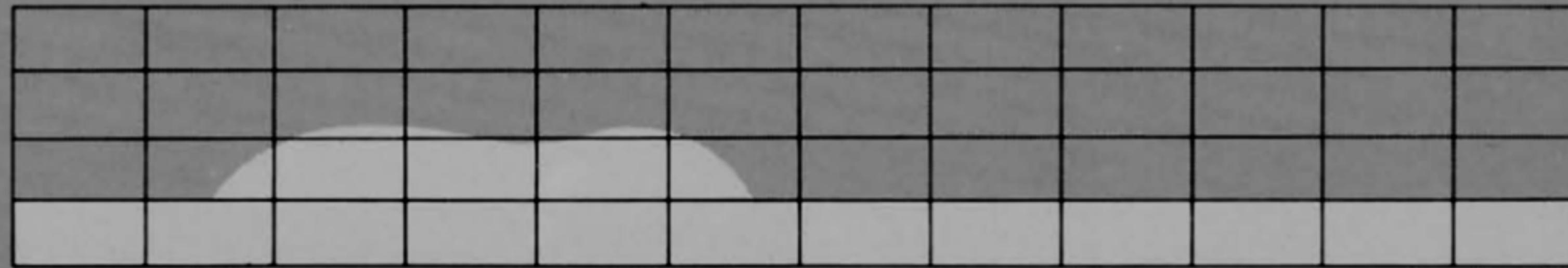
1873 -



A. Blue-sensitive

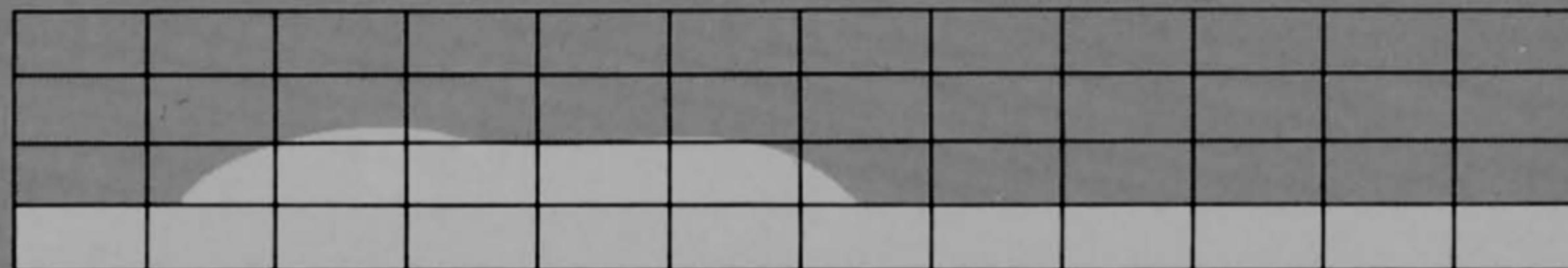


B. Orthochromatic

