



**HAL**  
open science

## Spectral Monte Carlo Image Denoising

Mathieu Noizet, Robin Rouphael, Stéphanie Prévost, Hervé Deleau, Luiz Angelo Steffenel, Laurent Lucas

► **To cite this version:**

Mathieu Noizet, Robin Rouphael, Stéphanie Prévost, Hervé Deleau, Luiz Angelo Steffenel, et al.. Spectral Monte Carlo Image Denoising. Journée Calcul et données, Sep 2023, Reims (51), France. hal-04716991

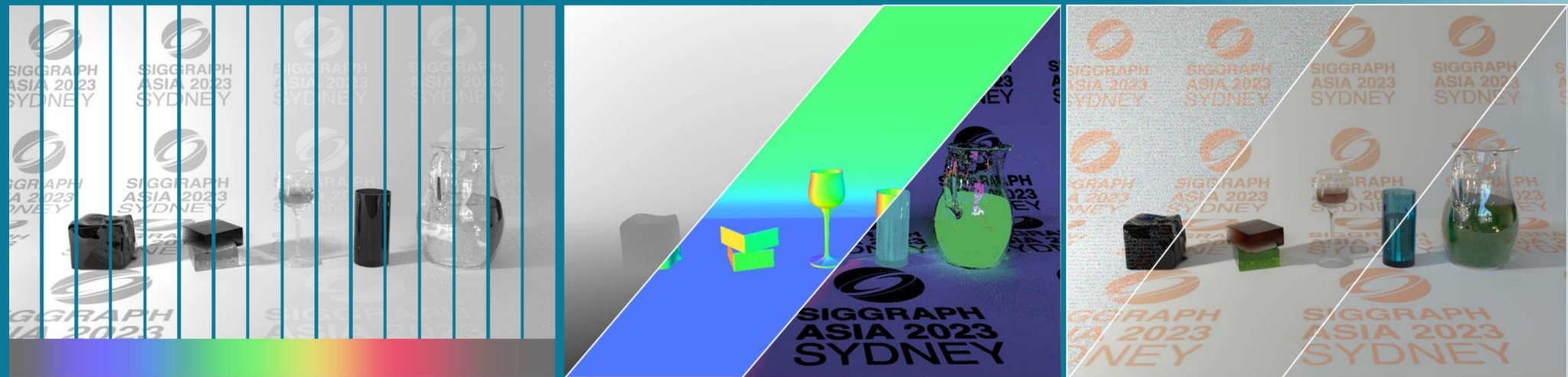
**HAL Id: hal-04716991**

**<https://hal.univ-reims.fr/hal-04716991v1>**

Submitted on 1 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Spectral Monte Carlo Image Denoising

M. Noizet<sup>1</sup>, R. Rouphael<sup>1,2</sup>, S. Prévost<sup>1</sup>, H. Deleau<sup>1</sup>, L. Steffenen<sup>1</sup>  
and L. Lucas<sup>1</sup>

