SPATIOTEMPORAL VARIANCE-GUIDED FILTERING: REAL-TIME RECONSTRUCTION FOR PATH-TRACED GLOBAL ILLUMINATION

Christoph Schied, Anton Kaplanyan, Chris Wyman, Anjul Patney, Chakravarty R. Alla Chaitanya, John Burgess, Shiqiu Liu, Carsten Dachsbacher, Aaron Lefohn, Marco Salvi

NVIDIA

KIT
Karlsruher Institut für Technologie

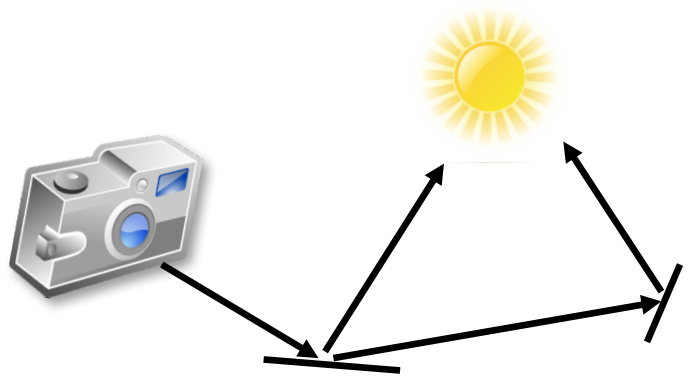
McGill UNIVERSITY

Université de Montréal

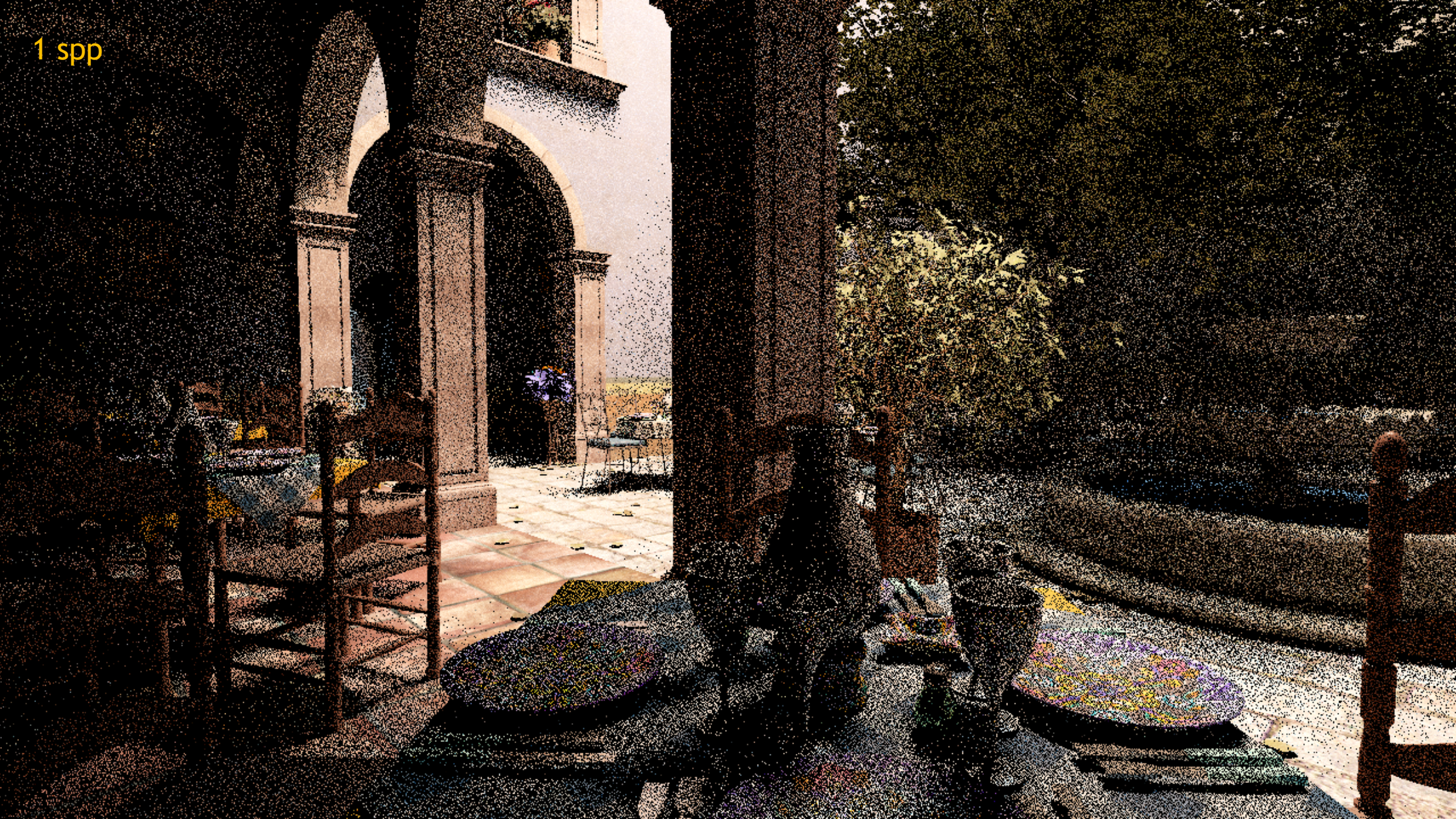

Reference

REAL-TIME PATH TRACING

- Two bounce global illumination requires 3 rays/sample
  - Sample primary hit with rasterizer
  - One ray to find indirect bounce
  - One shadow ray for direct and indirect each (next event estimation)

