NVIDIA

DLSS 2.0 - IMAGE RECONSTRUCTION FOR REAL-TIME RENDERING WITH DEEP LEARNING

Shiqiu (Edward) Liu, Sr. Research Scientist, August 6th 2020

ABOUT ME

7 years of real-time rendering research and development at NVIDIA

Key contributors to several next-gen rendering technologies

- DLSS, real-time ray tracing and denoising, VR Rendering

Developed technologies/algorithms that shipped in many titles and game engines

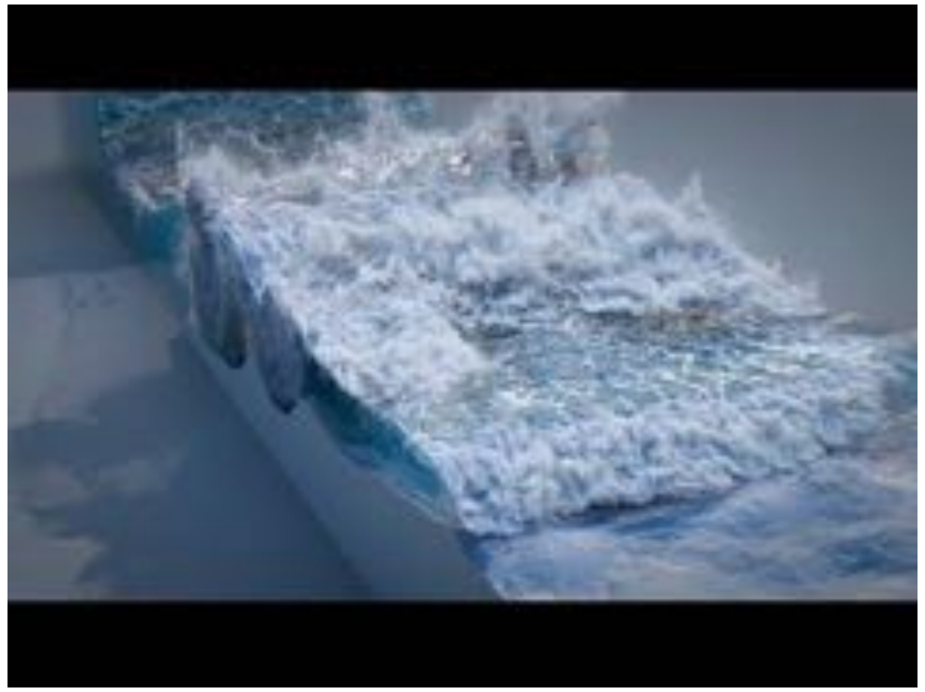
NEXT GEN GAMES NEED SUPER RESOLUTION

Ray tracing, physics, AR/VR, and higher resolution displays drive up GPU computing needs exponentially

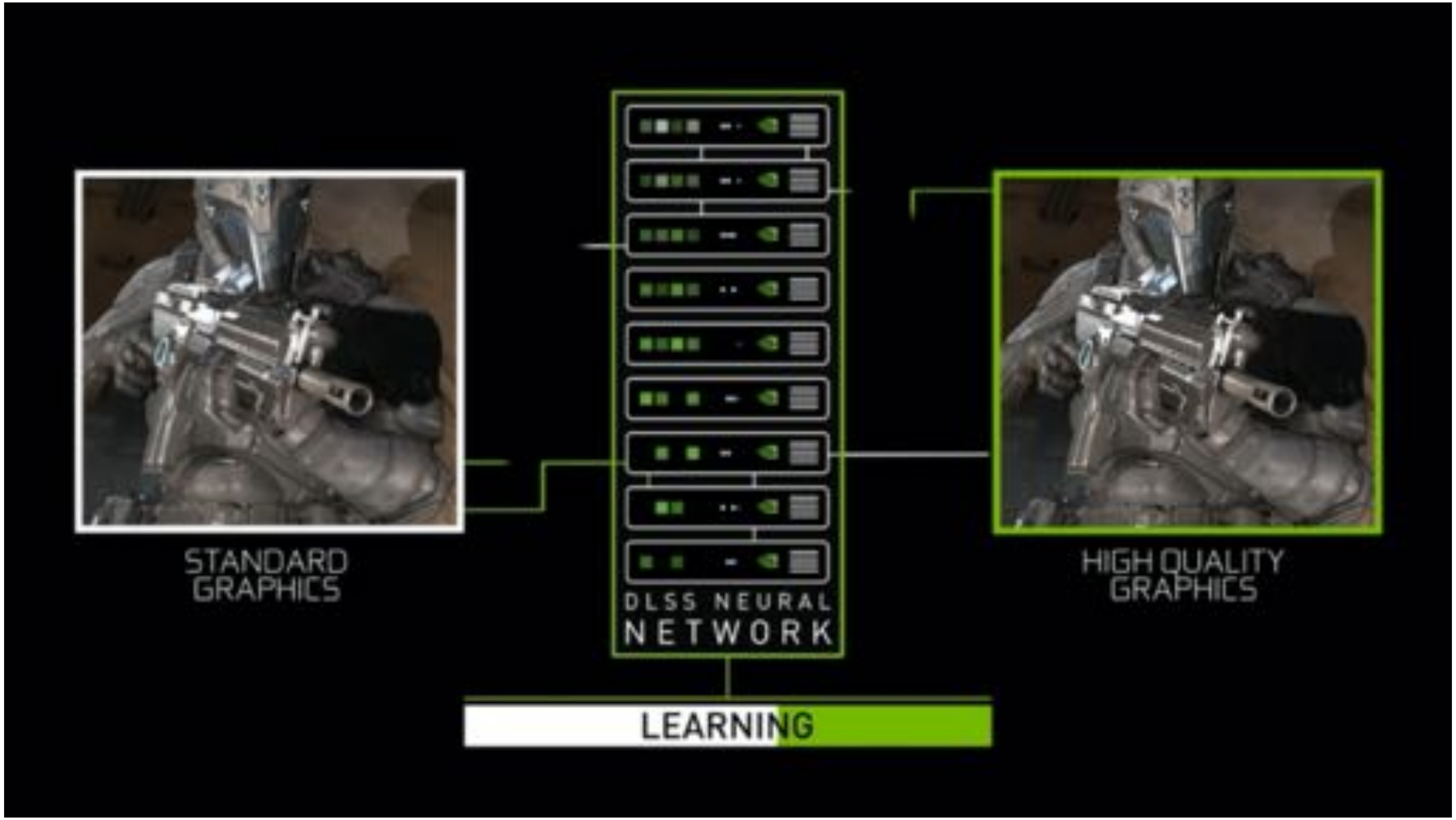
Ray tracing alone can demand many times the computing power of traditional rendering techniques

Super resolution technique become necessary

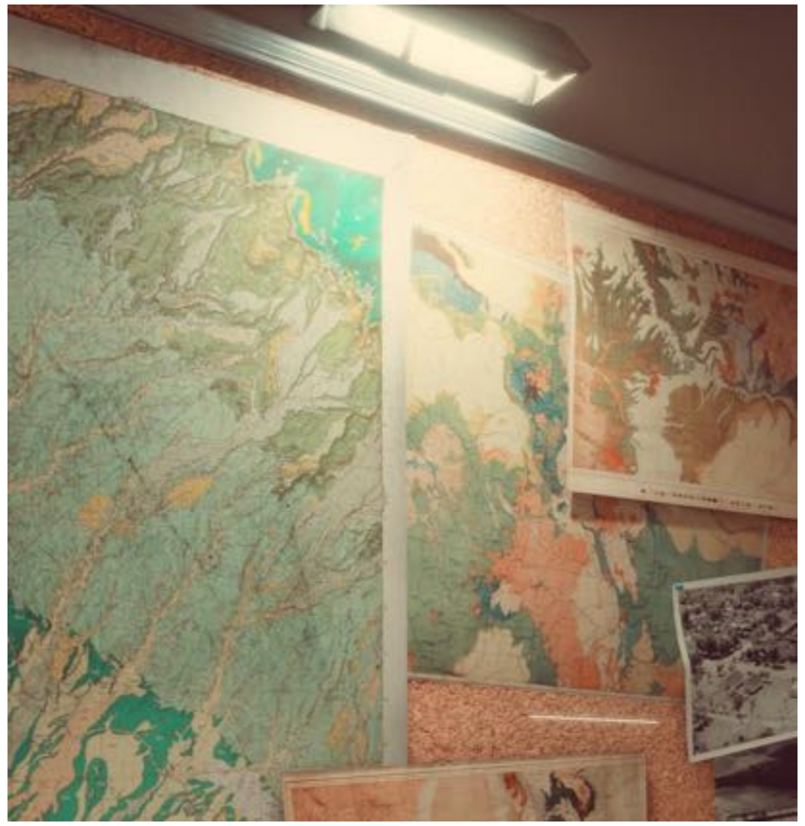
RTX GPUs have tensor cores to accelerate deep learning workloads



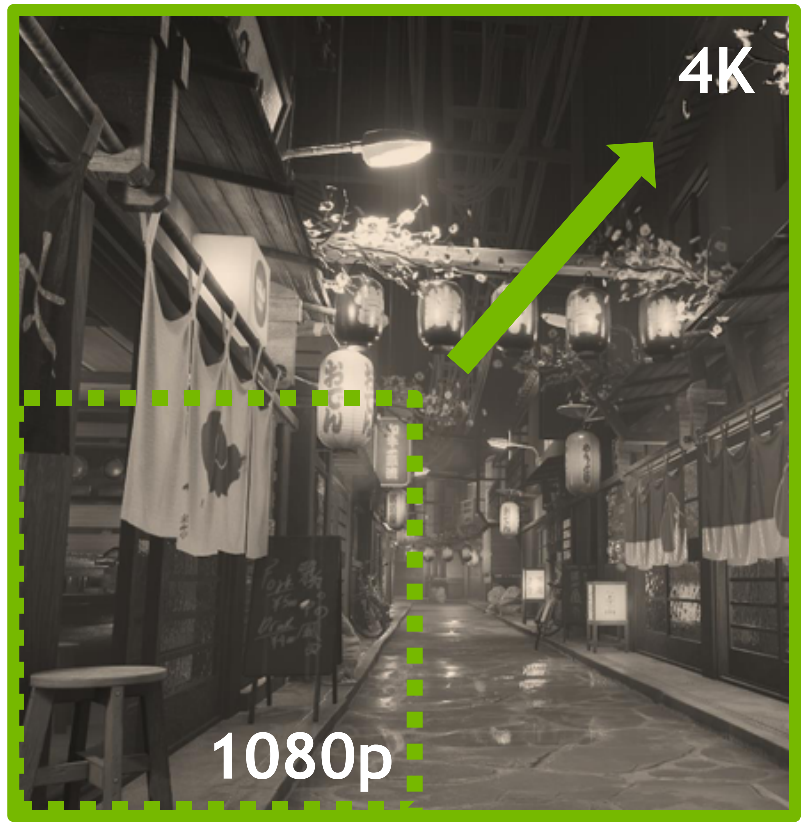
DLSS 1.0


STANDARD GRAPHICS
DLSS NEURAL NETWORK
HIGH QUALITY GRAPHICS
LEARNING

INTRODUCING DLSS 2.0



Great Image Quality
Details rival native resolution



4x Upscaling Ratio
540p to 1080p, 1080p to 4k

Generalized Model
One model to rule them all!