<sup>1</sup>Université de Reims Champagne-Ardenne, LICIIS, France

<sup>2</sup>United Visual Researchers, France

LICIIS  
LRC  
DIGIT

JCAD 2023

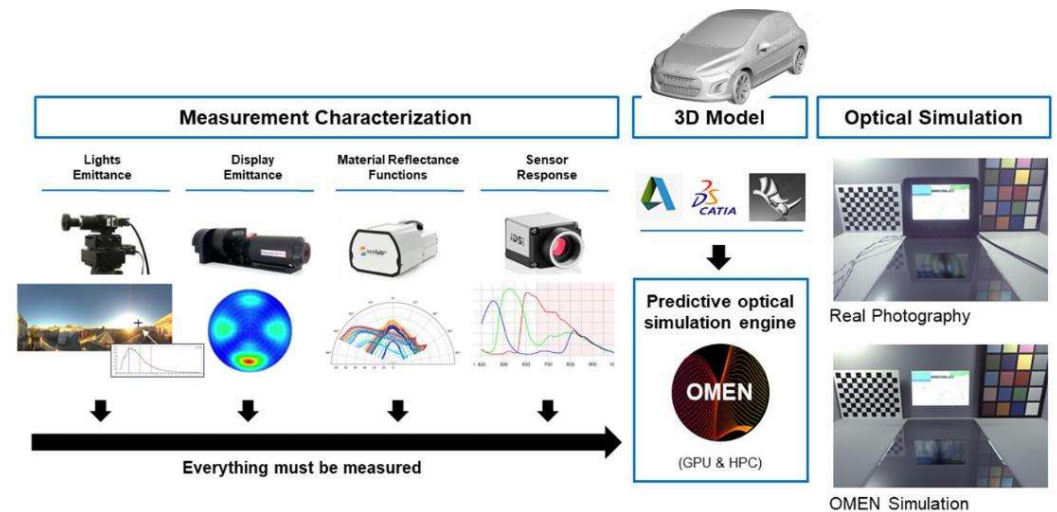
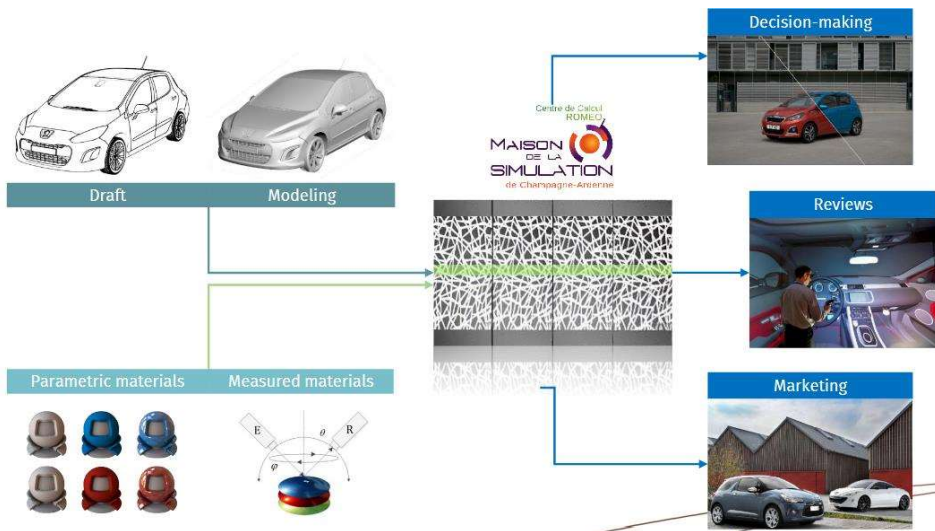