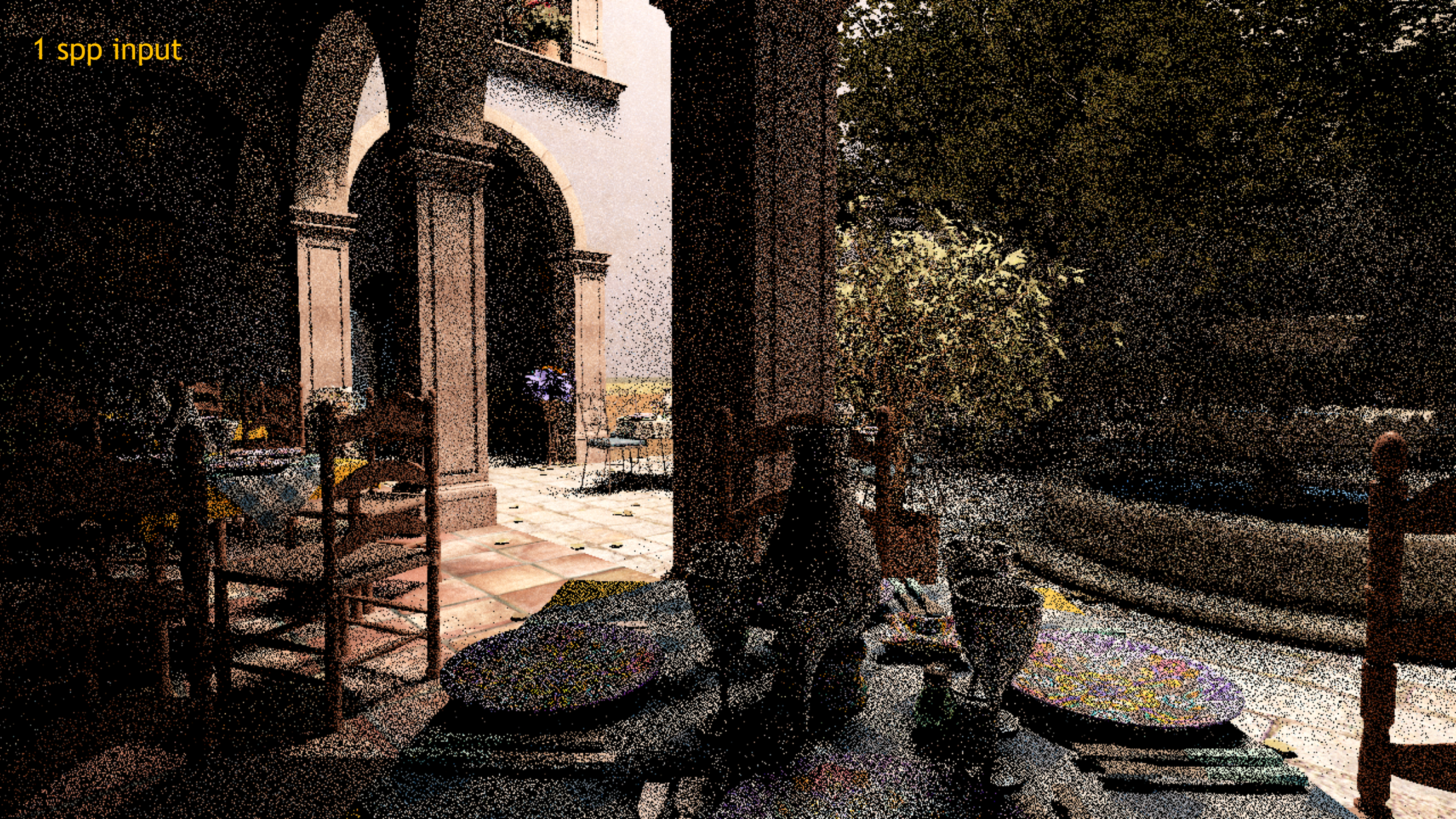


- Current ray-tracers achieve around 200 MRays/s
  - For real-time rendering we will have to survive at 1 spp for foreseeable time


1 spp

GOALS

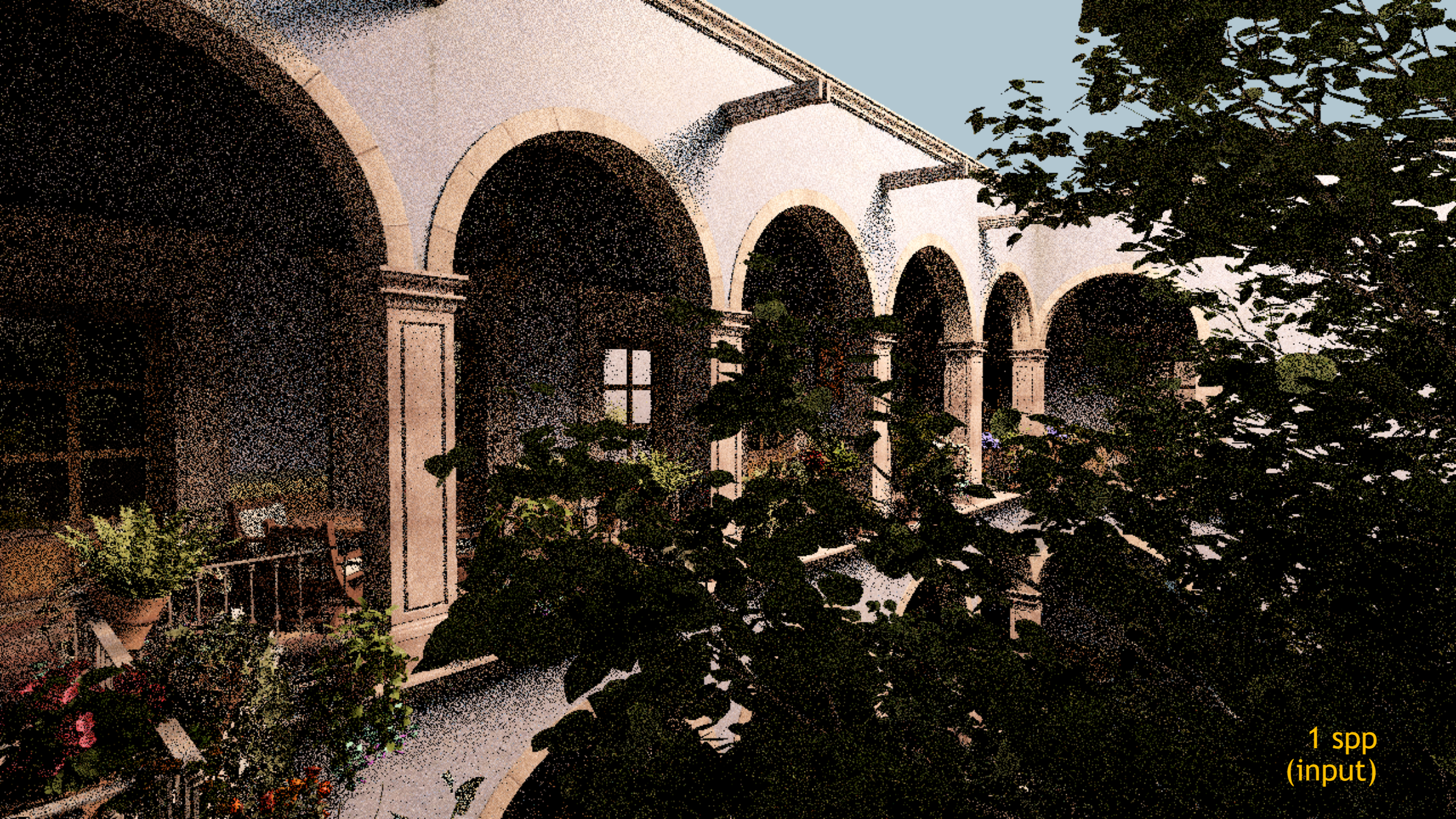
- Temporally-stable image reconstruction from 1 sample/pixel
- Preserve features and details where possible
  - Bias is acceptable!
- Filter needs to be very efficient


1 spp input


Our filter


Reference

PREVIOUS WORK


1 spp
(input)


Reference


Edge-avoiding à-trous wavelets (EAW)
[Dammertz et al. 2010]


SURE-based filter (SBF)
[Li et al. 2012]


Learning-based filter (LBF)
[Kalantari et al. 2015]


Our filter


Edge-Avoiding À-Trous
Wavelet Filter

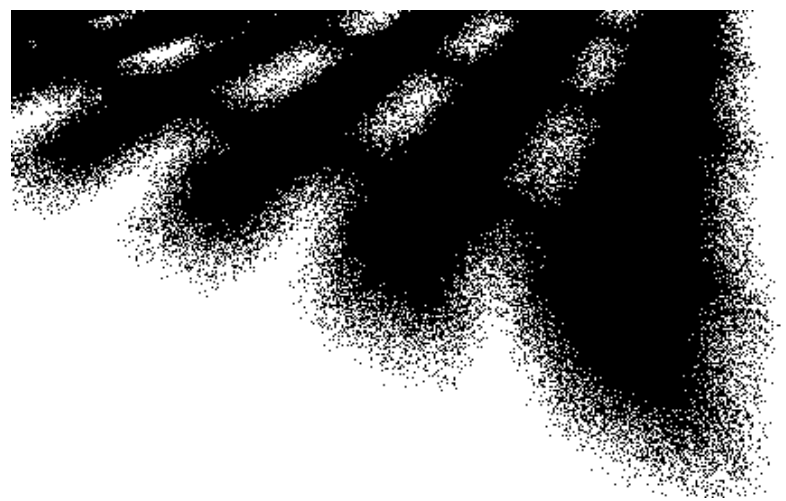
OUR FILTER

DESIGN PHILOSOPHY

- Assume noise-free G-buffer
- Exploit high frame rates, leverage temporal information
- All decisions performance-driven
  - Rules out many advanced methods (e.g., Fourier domain, regression models, NL-means, etc.)
- Filter agnostic wrt. light transport, bounces, number of lights, scene, etc.
  - Treat path tracer as a black box

