Pixels at Speed of Light Lessons in Deploying Computer Vision in Graphics & Games

Anjul Patney, Sr. Manager | CV4MR Workshop 2025

From Research to Reality

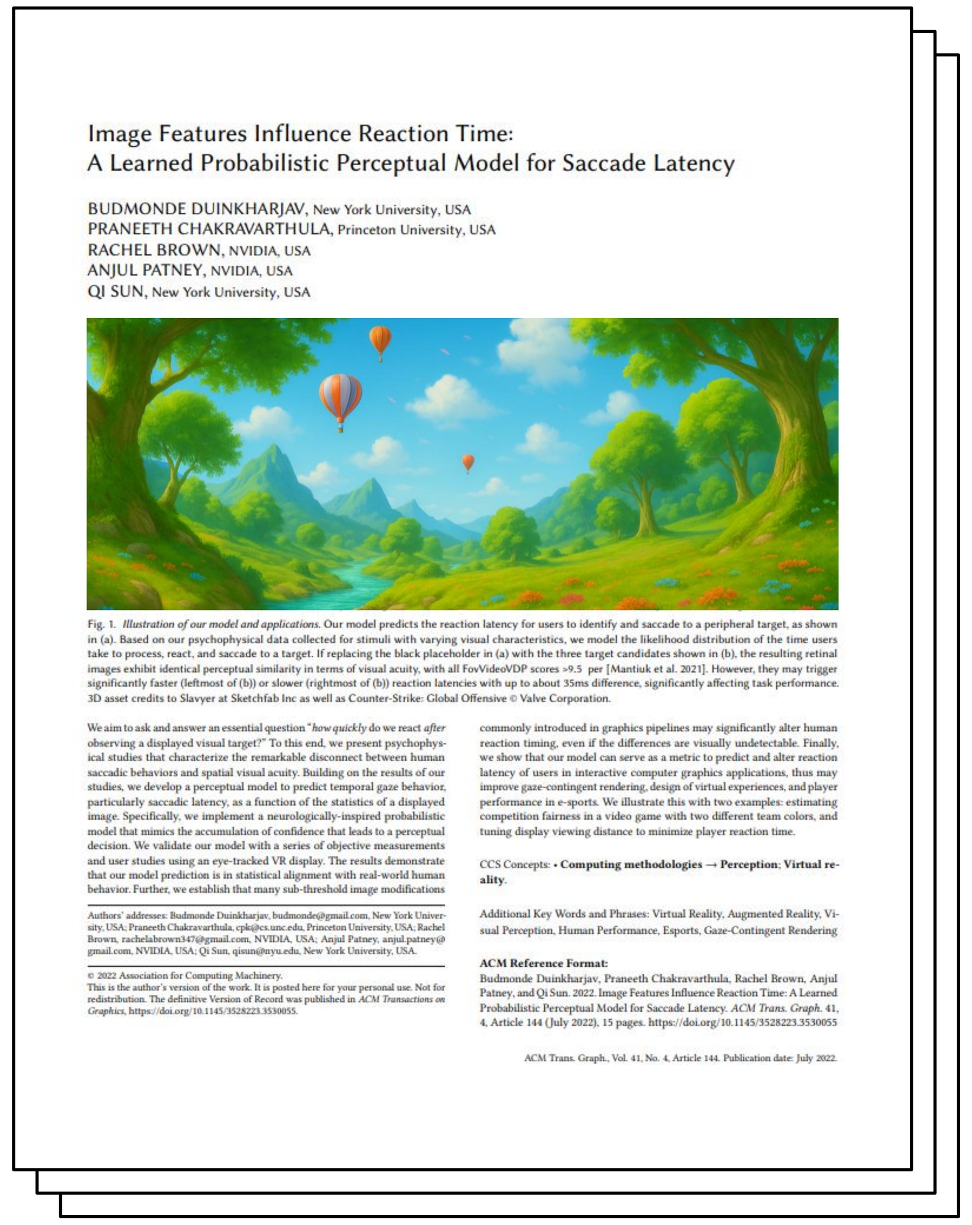


Image Features Influence Reaction Time:
A Learned Probabilistic Perceptual Model for Saccade Latency

BUDMONDE DUINKHARJAV, New York University, USA
PRANEETH CHAKRAVARTHULA, Princeton University, USA
RACHEL BROWN, NVIDIA, USA
ANJUL PATNEY, NVIDIA, USA
QI SUN, New York University, USA

Fig. 1. *Illustration of our model and applications.* Our model predicts the reaction latency for users to identify and saccade to a peripheral target, as shown in (a). Based on our psychophysical data collected for stimuli with varying visual characteristics, we model the likelihood distribution of the time users take to process, react, and saccade to a target. If replacing the black placeholder in (a) with the three target candidates shown in (b), the resulting retinal images exhibit identical perceptual similarity in terms of visual acuity, with all FovVideoVDP scores >9.5 per [Mantiuk et al. 2021]. However, they may trigger significantly faster (leftmost of (b)) or slower (rightmost of (b)) reaction latencies with up to about 35ms difference, significantly affecting task performance. 3D asset credits to Slavyer at Sketchfab Inc as well as Counter-Strike: Global Offensive © Valve Corporation.

We aim to ask and answer an essential question "*how quickly* do we react *after* observing a displayed visual target?" To this end, we present psychophysical studies that characterize the remarkable disconnect between human saccadic behaviors and spatial visual acuity. Building on the results of our studies, we develop a perceptual model to predict temporal gaze behavior, particularly saccadic latency, as a function of the statistics of a displayed image. Specifically, we implement a neurologically-inspired probabilistic model that mimics the accumulation of confidence that leads to a perceptual decision. We validate our model with a series of objective measurements and user studies using an eye-tracked VR display. The results demonstrate that our model prediction is in statistical alignment with real-world human behavior. Further, we establish that many sub-threshold image modifications commonly introduced in graphics pipelines may significantly alter human reaction timing, even if the differences are visually undetectable. Finally, we show that our model can serve as a metric to predict and alter reaction latency of users in interactive computer graphics applications, thus may improve gaze-contingent rendering, design of virtual experiences, and player performance in e-sports. We illustrate this with two examples: estimating competition fairness in a video game with two different team colors, and tuning display viewing distance to minimize player reaction time.

CCS Concepts: • **Computing methodologies → Perception**; **Virtual reality**.