1.5ms at 4K on 2080Ti
Works on all RTX GPUs at all resolution


720p

DLSS 1.0 720p to 1080p


720p

DLSS 1.0 720p to 1080p

DLSS 2.0 720p to 1080p

DLSS 1.0 720p to 1080p

DLSS 2.0 720p to 1080p


1080p


DLSS 2.0 720p to 1080p


1080p

DLSS 2.0 720p to 1080p


540p

DLSS 2.0 540p to 1080p


540p

DLSS 2.0 540p to 1080p


540p
DLSS2.0
720p
DLSS1.0
1080p TAA
540p TAA


540p
DLSS2.0
720p
DLSS1.0
1080p TAA
540p TAA


540p - 89fps


1080p - 48fps


540p to 1080p w/ DLSS2.0 - 86fps

1080p - 48fps

540p to 1080p w/ DLSS2.0 - 86fps

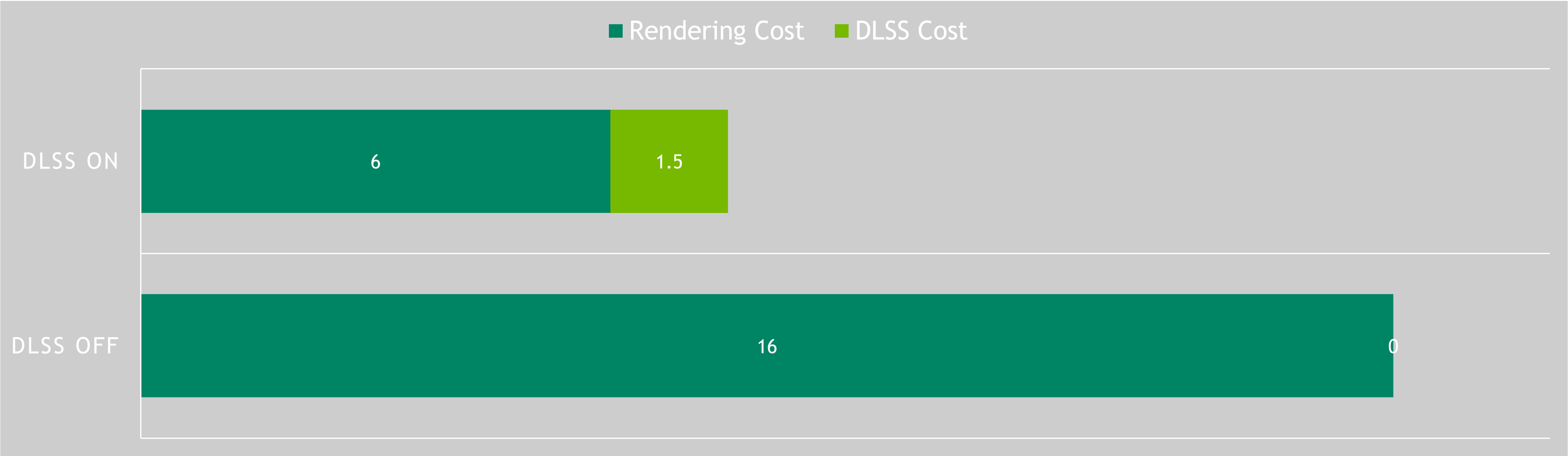

32spp Reference 1080p


540p to 1080p w/ DLSS2.0 - 86fps


32spp Reference 1080p


540p DLSS2.0
1080p TAA
32spp Reference
540p TAA

DLSS 2.0 - ACCELERATED RENDERING



DLSS 2.0 PERFORMANCE BOOSTS

Performance Mode – 1080p to 4K

“DLSS is impressive to the point where I believe you'd be nuts not to use it.”

— Digital Foundry

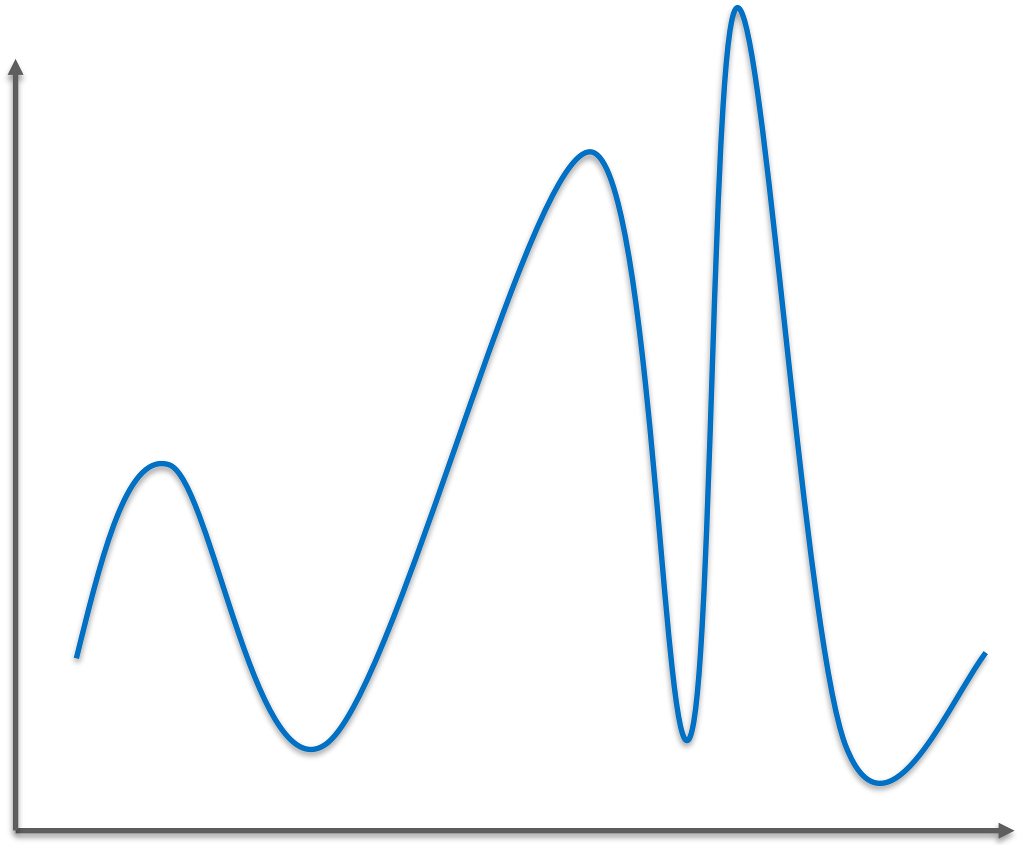
“The upscaling power of this new AI driven algorithm is extremely impressive… it’s basically a free performance button.”