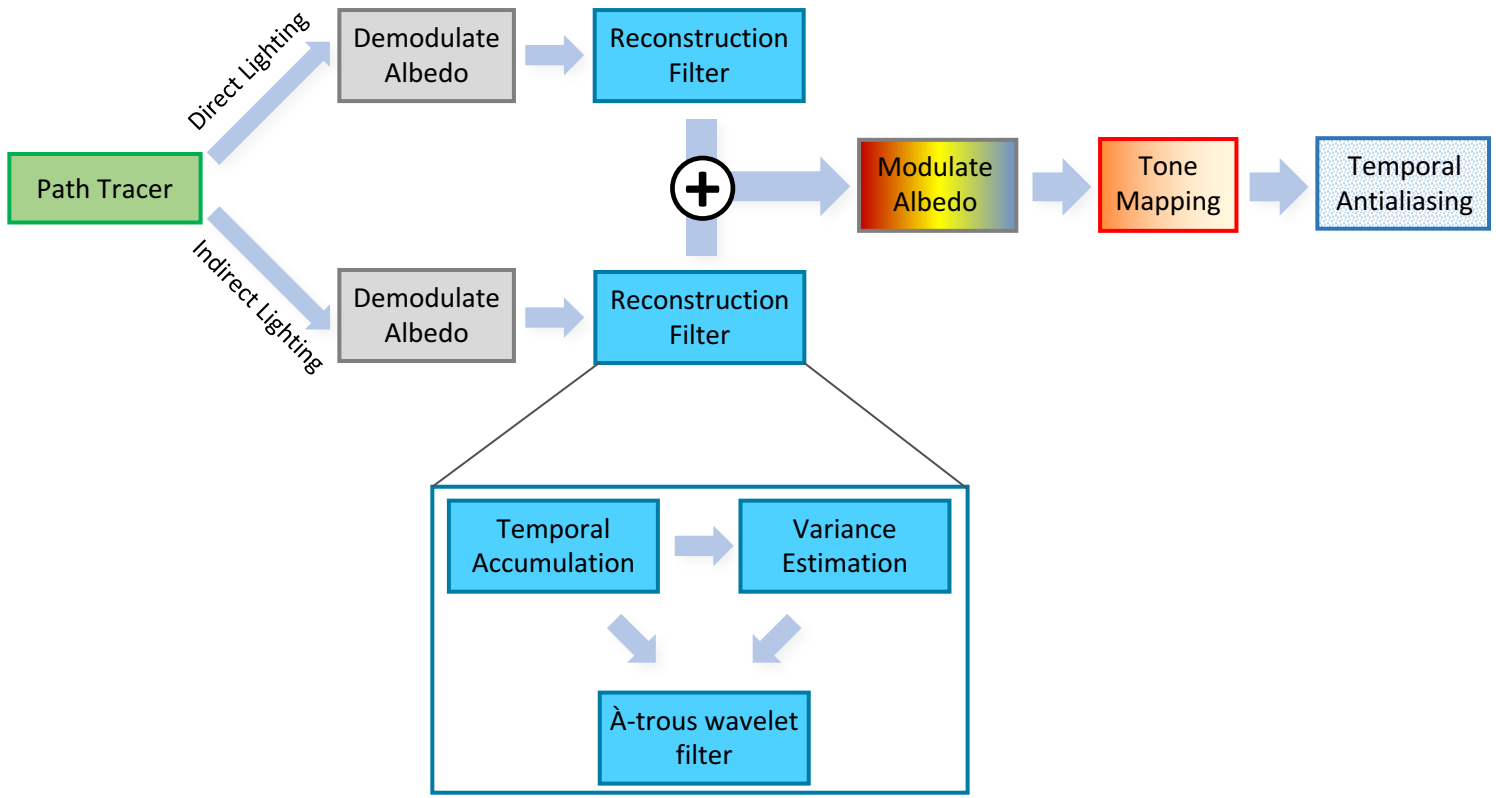
MAIN IDEAS BEHIND OUR FILTER

- Locally analyze image over time to control filter
  - Temporally unstable signal → blur more
  - Temporally stable signal → blur less



- Filter image hierarchically
  - Start with small filter footprint
  - Estimate temporal stability after each spatial filter iteration
  - Strong blur more likely in early iterations
    - Limit bias to small filter footprints

OVERVIEW



TEMPORAL ACCUMULATION

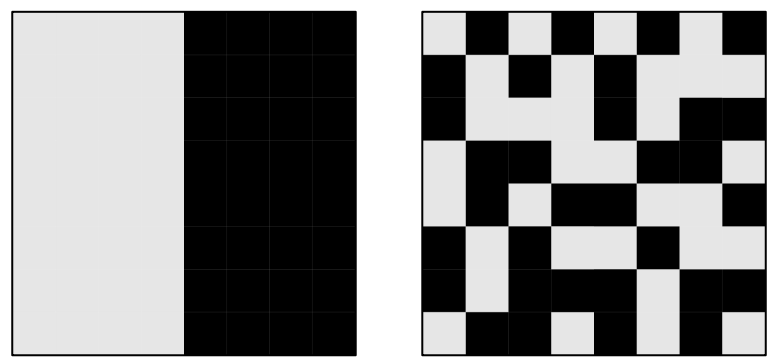
- Continuous sample accumulation via exponential moving average:

$$C'_i = \alpha \cdot C_i + (1 - \alpha) \cdot C'_{i-1}$$

- Use motion vectors to back-project samples from previous frame
  - Use depth, normals, and mesh ID tests to accept or reject individual filter taps

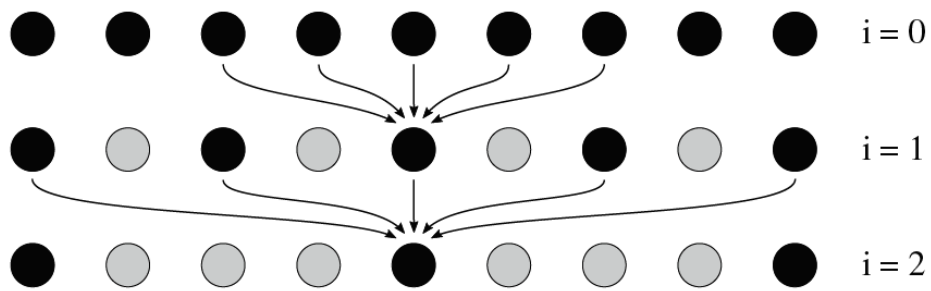
VARIANCE ESTIMATION

- Only consider luminance for efficiency reasons
- Temporally accumulate first two moments of luminance μ'_{1_i} μ'_{2_i}
- Estimate variance as $\sigma'^2_i = \mu'_{2_i} - \mu'^2_{1_i}$
- Fallback to spatial estimate of variance when filter lacks temporal history
  - 7x7 cross-bilateral filter (reuse edge-stopping functions of wavelet filter)



EDGE-AVOIDING À-TROUS WAVELETS

- Each iteration performs convolution with filter kernel h
- Double the filter size in each iteration
  - Undecimated wavelets, i.e. full resolution in each iteration
  - Skip more values each iteration
- Edge-stopping functions $w(p, q)$ based on depth, normals and luminance of pixels p, q
- We compute a new variance estimate at each iteration