Additional Key Words and Phrases: Virtual Reality, Augmented Reality, Visual Perception, Human Performance, Esports, Gaze-Contingent Rendering

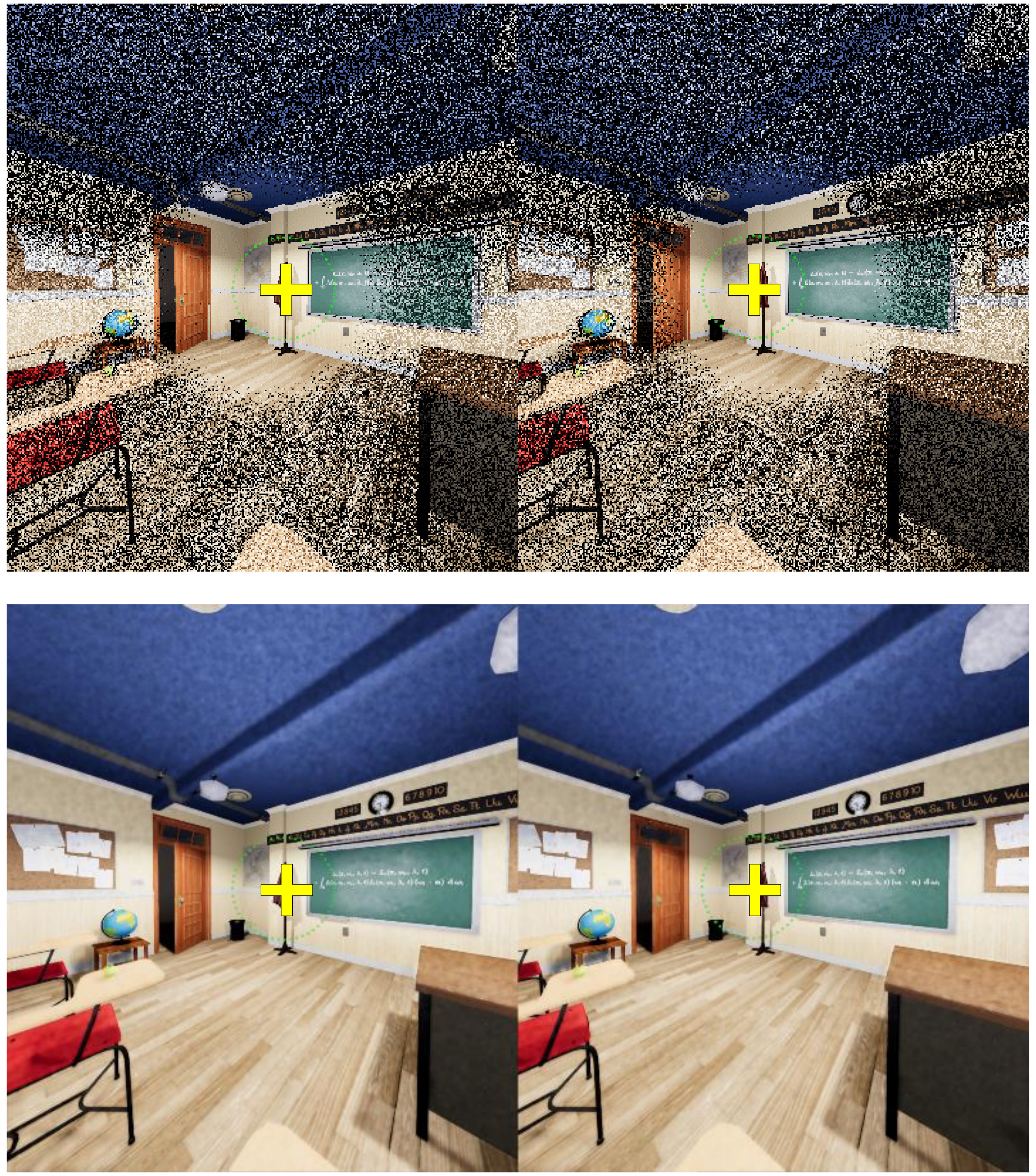
Authors' addresses: Budmonde Duinkharjav, budmonde@gmail.com, New York University, USA; Praneeth Chakravarthula, cpk@cs.unc.edu, Princeton University, USA; Rachel Brown, rachelabrown347@gmail.com, NVIDIA, USA; Anjul Patney, anjul.patney@gmail.com, NVIDIA, USA; Qi Sun, qisun@nyu.edu, New York University, USA.

ACM Reference Format:
Budmonde Duinkharjav, Praneeth Chakravarthula, Rachel Brown, Anjul Patney, and Qi Sun. 2022. Image Features Influence Reaction Time: A Learned Probabilistic Perceptual Model for Saccade Latency. *ACM Trans. Graph.* 41, 4, Article 144 (July 2022), 15 pages. https://doi.org/10.1145/3528223.3530055

Introduction

- Leading a research-to-product team in the AI4G org at NVIDIA
- Career researcher in real-time computer graphics, AR/VR graphics, and AI for graphics
- Building real-life technologies that utilize AI to provide efficient and engaging virtual experiences

