AI时代的图形流水线探讨

浙江大学 · CAD&CG
霍宇驰
2024.4.3

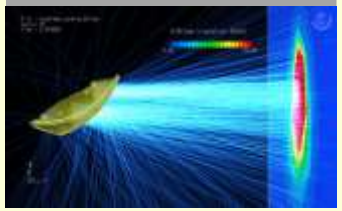
1. Classic Graphics Pipeline

Connecting human, machine, and physics

Physical  → Modeling → Digital  → Rendering → Human 

Applications

Design 

Optics 

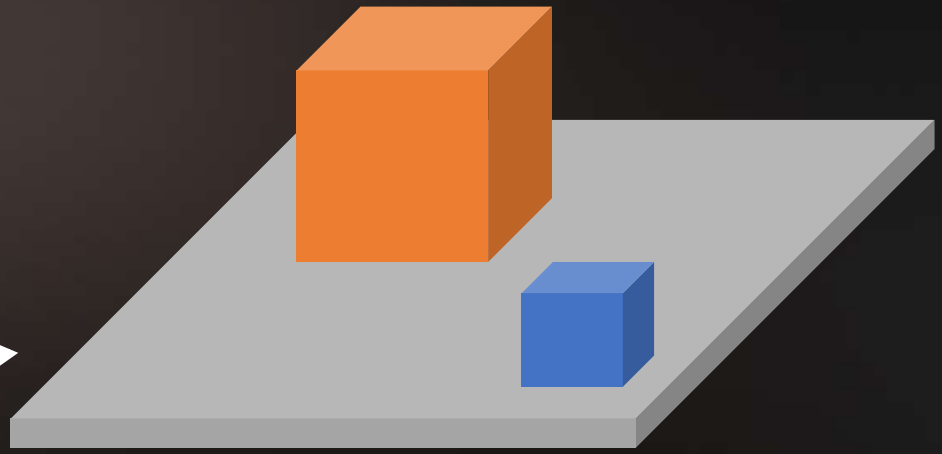
XR 

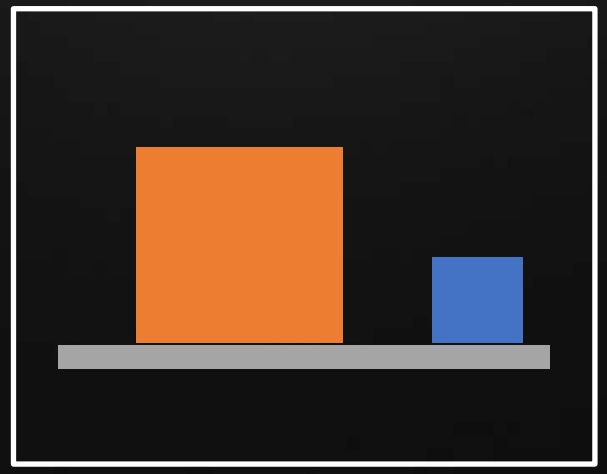
Intelligence 

Metaverse 

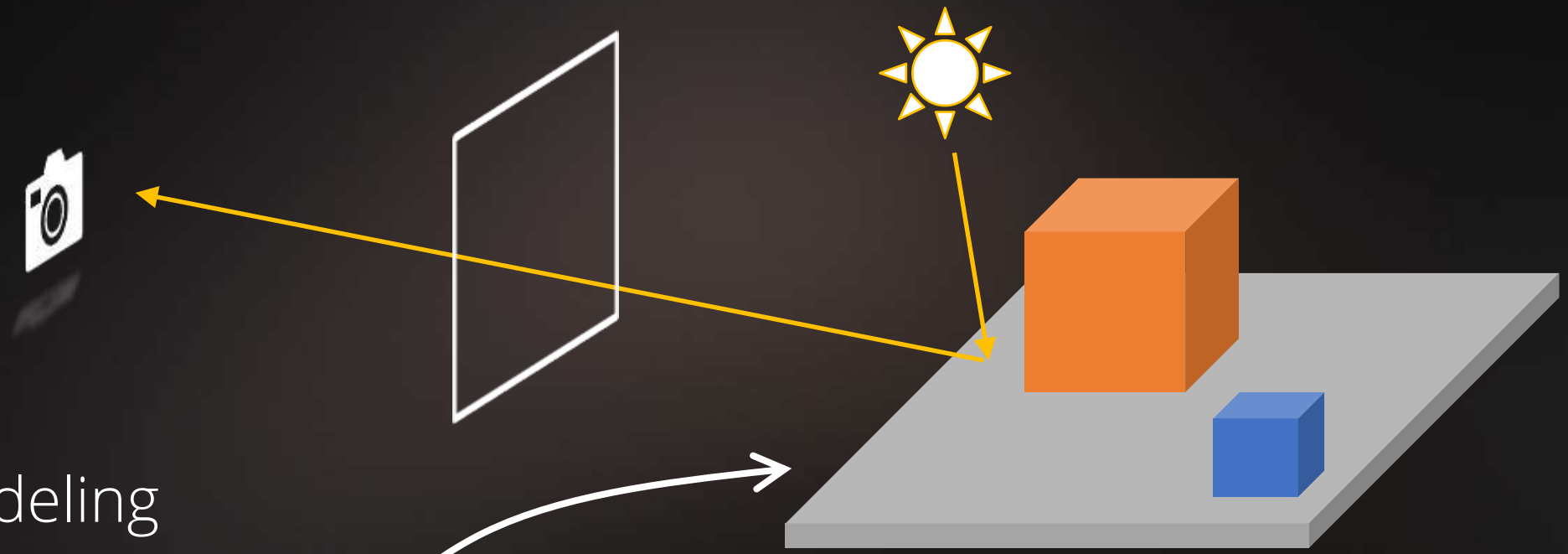
Entertainment 

Modeling

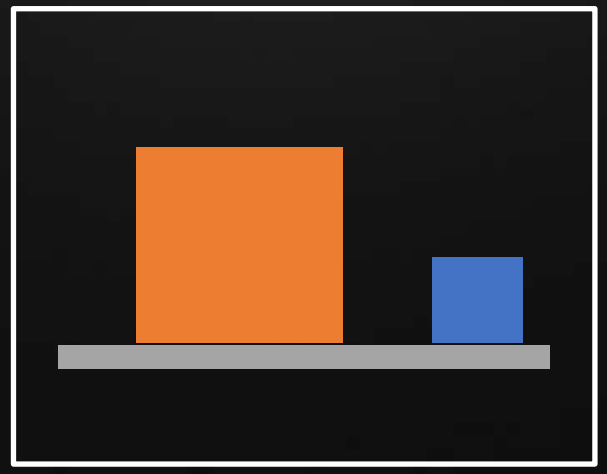




Rendering



Modeling

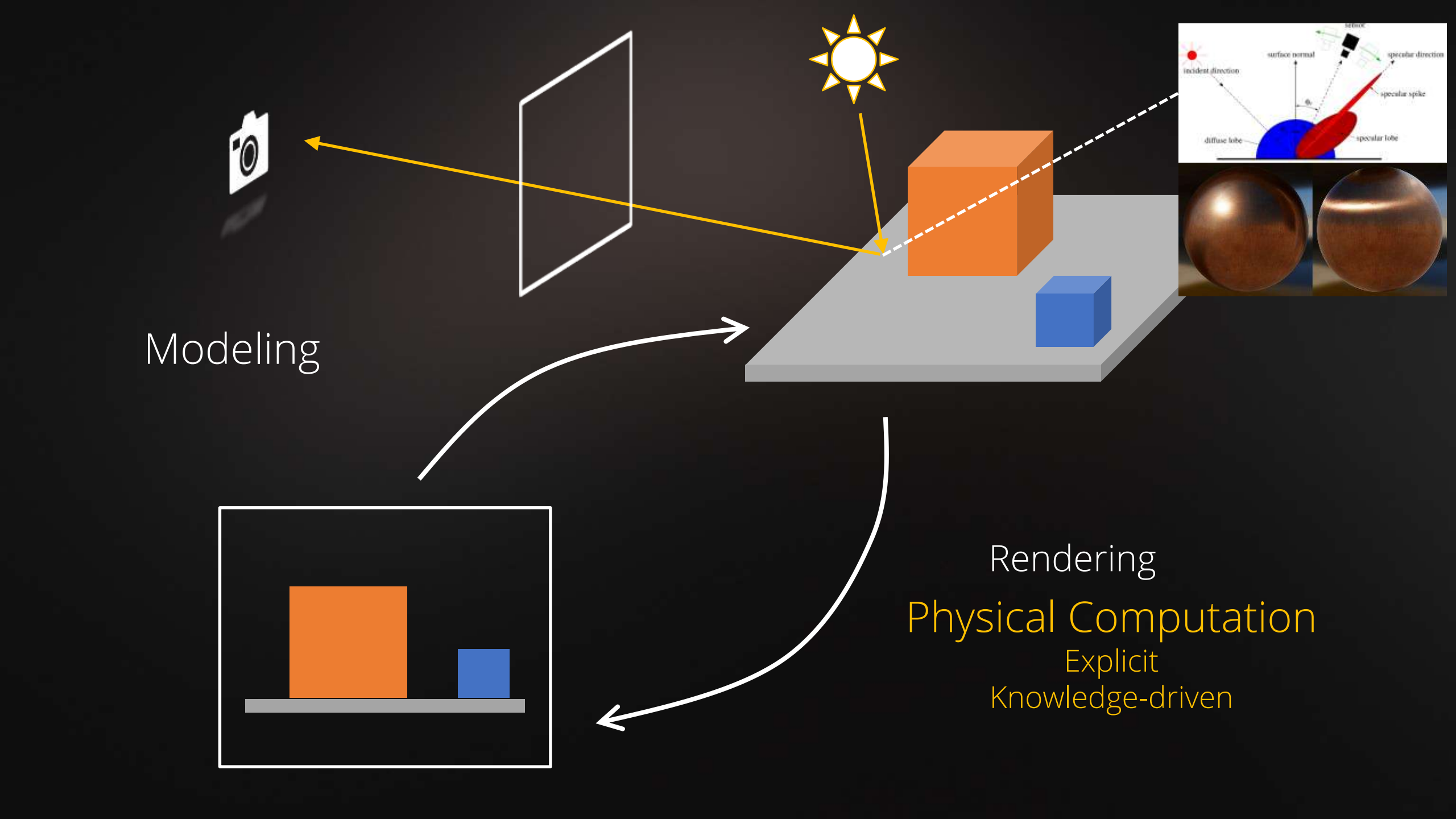


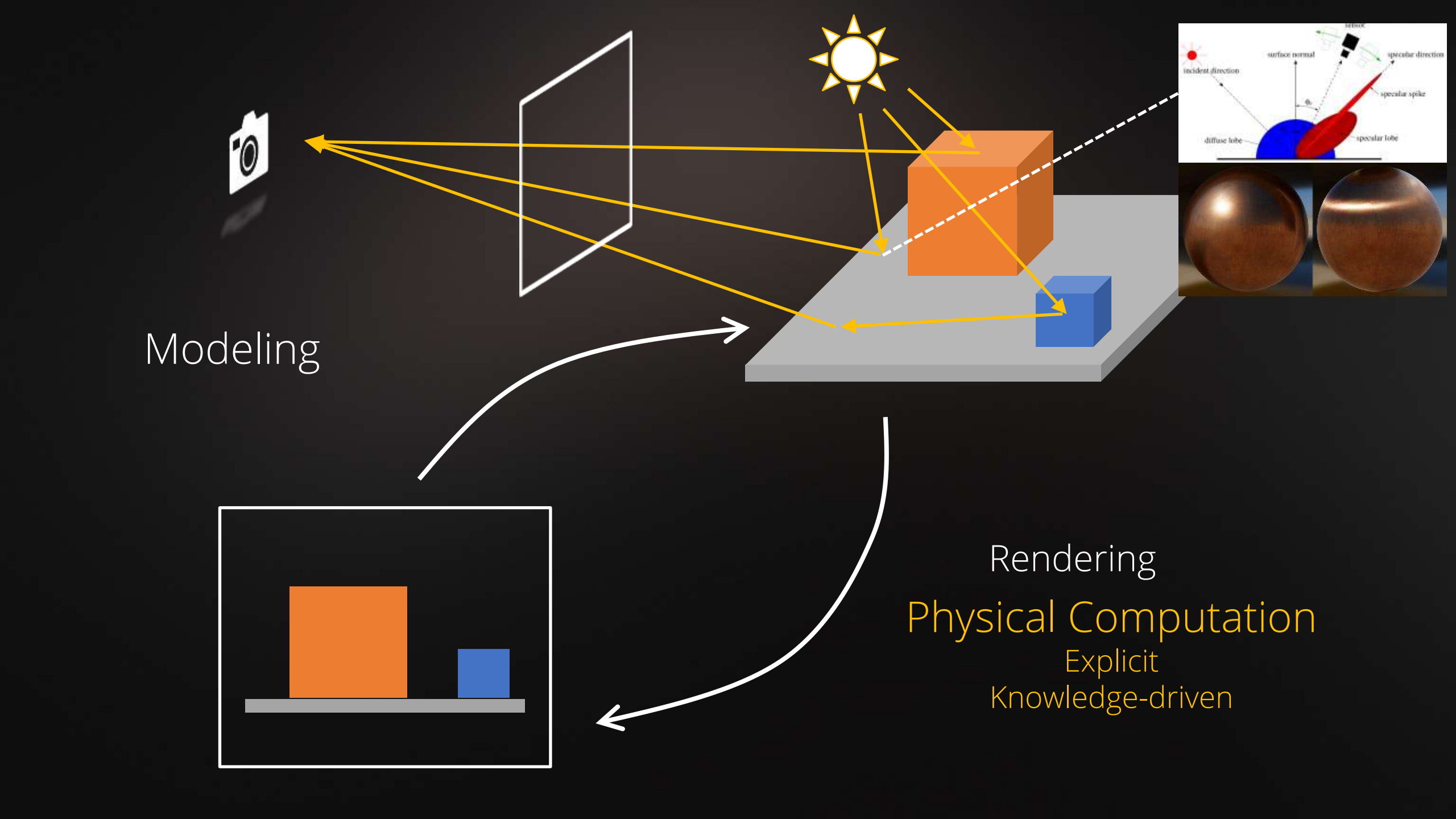
Rendering

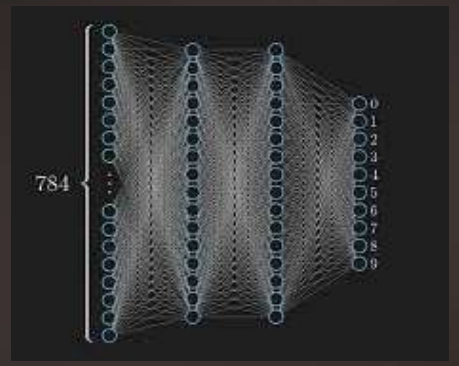
Physical Computation

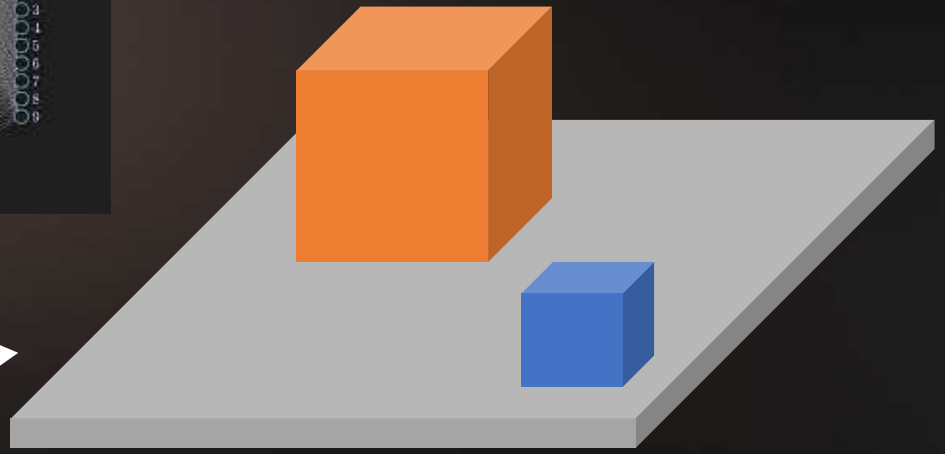
Explicit

Knowledge-driven


Modeling
Rendering
Physical Computation
Explicit
Knowledge-driven
surface normal
specular direction
incident direction
specular spike
diffuse lobe
specular lobe


Modeling
Rendering
Physical Computation
Explicit
Knowledge-driven
surface normal
specular direction
incident direction
specular spike
diffuse lobe