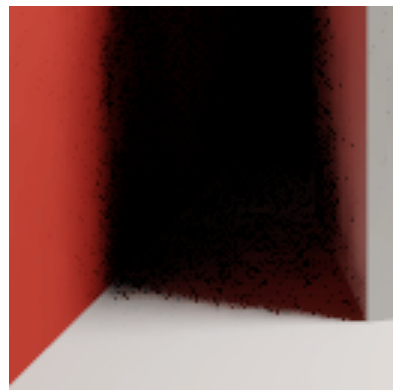
$$\hat{c}_{i+1}(p) = \frac{\sum_{q \in \Omega} h(q) \cdot w(p, q) \cdot \hat{c}_i(q)}{\sum_{q \in \Omega} h(q) \cdot w(p, q)}$$

LUMINANCE EDGE-STOPPING FUNCTION

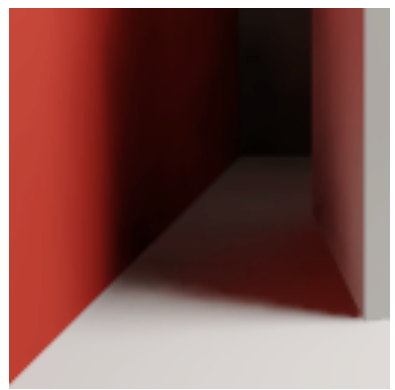
- Responsible for preserving features present in the input samples
  - Automatically adapts to all scales/iterations
- Luminance delta is re-normalized by local standard deviation
  - Variance estimate is pre-filtered with 3x3 Gaussian kernel

$$w(p,q) = \exp\left(-\frac{|l_i(p) - l_i(q)|}{\sqrt{g_{3\times 3}\left(Var\left(l_i(p)\right)\right)}}\right)$$

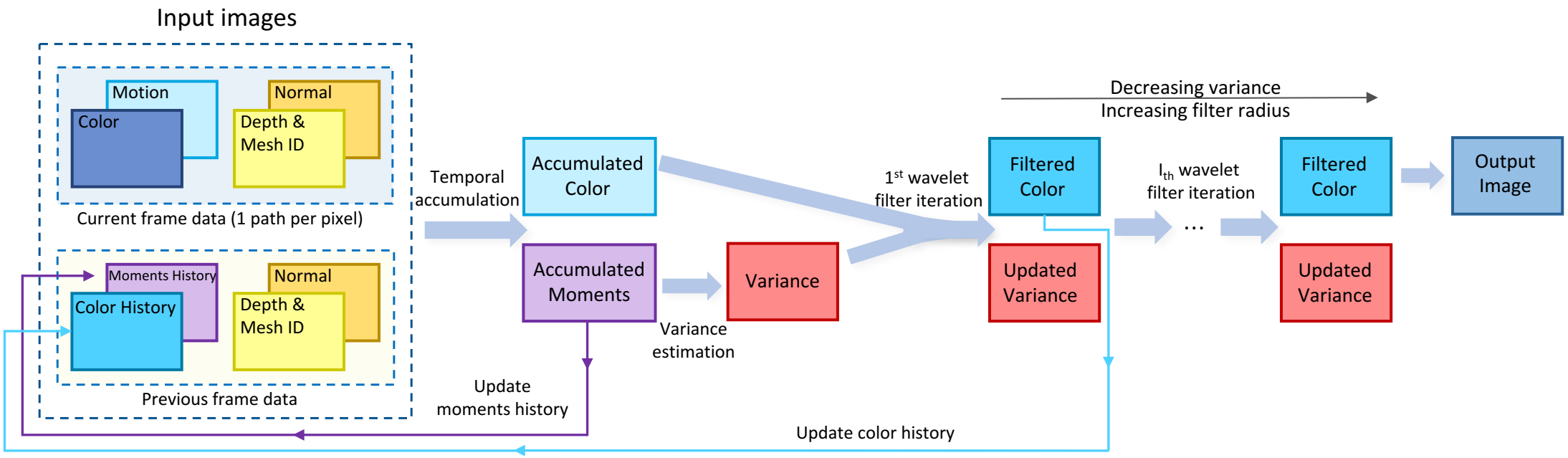
no pre-filter



with variance pre-filter



FILTER OVERVIEW



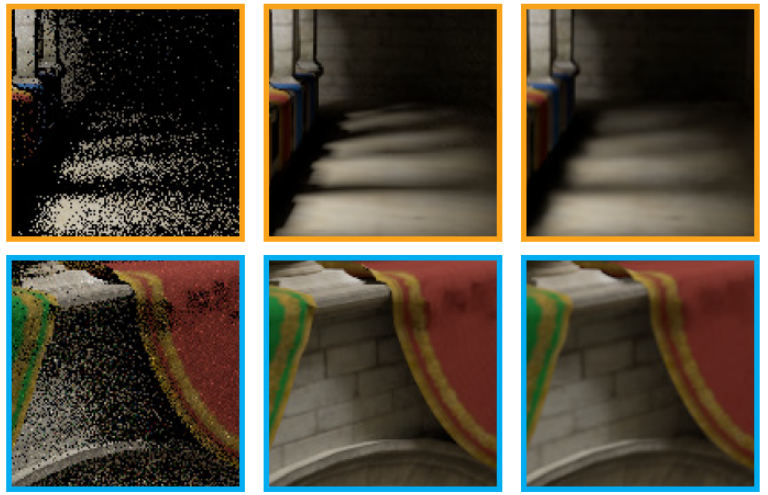
RESULTS

EVALUATION

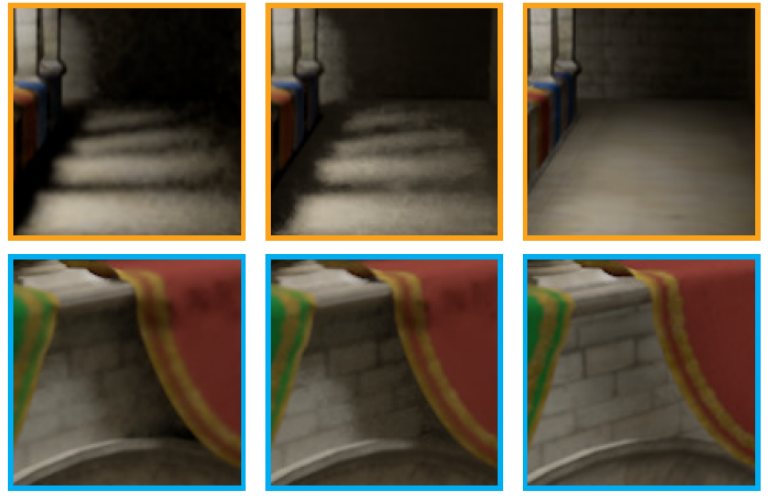
- We compare to previous work augmented with temporal accumulation and TAA
- Images are captured from animated sequences
- Looped Halton sequence, repeated after 16 frames
- Filter implemented using OpenGL in the Falcor research framework