Gaze-Contingent Virtual Reality Graphics



Foveated Rendering

Redirected Walking

FovVideoVDP: Wide Field-of-View Video Quality Metrics

Reference

Test

FovVideoVDP

Visual Difference Map

DLSS – Real-Time AI Super Resolution for Games

Native 720p Rendering

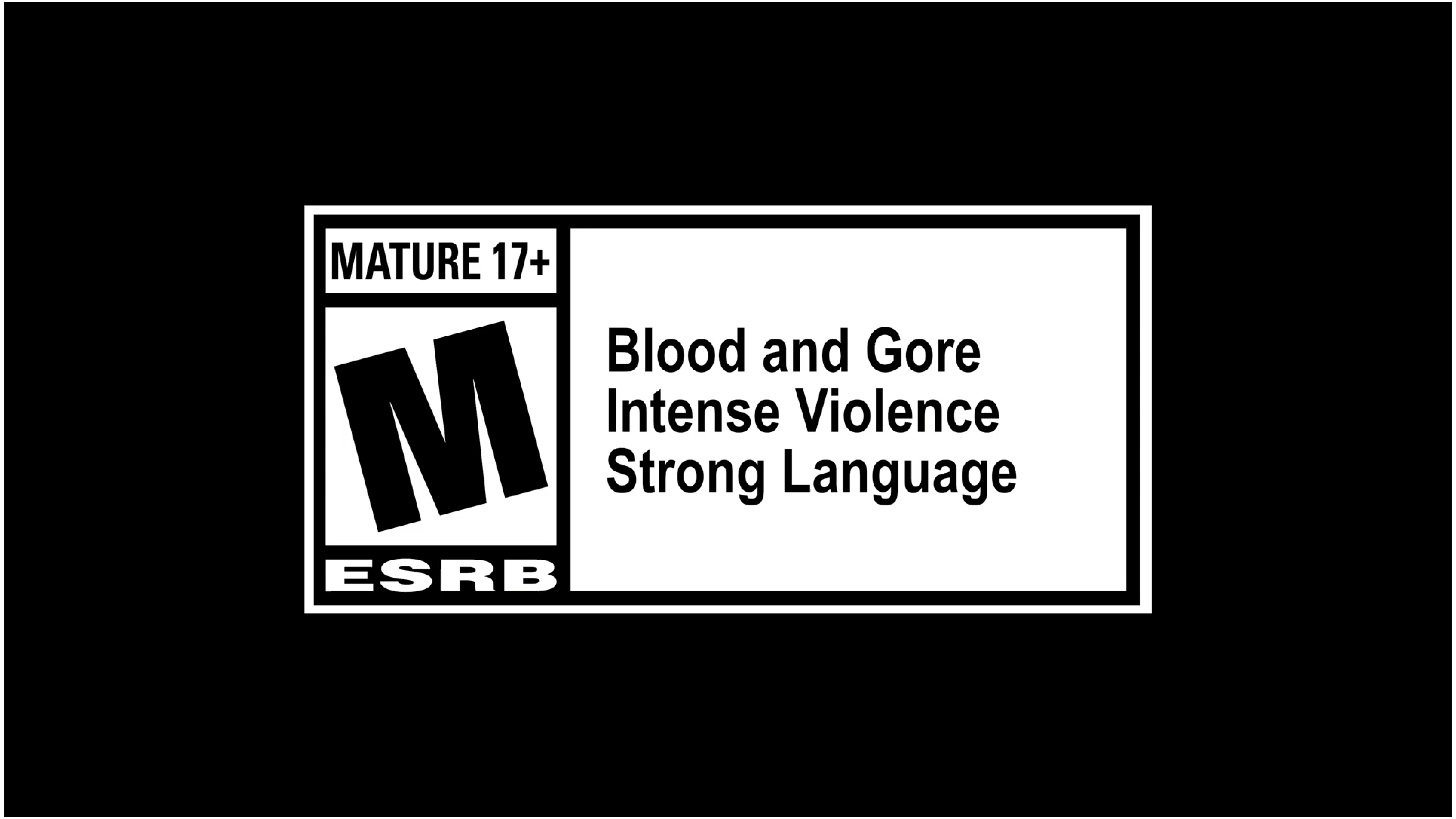
Source: Talk by Edward Liu, GTC 2020 titled
DLSS 2.0 – Image Reconstruction for Real-Time Rendering
with Deep Learning

DLSS 2.0 Output

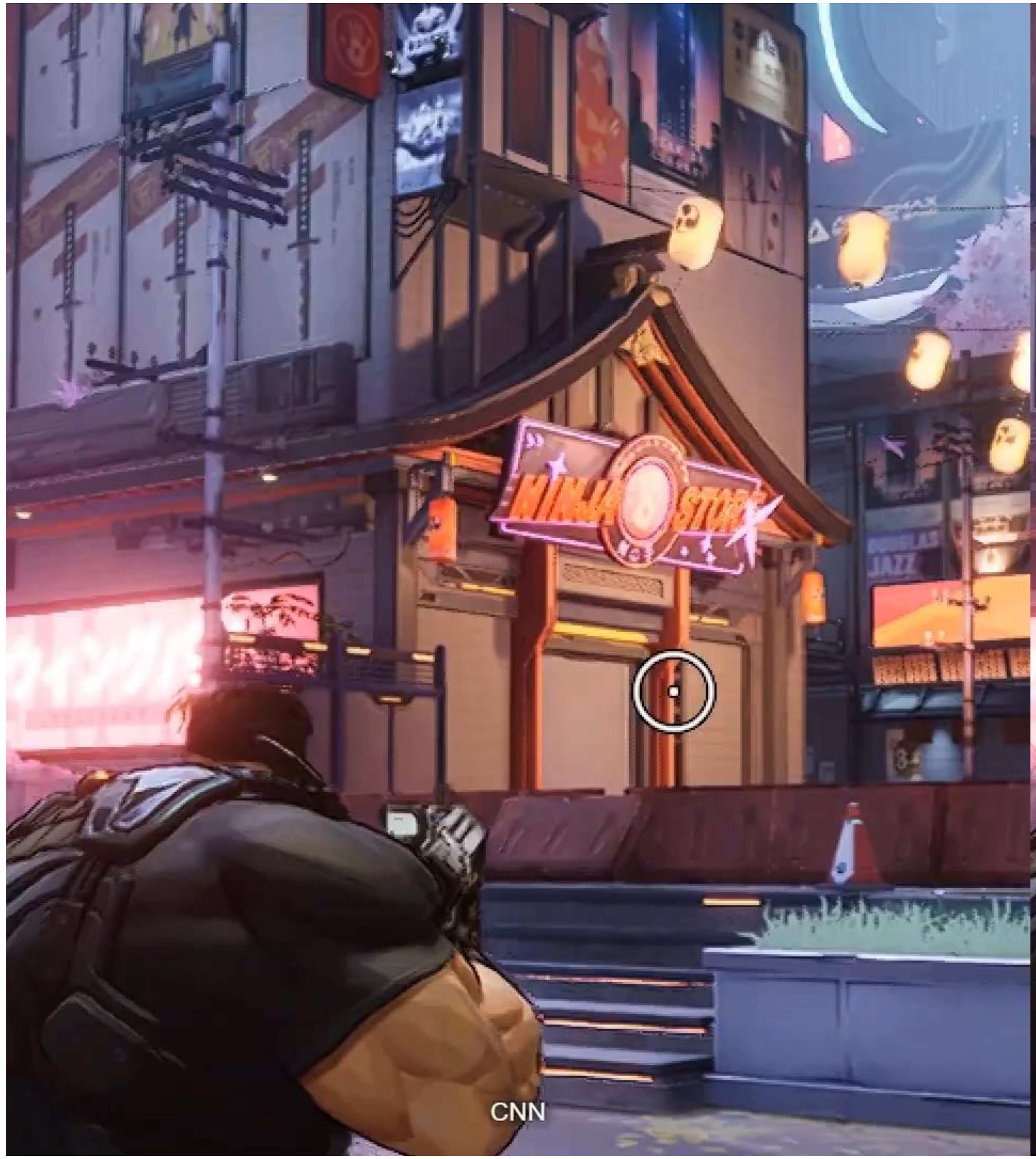
Source: Talk by Edward Liu, GTC 2020 titled
DLSS 2.0 – Image Reconstruction for Real-Time Rendering
with Deep Learning