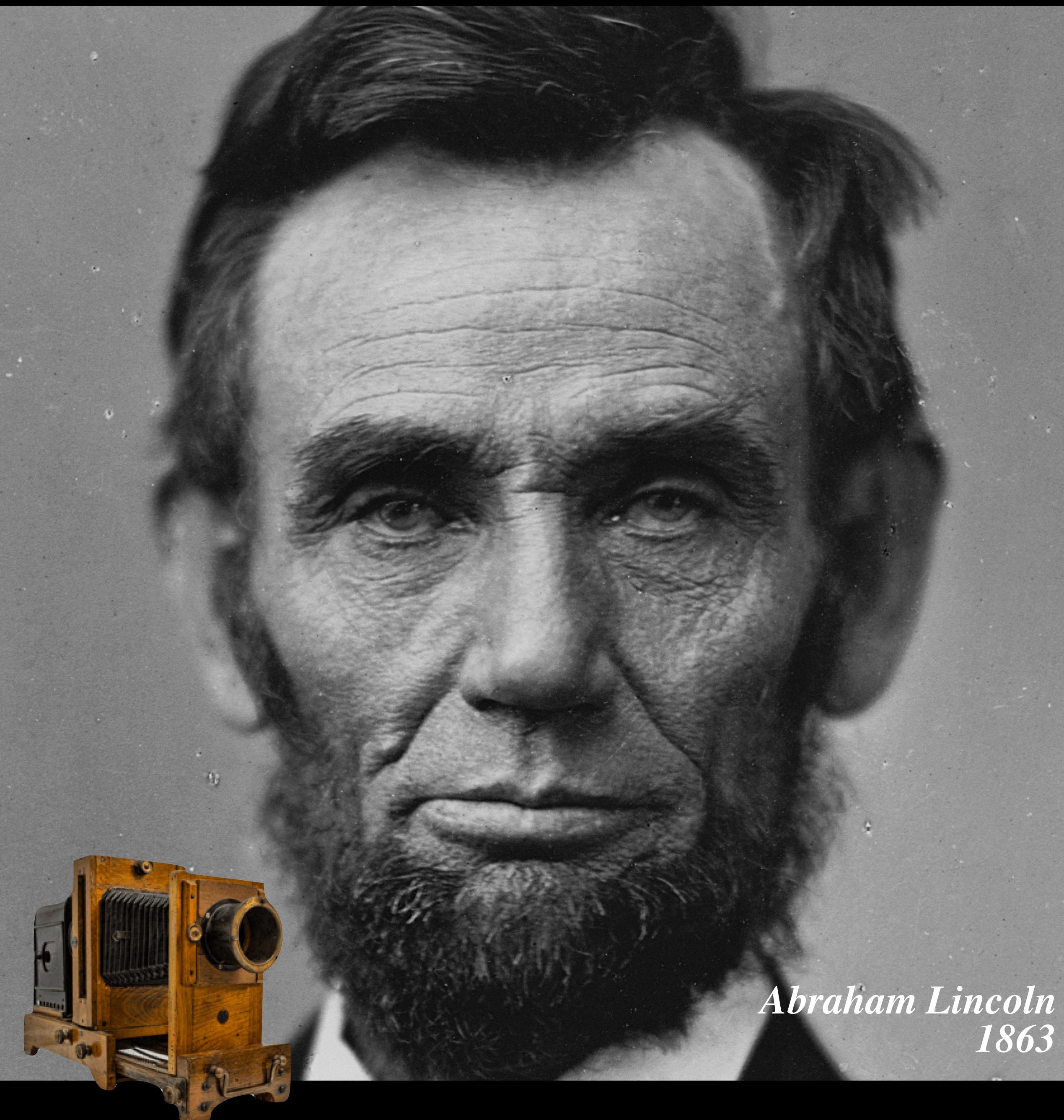


1873 -

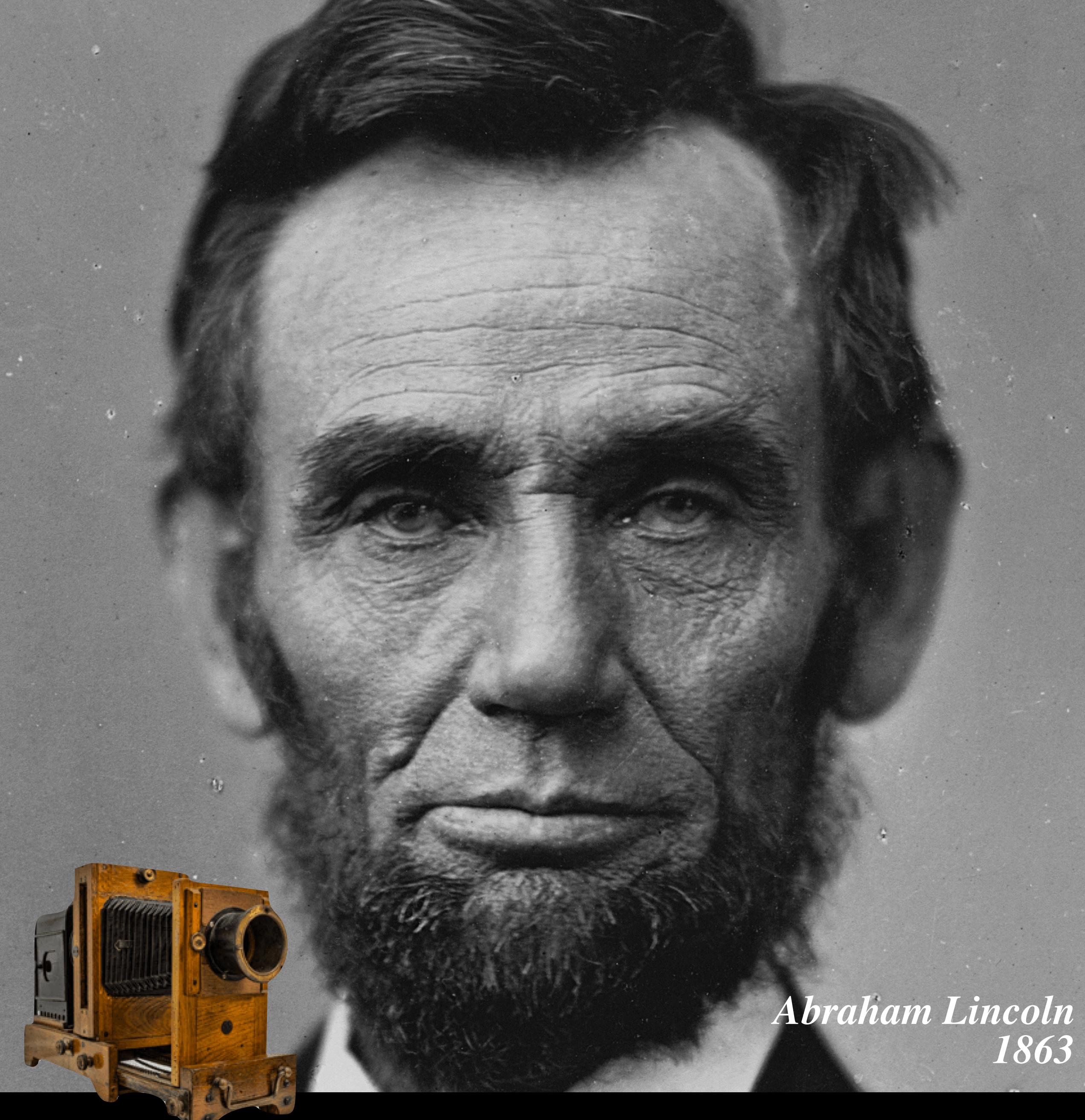
C. Panchromatic



1907 -