ours

1spp | reference | ours



SBF | LBF | EAW

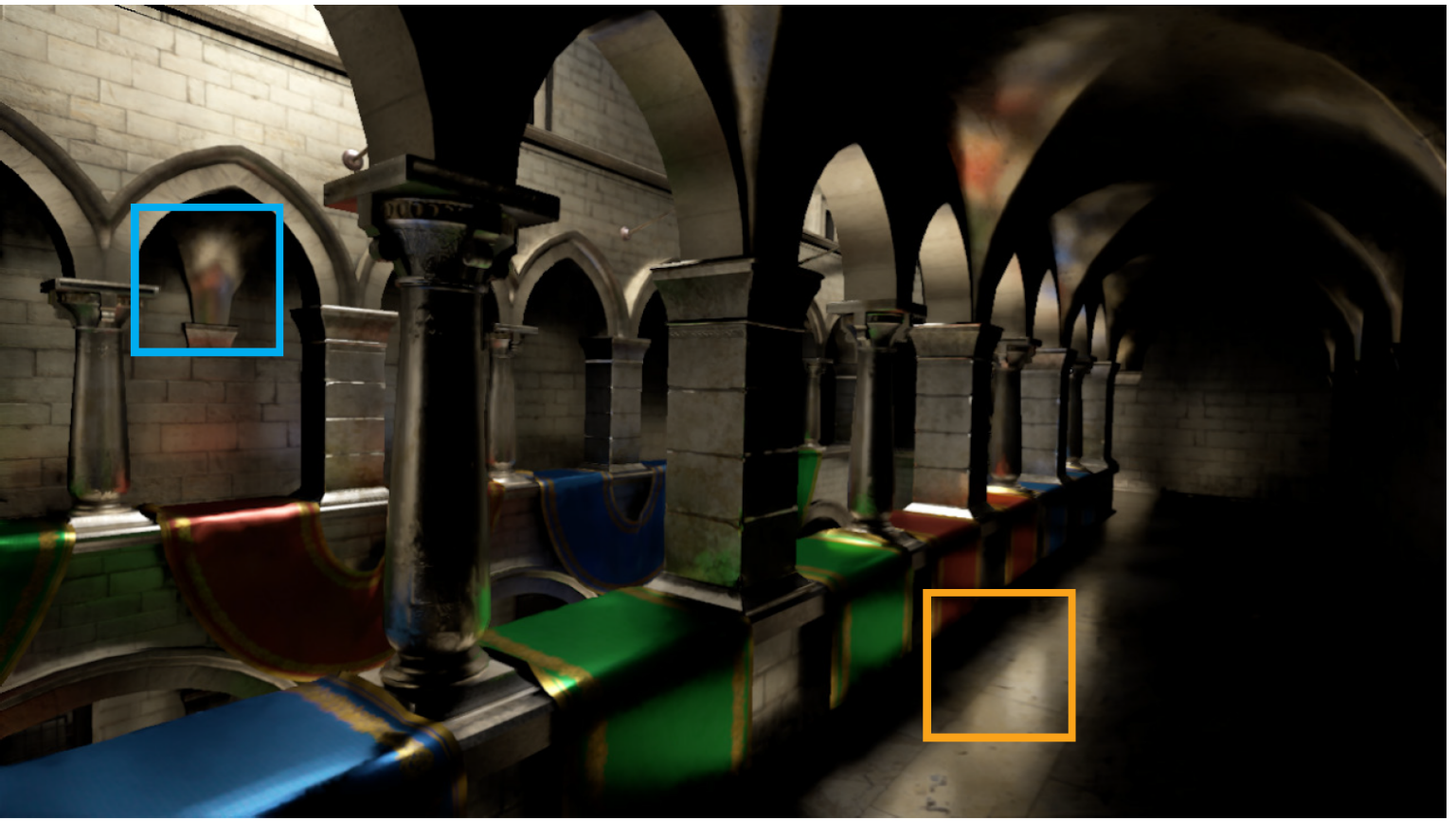


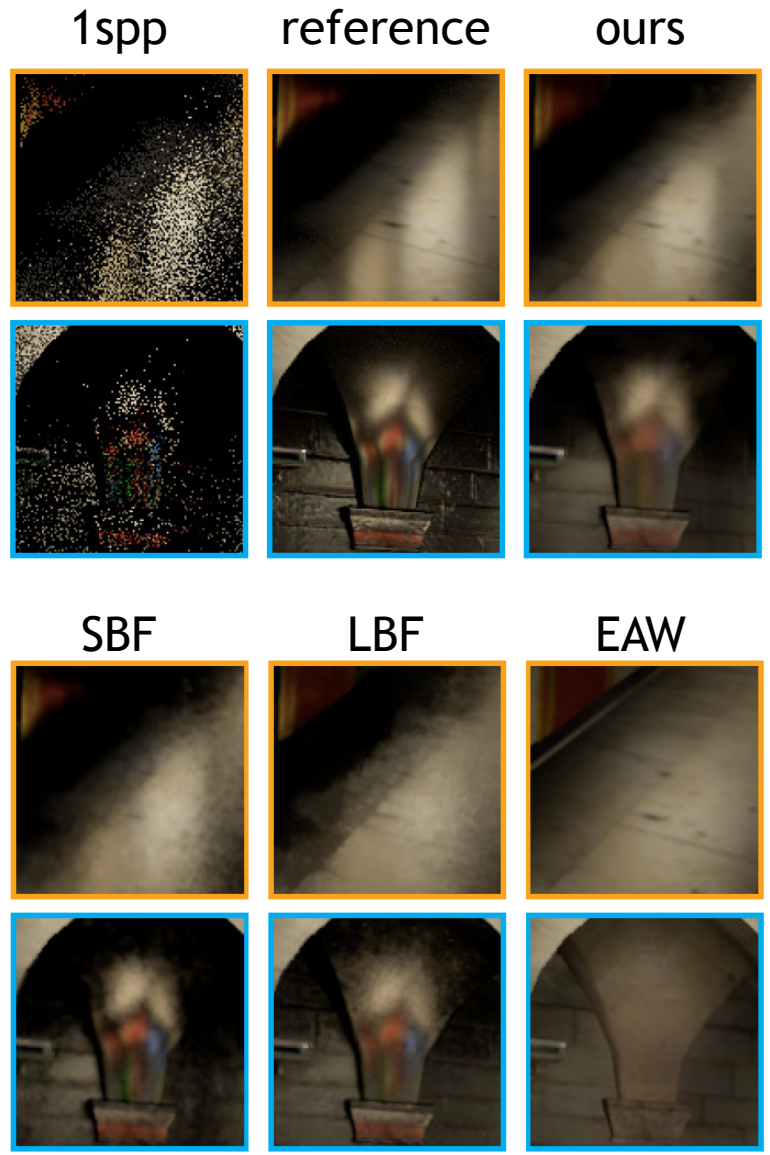
ours

1spp | reference | ours

SBF | LBF | EAW