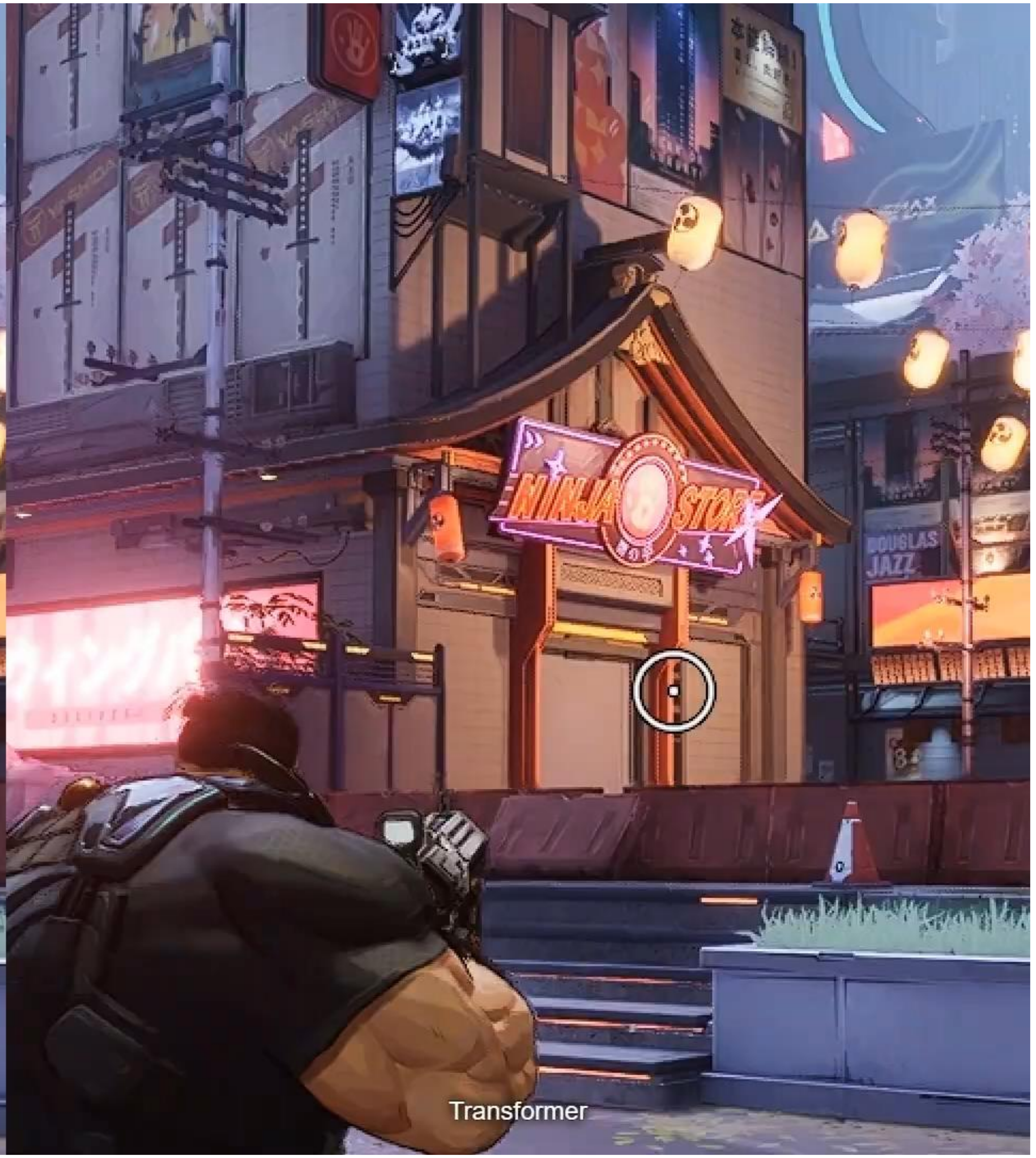
Source: Talk by Edward Liu, GTC 2020 titled
DLSS 2.0 – Image Reconstruction for Real-Time Rendering
with Deep Learning