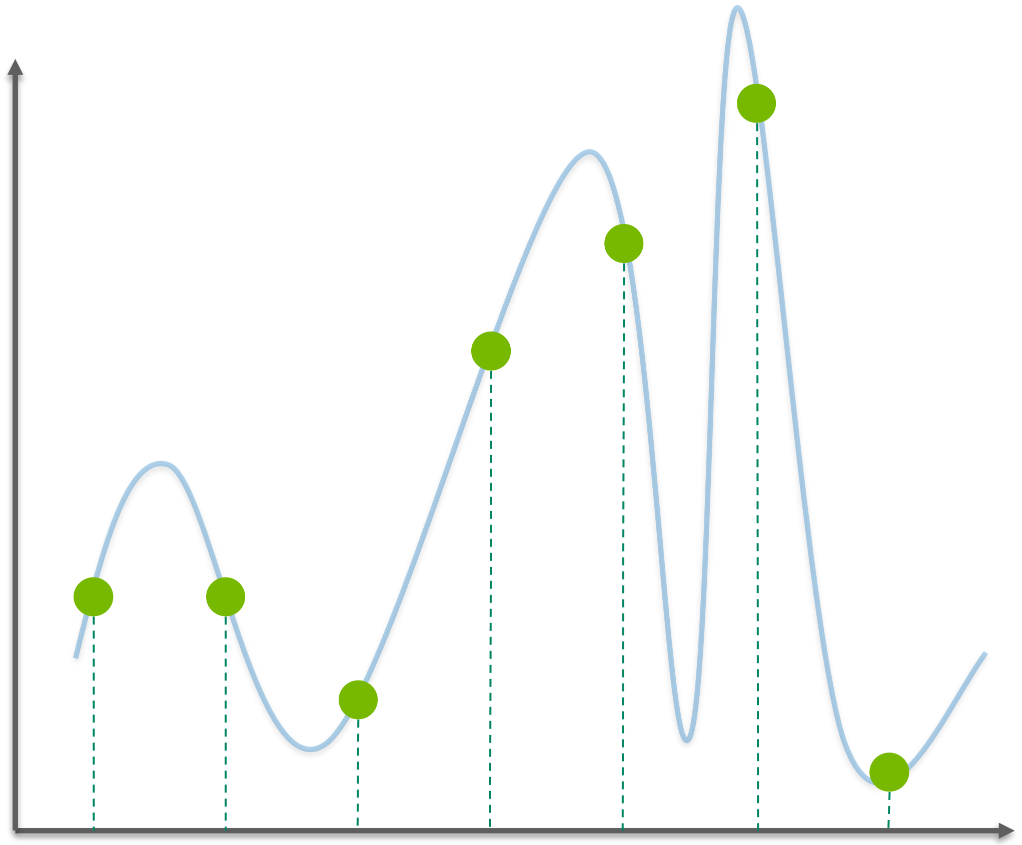
— Hardware Unboxed

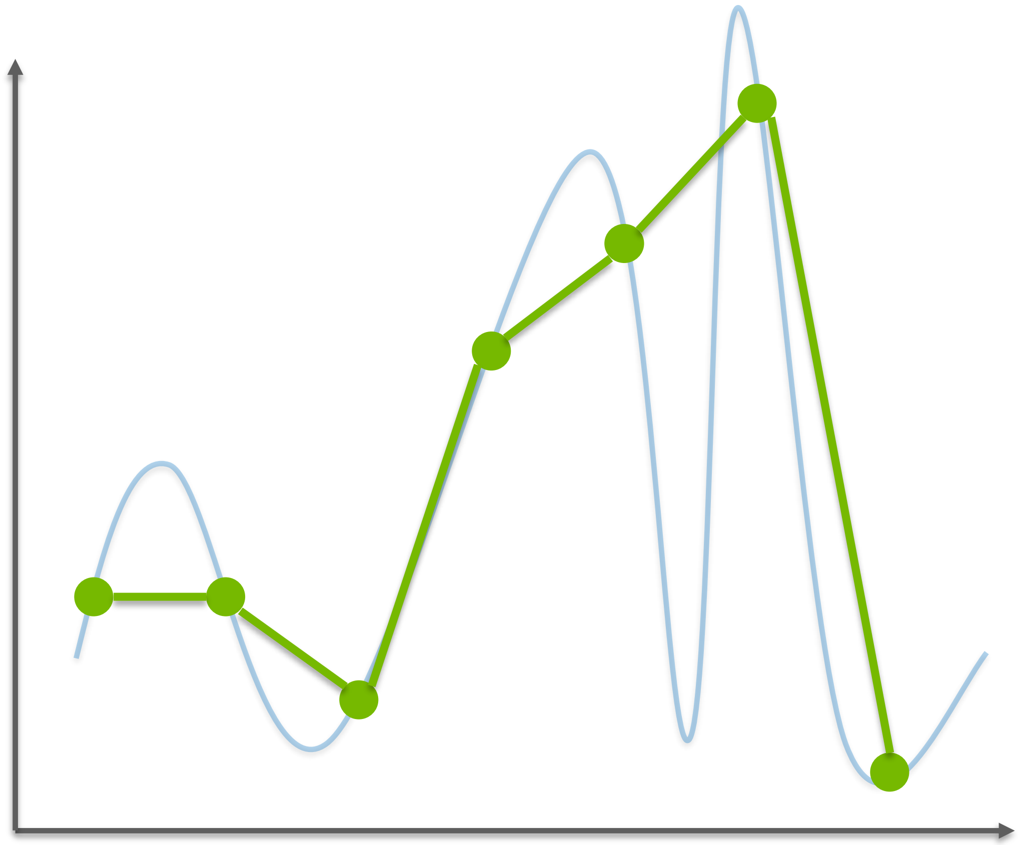
SHIPPING IN THE FOLLOWING TITLES

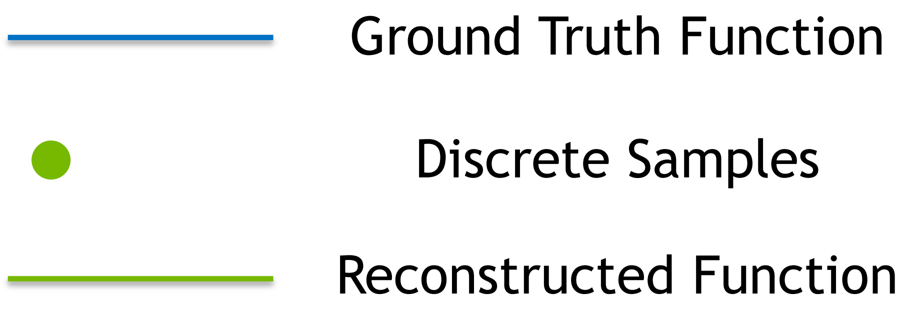
CHALLENGES IN IMAGE SUPER-RES FOR REAL-TIME RENDERING

RECONSTRUCTION 101



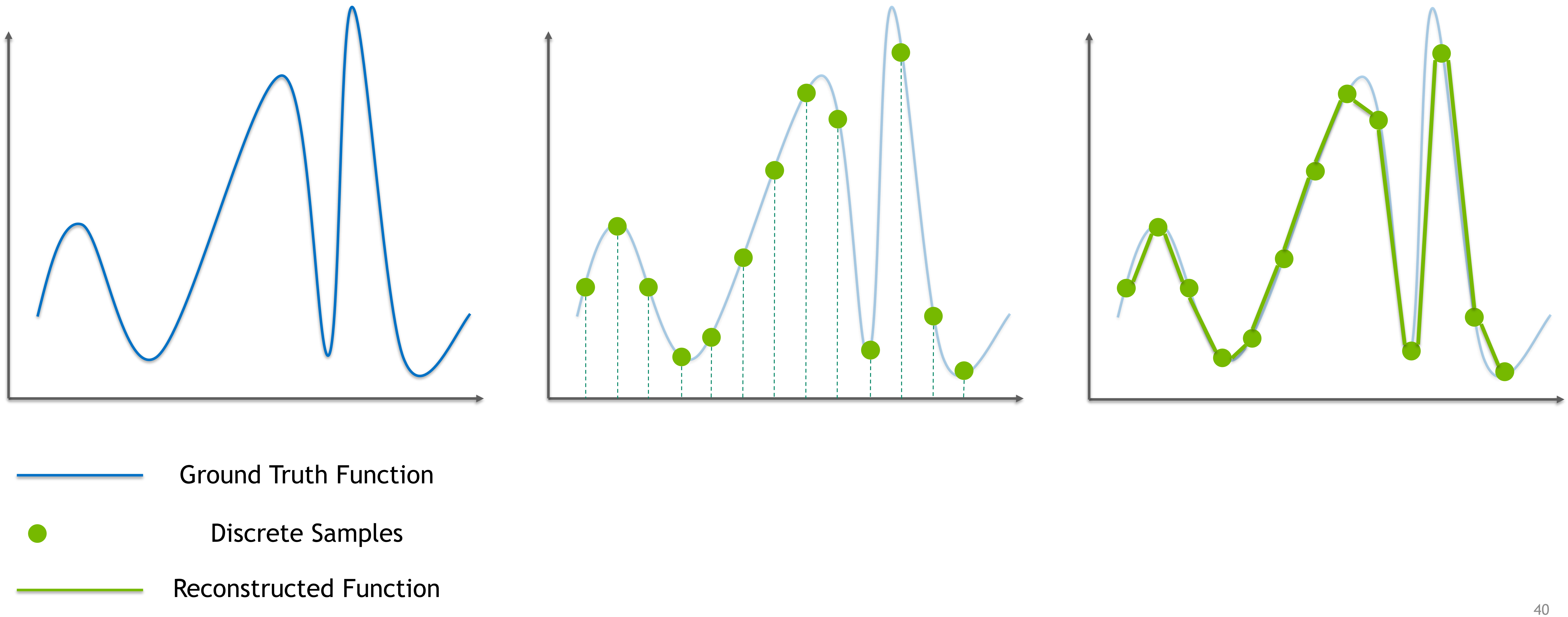




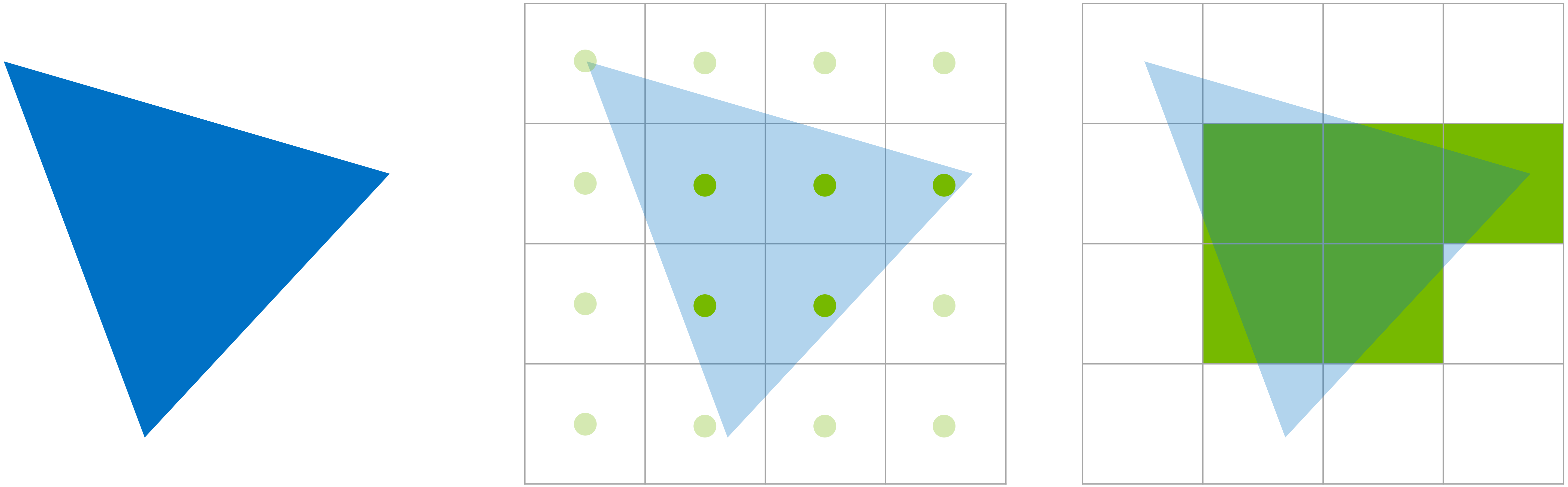


RECONSTRUCTION 101

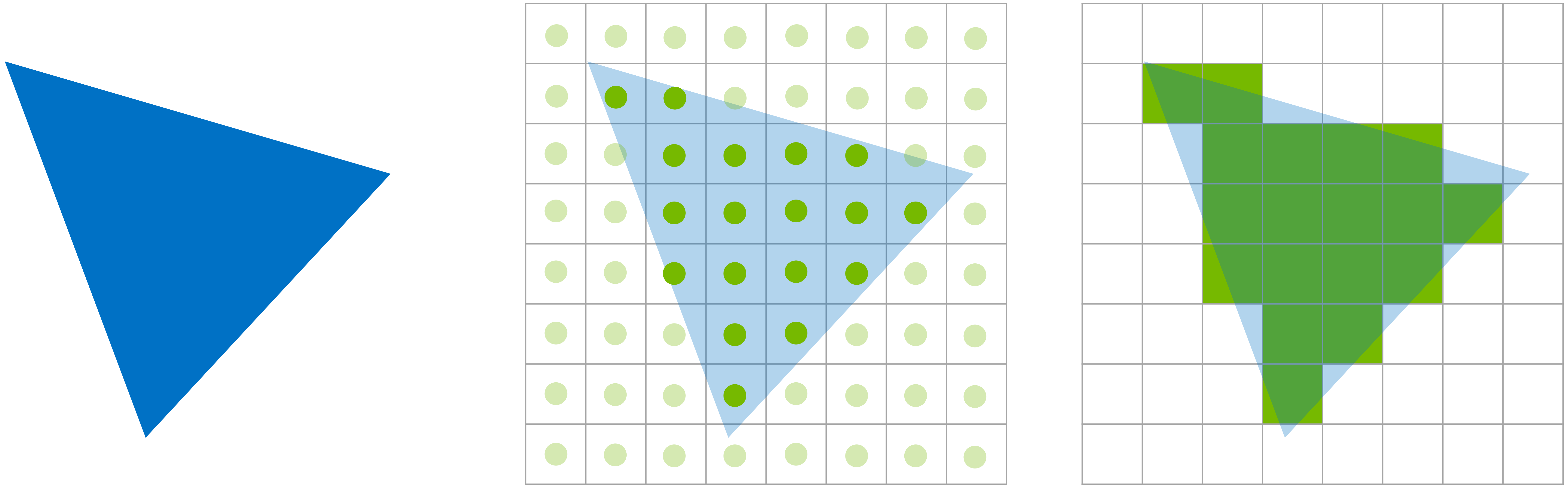
Double the sampling rate



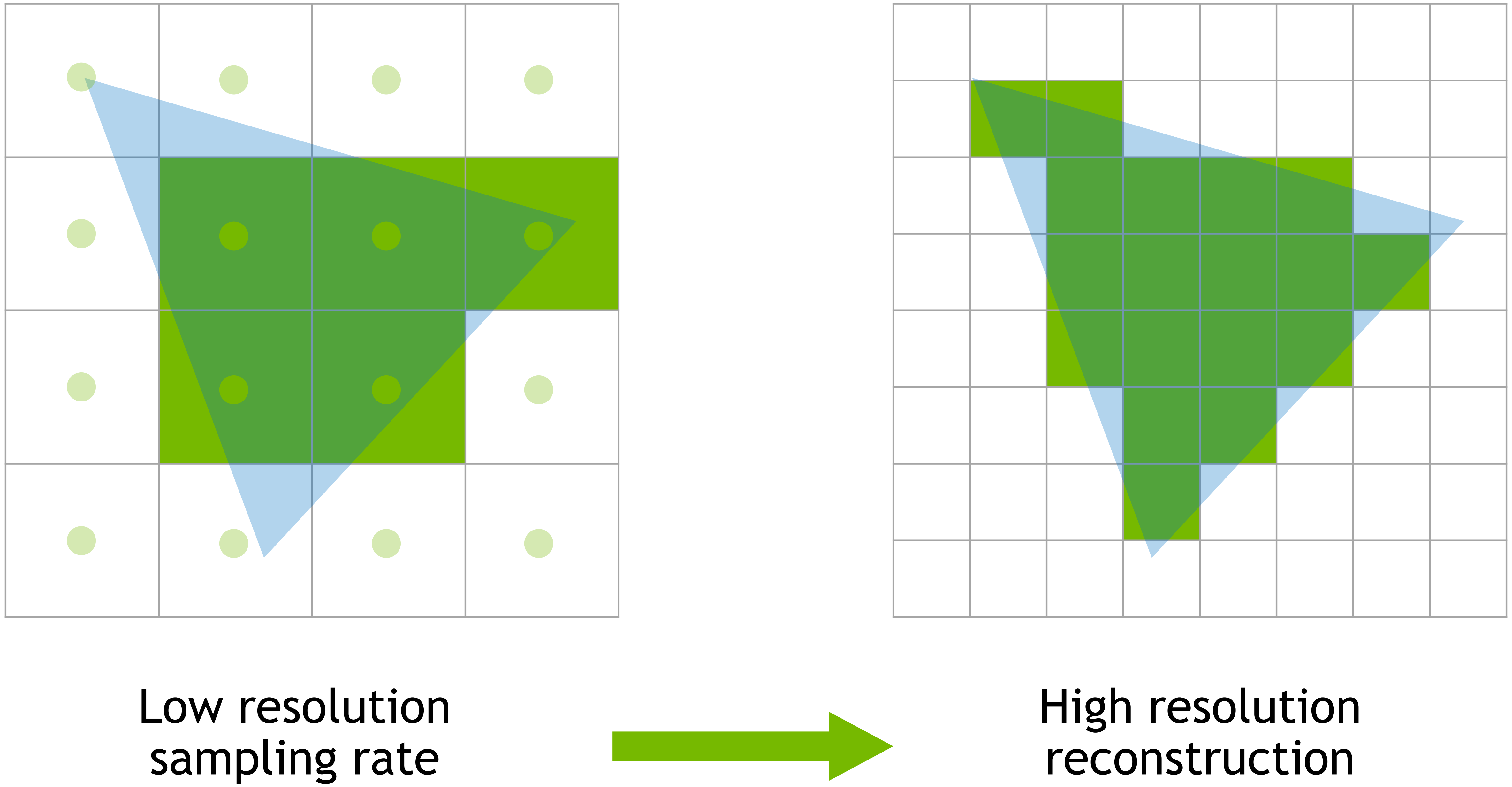
RECONSTRUCTION 101



RECONSTRUCTION 101



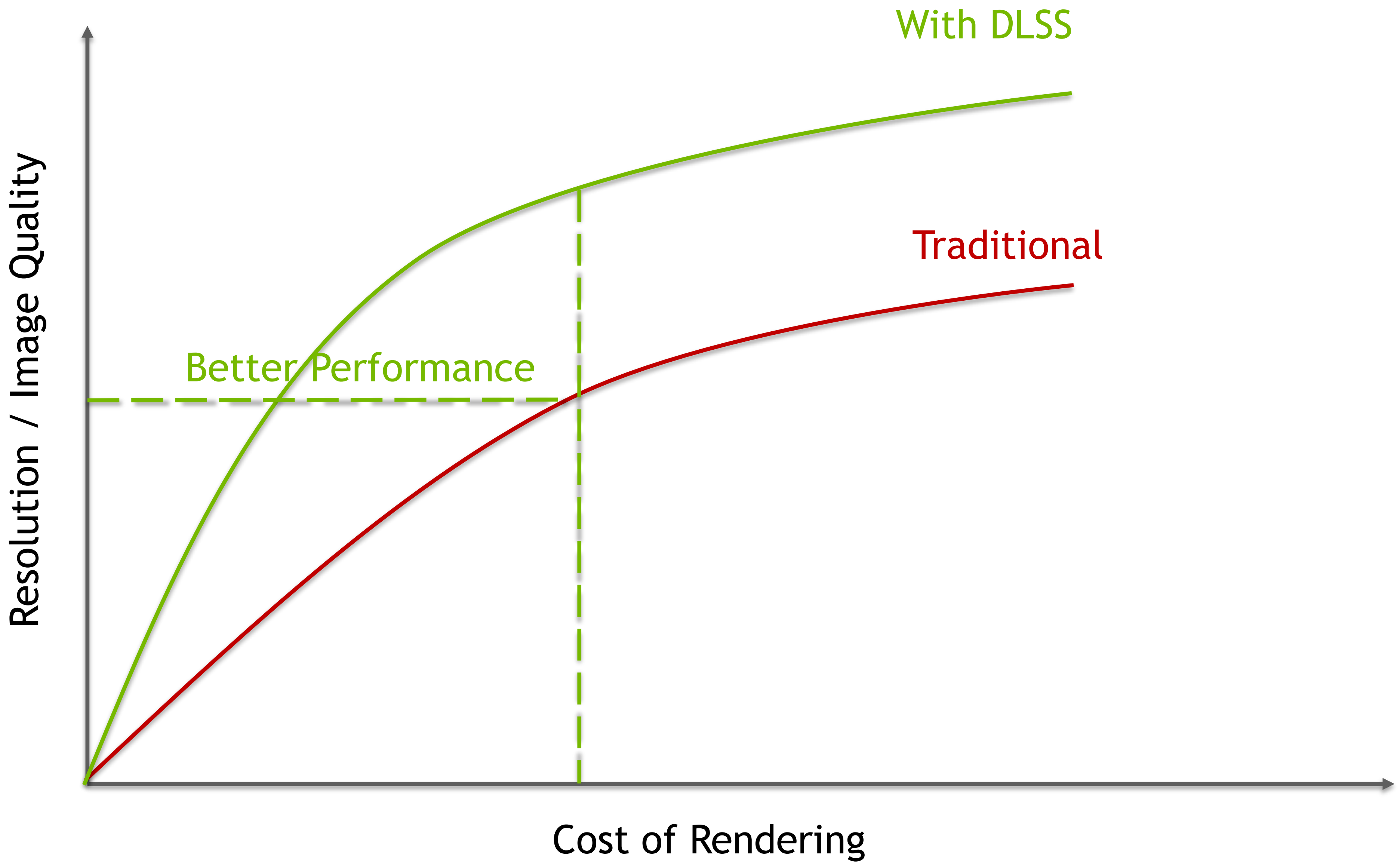
DLSS PROBLEM STATEMENT



Low resolution sampling rate

High resolution reconstruction

DLSS PROBLEM STATEMENT



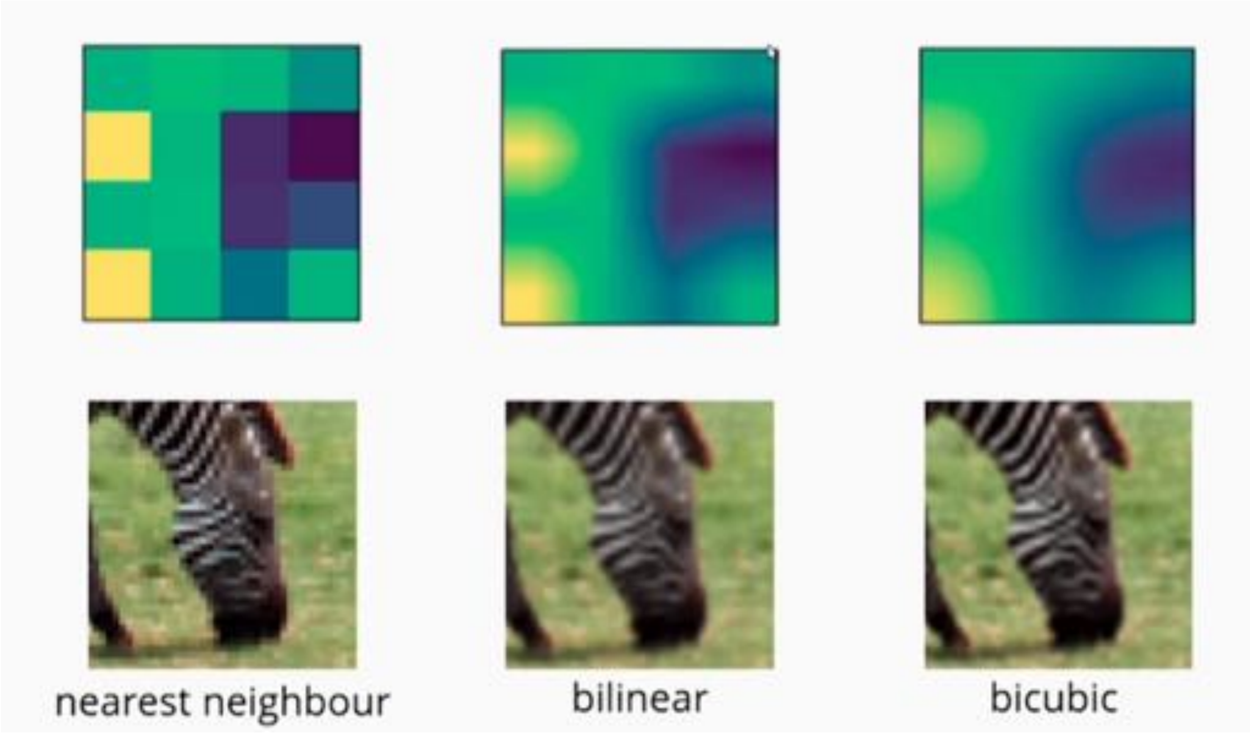
SINGLE IMAGE SUPER-RES

Previous work

Reconstruct high resolution image by interpolating the low-resolution pixels

- Common choices are bilinear, bicubic, lanczos
- Contrast aware sharpening

deep neural networks can hallucinate new pixels conditioned on existing pixels based on priors or training data

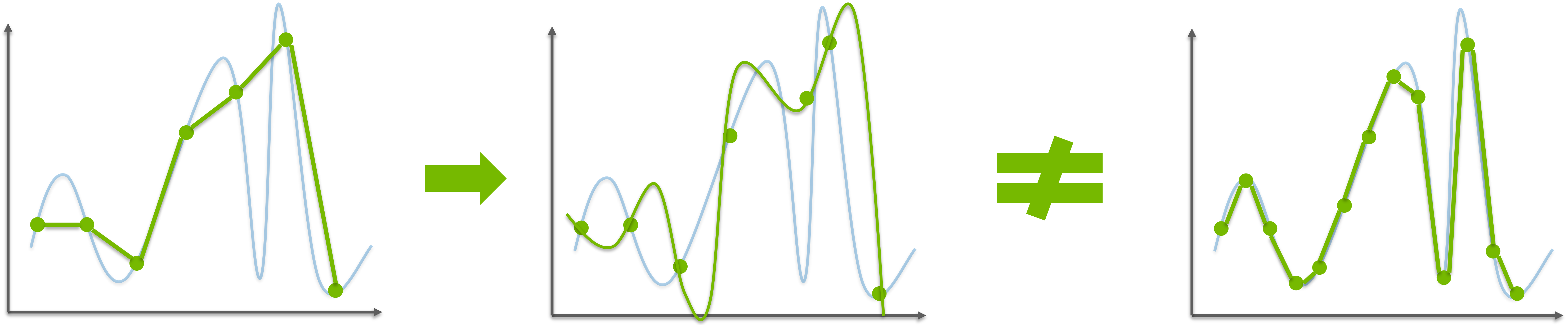


[Ledig et al. 2017]

SINGLE IMAGE SUPER-RES