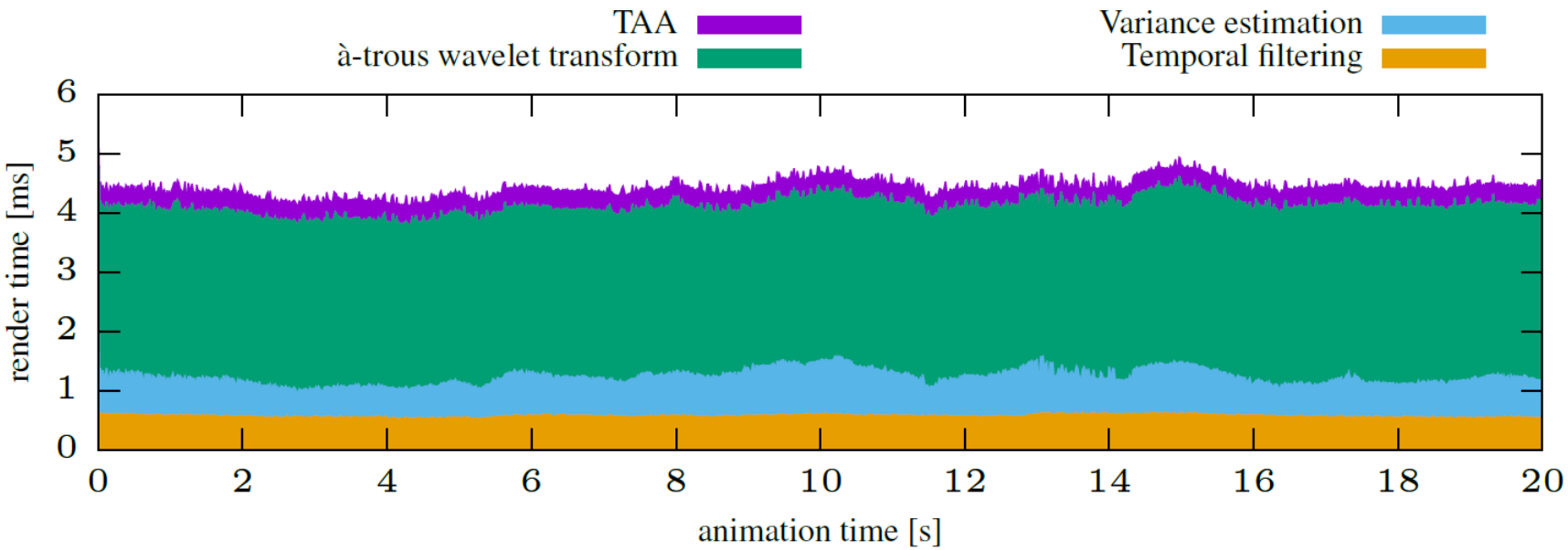
ours

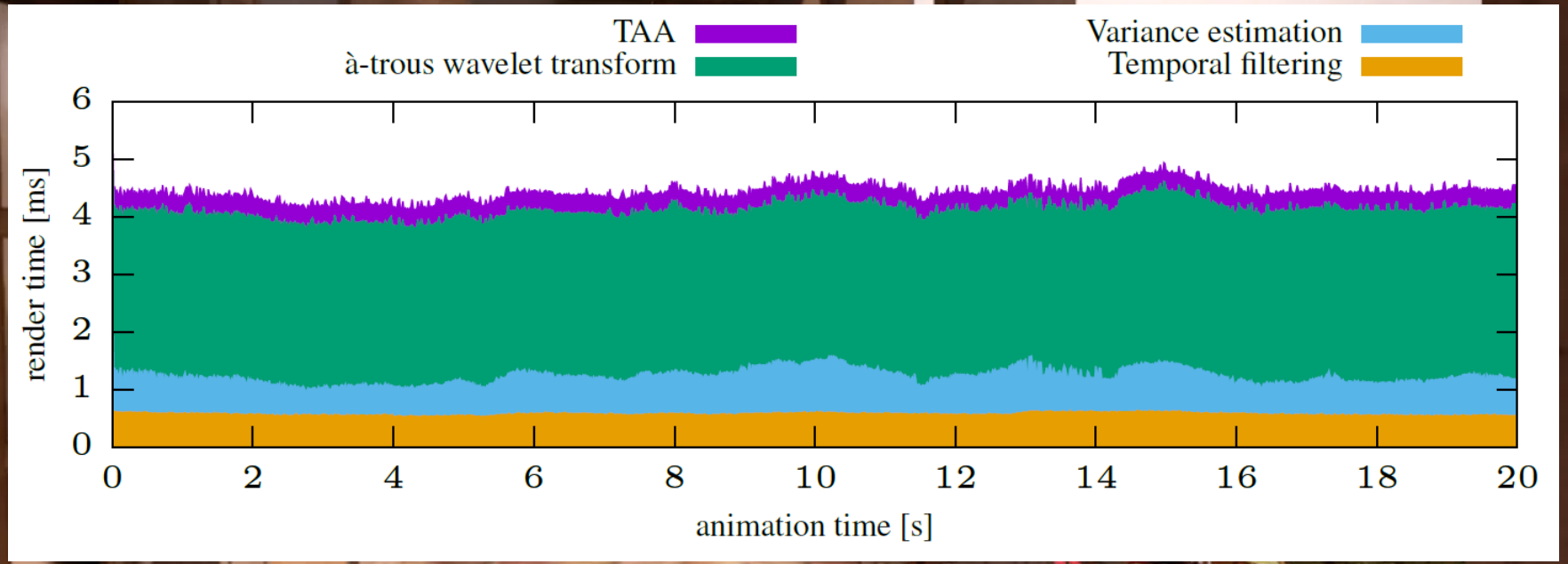




PERFORMANCE

San Miguel, 720p

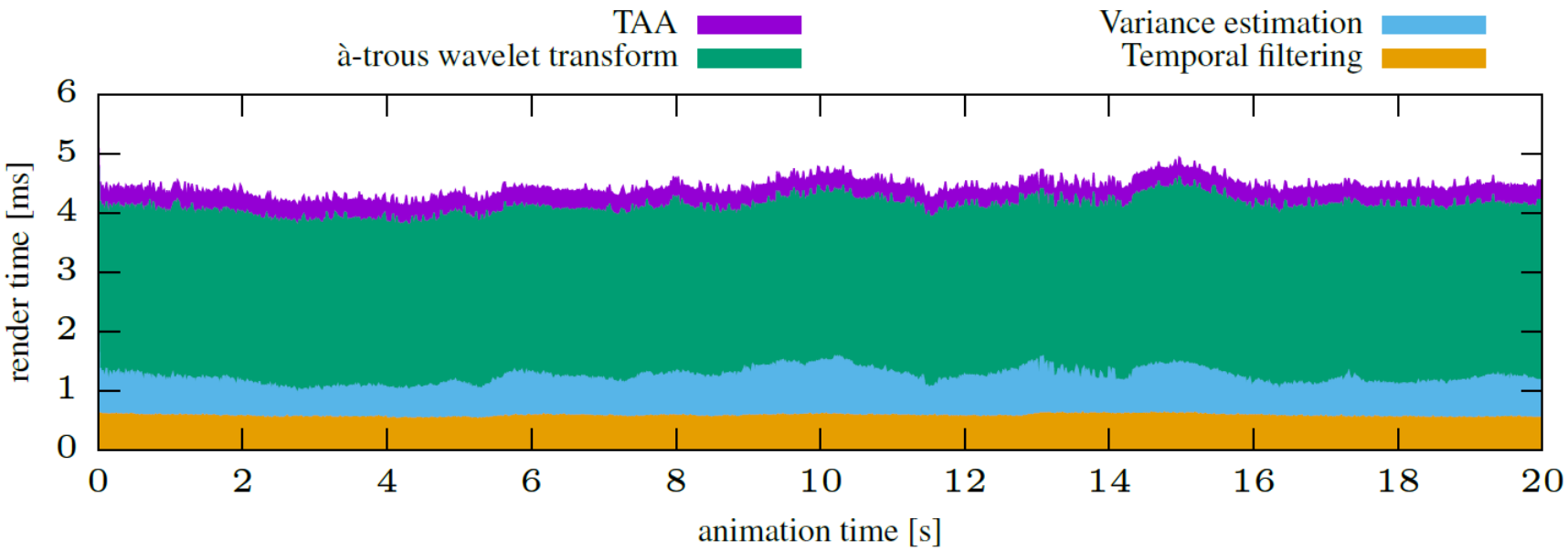




Our Spatiotemporal Variance-Guided Filter (SVGF)