DLSS – Real-Time AI Super Resolution for Games

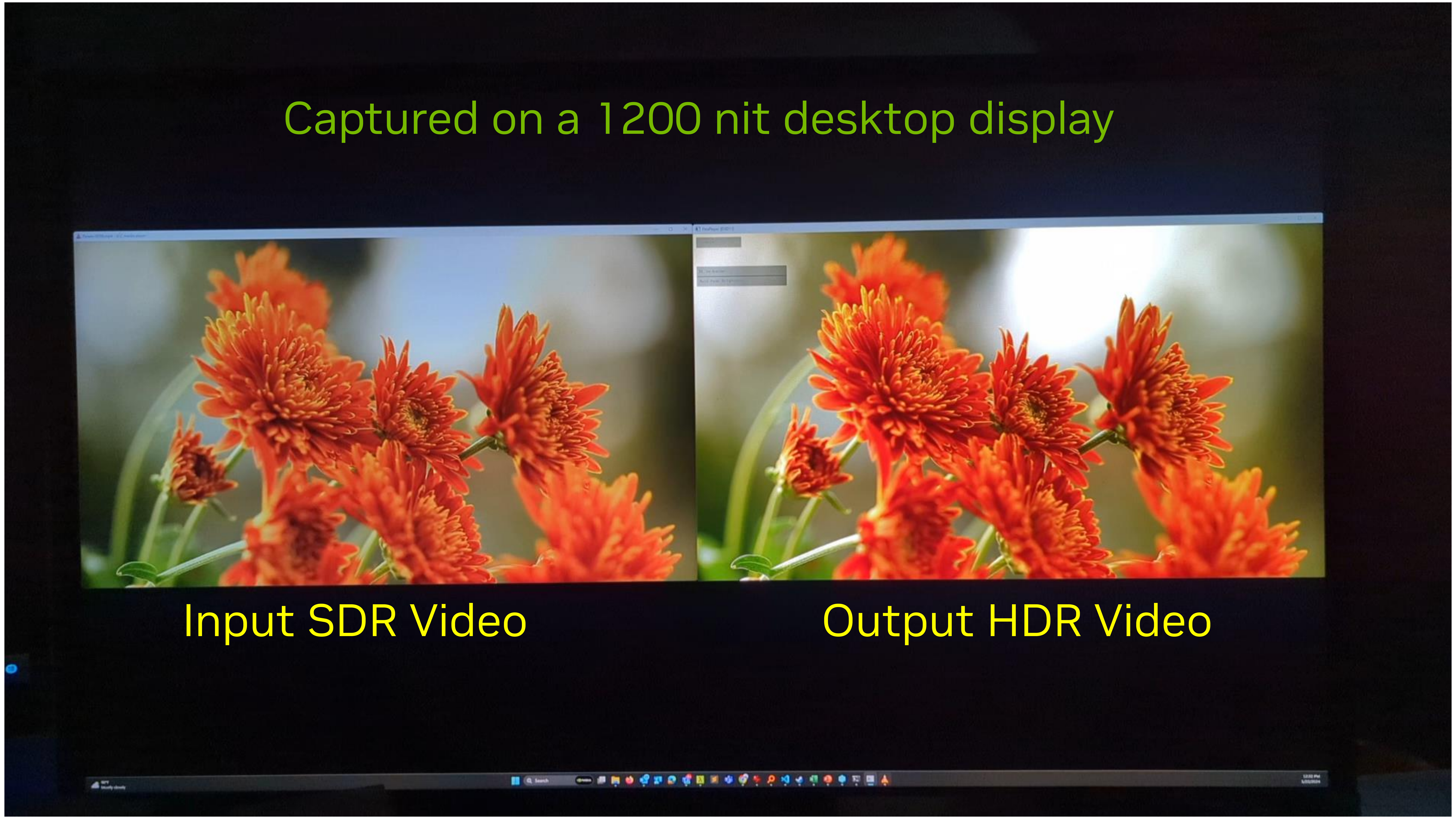


DLSS – Real-Time AI Super Resolution for Games



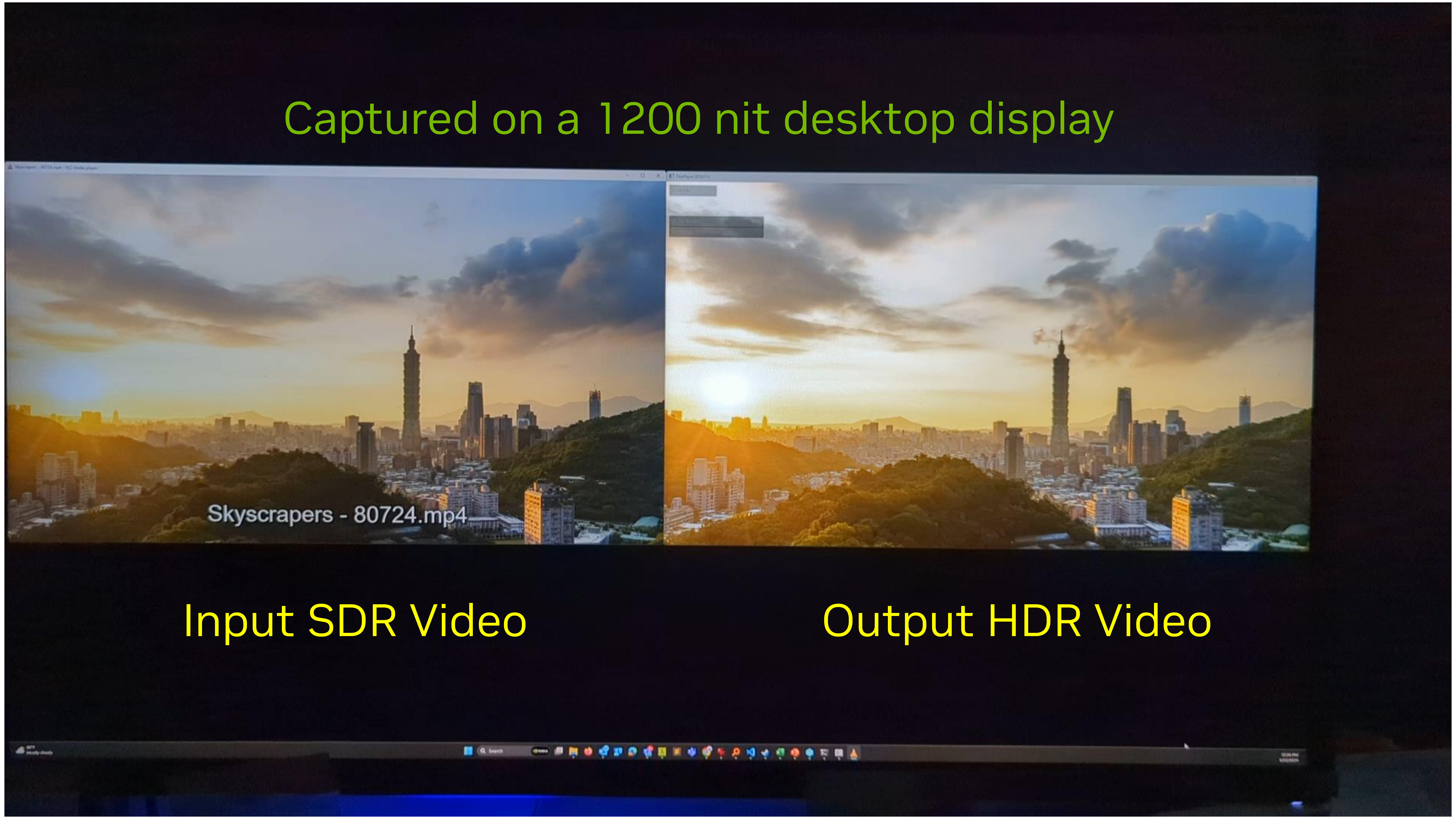


RTX HDR - Live



Video by Raman Oza from Pixabay

We use AI to live remaster video playback from SDR to HDR