# SUMMARY

- General context
- Previous works
- Contributions and results
- Conclusion

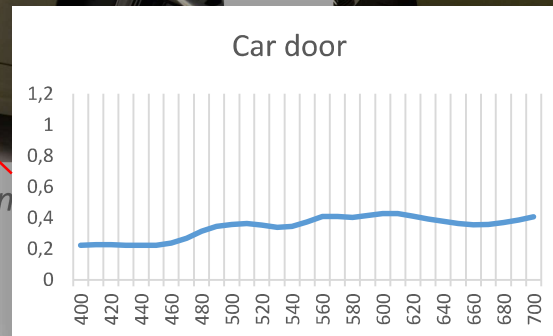
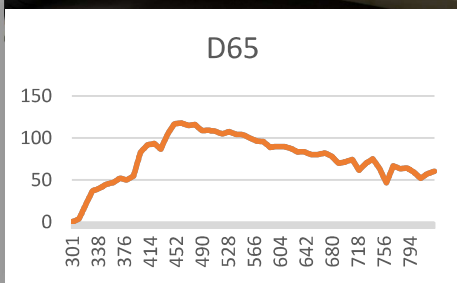
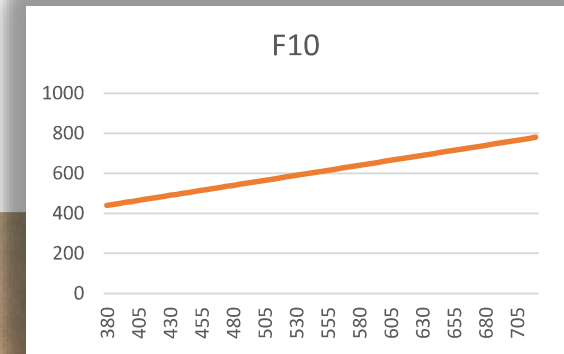
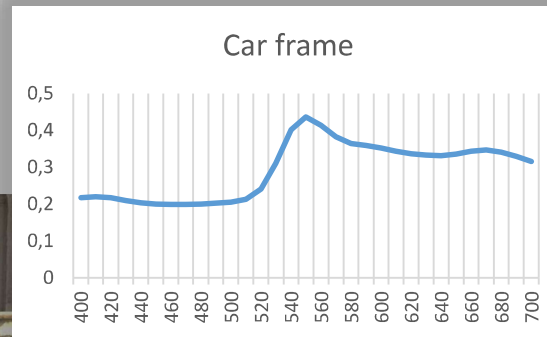
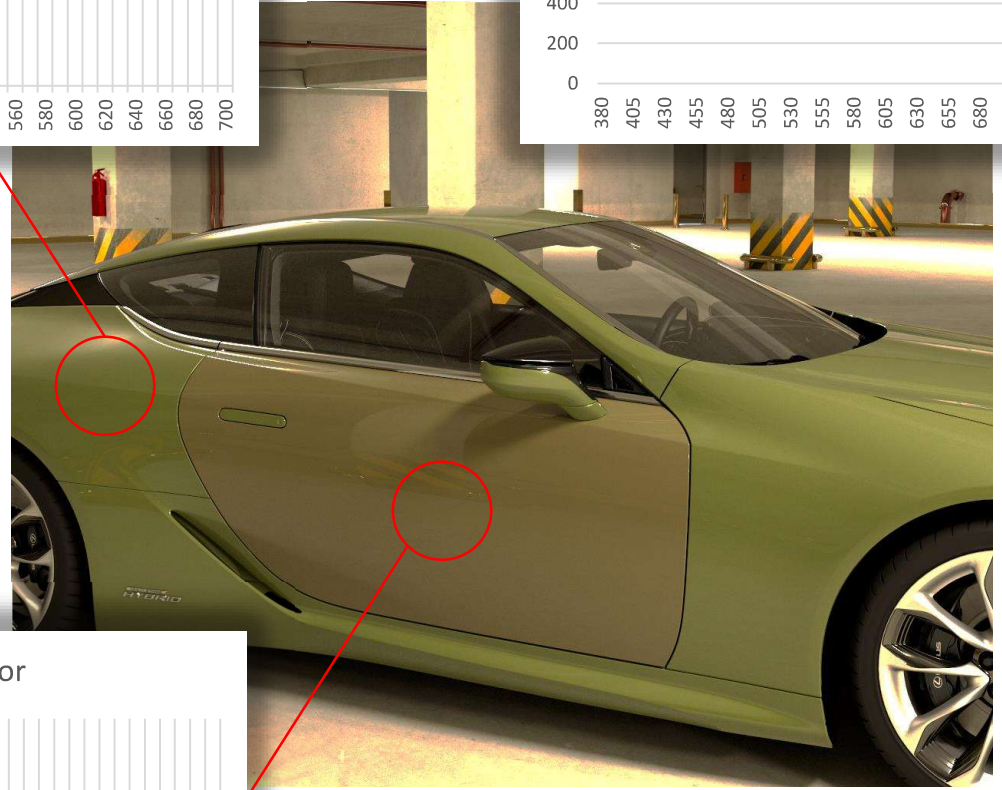
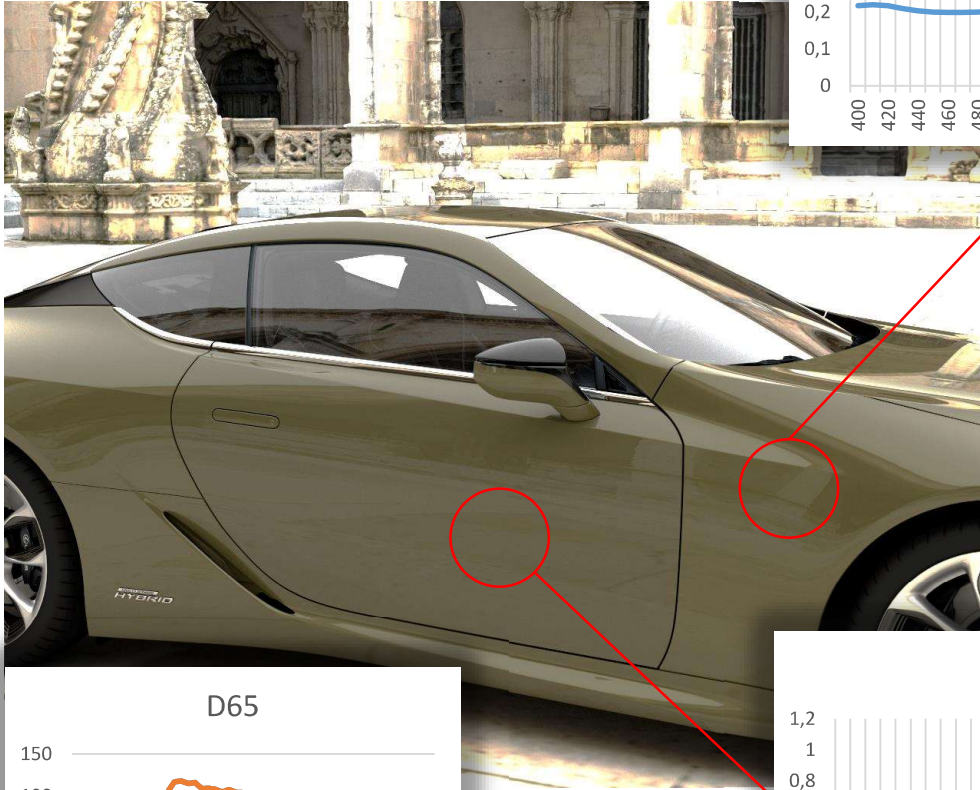
# GENERAL CONTEXT