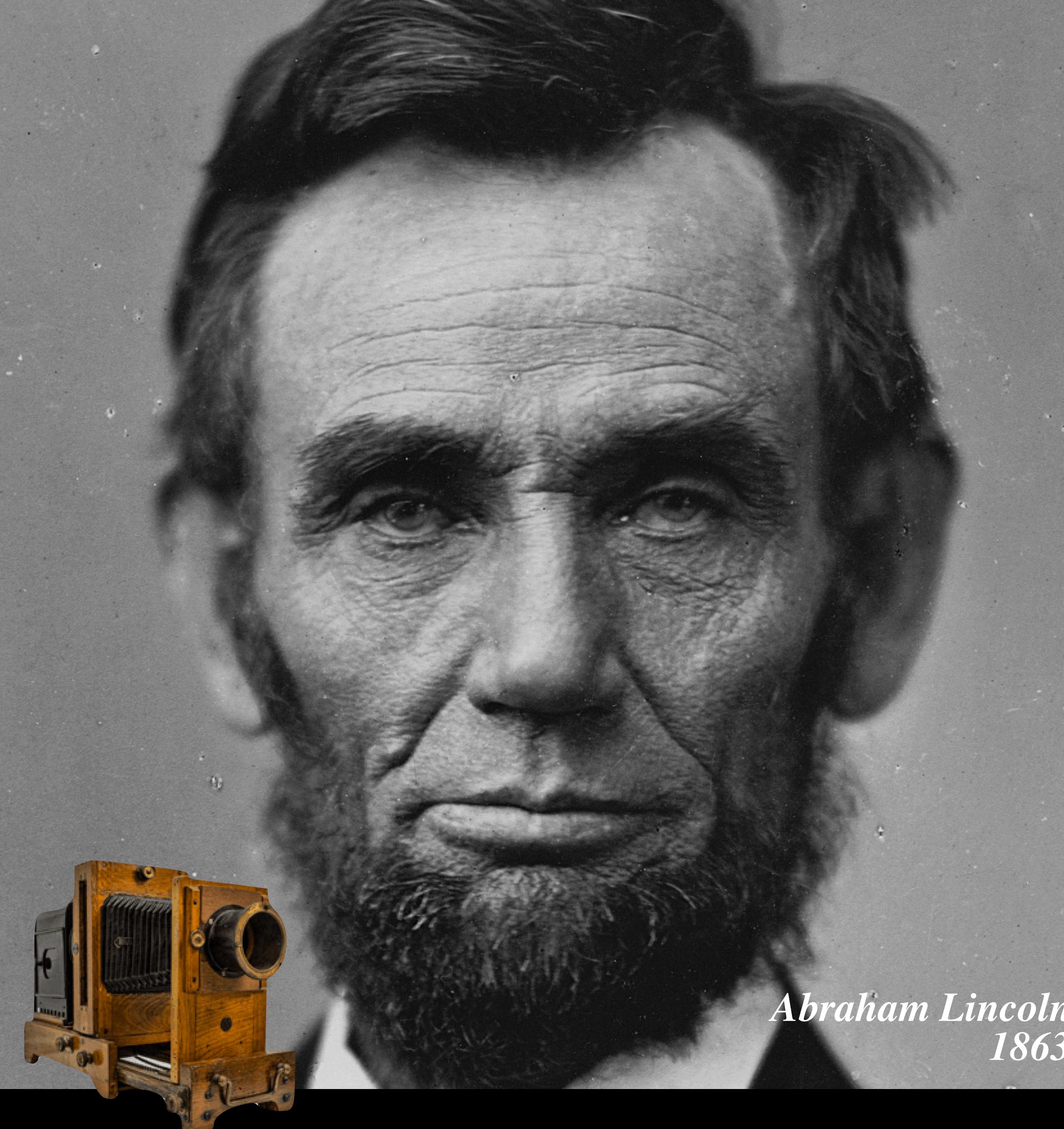


Abraham Lincoln
1863



Abraham Lincoln
1863

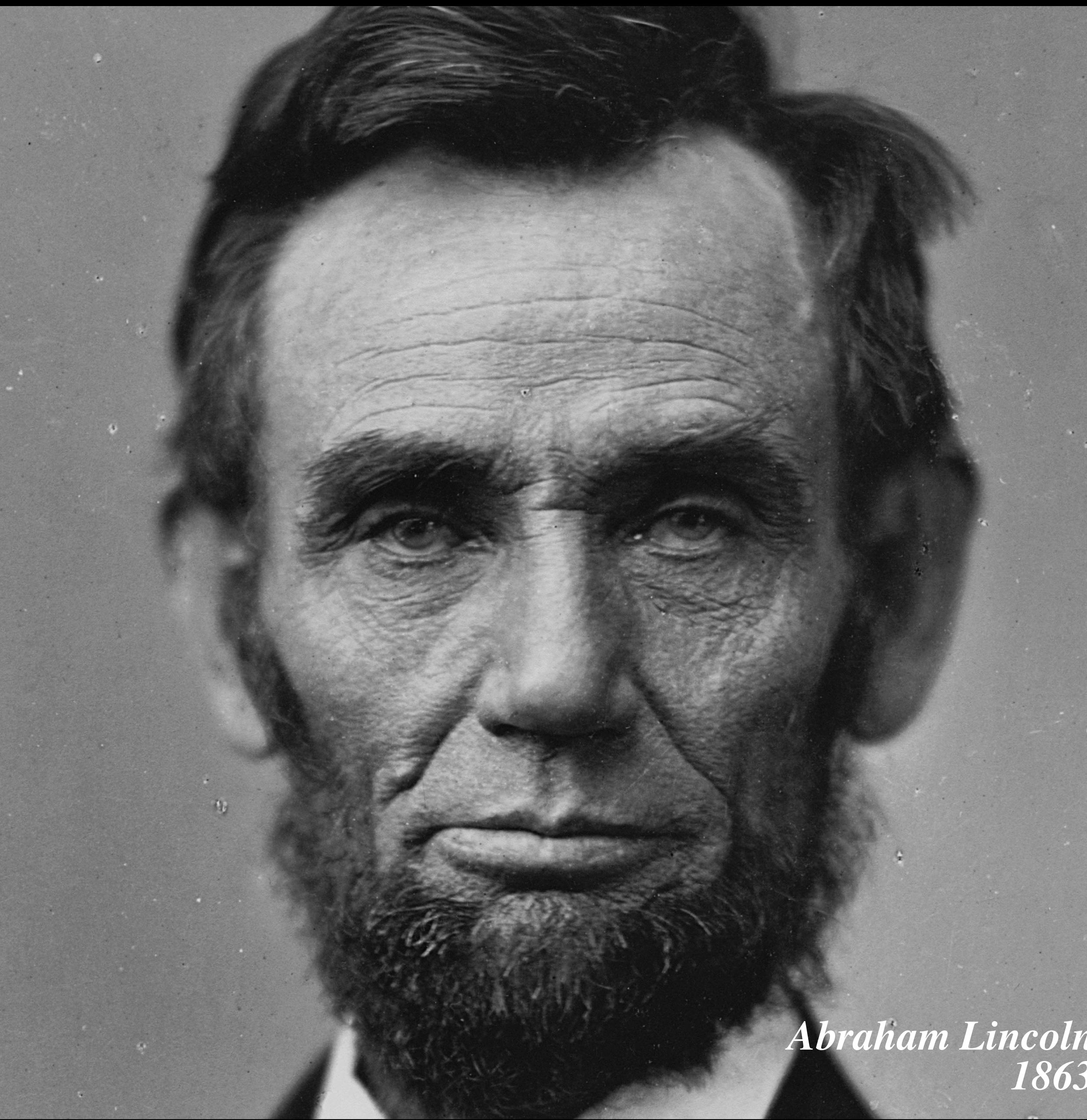