Video by Timo Volz from Pixabay

RTX Dynamic Vibrance

RTX Dynamic Vibrance OFF

RTX Dynamic Vibrance ON

From Research to Reality

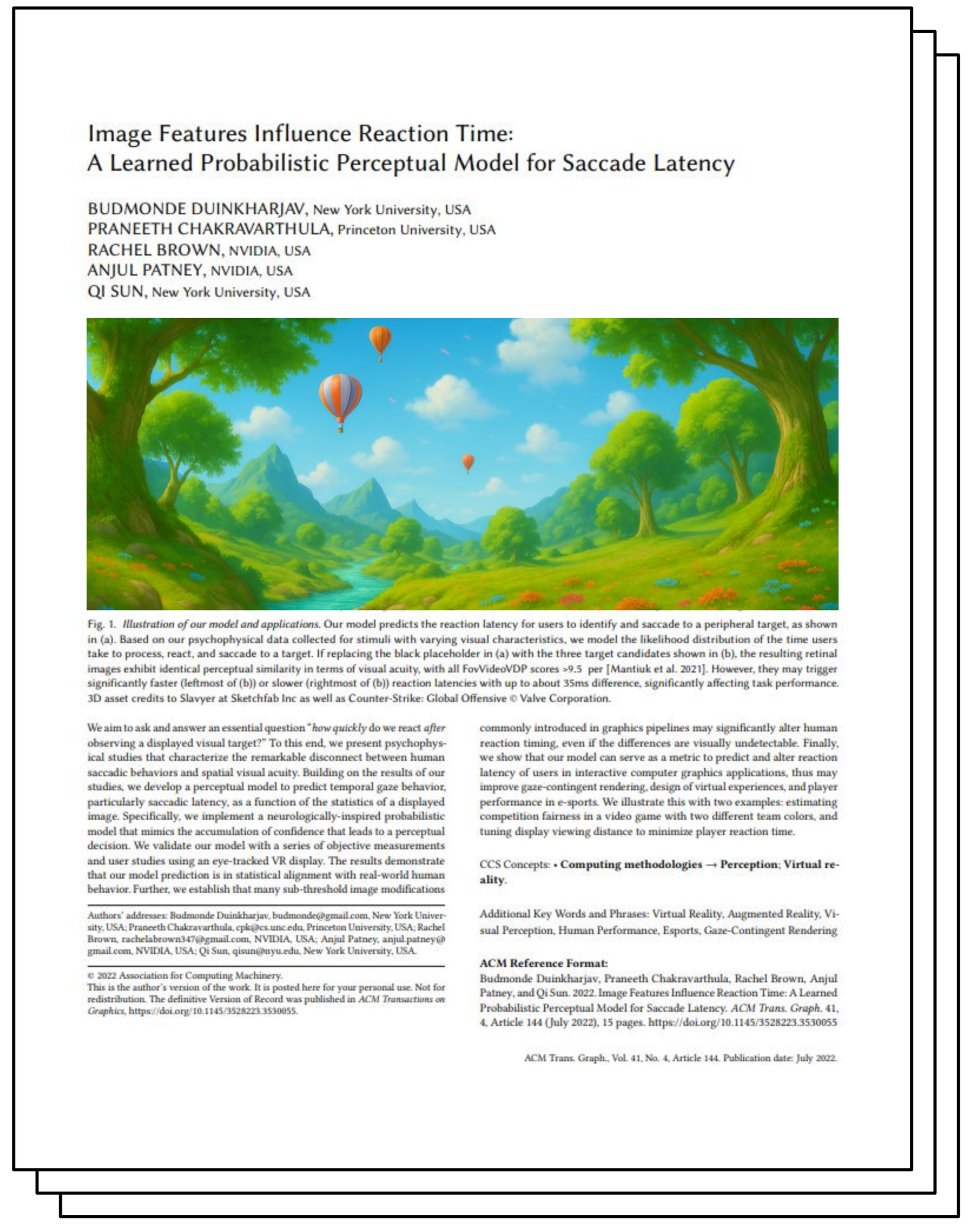


Image Features Influence Reaction Time:
A Learned Probabilistic Perceptual Model for Saccade Latency

