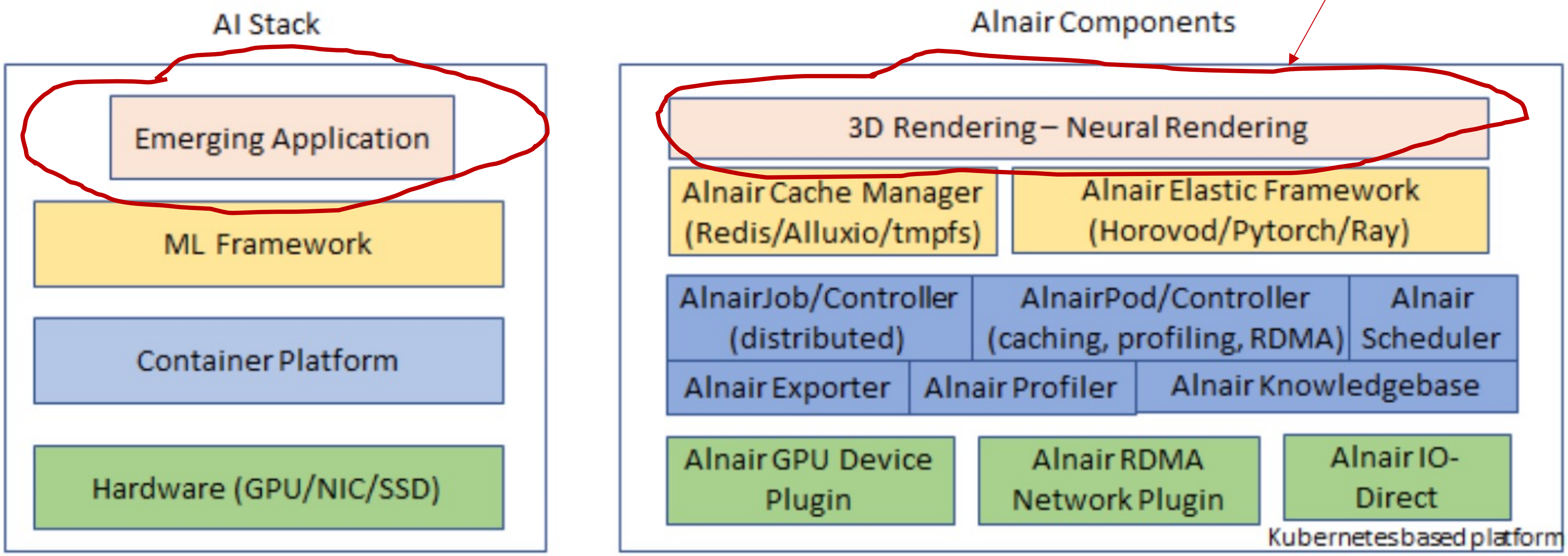
4D Facial Avatar From a Monocular Video

AI Team: Yaohui Ding
Oct 2022

Goal of Alnair Project

Accelerate AI training across multi-layer of AI stack

This demo works in application layer



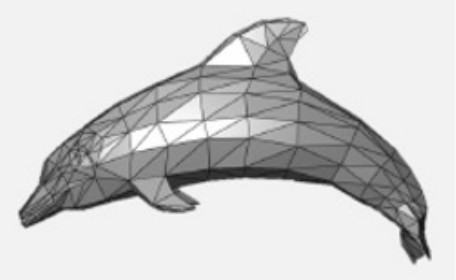
- Image credits to Zhaobo

Background: Neural Rendering and NeRF

Rendering: the process of transforming a scene definition including cameras, lights, surface geometry and material into a simulated camera image.

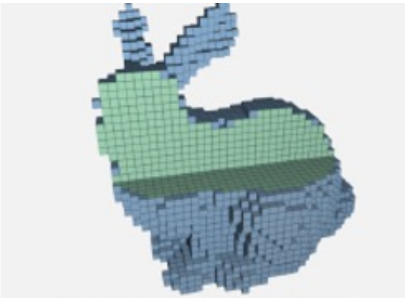
- MIT EECS 6.837

Step 1. Represent a scene via polygonal meshes, point clouds, voxelization, density …