Resulted images lack details compared to native high-resolution images

Images may be inconsistent with native rendering because of hallucination, and temporally unstable



Reconstructed with linear interpolation | Reconstructed with DL or other interpolation | High res samples

SINGLE IMAGE SUPER-RES

Resulted images lack details compared to native high-resolution images

Images may be inconsistent with native rendering because of hallucination, and temporally unstable

DL Upscaled
720p to 1080p

≠

Native
Rendering
1080p


1080p with TAA

540p to 1080p DLSS2.0

540p Bicubic Upsampled to 1080p


540p to 1080p with ESRGAN


540p DLSS2.0
1080p TAA
540p ESRGAN
540p TAA Bicubic

MULTI-FRAME SUPER-RES

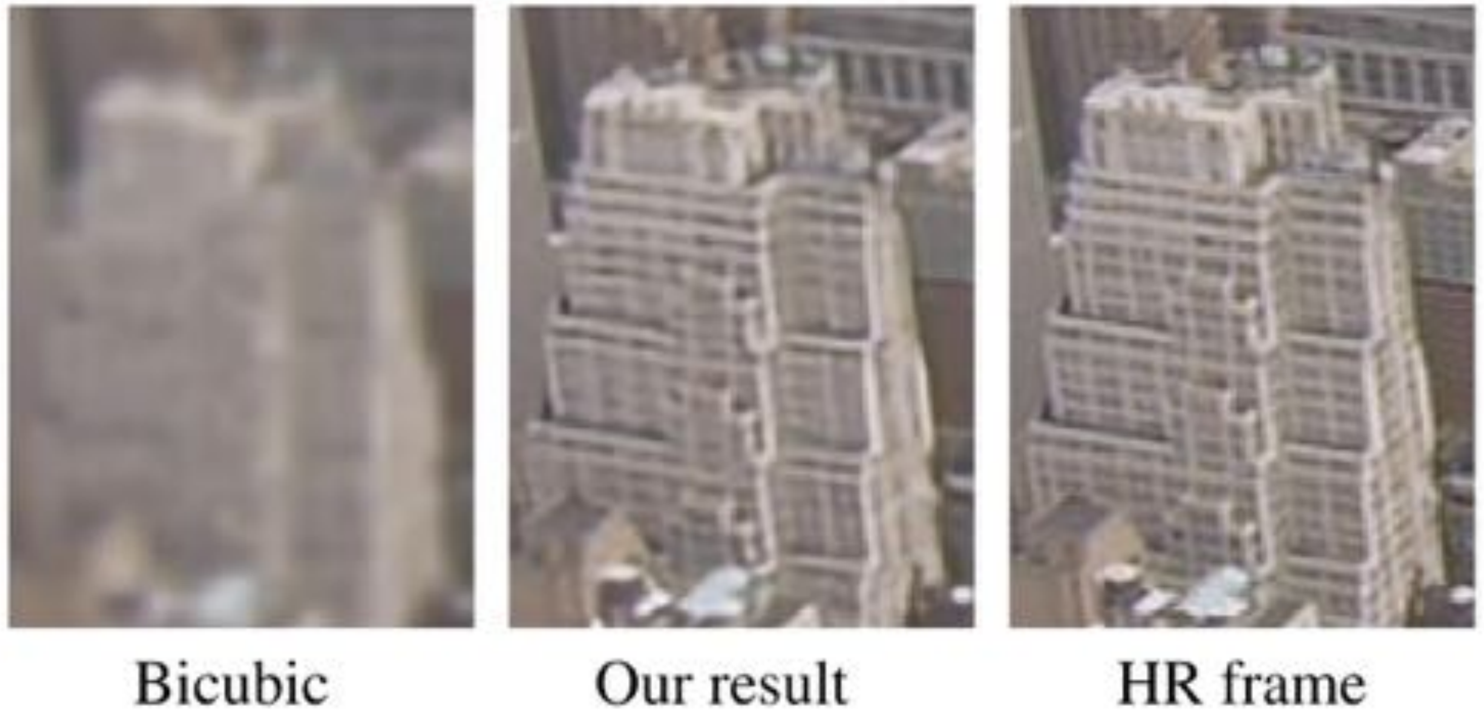
Previous work

Less ill-posed than single image super-res, restore true optical details better

Designed for videos or burst mode photography, not leveraging rendering specific information

- Optical flow vs. geometric motion vector
- Pixels vs samples
- Using frames in the future

[Wronski et al. 2019]

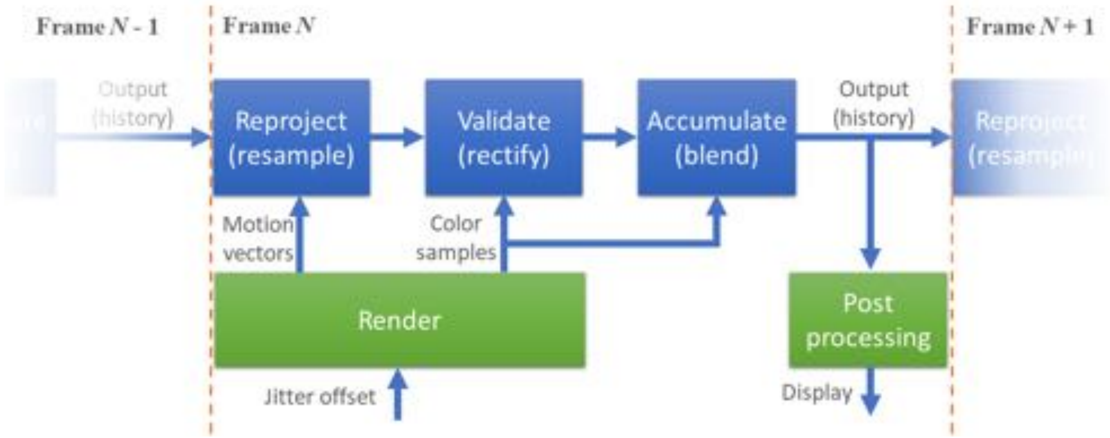


[Sajjadi et al. 2018]