specular lobe



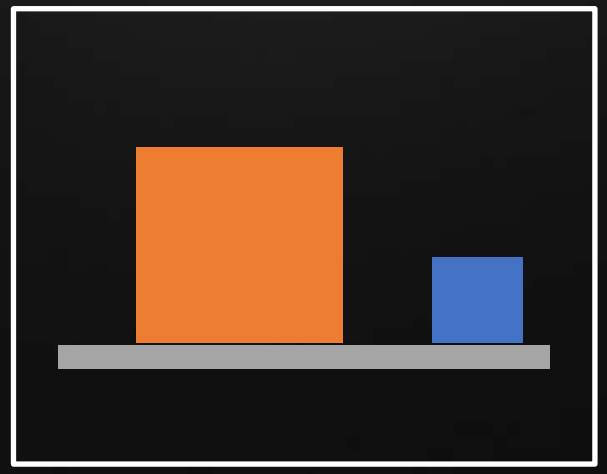


Modeling

Neural Computation

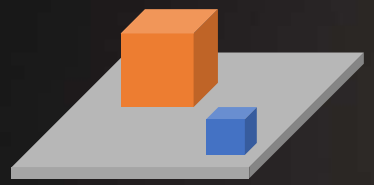
Implicit

Data-driven



Rendering

Fusing Knowledge and Data



3D world

Neural Rendering

Physical Rendering

Computer Vision

Differentiable Rendering

2D Image

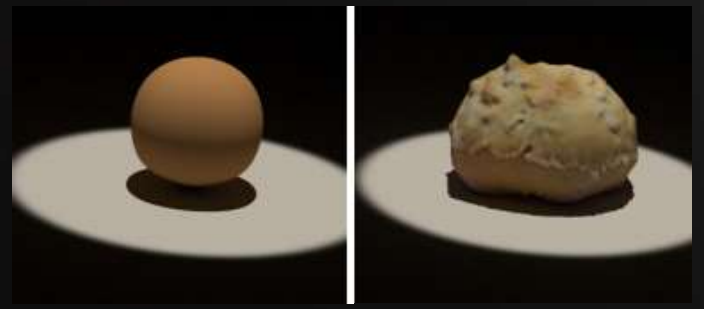
Neural Computation

Physical Computation

Neural Rendering

Physical Rendering

Computer Vision



Differentiable Rendering

The Graphics Pipeline

The Graphics Pipeline

Reconstruction

Generation

Interaction

Rendering

Computing

The Graphics Pipeline with AI

Reconstruction

Generation

Interaction

Rendering

Computing

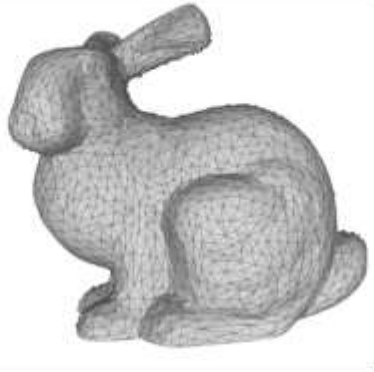
2. The New Wave

Graphics Pipeline Beyond Triangles

Point cloud

$\{\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_N\}$

Unordered set of points



Mesh

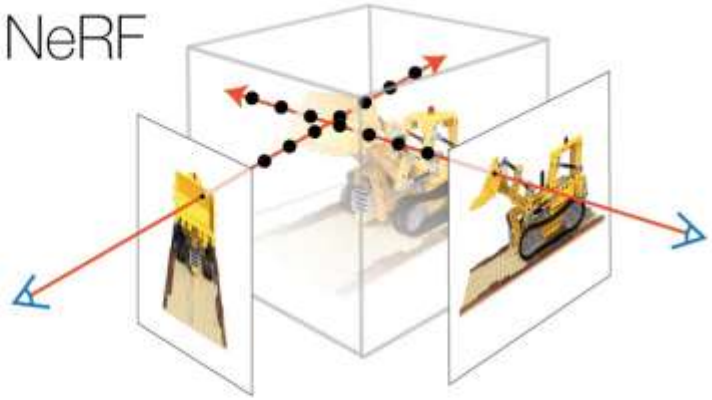
Vertices + Faces

Voxel

$V[x, y, z] \in [0,1]$

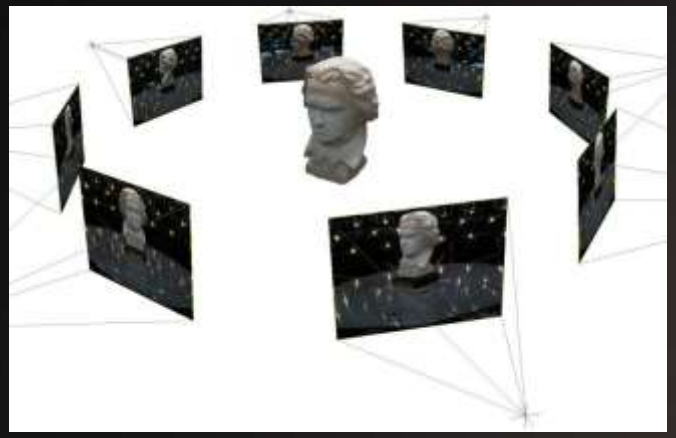
Signed distance function (SDF)