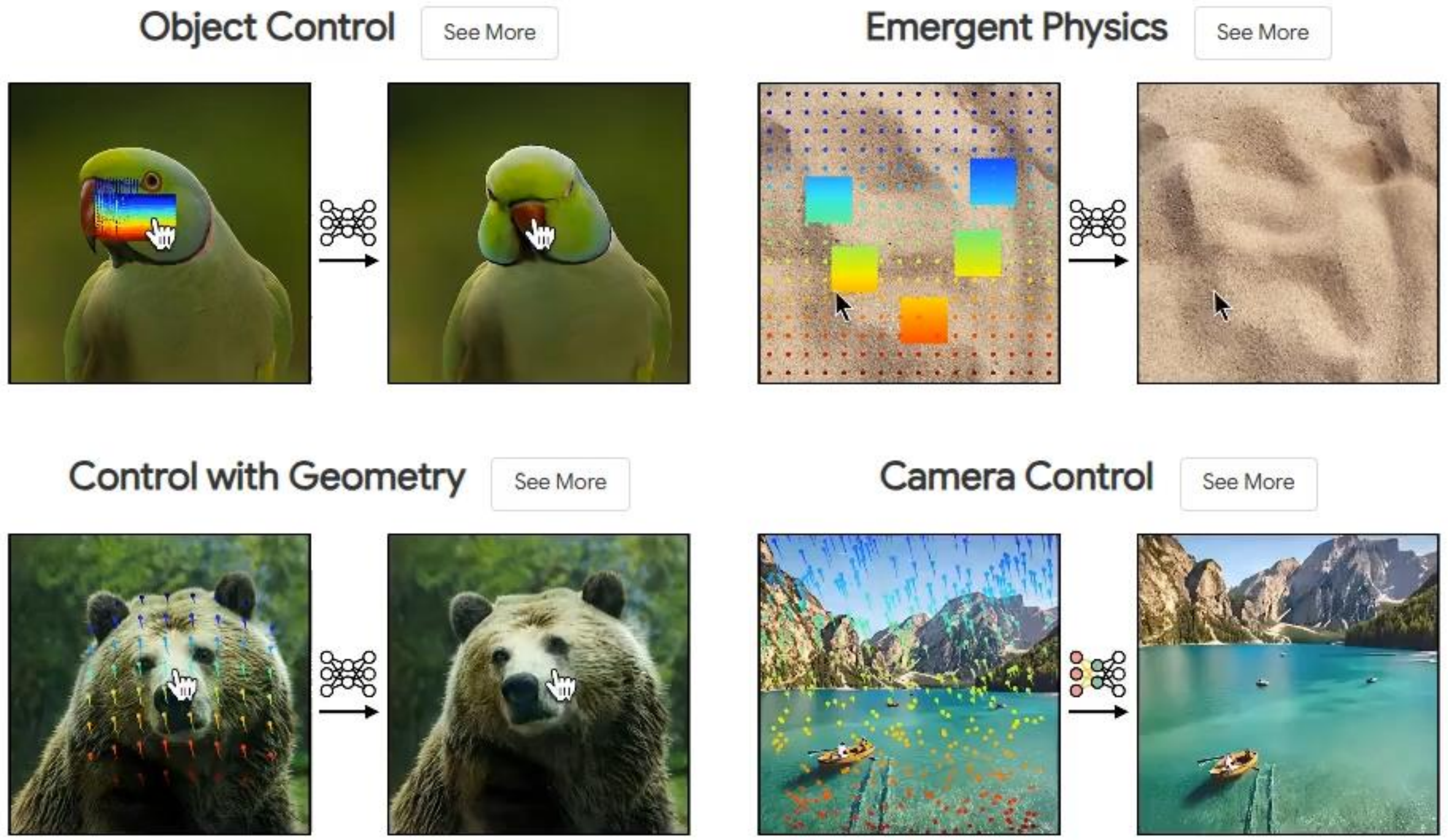
BUDMONDE DUINKHARJAV, New York University, USA
PRANEETH CHAKRAVARTHULA, Princeton University, USA
RACHEL BROWN, NVIDIA, USA
ANJUL PATNEY, NVIDIA, USA
QI SUN, New York University, USA

Running offline, or 3-seconds per frame on a 800W GPU

Must run live in ms per frame

Ongoing AI Research will Redefine all Visual Experiences

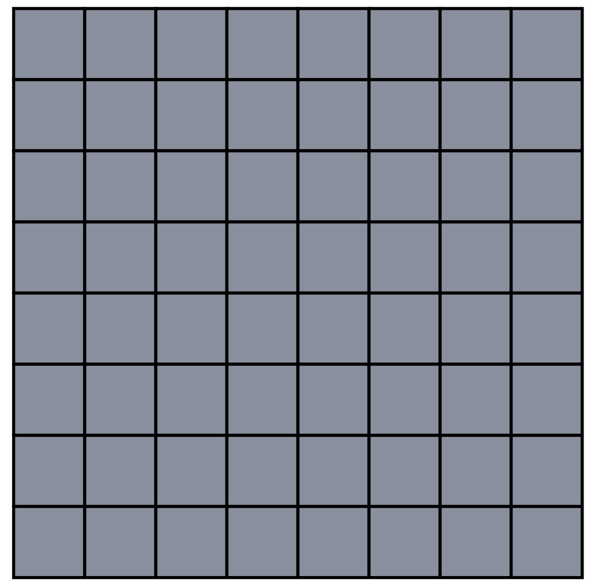
- All pixels, MR or otherwise will be generated soon
- How do we build technologies that get there?



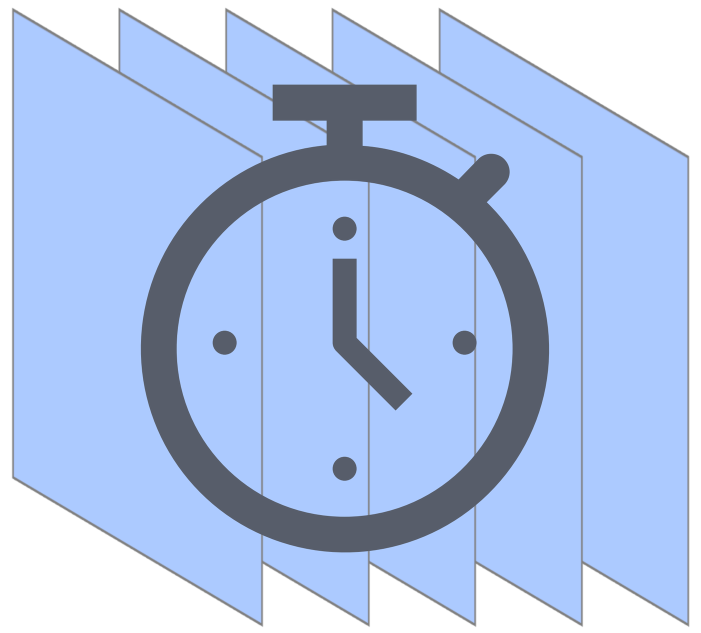
Motion Prompting: Controlling Video Generation with Motion Trajectories
Geng et al., CVPR 2025 (Oral)