PERFORMANCE

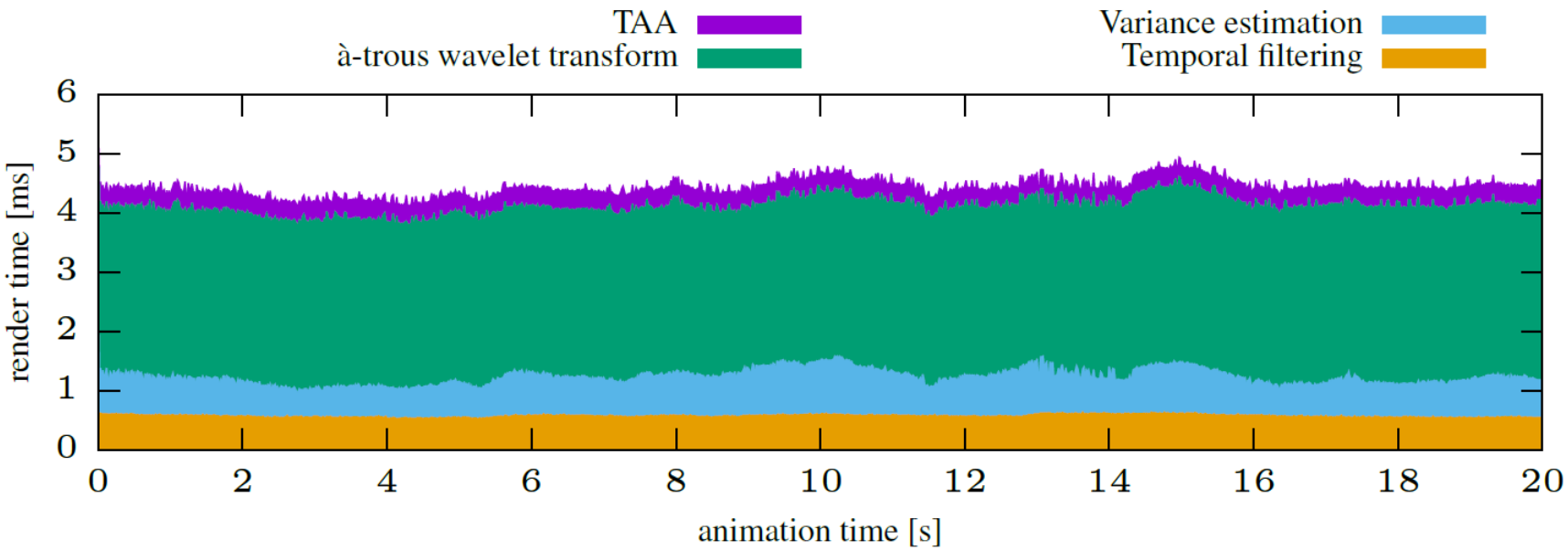
San Miguel, 720p



- Filter is unoptimized
- Performance scales linearly with screen resolution

PERFORMANCE

San Miguel, 720p



LIMITATIONS

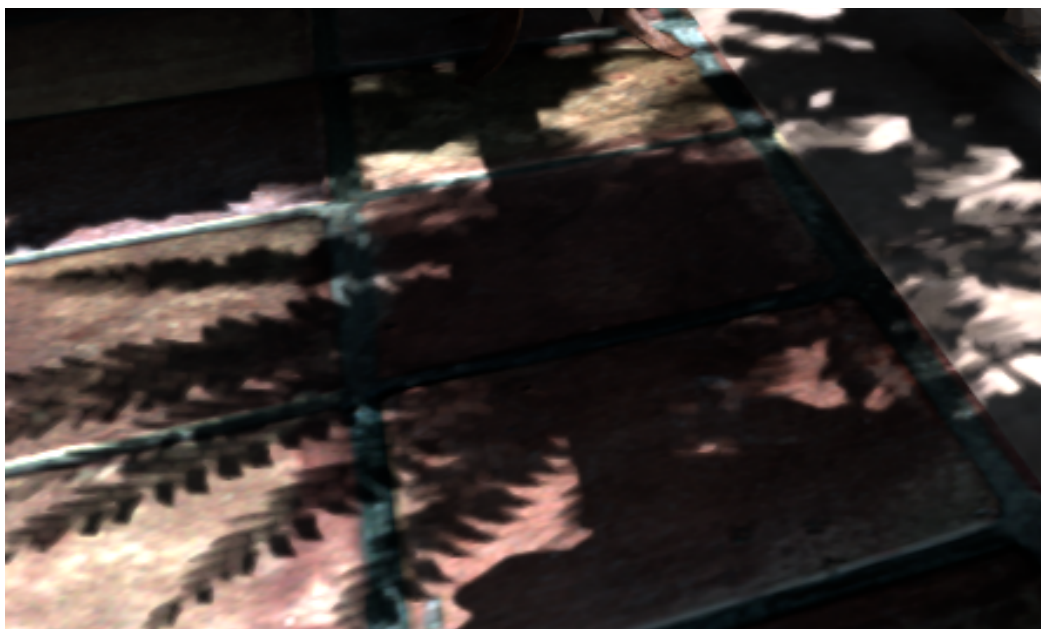
LIMITATIONS

Over-blurred specular reflections

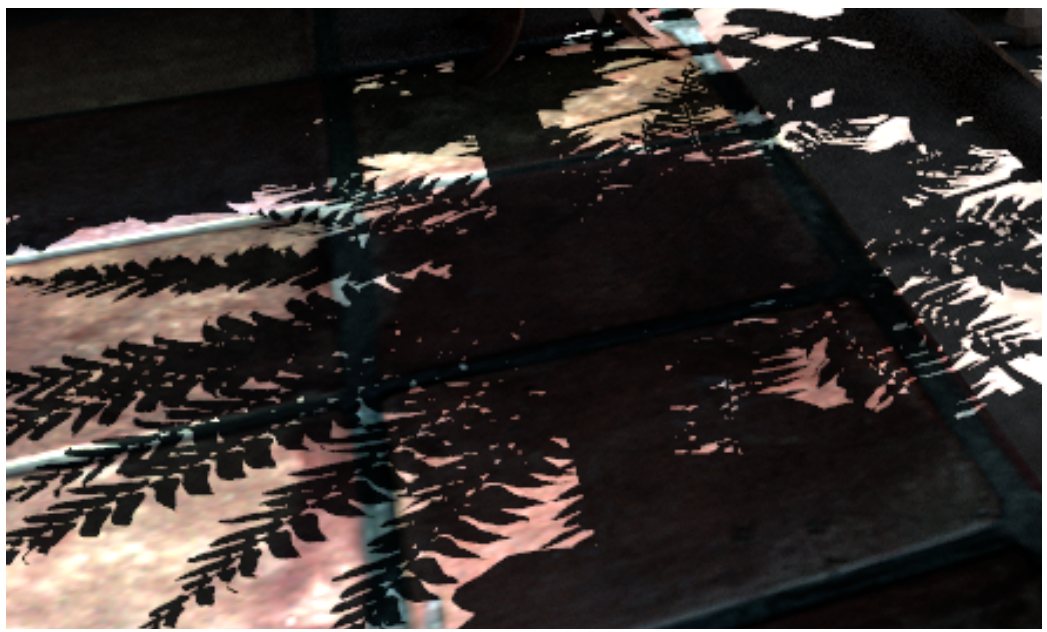
LIMITATIONS

Over-blurring of stable features under motion

ours



reference



LIMITATIONS

Temporal Lag

LIMITATIONS

- Filter does not work with stochastic primary visibility (motion and defocus blur)
  - Temporal accumulation difficult
  - Assumption that surfaces nearby in world-space are close in screen-space breaks
- High-frequency geometry limits applicability of filter
  - Filter cannot find similar surfaces for filtering

CONCLUSION

- Important step towards real-time path tracing
- Better quality than previous work, but:
  - Some temporal instabilities left
  - Temporal lag
- Unoptimized filter implementation close to performance target