SPATIAL-TEMPORAL SUPER SAMPLING

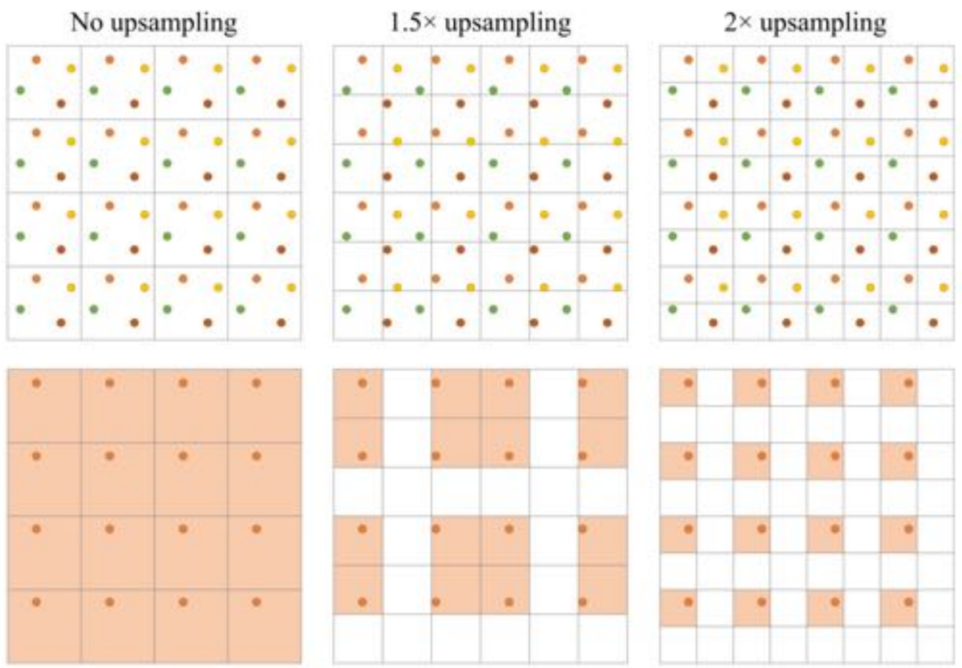
Previous work

Temporal Antilasing (TAA)



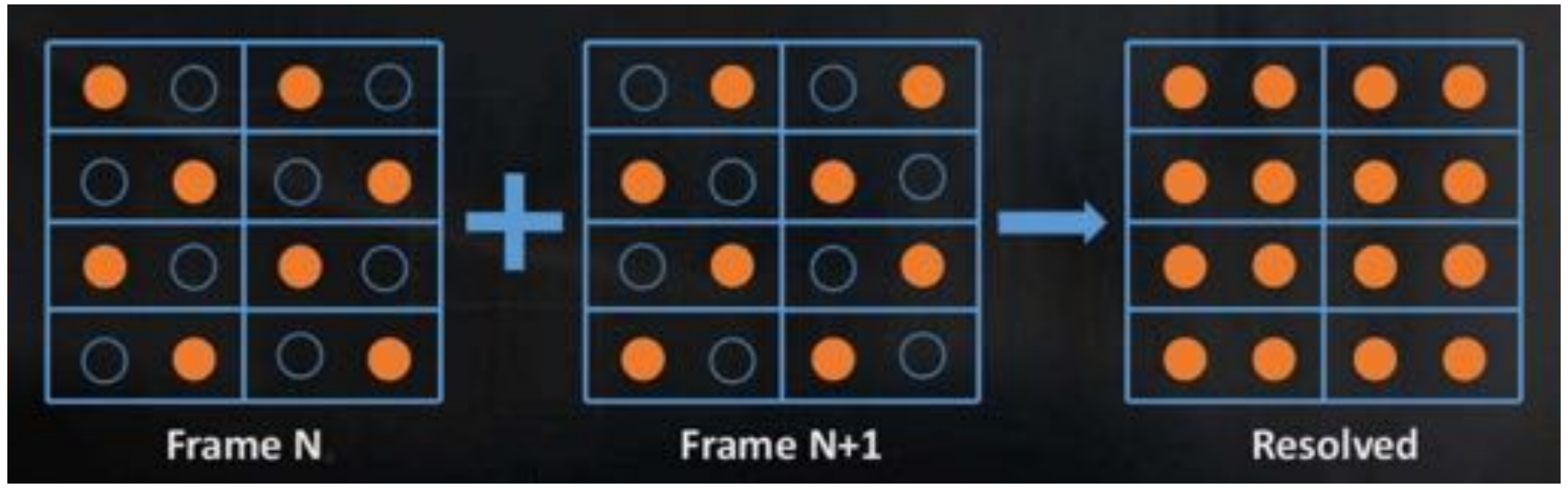
[Yang09, Lottes11, Sousa11, Karis14, Salvi16]

Temporal Upsampling



[Yang09, Herzog10, Malan12, Valient14, Aalto16, Epic18]

Checkerboard Rendering (CBR)



[ElMansouri16, Carpentier17, Wilidal17]

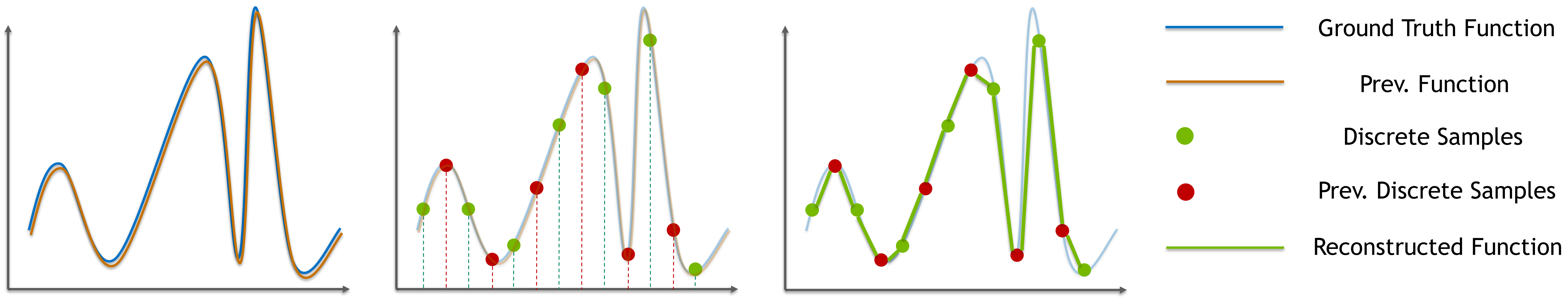
References can be found in <A Survey of Temporal Antialiasing Techniques>, Yang et al.

SPATIAL-TEMPORAL SUPER SAMPLING

Reconstruct high resolution image using samples from across multiple frames

Effective sampling rate drastically increased

Reconstructed image much closer to ground truth

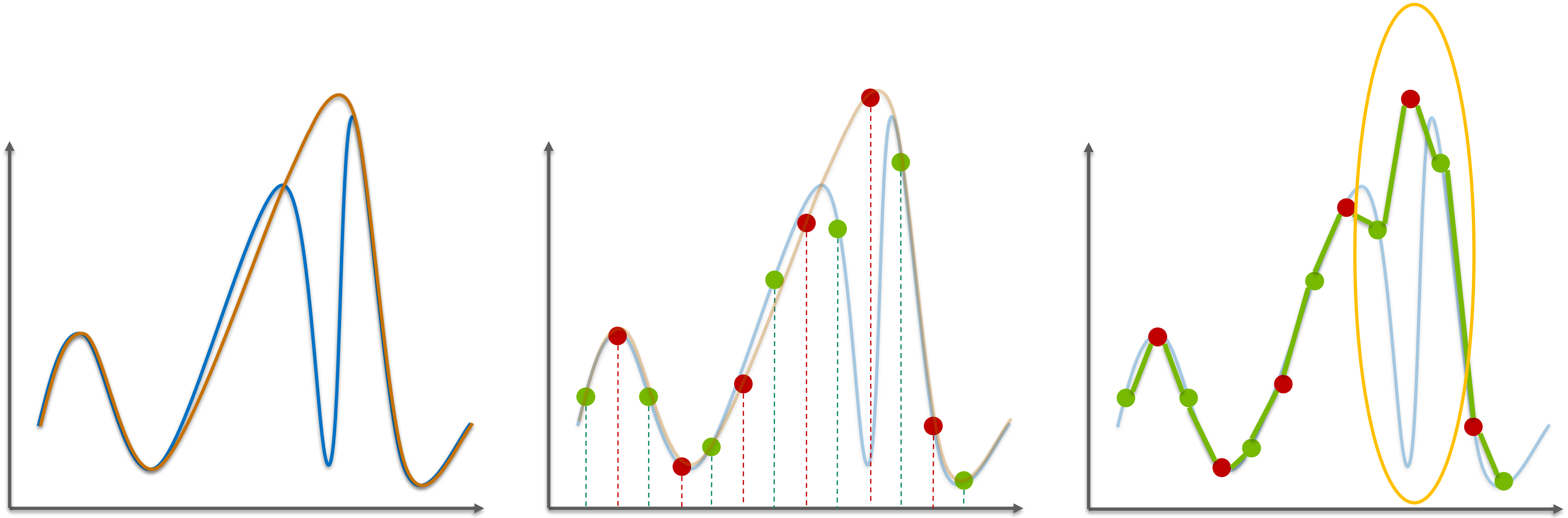


SPATIAL-TEMPORAL SUPER SAMPLING

The devil's in the details

Samples from previous frames might no longer be correct due to content changes

Using samples from previous frame naively might lead to artifacts like ghosting



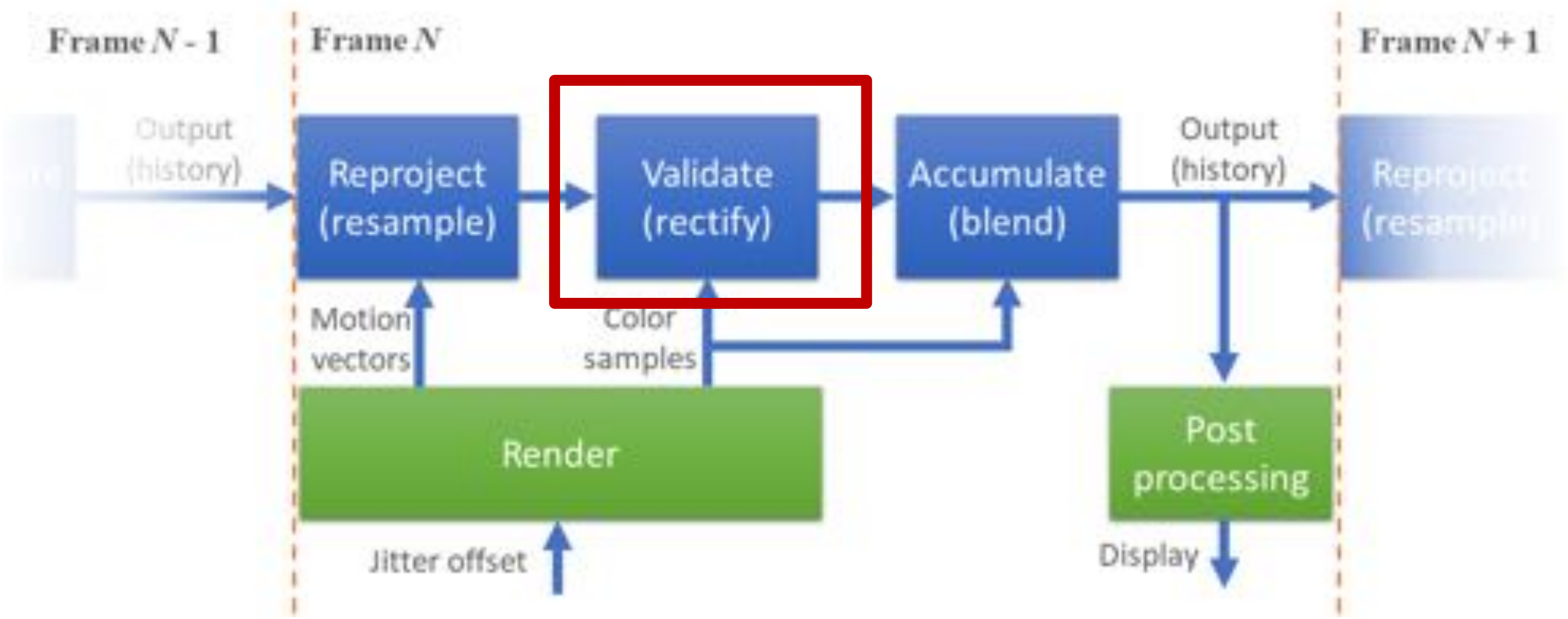
SPATIAL-TEMPORAL UPSAMPLING

History Rectification

Traditional spatial-temporal upsampling algorithms leverages heuristics to rectify invalid samples from previous frames