*Abraham Lincoln
1863*



Our Result

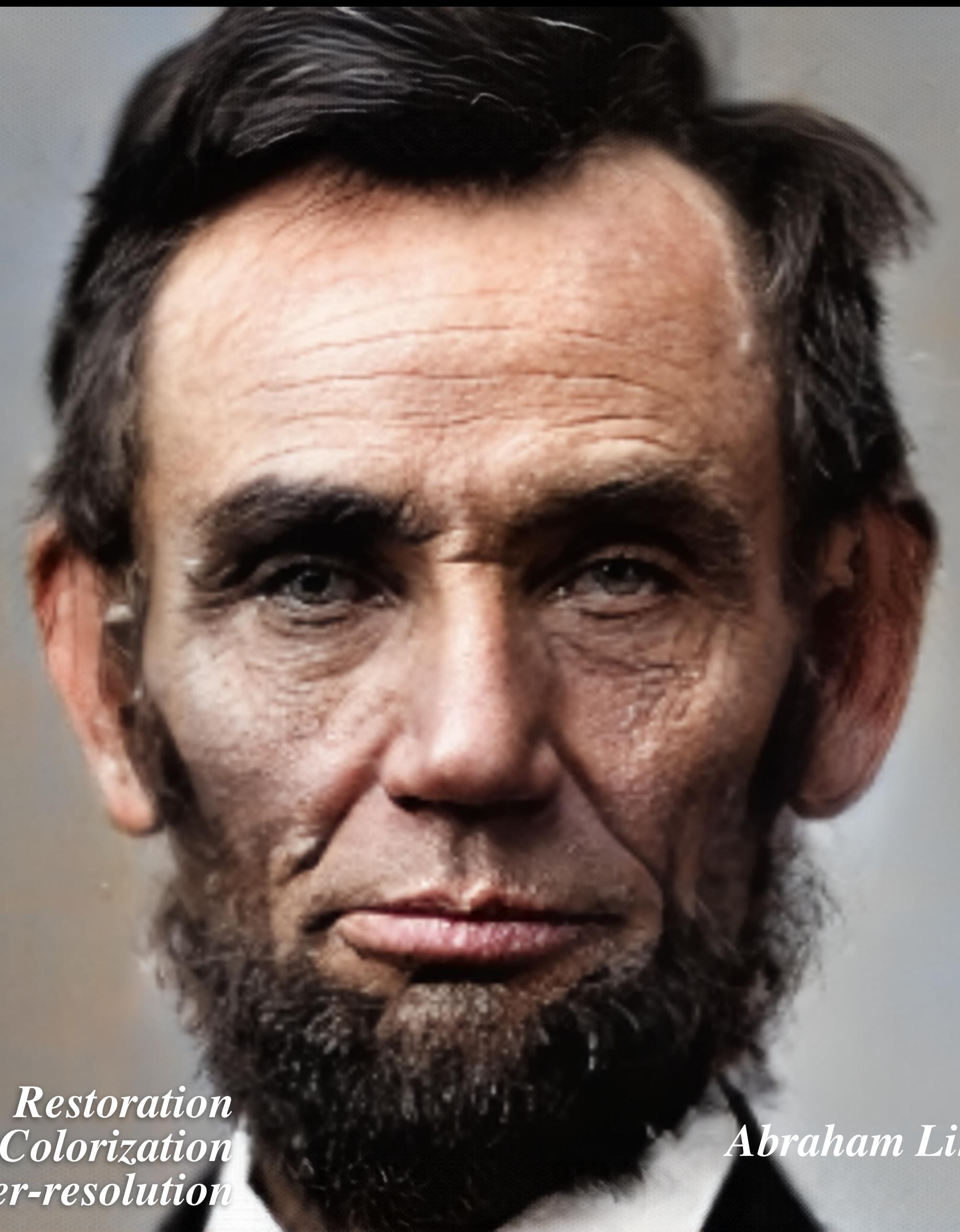


*Abraham Lincoln
1863*



Our Result