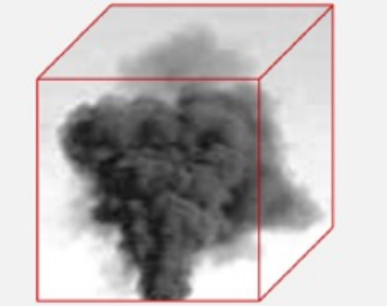


Polygonal Meshes

Pointclouds



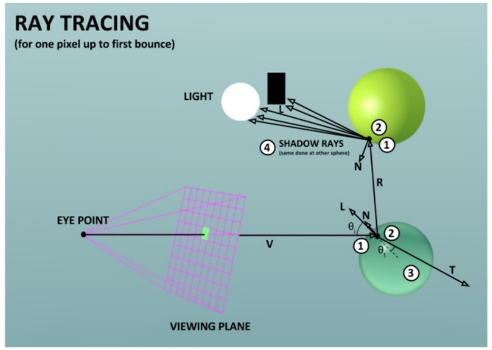
Occupancy / Voxelization



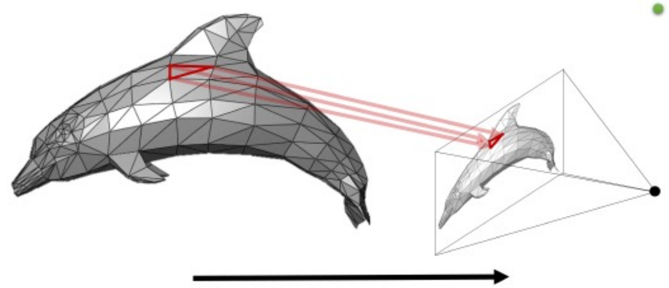
Density

In traditional rendering, these are modeled by artists or scanned by special equipment

Step 2. Calculate the appearance using these representation



Forward Rendering (e.g., rasterization)

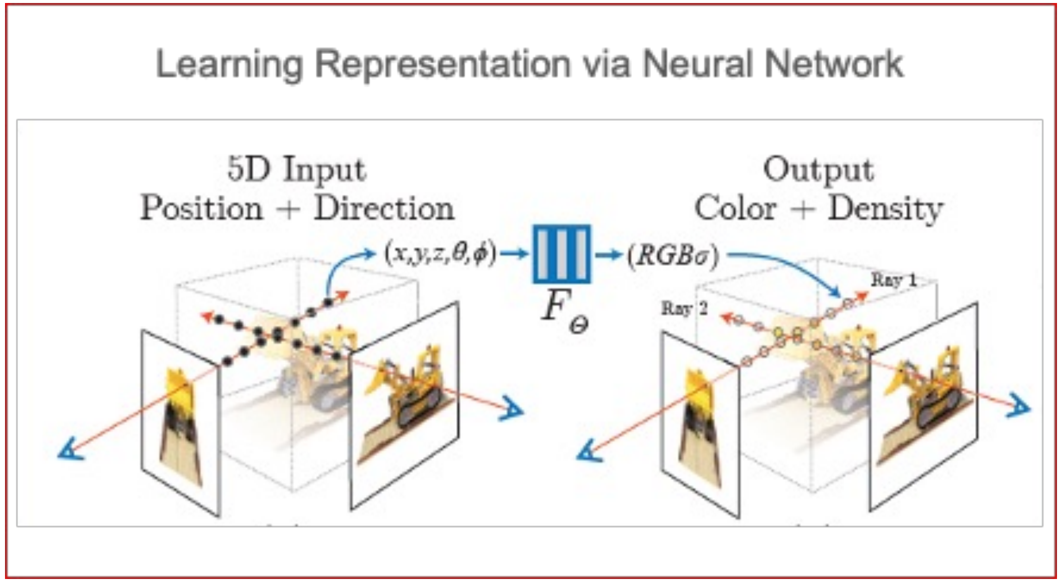


let us call them shading functions

FUTUREWEI Technologies

NeRF: Neural Radiance Fields

Key of NeRF: a new type of representation that is learned via Neural Network



Input:
images from a camera

Training:
representations are learned via a Neural Net

Output:
images generated by computer

Because this object (lego truck) is represented so well via NeRF, we can synthesize high-fidelity images of this object from arbitrary angles, a.k.a Novel View Synthesis

FUTUREWEI Technologies

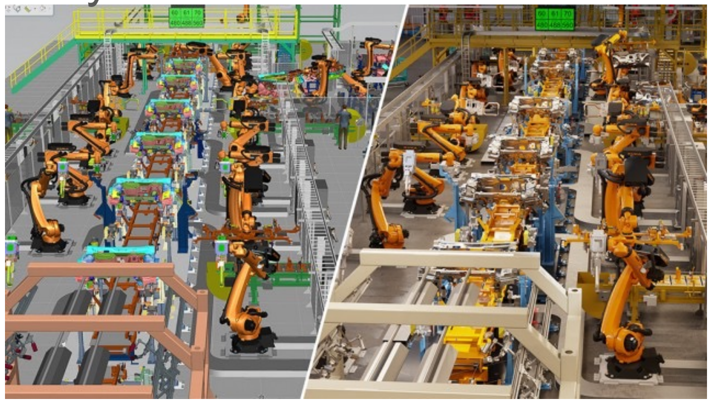
NeRF is Exploring, in both Academic and Industry

- NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis (ECCV2020 Best paper honorable mention)
- GIRAFFE: Representing Scenes as Compositional Generative Neural Feature Fields (CVPR2021 Best paper)
- MiP-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields (ICCV2021 Best paper honorable mention)
- Ref-NeRF: Structured View-Dependent Appearance for Neural Radiance Fields (CVPR2022 Best paper honorable mention)

Search "3D" in paper titles of CVPR2022:
564 papers (total 2,064)

Google 3D map

synthesized real world



Our Alnair platform will accelerate NeRF-related rendering as a new category of AI workload

FUTUREWEI Technologies

NeRF vs. Traditional Rendering

Pro:

- Simpler inputs, no need for mesh design, depth cameras, or fancy sensors
- Higher quality as a volumetric rendering, e.g. for rendering hair
- End-to-end training, just like other CV algorithms using GPUs

Con:

- Slow training and rendering
- One scene one model
- Only for static scenes?
- etc.

4D Facial Avatar from a Monocular RGB Video

Input

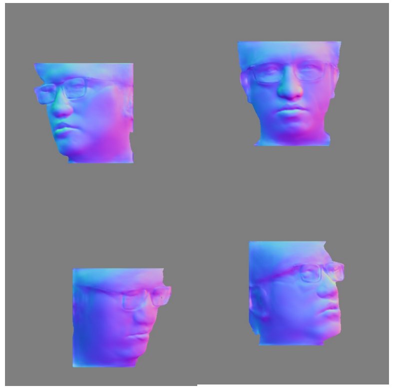
13 sec video using iphone camera

Pre-process

face normal

face segment

background mat



NeRF-based rendering

FLAME model

head tracker

NeRF-based Optimization

Landmarks, shape, appearance, lighting condition param, camera param, etc.

4D facial Avatar
Demo 1,2,3.

FUTUREWEI Technologies

Demo 1. Manually change expr and pose

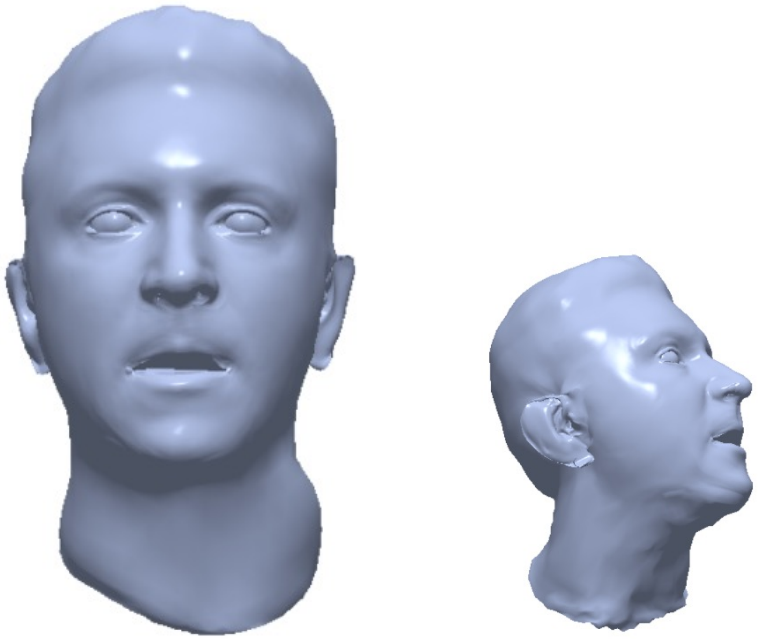
Out[13]:

jaw		0.50
rot0		0.00
rot1		3.14
rot2		0.00
neck0		0.00
neck1		0.00
neck2		0.00

e0		3.00
e1		0.00
e2		1.60
e3		0.00
e4		2.00

```
CPU times: user 8 µs, sys: 4 µs, total: 12 µs
Wall time: 25.7 µs
```