However common rectification heuristics often trade off between different artifacts:

Blurriness, temporal instability, even moire pattern vs. lagging and ghosting



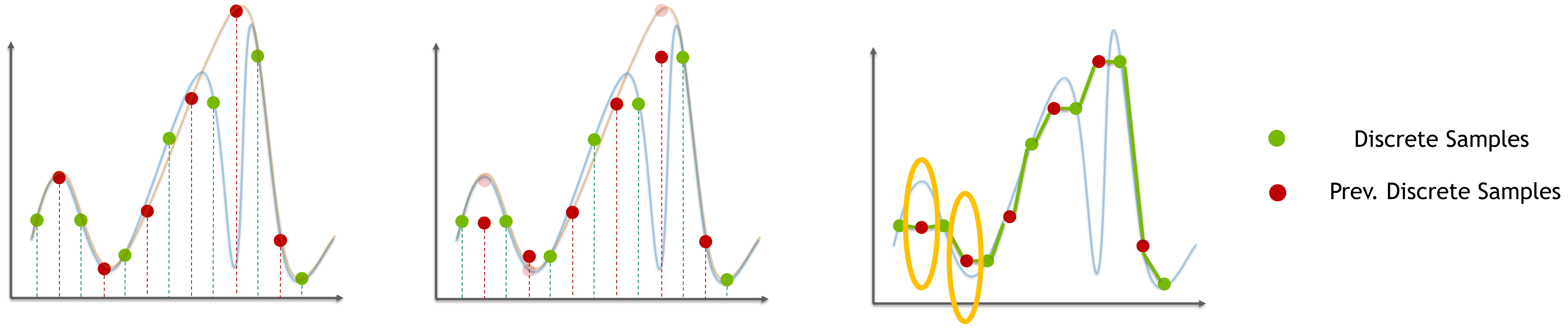
[Yang et al. 2020]

NEIGHBORHOOD CLAMPING

Most commonly used sample rectification technique [Karis14], [Salvi16]

Clamp previous frames samples to the min/max of the neighboring current frame samples

Resulted in loss in details in the reconstructed image



Ghosting Happens without History Rectification

NEIGHBORHOOD CLAMPING

Blurriness/Losing details

1spp Input

Reconstruction with clamping

Reconstruction without clamping

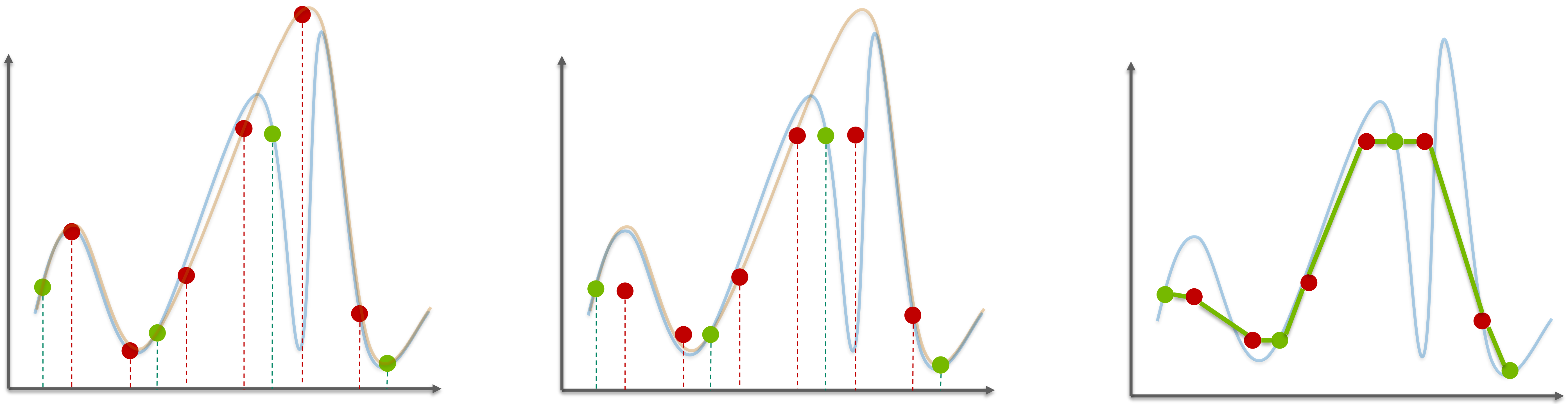
[Yang et al. 2020]