NERFs

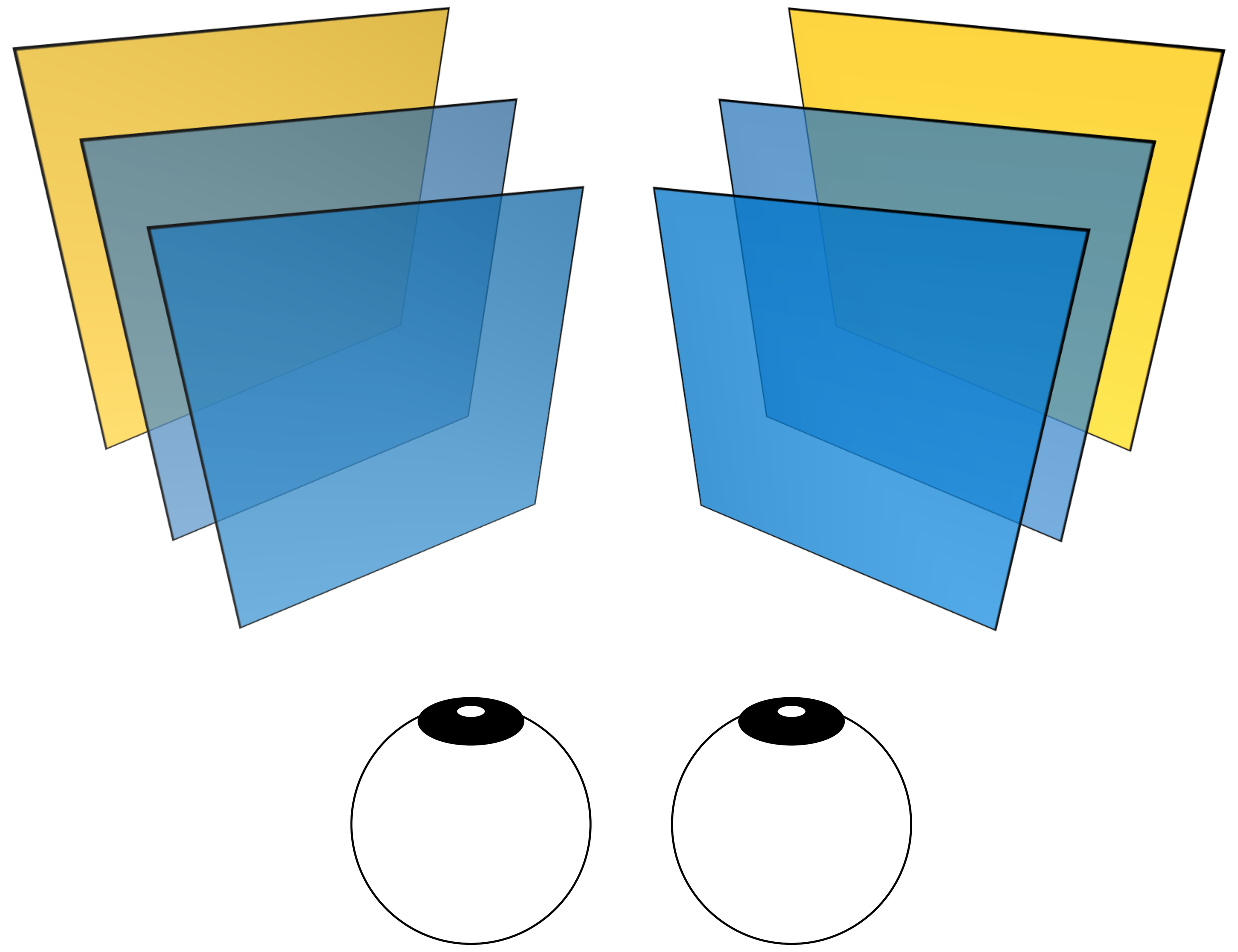
Mixed Reality Pixels are Expensive



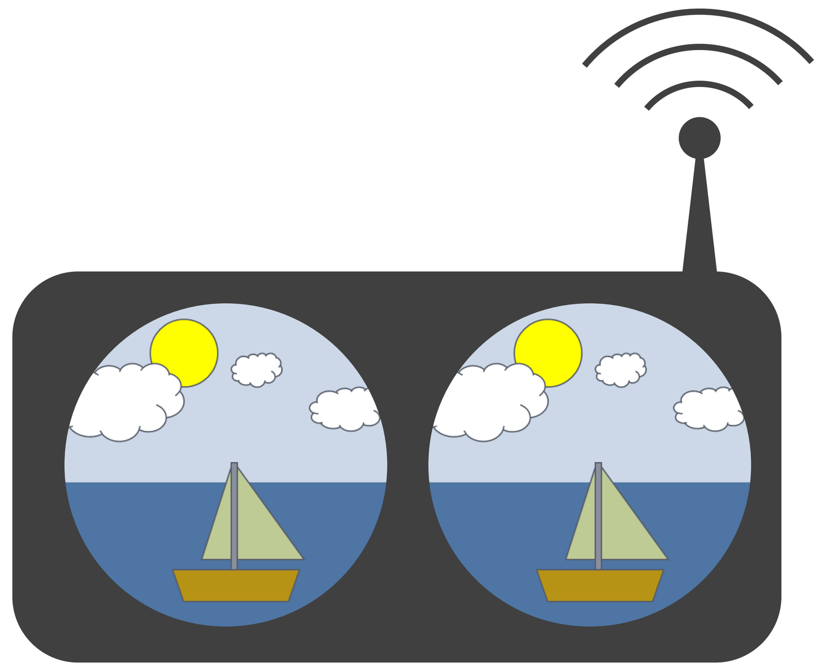
High Resolution



High Framerate



Multiple Views



Untethered Displays

What does it take to bring cutting-edge AI research to real-life MR experiences?

Look at CV4MR papers already doing this stuff

Last Mile Research

- We have repeatedly found that a significant component of research happens during production
- Experiments with real use cases reveal problems and challenges unaddressed in academic research
- Outcomes of these experiments frequently decisive
- Two major categories
  - Will it work as expected (correctness)?
  - Will it be fast enough (efficiency)?

Correctness - Will it Work?

- Many ideas work well on early/offline test-cases, but
- Examples-
  - UI/HUD pixels
  - Sniper scopes
  - Subtle artifacts not apparent without interactivity
- Need more examples here

Efficiency - Will it Work Fast Enough?

- Important to define latency requirements for MR applications
- Research questions
  - What are the true runtime overheads
  - How should we re-architect the algorithm (or distil, or optimize)

Add an example use case + image/video

The Power of Interactive Proof-of-Concept Experiences

- You need to see your pixels
- You need to know that you can answer most last-mile research questions
- Drives mindshare and production decisions
- “A demo is worth a thousand videos”

Add an example use case + image/video

Some Open Challenges in CV4MR

- Near speed-of-light implementation of AI models
- Temporal stability and consistency of generative AI
- Reliable, low-latency remote/split rendering
- Perceptual models of image and video quality

Summary and Call to Action

- It is an exciting time for CV4MR – generative AI meets mixed-reality
- Last mile research is essential to realize these advances
- Build more prototypes – see your pixels
  - Solve some last-mile challenges
- If my mission resonates with you, reach out! apatney@nvidia.com

Thank you

apatney@nvidia.com

anjulpatney.com

Extra Content