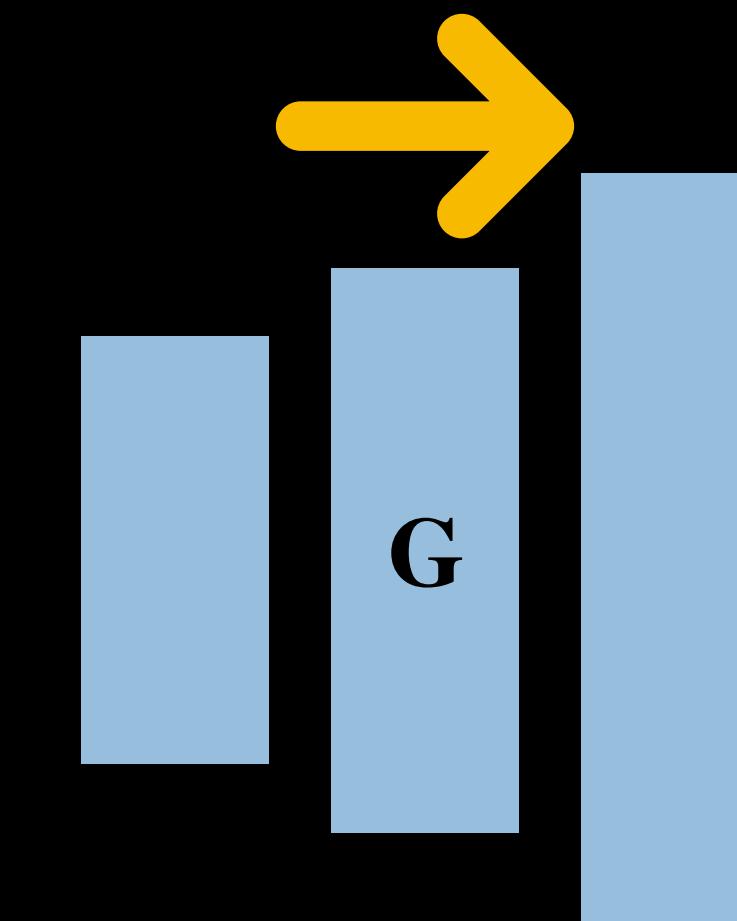
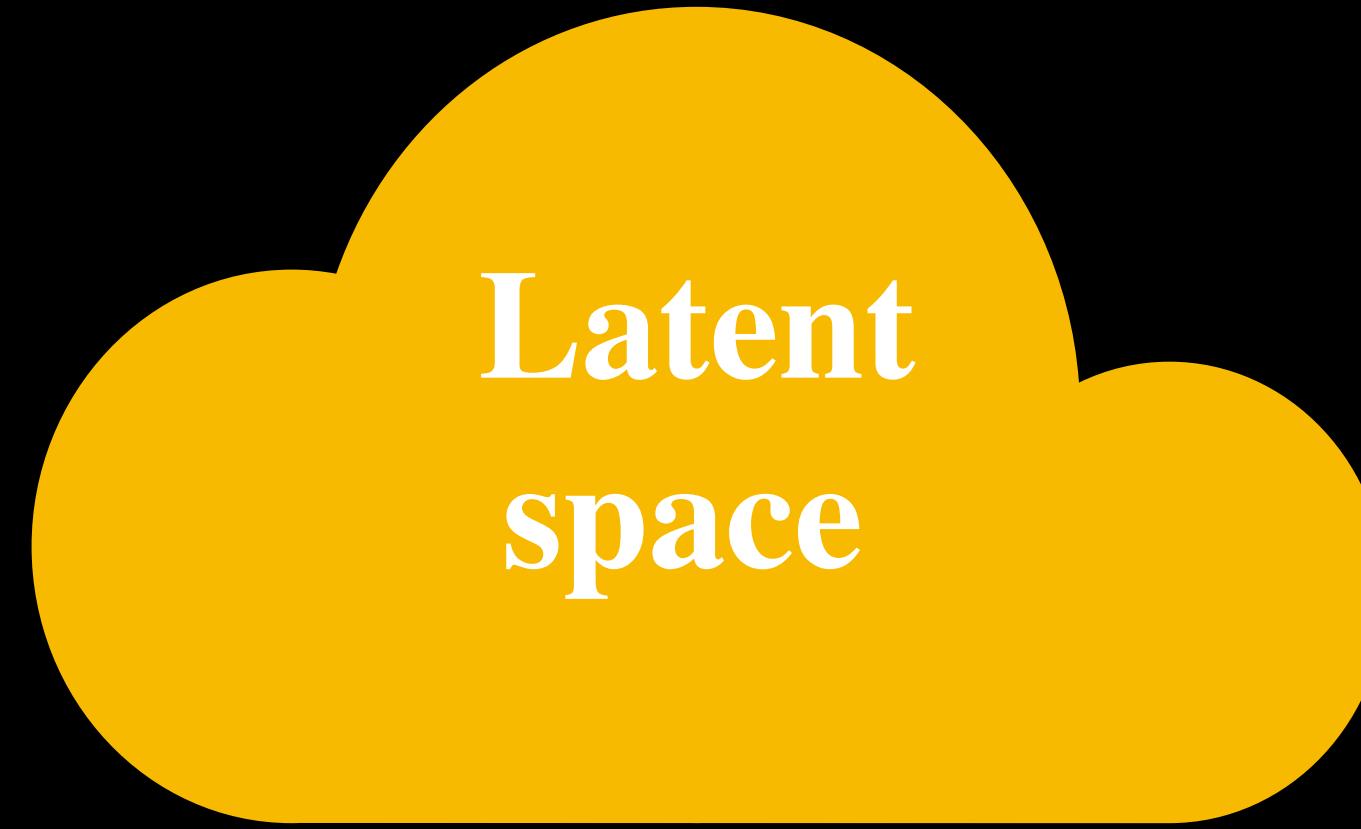
+ Restoration
+ Colorization
+ Super-resolution