$\{\mathbf{p} \mid f(\mathbf{p})\} = 0$



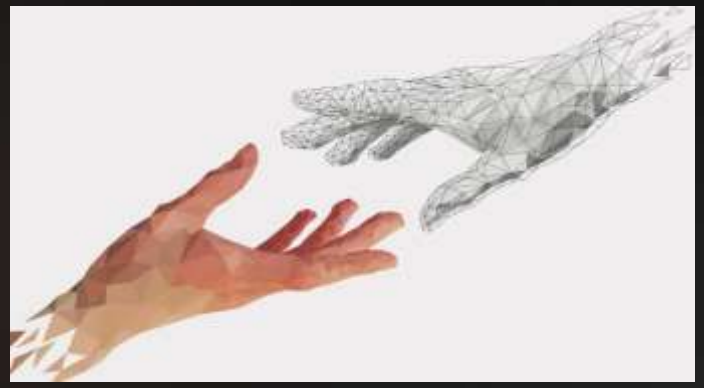
Neural Radiance Field

Network-based Field

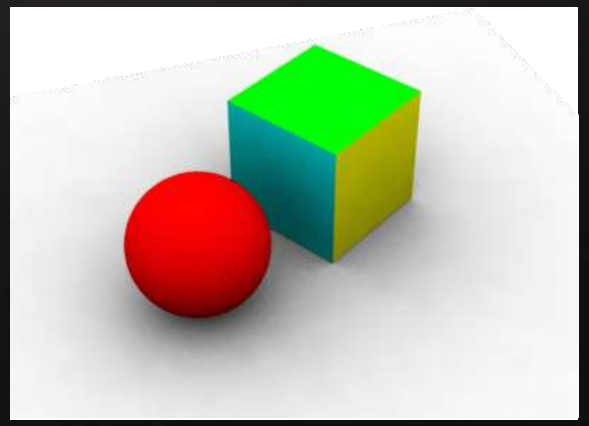


Reconstruction

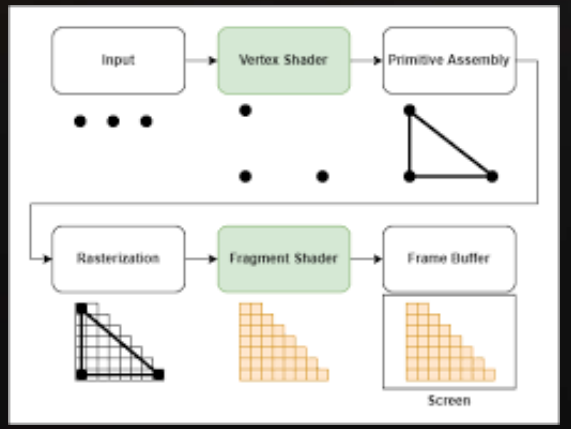
Generation



Interaction



Rendering



Computing

Reconstruction	Single-View（Mat.+Geo.+Light）[SIGGRAPH ASIA 22] Multi-View（ Mat.+Geo.+Light ）[CVPR 23]
Generation	Geometry + Light AIGC [ACM TOG 23] Gesture AIGC [IJCAI 23]
Interaction	NeRF Editing [ICCV 23] Shader Editing [SIGGRAPH 22]
Rendering	Super Resolution [SIGGRAPH ASIA 23] Frame Prediction [SIGGRAPH ASIA 23]
Computing	Object-Oriented Neural Rendering [ACM TOG 23] Rendering for AI [Nature Communications 23]

Reconstruction

- Implicit Neural Representations
  - NeRF - network-based radiance field, high image quality, low geometry quality
  - SDF - network-based geometry field, lower image quality, higher geometry quality
  - Plenoctree - spherical harmonics-based radiance field, no neural networks
  - TensoRF – tensor-based radiance field, few artifacts
  - Instant NGP - voxel-based radiance field, real-time speed
  - Gaussian Splatting - point cloud-based radiance field, fastest till now
  - Adaptive Shell - network-based shell field, high geometry and image quality
- Speed
  - Offline → interactive → real-time → large scale
- Scalability
  - Object level → scene level → city level → dynamics
- Components
  - Appearance → geometry → materials → lighting → semantics
- Materials
  - Diffuse → specular → all frequency

Generation

- Generate 3D Objects
  - Pseudo 3D - conditional generation model
  - Primitive to 3D - translate primitives to image
  - 2D to 3D - leverage pretrained 2D large models
  - 3D foundation - trained from 3D dataset
  - Procedure generation - generate from template
- Scalability
  - Object generation → scene generation
- Component
  - Object → texture → geometry → lighting → physically-based generation
- Animation
  - Rigid transformation → pose → key frames → deformable
- Modality
  - Text → example → mask → multi-modal

Interaction

- Editing
  - Mesh policy - control via a mesh policy
  - Pixel level - project pixel to 3D
  - Decomposition - shape and color
- Modality
  - Text → mouse → mask → gesture → multi-modal
- Control
  - Transformation → pose → rigging
- Physics
  - Rigid → collision → deformable
- Speed
  - Offline → interactive → real-time

Seal-3D: Interactive Pixel-Level Editing for Neural Radiance Fields

Rendering

- Acceleration
  - Denoising - smooth path tracing
  - Super Resolution - increase render resolution
  - Frame Prediction - extrapolate new frames
  - Neural Rendering - implicit representation
  - Caching - reuse lighting result
  - IBL - lighting on image space
- Effect
  - Diffuse → glossy → specular
- Space
  - Image → temporal → path → feature
- Speed
  - Offline → interactive → real-time
- Generality
  - Effect → view → lighting → scene

Real-Time Monte Carlo Denoising with
Weight Sharing Kernel Prediction Network

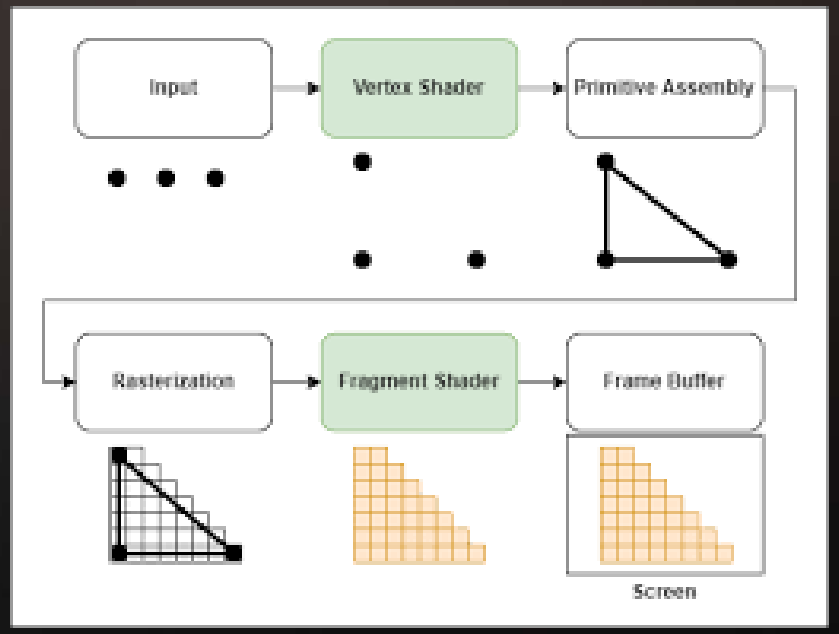
Submission ID: 1067

Result:

UE4 Scene
Kite Boy

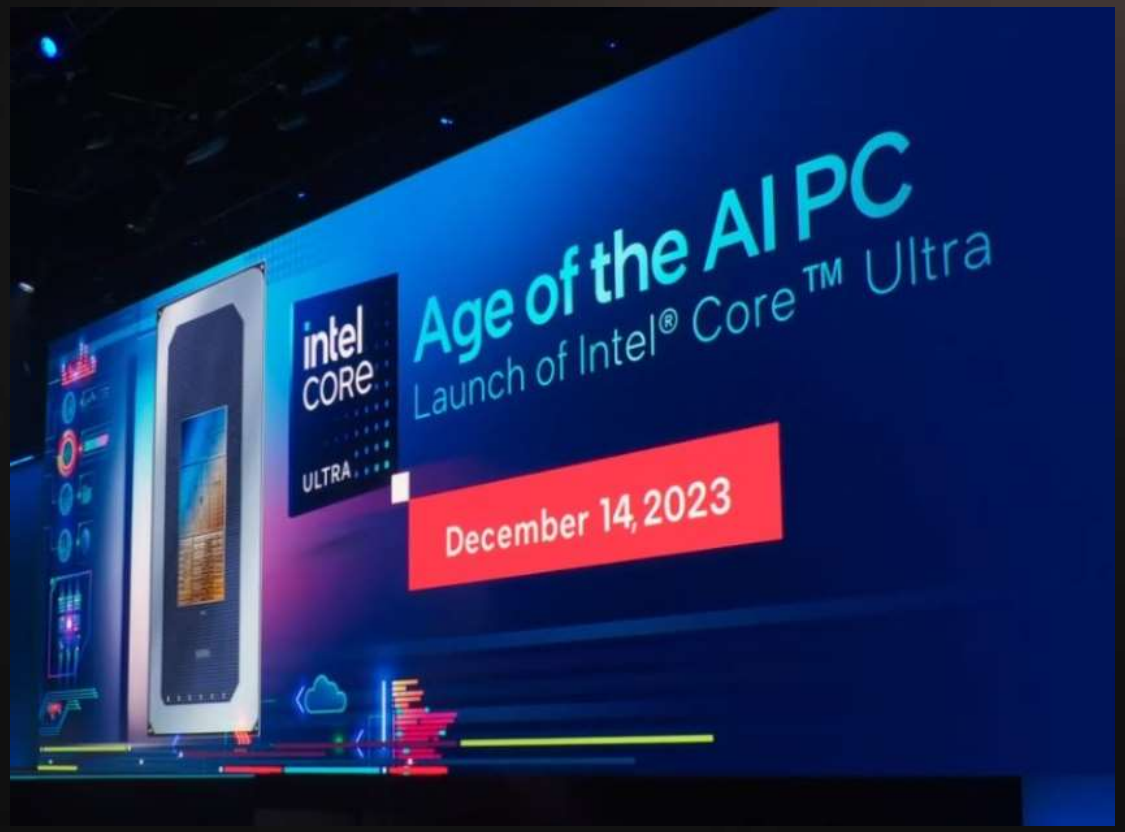
4X4 upsampling

Computing



Global Neural Rendering Pipeline

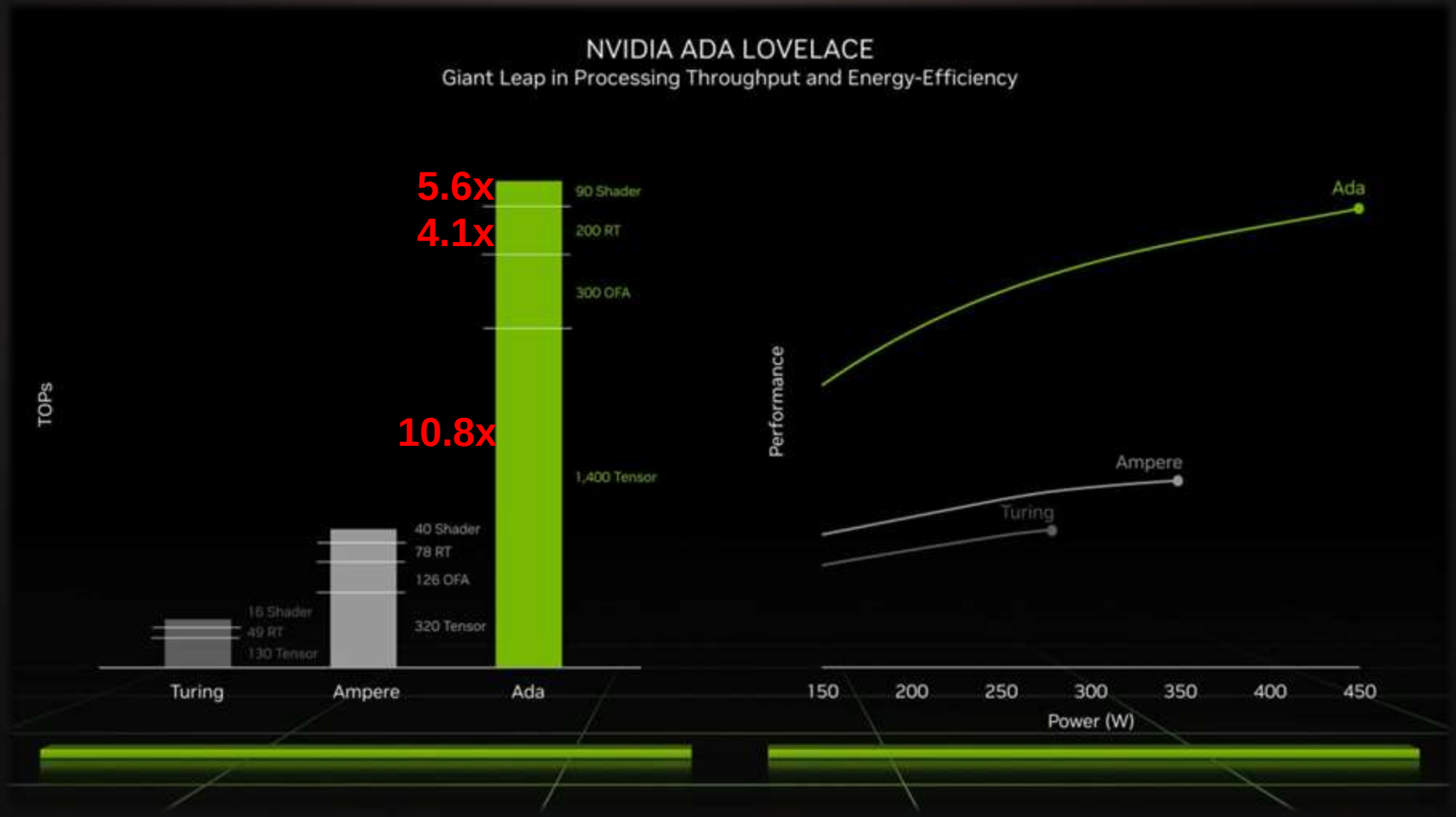
Age of the AI PC & Mobile



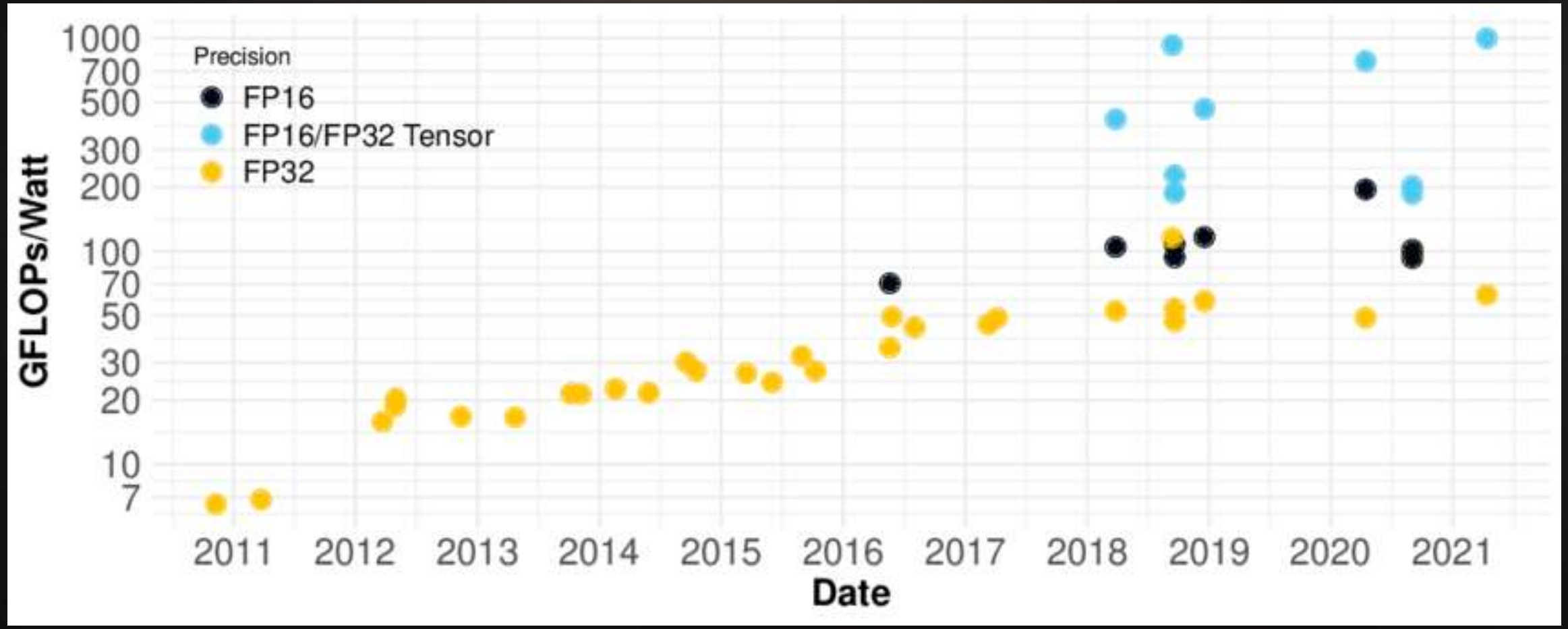
英特尔“AI PC 加速计划”将在 2025 年前为超过 1亿台 PC 带来人工智能（AI）特性

Windows 12将于今年秋季发布，windows键盘将迎来30年来最重大的变化

The Growth

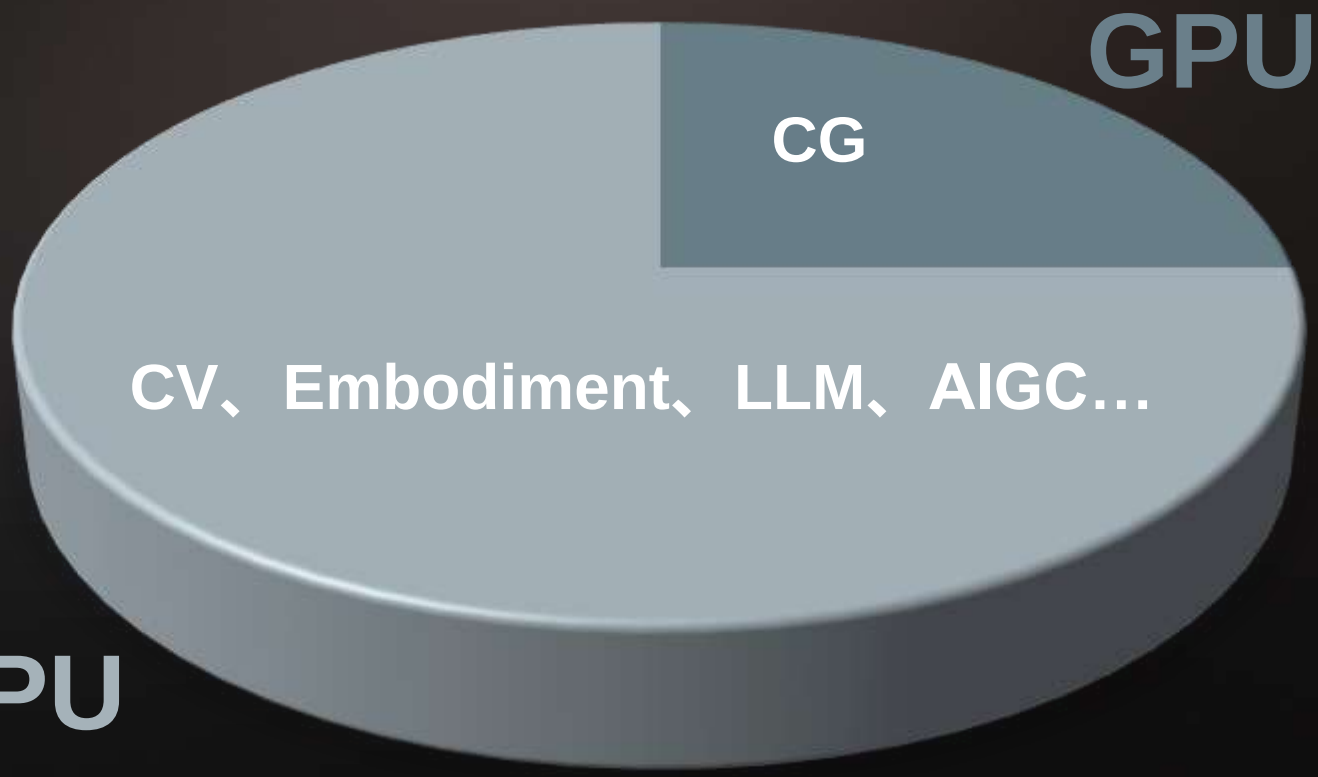

NVIDIA ADA LOVELACE
Giant Leap in Processing Throughput and Energy-Efficiency
TOPs
5.6x
4.1x
10.8x
90 Shader
200 RT
300 OFA
1,400 Tensor
40 Shader
78 RT
126 OFA
320 Tensor
16 Shader
49 RT
130 Tensor
Turing
Ampere
Ada
Performance
Ada
Ampere
Turing
150
200
250
300
350
400
450
Power (W)