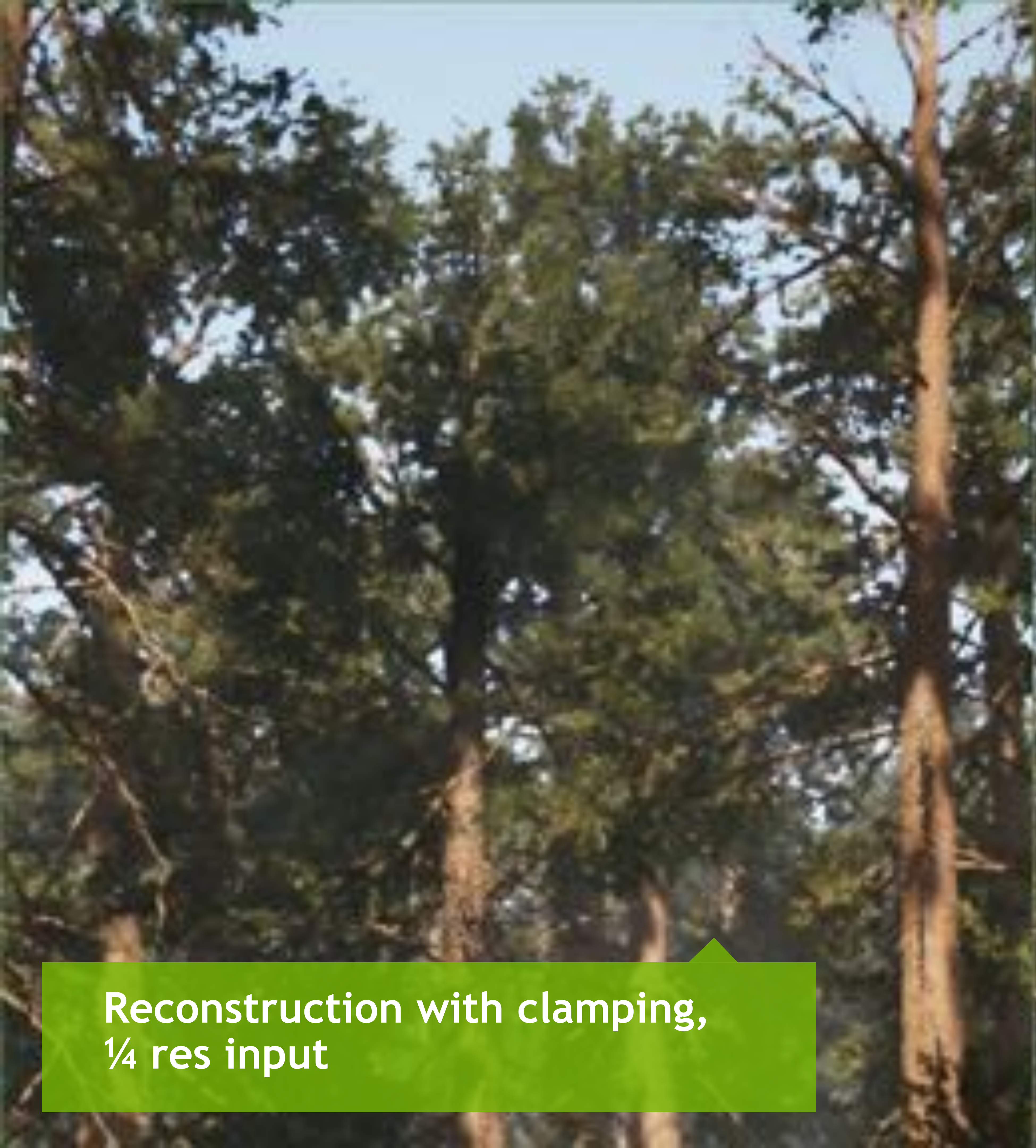
NEIGHBORHOOD CLAMPING

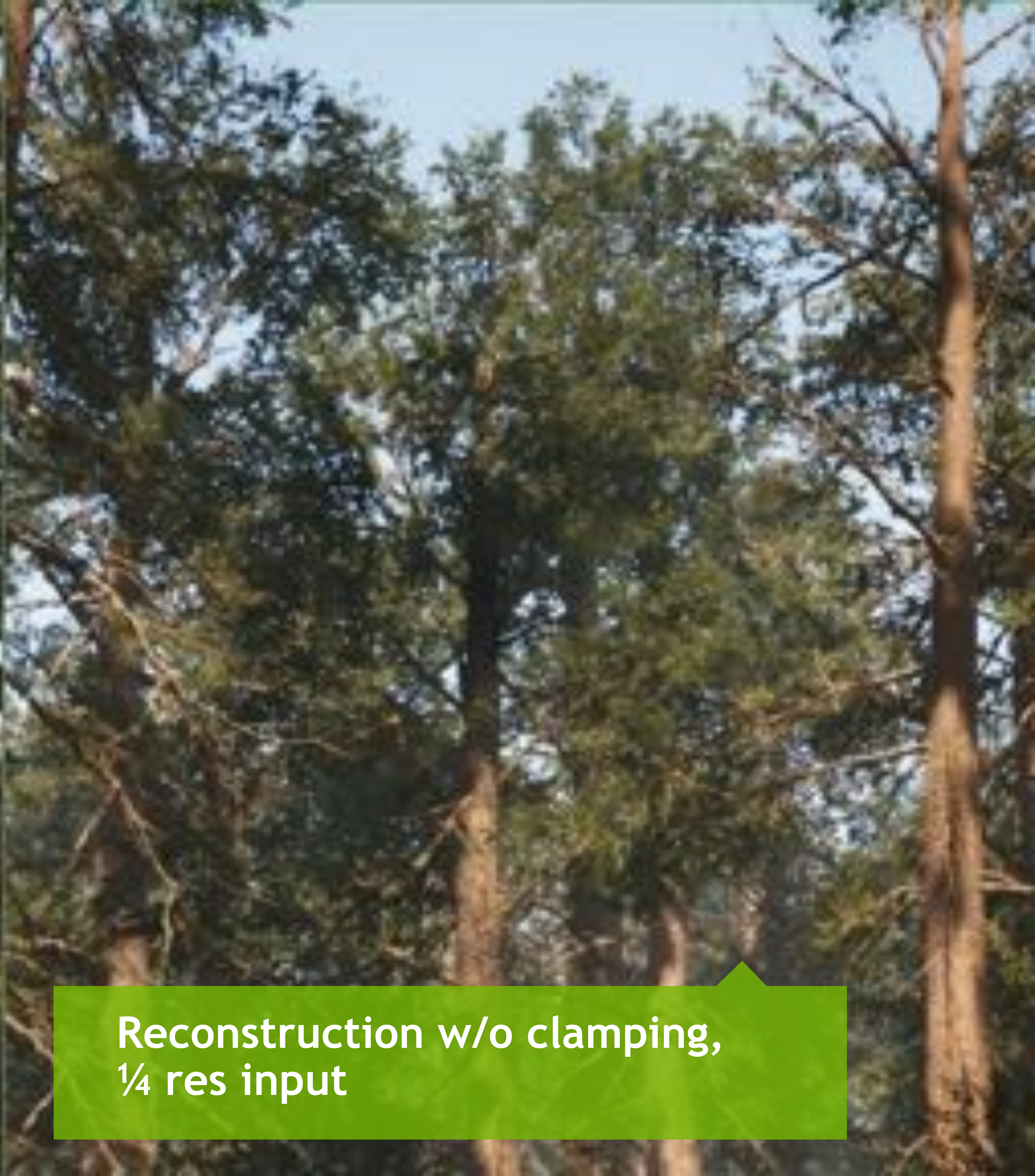
Blurriness/Losing detail

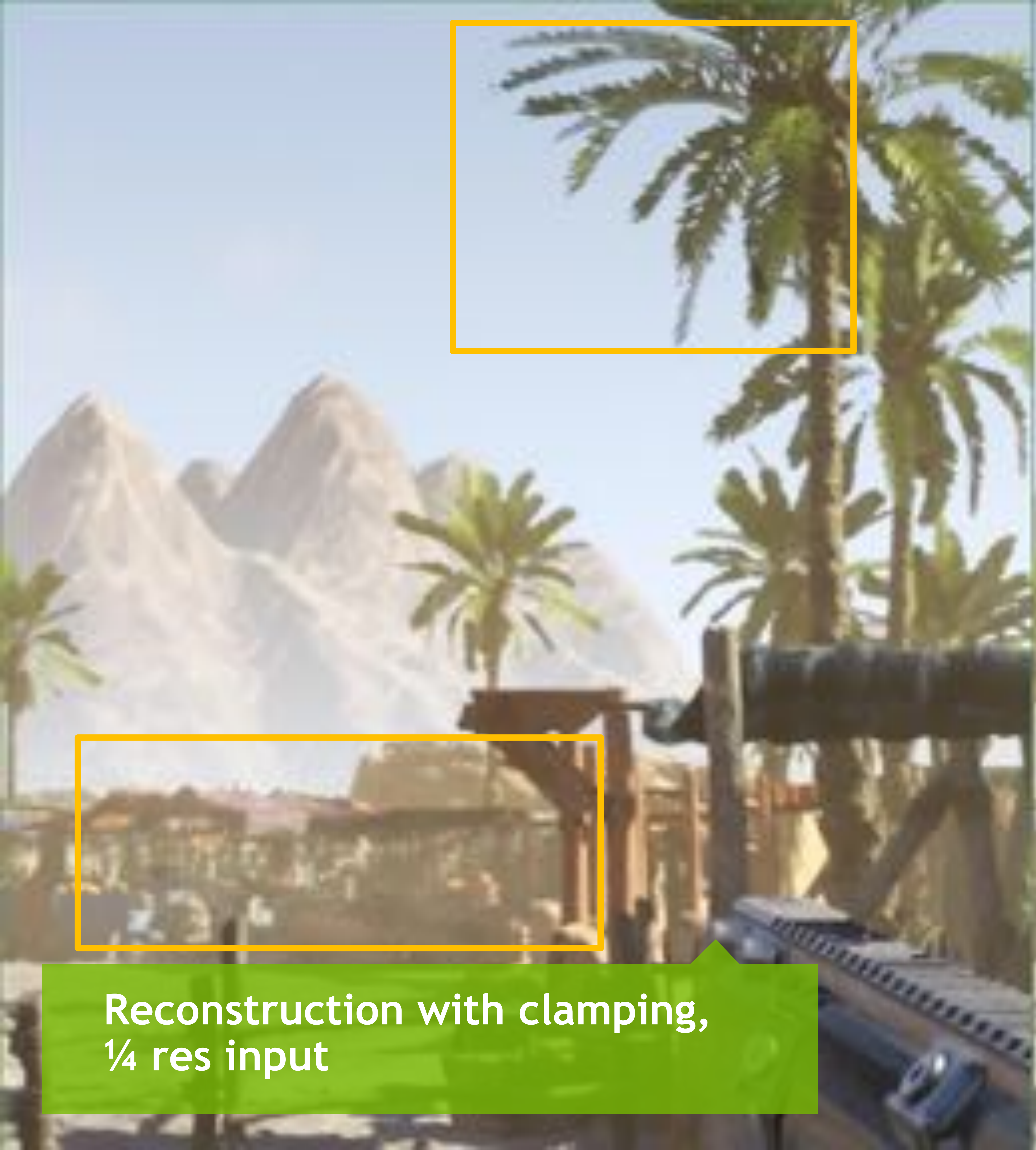
When perform temporal upsampling, clamping introduces more loss in detail

Since bounding boxes are calculated from a low-resolution image

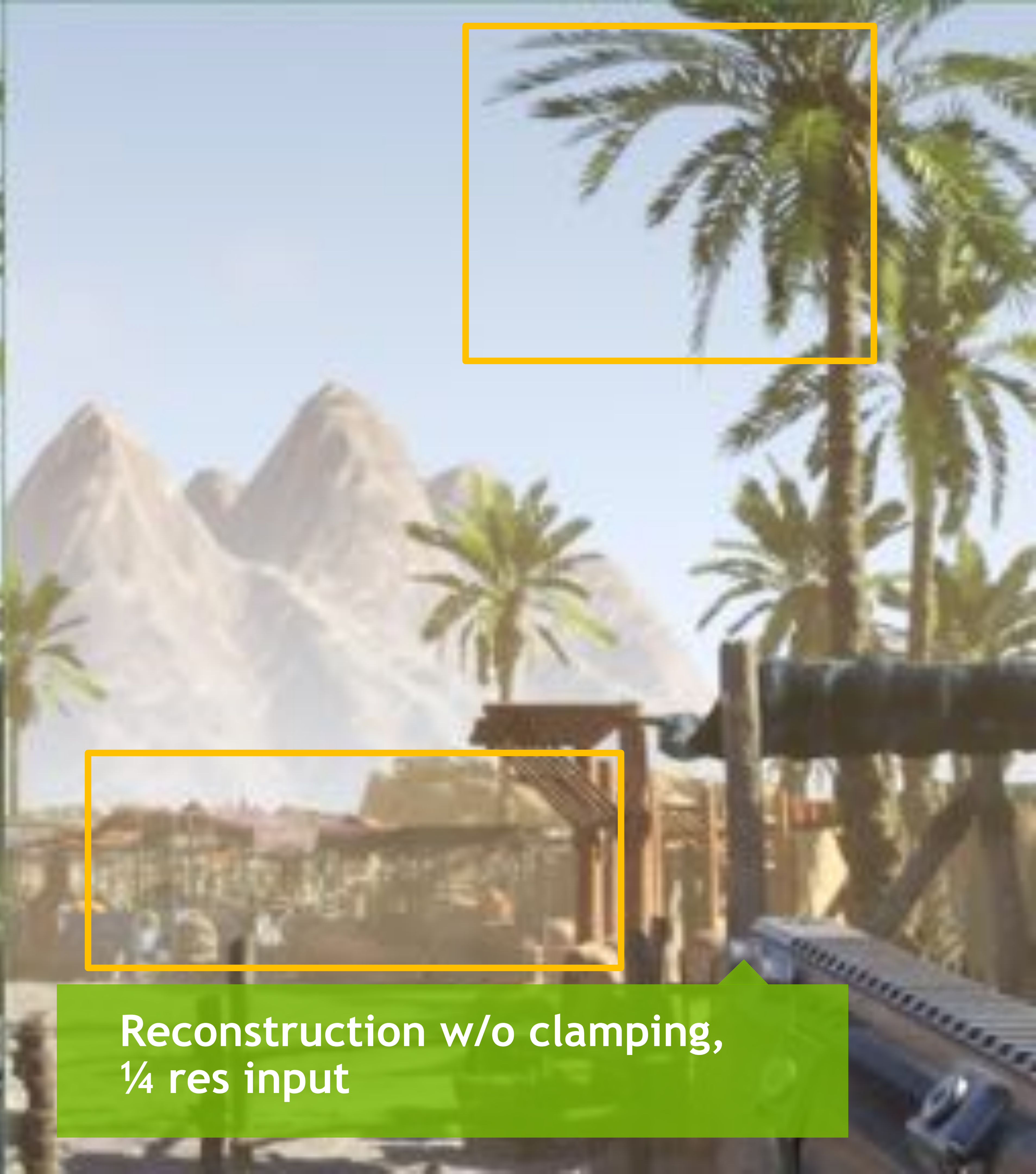



Reconstruction with clamping,
¼ res input


Reconstruction w/o clamping,
¼ res input



Reconstruction with clamping, ¼ res input



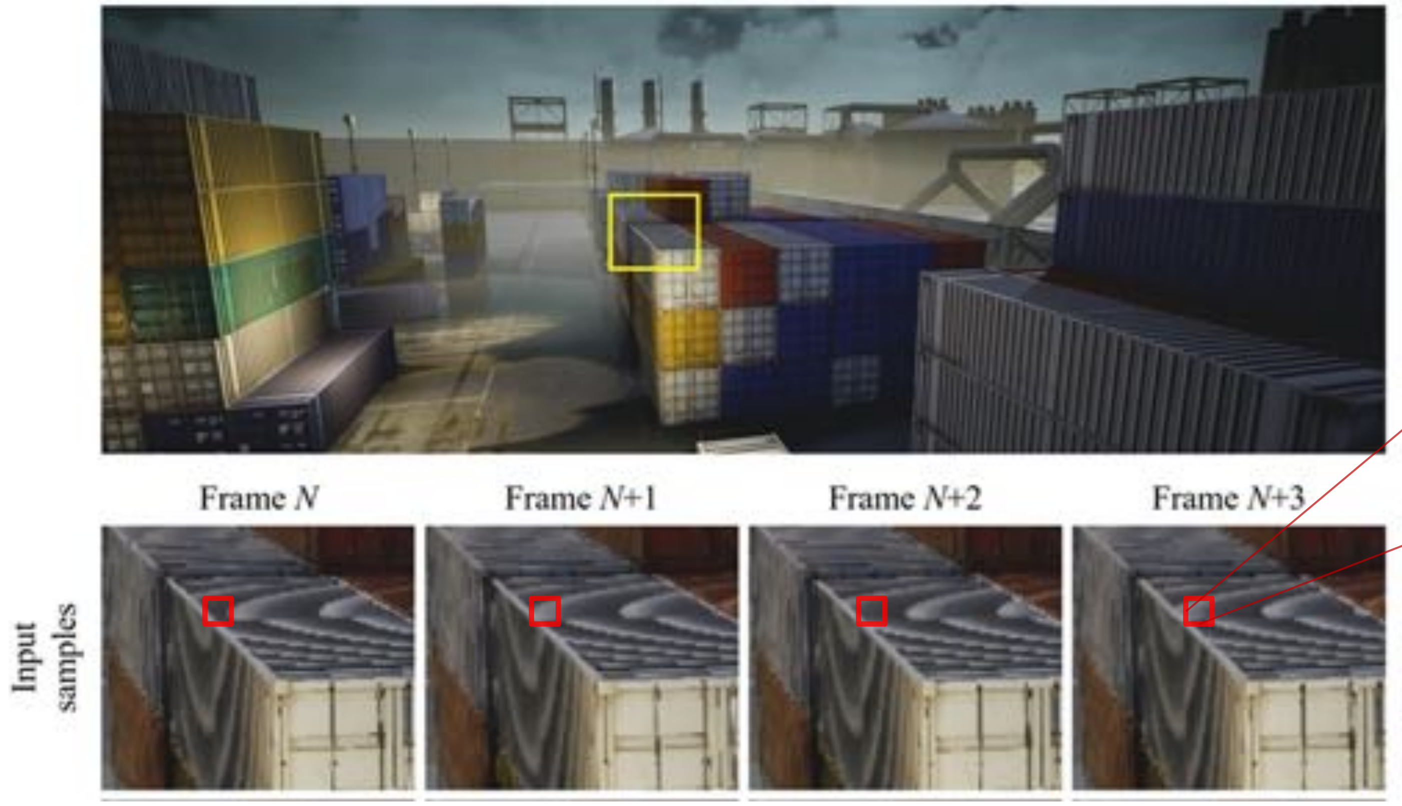
Reconstruction w/o clamping, ¼ res input

Temporal Instability
1080p TAA with clamping

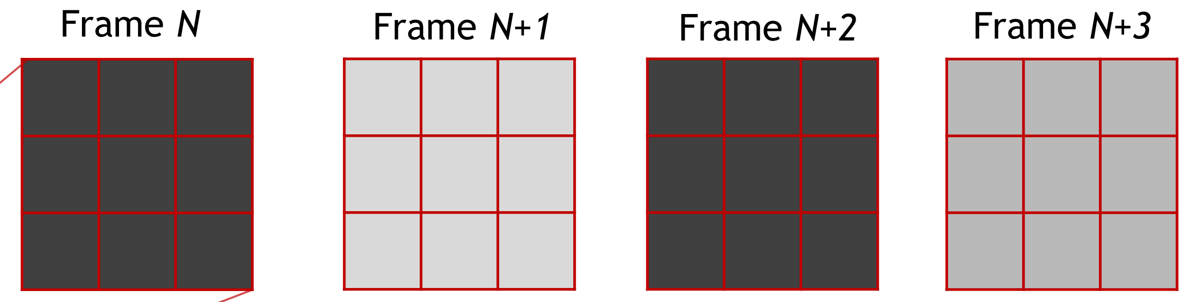
Temporal Instability
540p to 1080p TAAU with Clamping