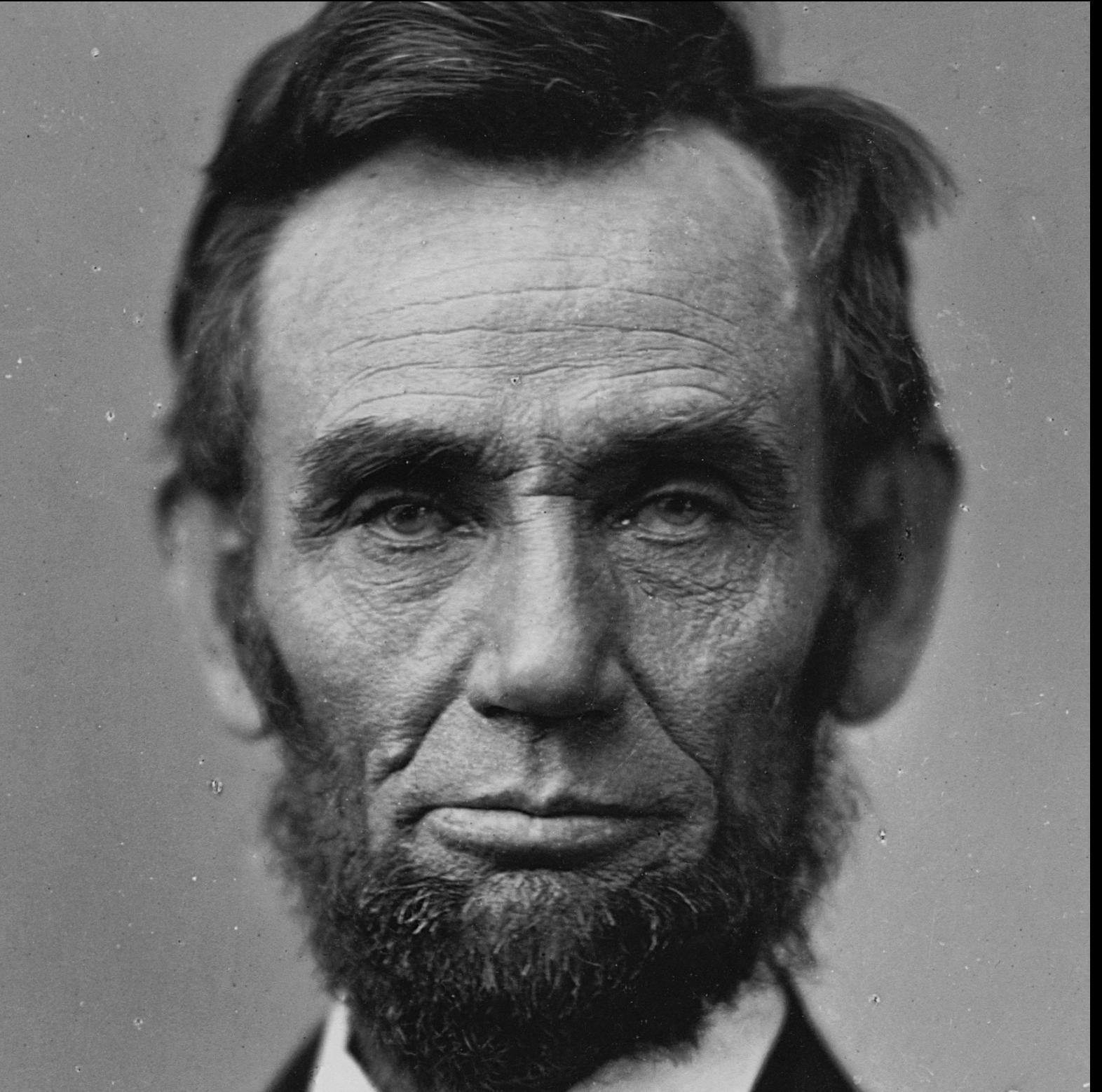
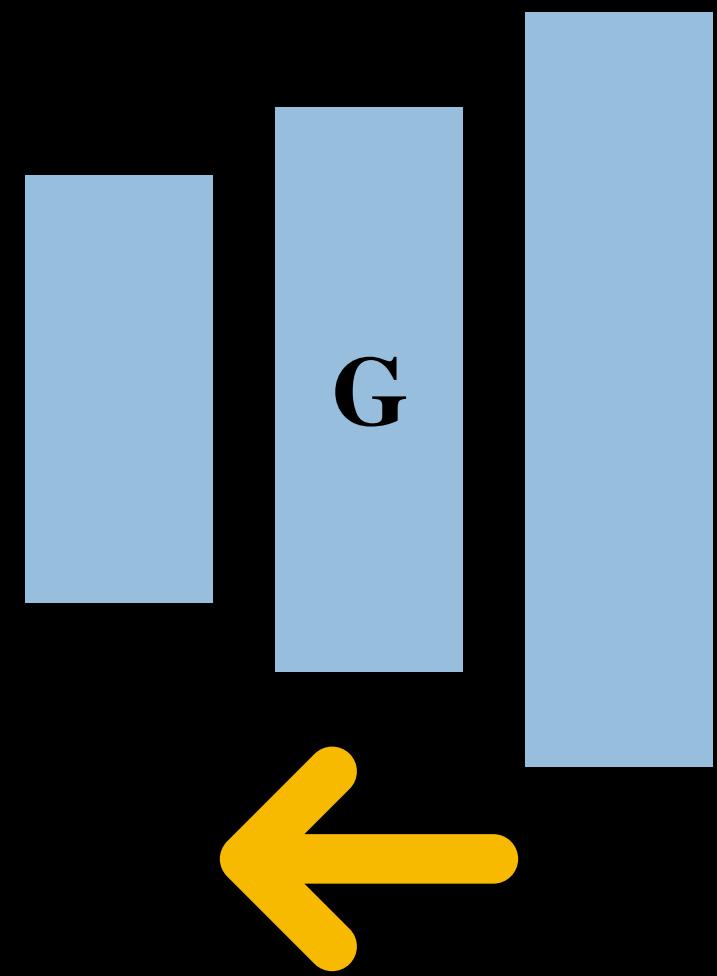