The Efficiency

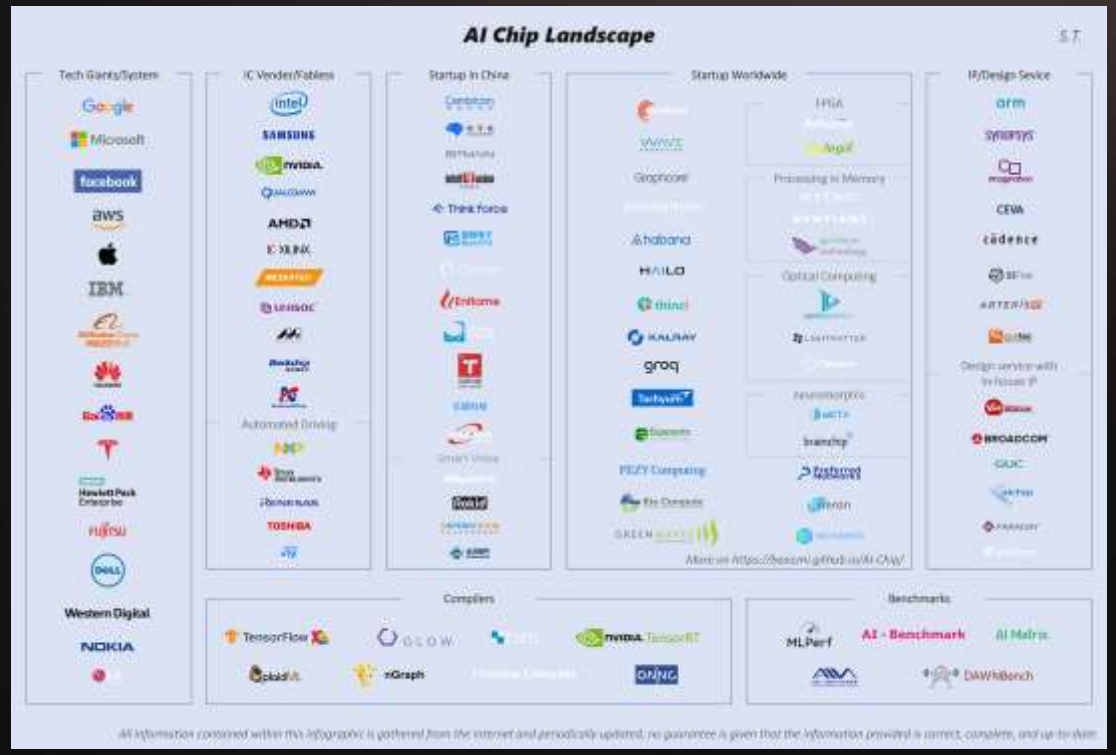


Nvidia显卡的算力/功耗比

The Generality



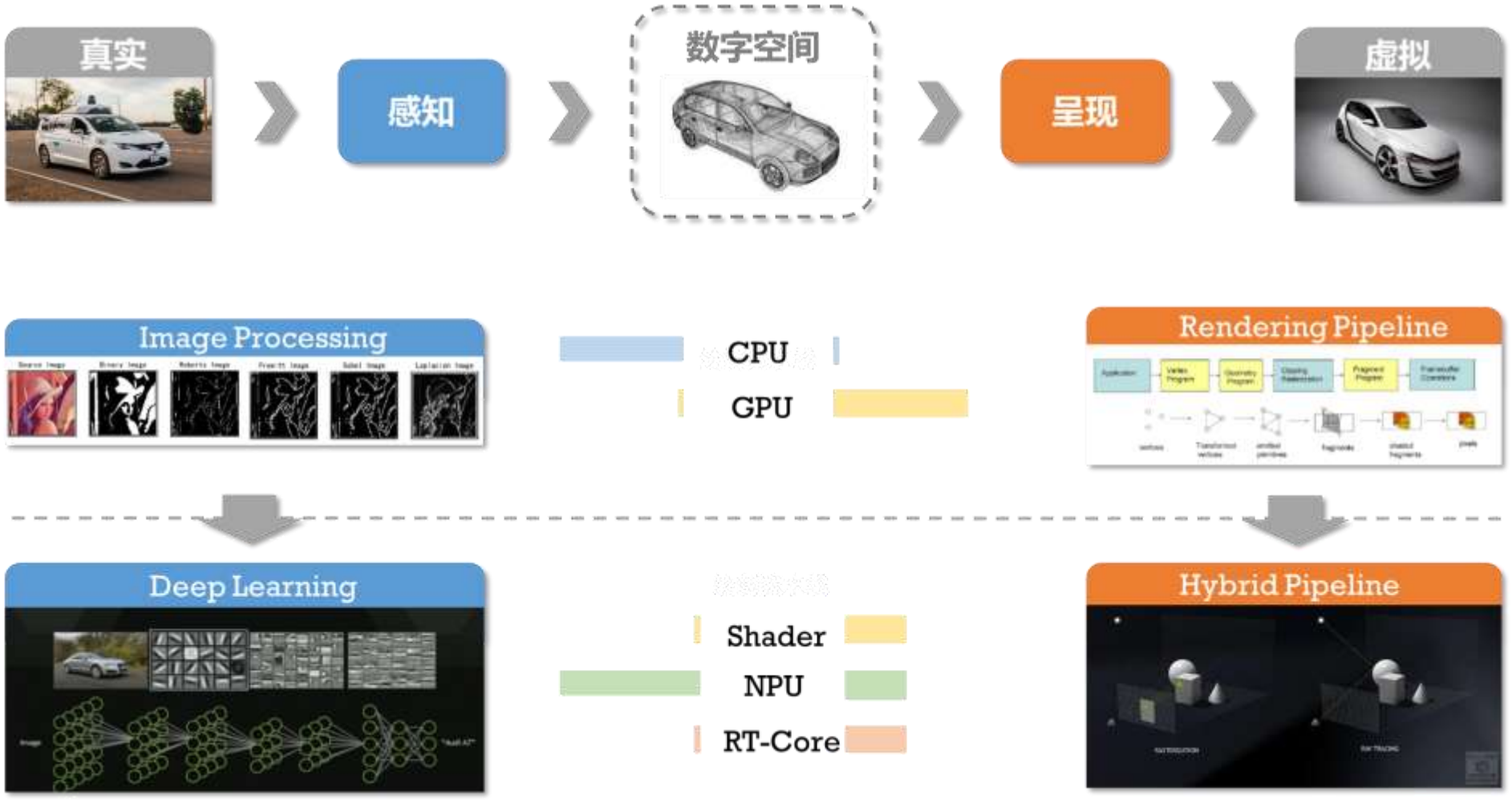
The Community



AI

GPU

How to Render in AI Age ?



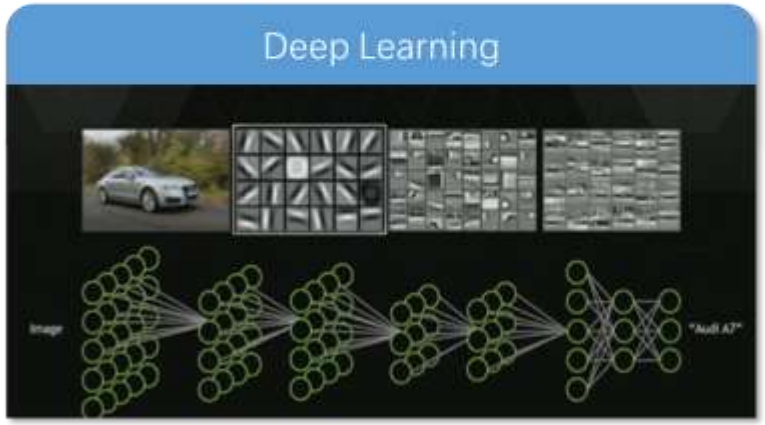
第一代

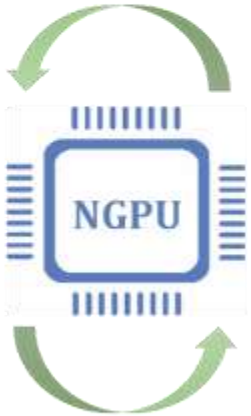
- Shader
- 局部光照

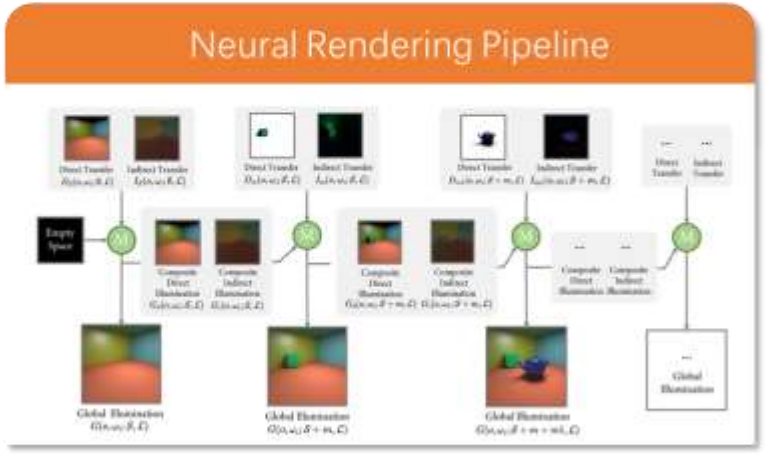
第二代

- RT-Core
- 全局光照

How to Render in AI Age ?







Neural Graphics Processing Unit

下一代

- NPU
- 神经绘制

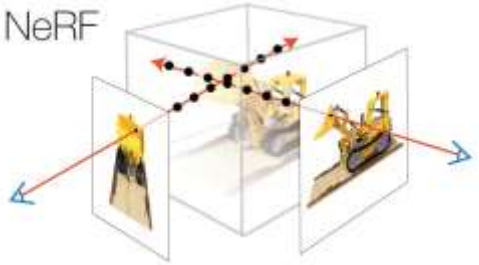
How? Build a Pipeline with Neural Representation

Point cloud

Mesh

Voxel

Signed distance function (SDF)