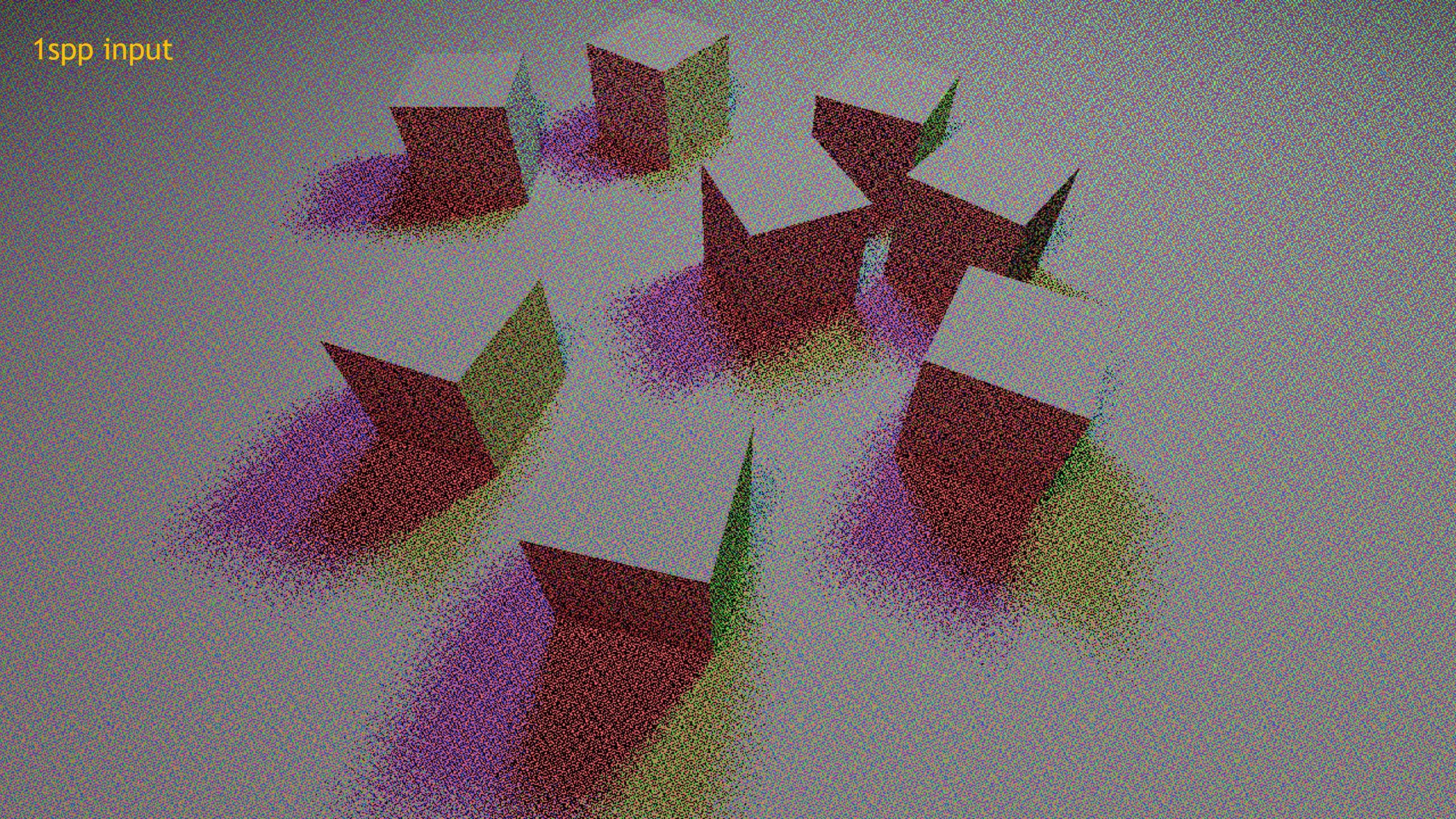
QUESTIONS?

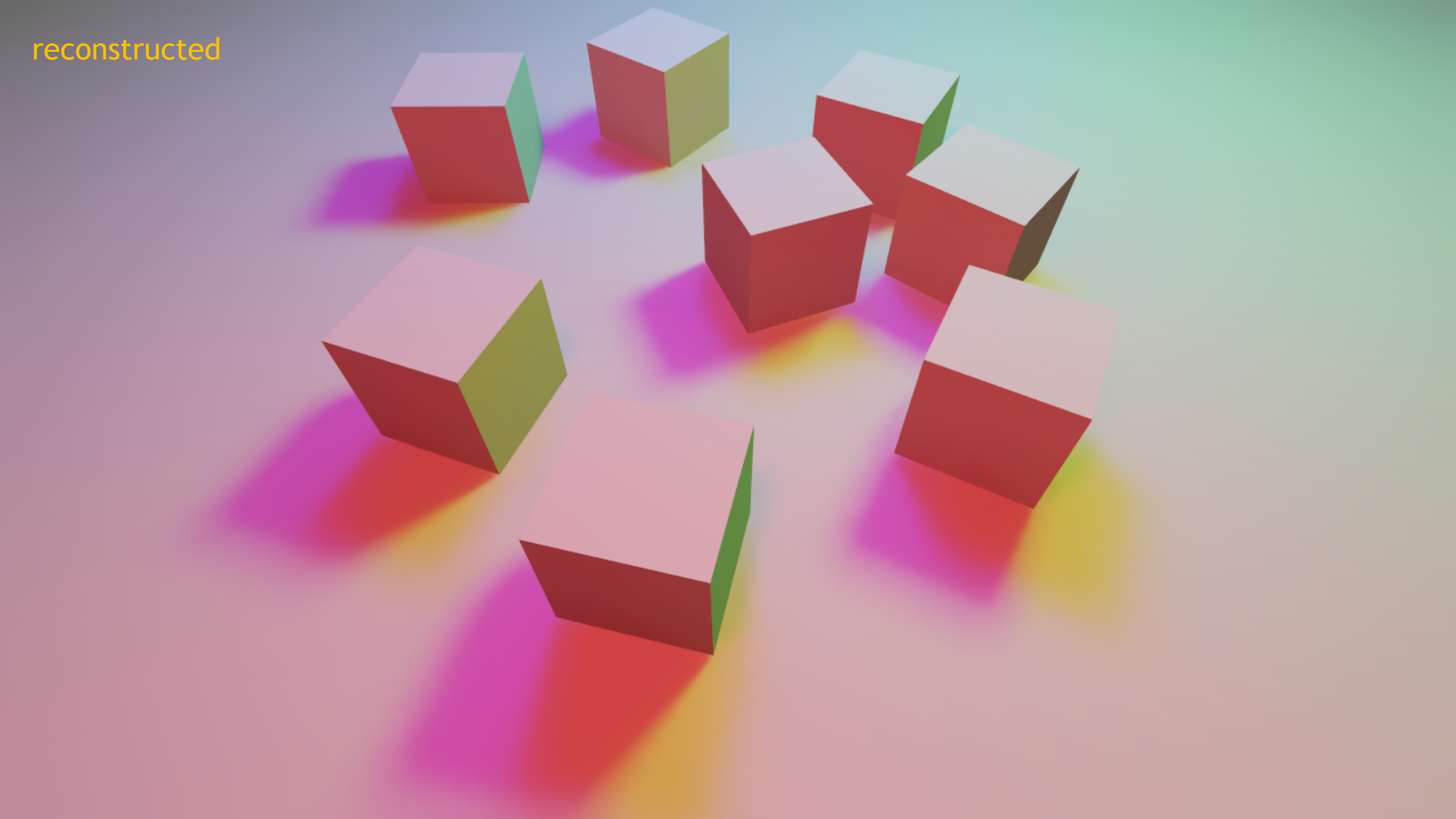
- Thanks to
  - Petrik Clarberg for help in the writing
  - Nir Benty for the Falcor research framework
  - Nicholas Hull for the clean-up of the San Miguel scene


1spp input


reconstructed


reference


1spp input


reconstructed

reference