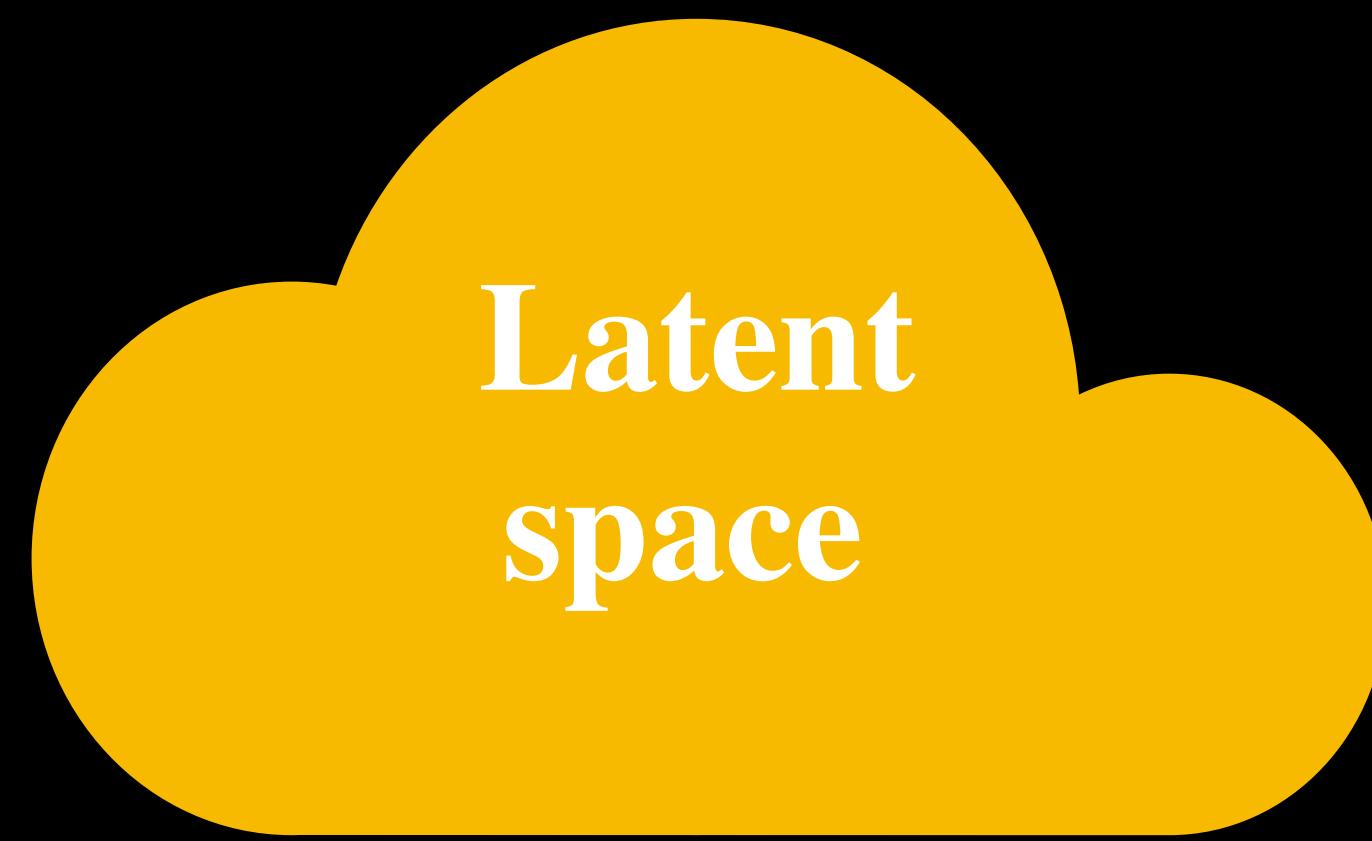
*Abraham Lincoln
1863*