Neural Radiance Field

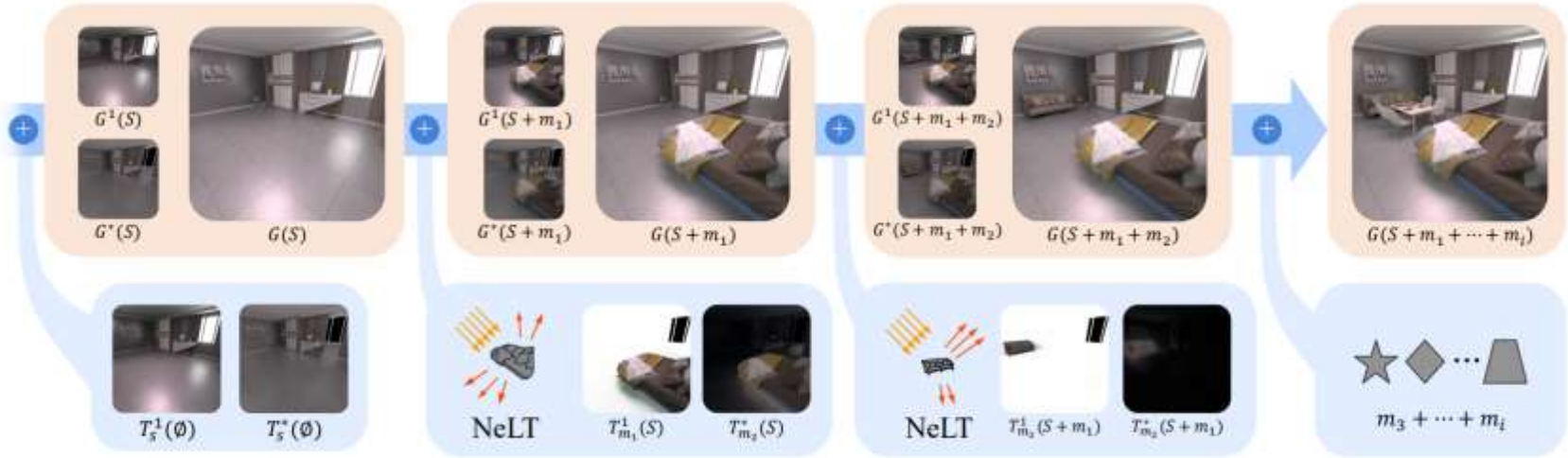
Shadow

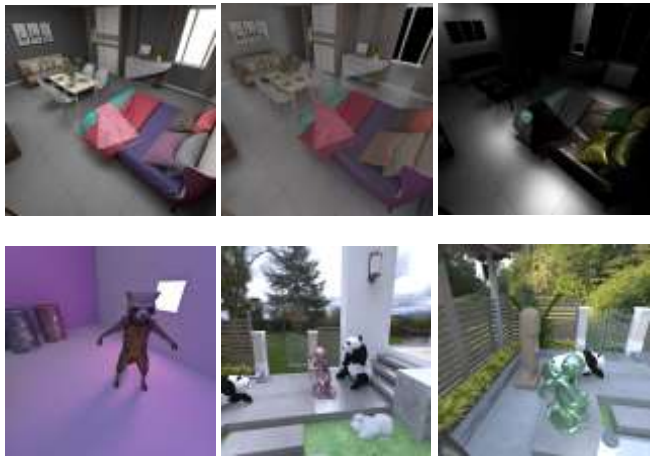
Reflection

Bleeding

⋮

1st : Object-Oriented Pipeline (TOG23)



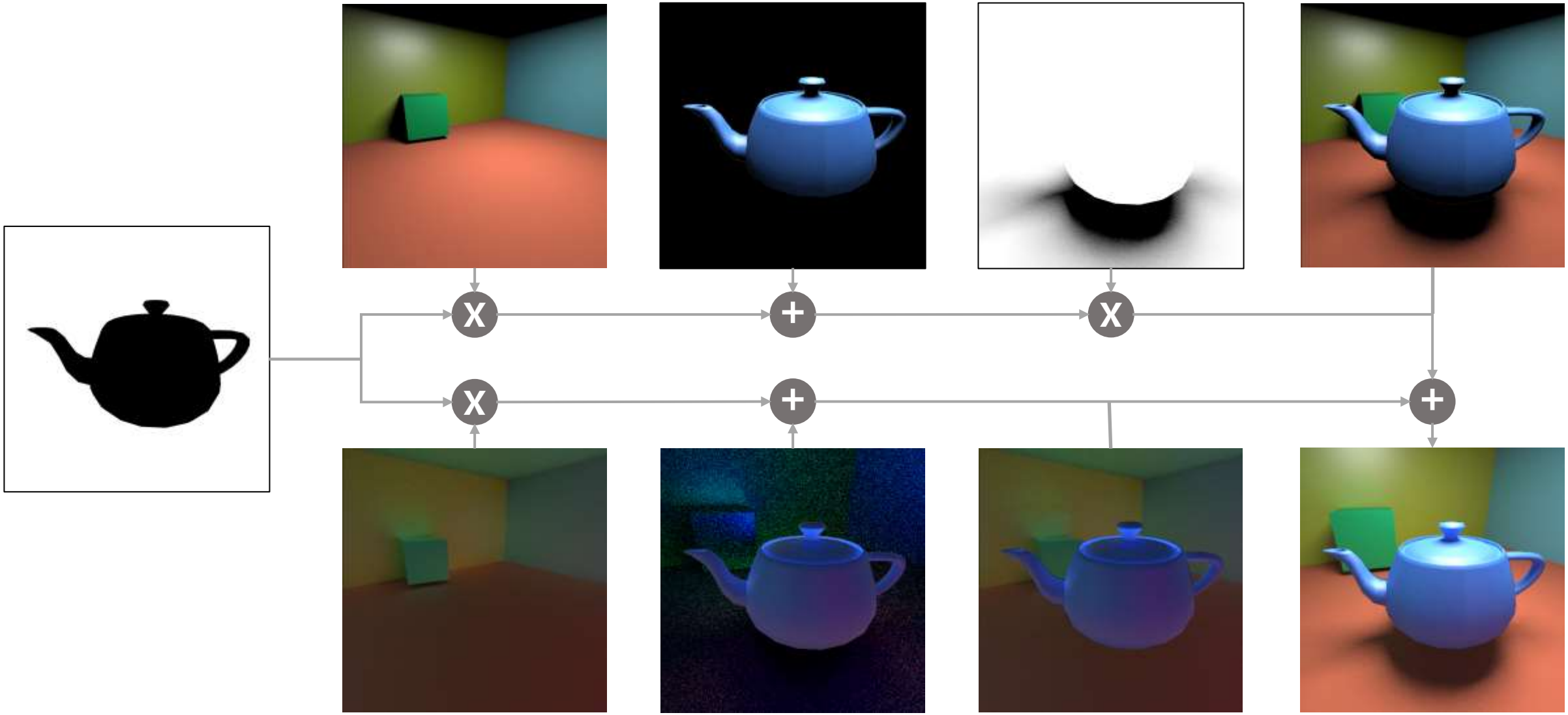


The Object-Oriented Pipeline

直接光传输变化

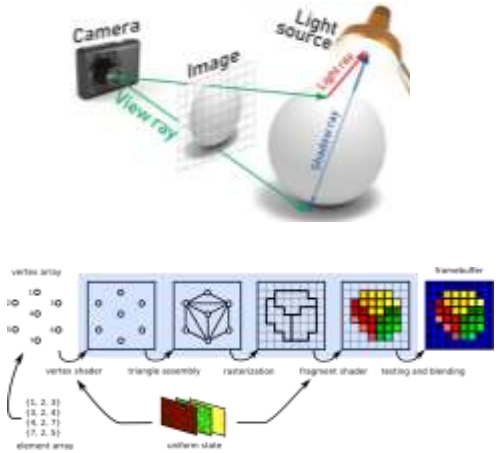
间接光传输变化

The Object-Oriented Pipeline

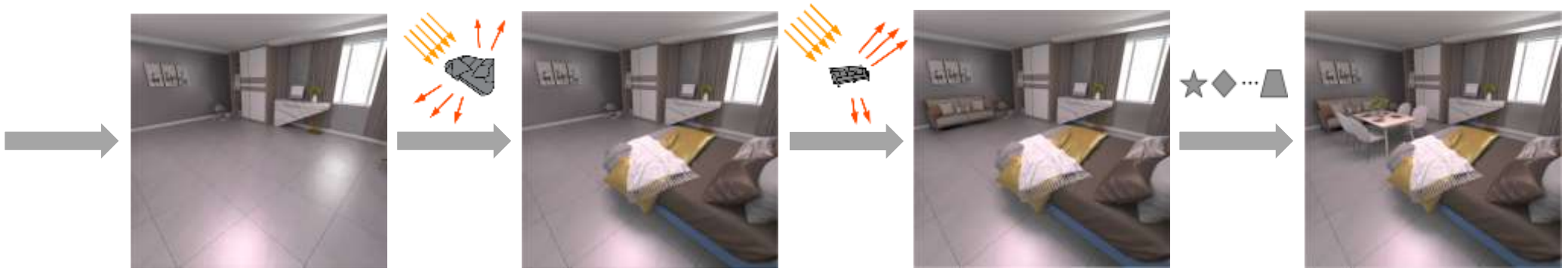


The Object-Oriented Pipeline

传统绘制



神经绘制



Compare Path Tracing

Ground truth

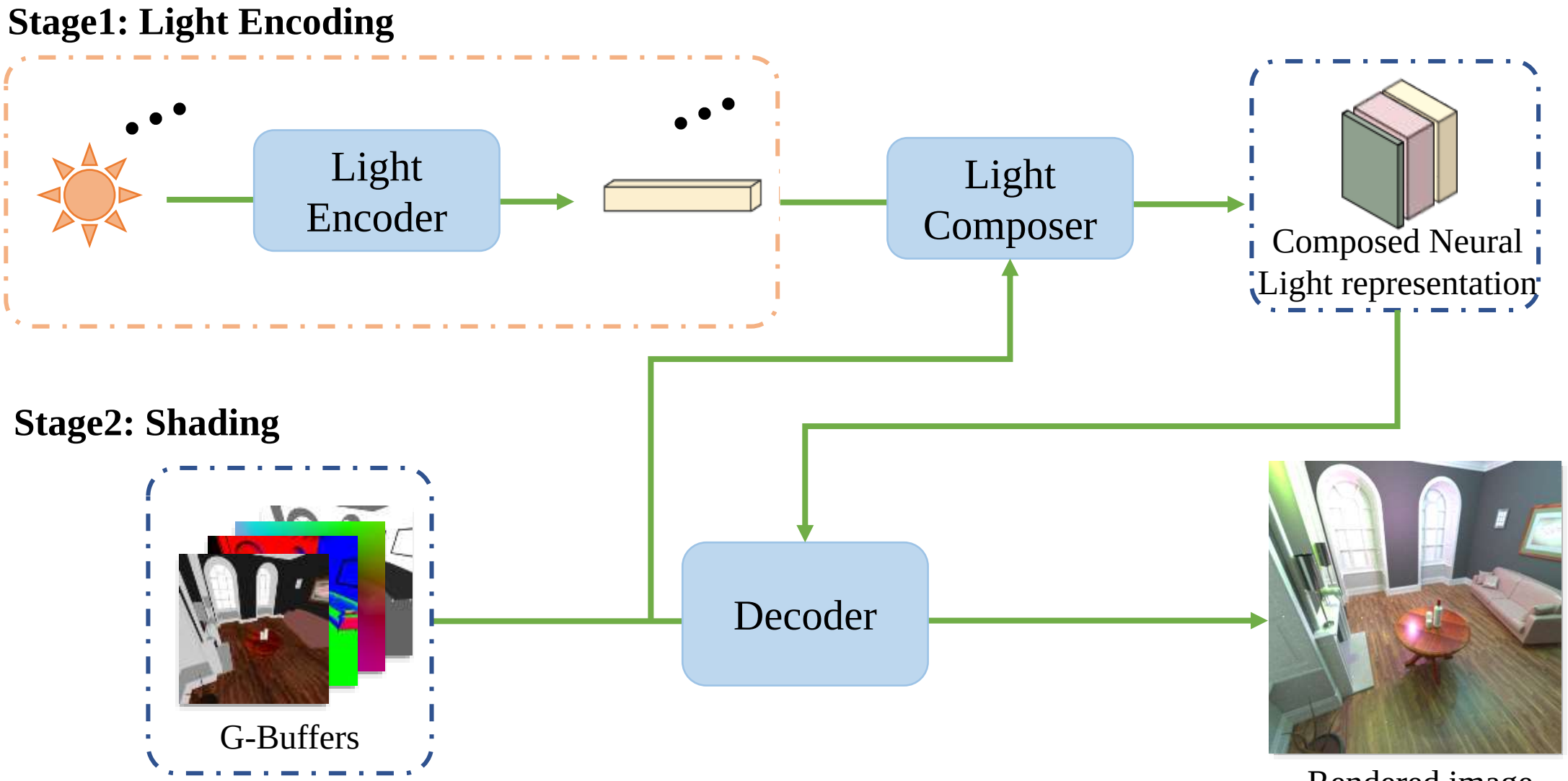
Ours

AFGSA [Yu et al. 2021]