NEIGHBORHOOD CLAMPING

Temporal Instability and Moire

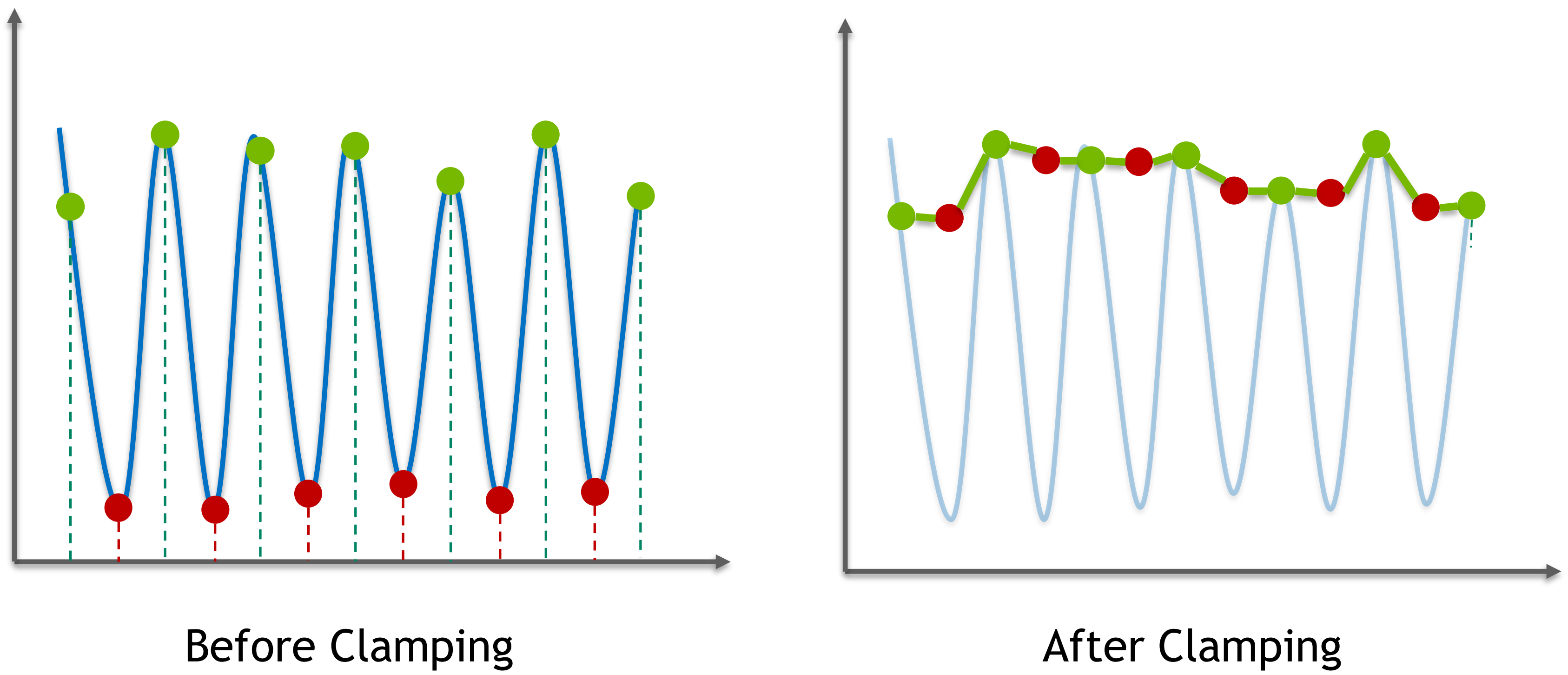


[Yang et al. 2020]



Discrete Samples

Prev. Discrete Samples



540p to 1080p TAAU w/o Clamping

REAL-TIME SUPER-RES CHALLENGES

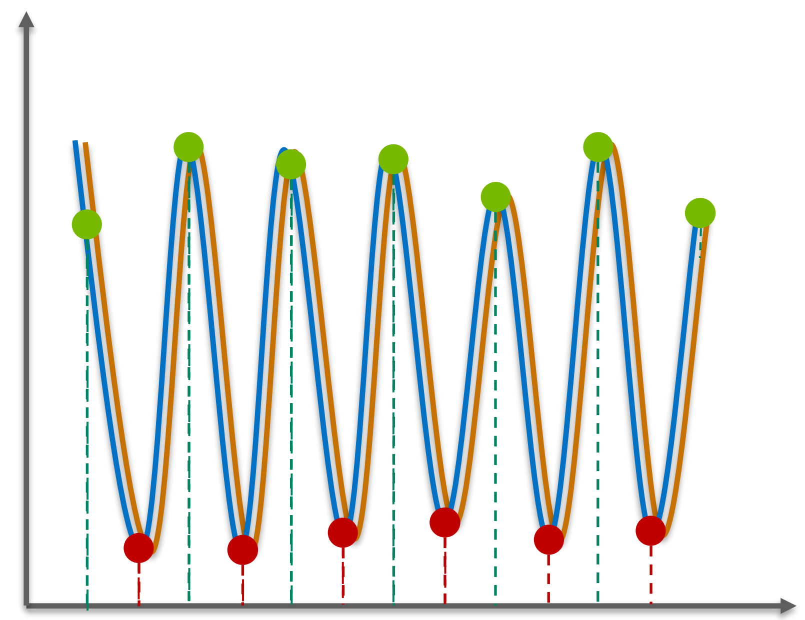
- Single frame approach
  - Blurry image quality
  - Inconsistent with native rendering
  - Temporally unstable
- Multi-frame approach
  - Heuristics to detect and rectifies changes across frames
  - Limitation in heuristics causing blurriness, temporal instability and ghosting

DLSS 2.0: DL BASED MULTI-FRAME RECONSTRUCTION

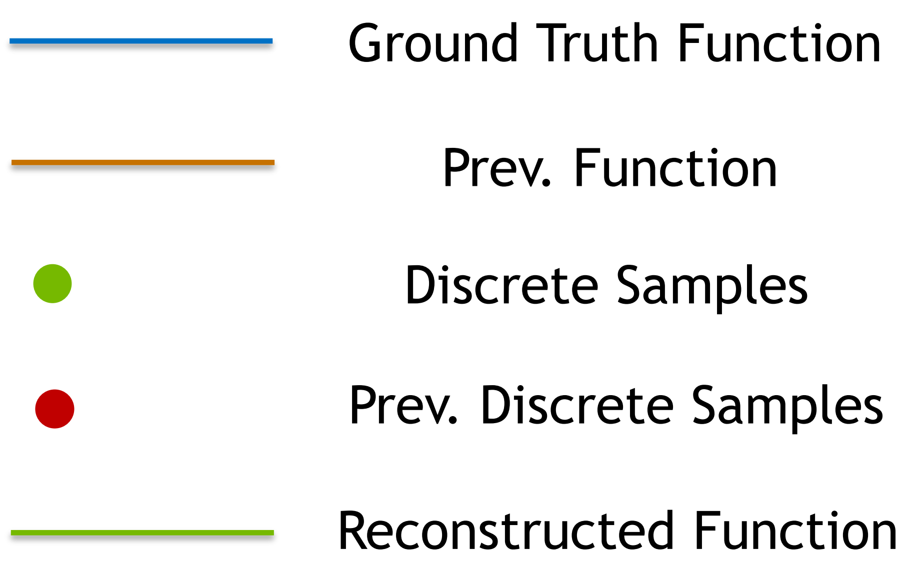
DLSS uses a neural network trained from tens of thousands of high-quality images

Neural networks are much more powerful than handcrafted heuristics

Much higher quality reconstructions using samples from multiple frames



Multi-frame samples and GT function



DLSS 2.0: DL BASED MULTI-FRAME RECONSTRUCTION

DLSS uses a neural network trained from tens of thousands of high-quality images

Neural networks are much more powerful than handcrafted heuristics

Much higher quality reconstructions using samples from multiple frames

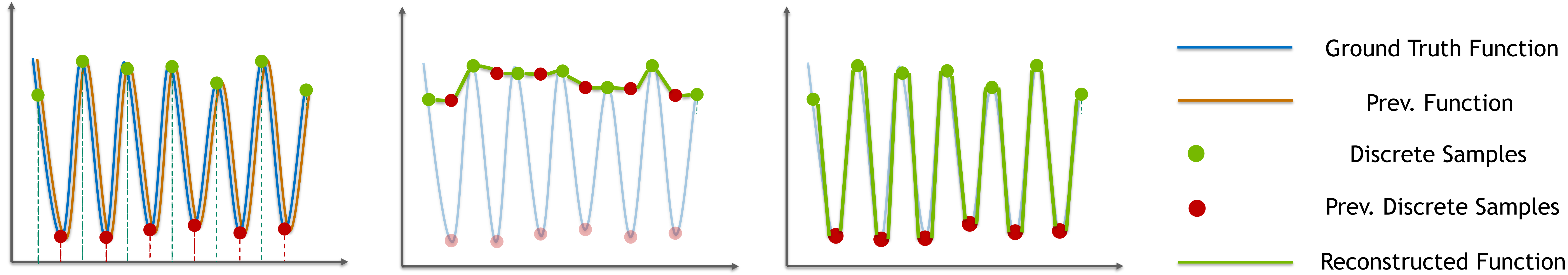
Multi-frame samples and GT function

DLSS 2.0: DL BASED MULTI-FRAME RECONSTRUCTION

DLSS uses a neural network trained from tens of thousands of high-quality images

Neural networks are much more powerful than handcrafted heuristics

Much higher quality reconstructions using samples from multiple frames



Multi-frame samples and GT function

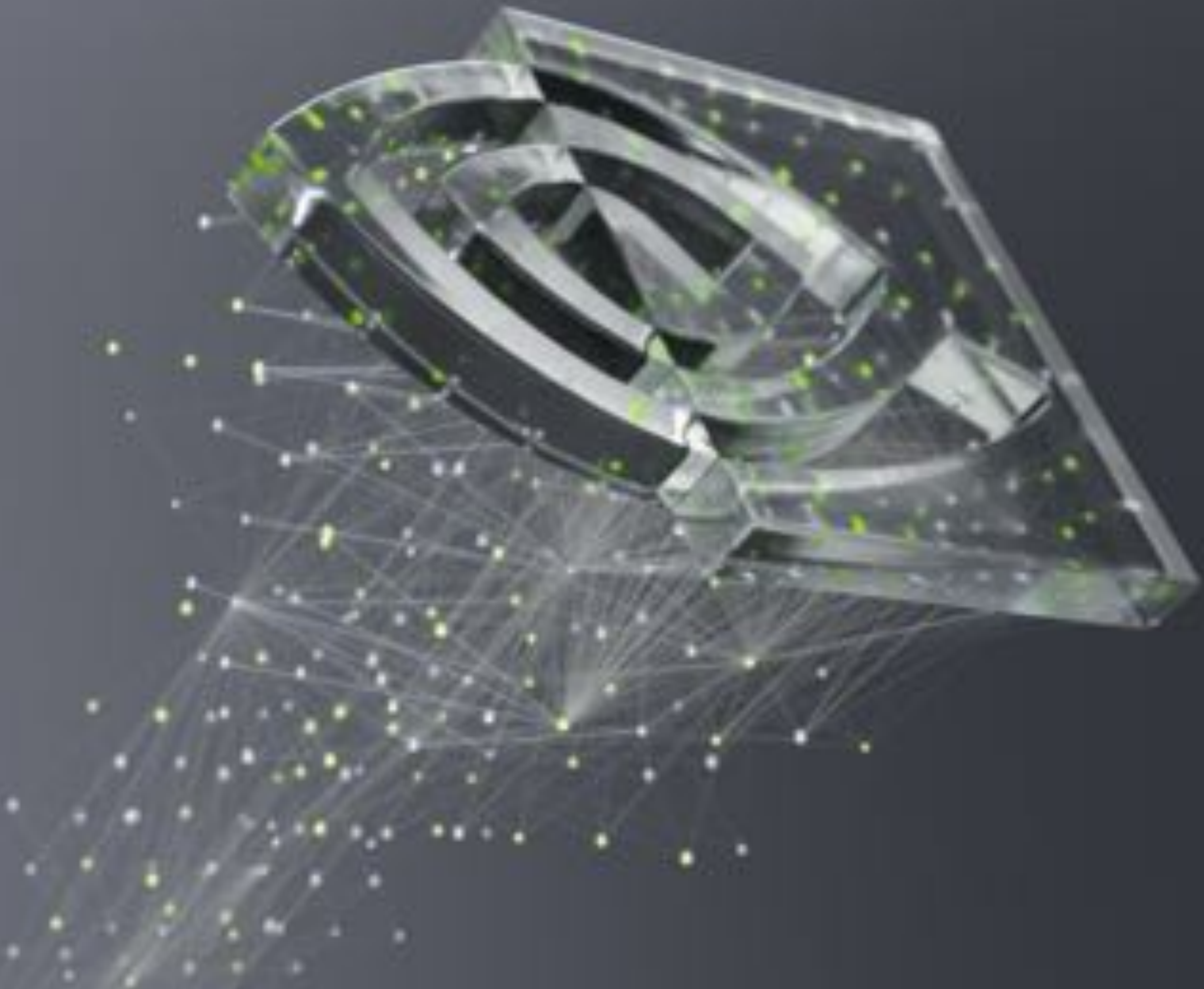
Non-DL reconstruction

DLSS reconstruction

DLSS 2.0
540p to 1080p


1080p TAA
540p DLSS 2.0
540p TAAU


1080p TAA
540p DLSS 2.0
540p TAAU



Thank you!

: @edliu1105

知 : @文刀秋二