4D Avatar



shaded mesh

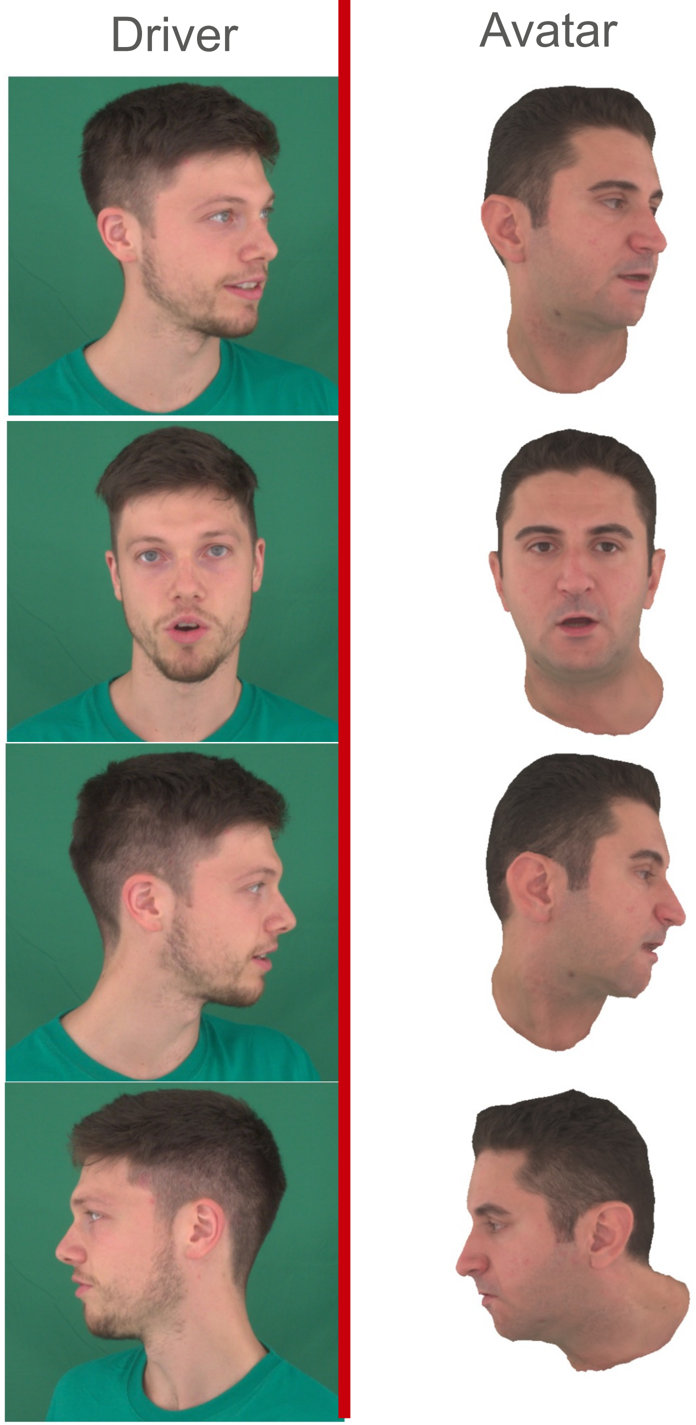
FUTUREWEI Technologies

Demo 2. Frame-by-Frame Reenactment

```python
rgba_driver, rgba_target = frame_reenact(target_model, target_tracking_results,
                                          driving_tracking_results,
                                          driving_frame_id=9)
## show range
fig, axes = plt.subplots(ncols=2, figsize=(20,10))
axes[0].imshow(cv2.cvtColor(rgba_driver, cv2.COLOR_BGR2RGB))
axes[1].imshow(rgba_target[0,:3].cpu().permute(1,2,0))

[a.axis("off") for a in axes]
plt.show()
```

The expression and pose from the driver (person on the left) is transferred/projected to the avatar (synthesized head on the right)



Driver Avatar

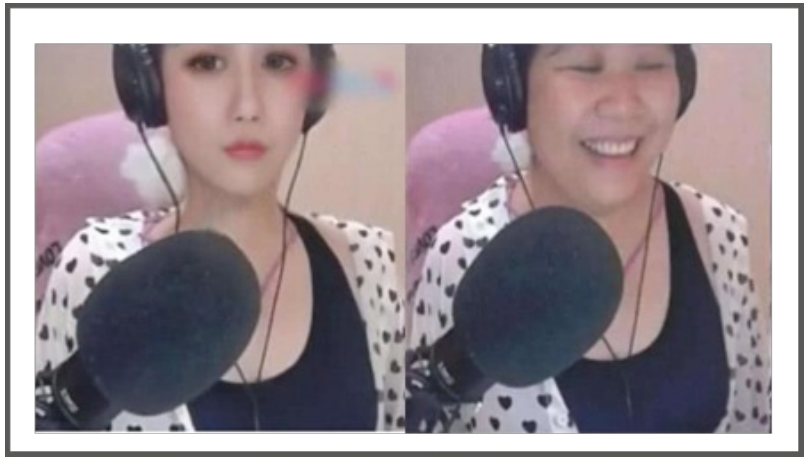
Demo 3. Video-to-Video Reenactment

Input a video, the expression and pose from the person in the video, can be transferred/projected to any synthesized head

Potential Application Scenarios

Teleconference, video chat, or a video version of DeepFake

Our 4D Avatar vs. **Facebook Horizon** vs. **Camera filter**



cartoon characters are much easier

only works under certain condition

FUTUREWEI Technologies

Thank You.

FUTUREWEI Technologies