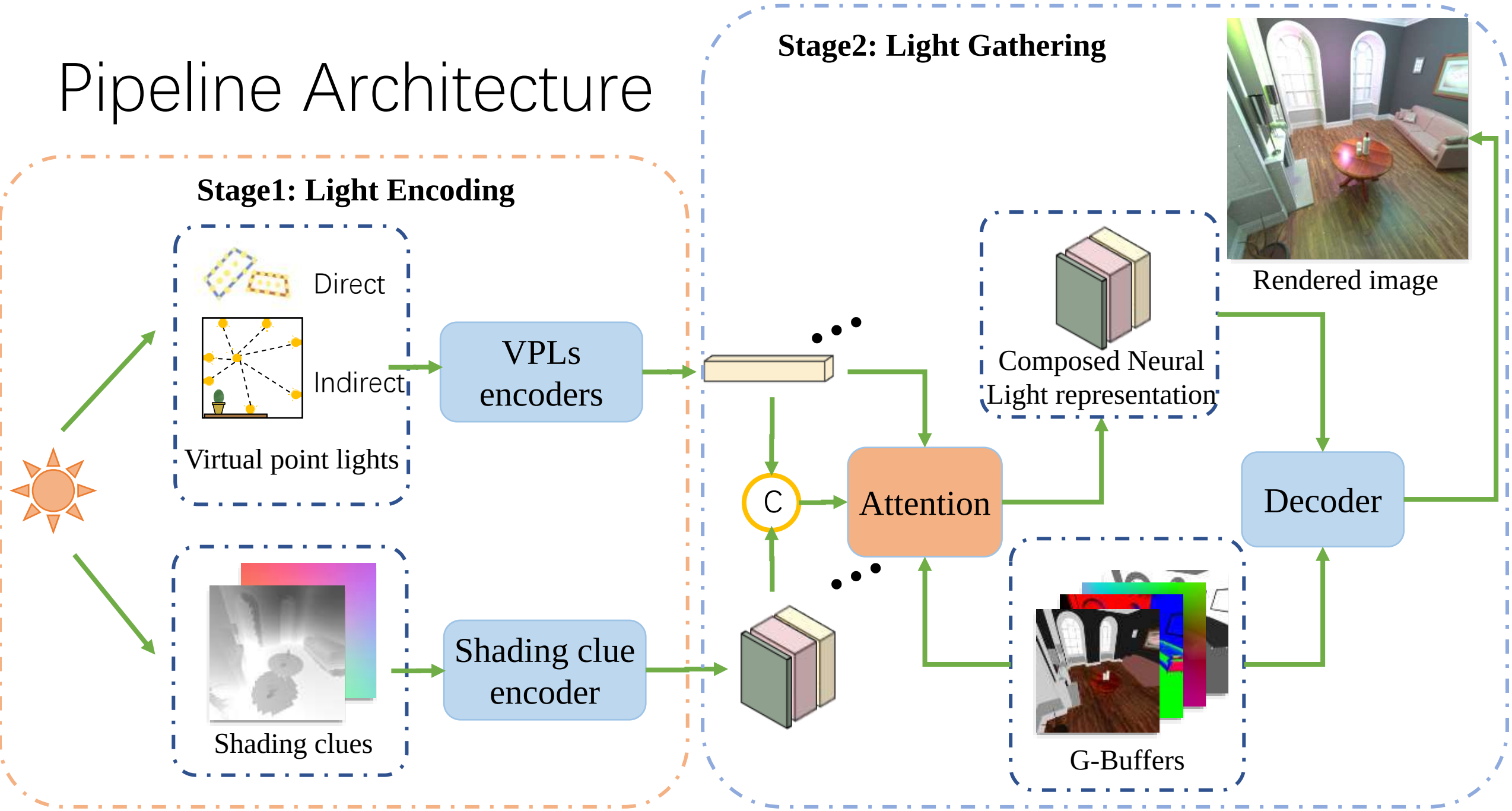
Dynamic Lighting & Material

Scene Pipeline

2nd : Light-Oriented Pipeline (SIG24)