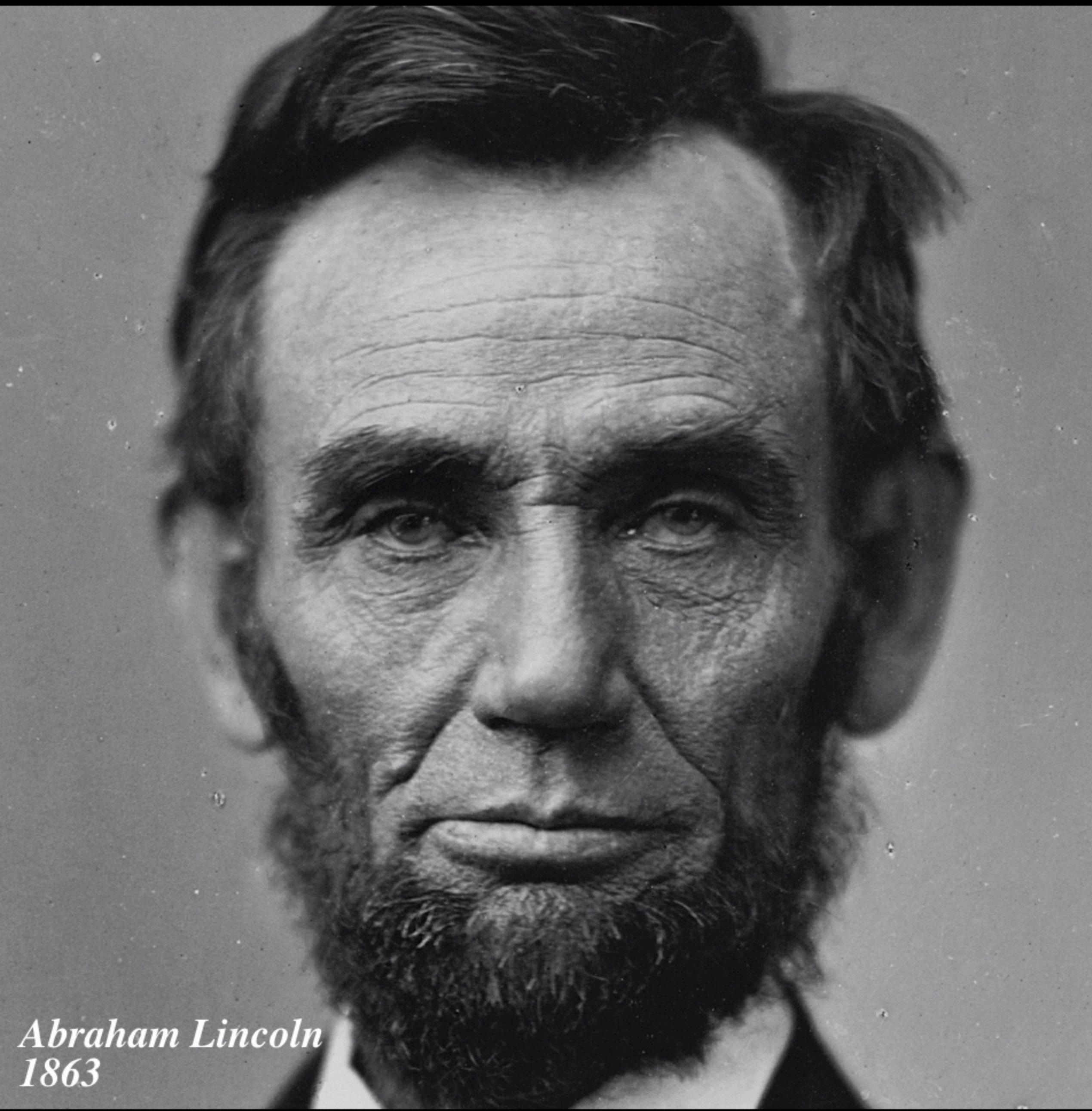
Our Result



[Karras et al. "Analyzing and Improving the Image Quality of StyleGAN." CVPR 2020]



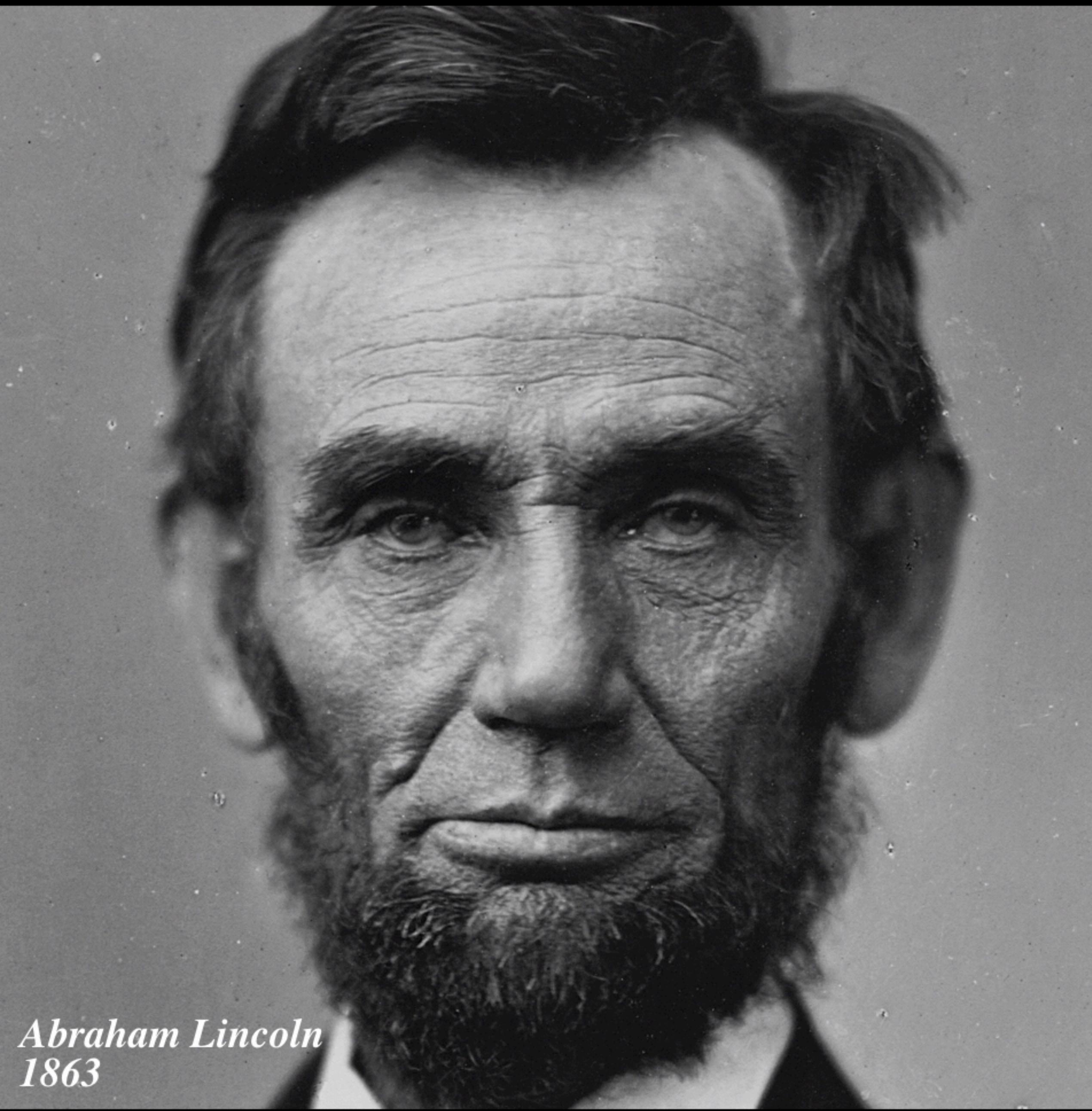
<https://time-travel-rephotography.github.io/>



Abraham Lincoln
1863



Ours



Abraham Lincoln
1863



Ours