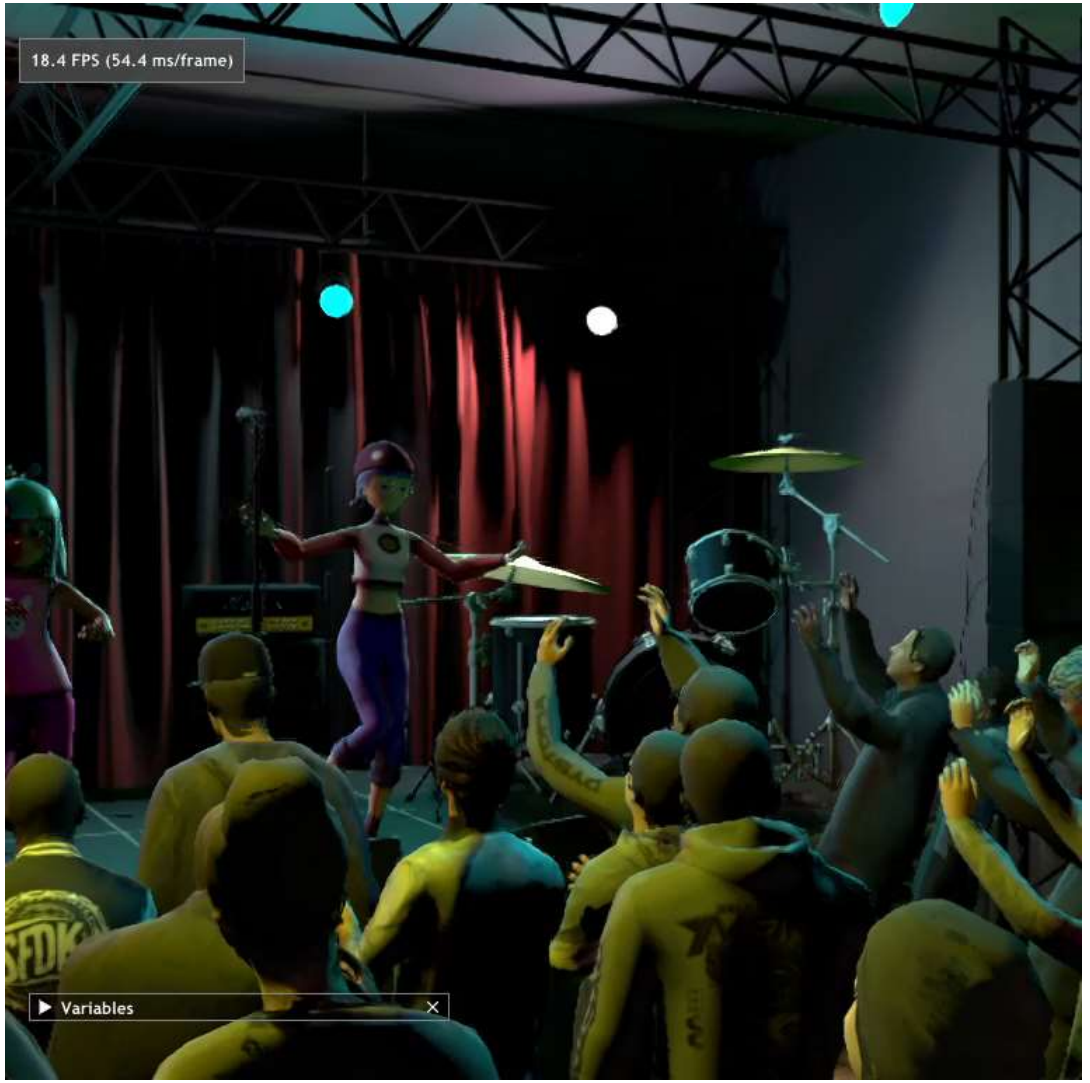


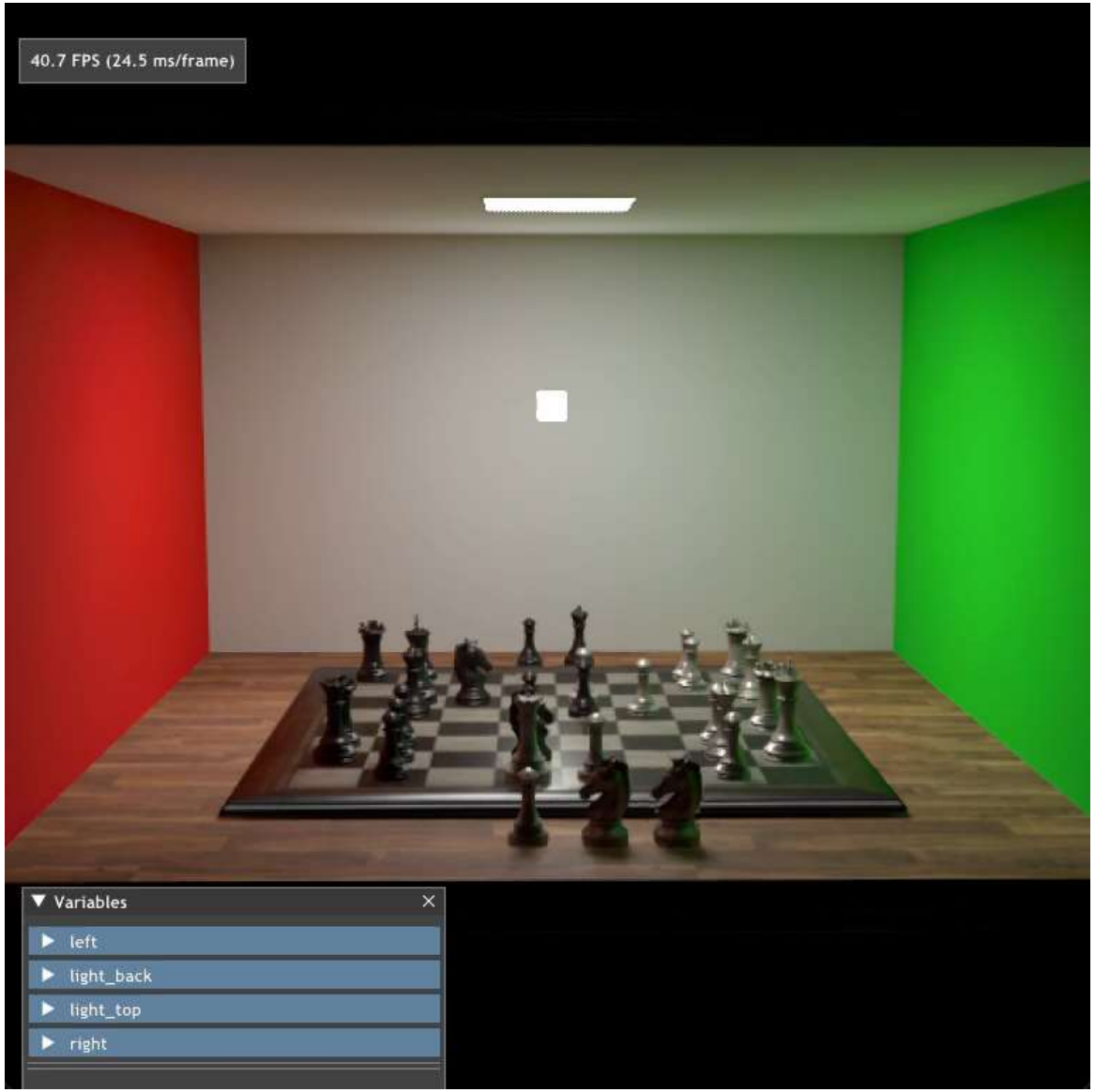
Pipeline Architecture

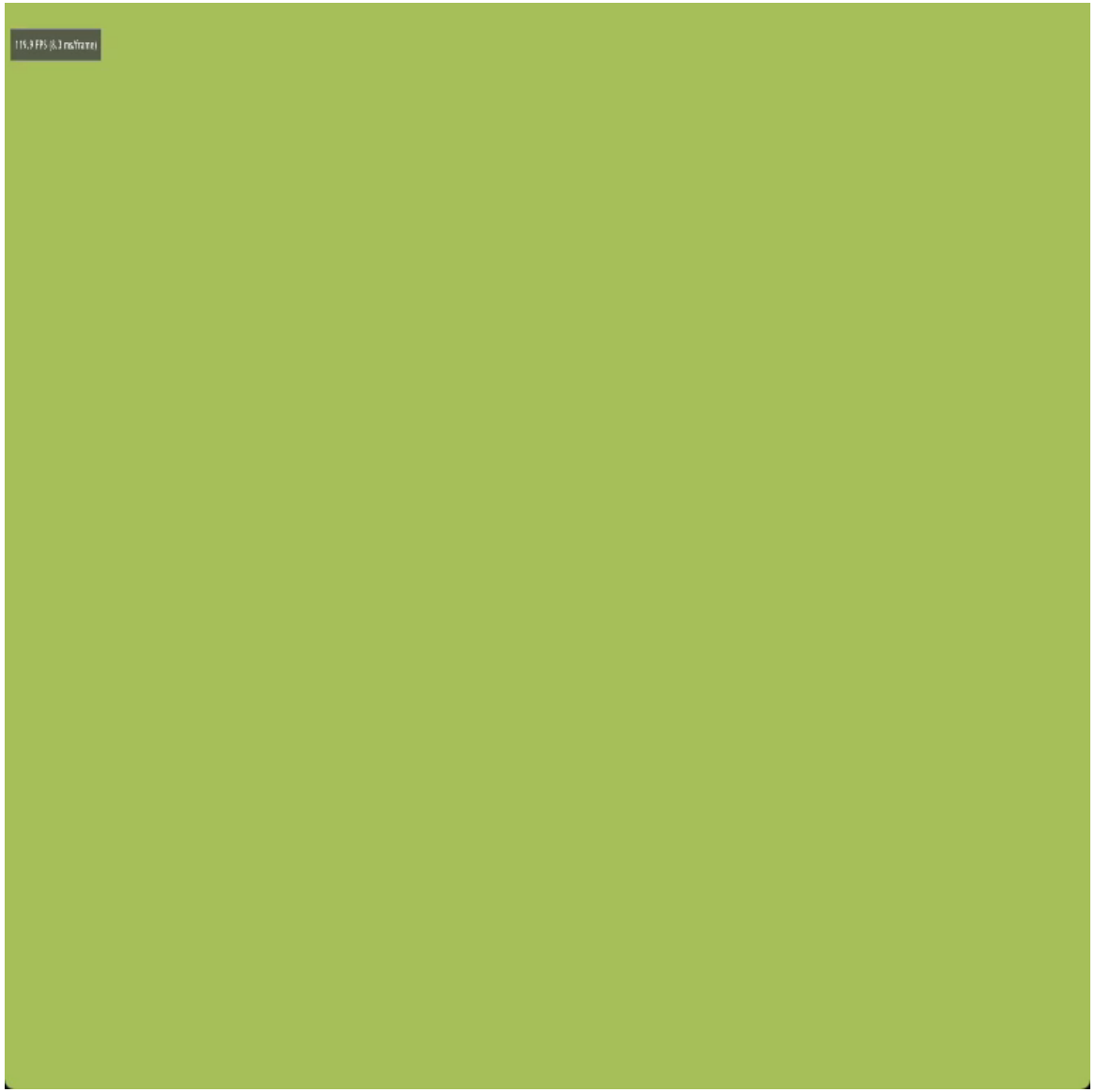


Compare Path Tracing

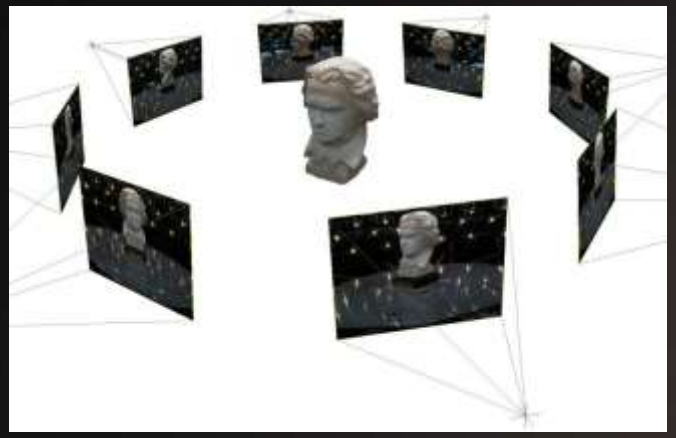

34.9 FPS (28.7 ms/frame)
Variables
Floor
Sofa
Walls
cross
hexagon
star


18.4 FPS (54.4 ms/frame)
Variables

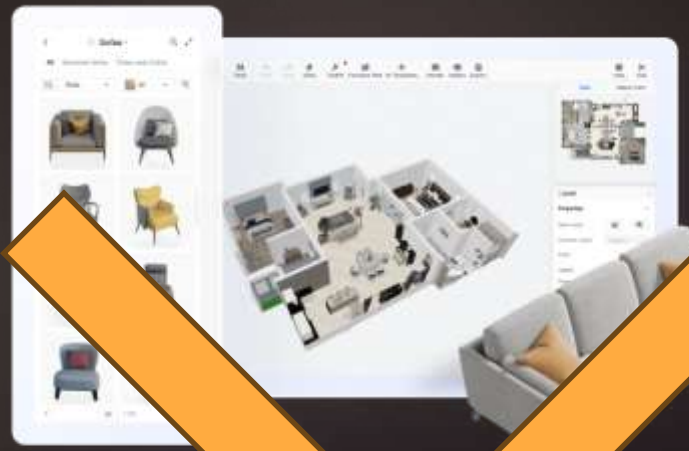

40.7 FPS (24.5 ms/frame)
Variables
left
light_back
light_top
right



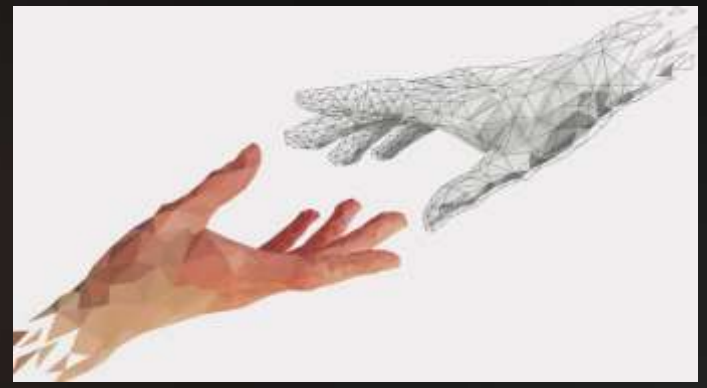
3. SORA



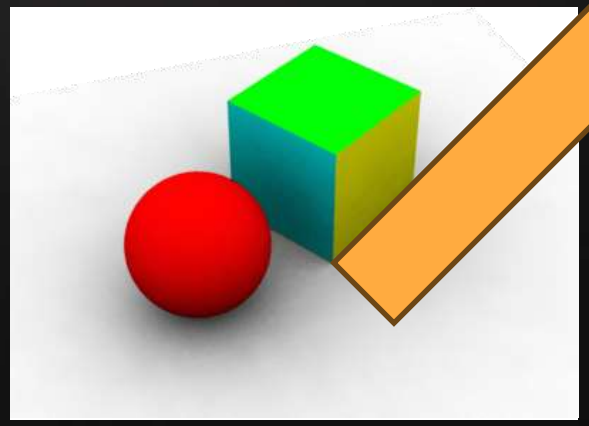
Reconstruction



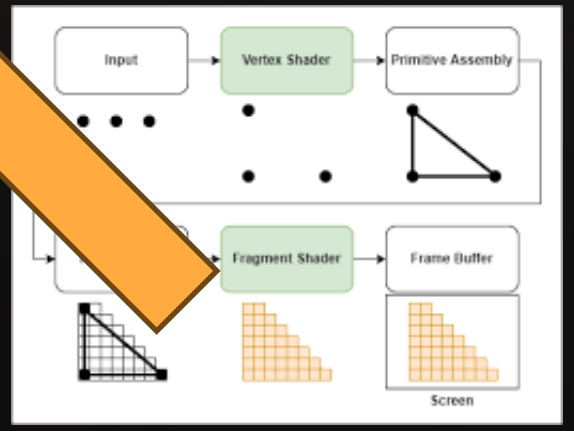
Ge



Interaction



Rendering



Computing

Graphics Pipeline

SORA

分阶段

模糊

THANKS

浙江大学 · CAD&CG　霍宇驰