- Predictive visualisation – in interactive time – of complex materials for industry (CA<sup>2</sup>O)
  - ANR LUCE PRCE 2021-2024
  - Optical simulation with Spectral information
    - Generate predictive image for virtual make-up
    - Time consuming
    - All light phenomena (metarism, polarization, etc.)





# METARISM

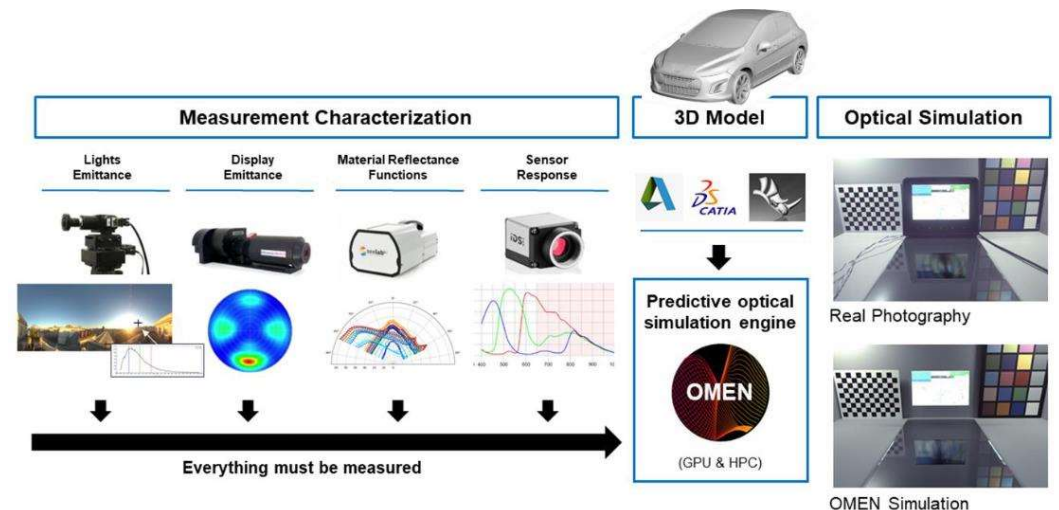
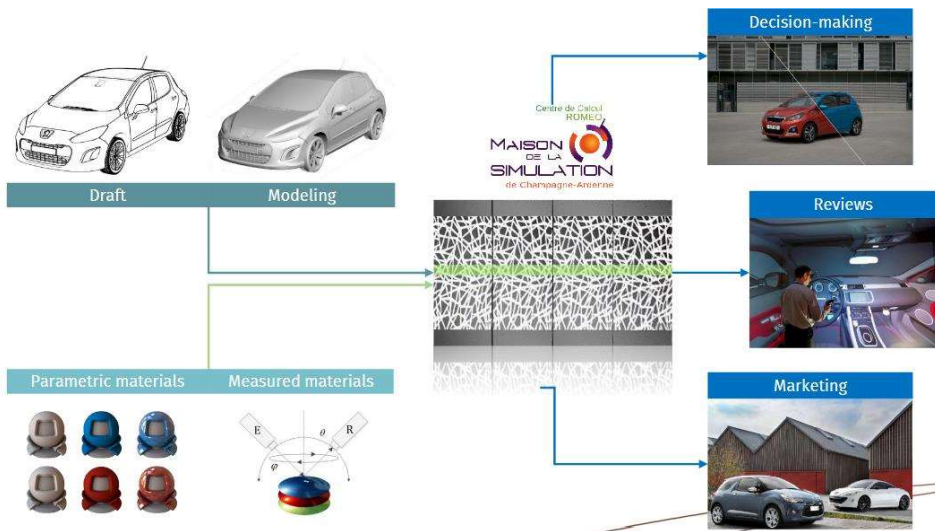


18 octobre 2023

JCAD 2023

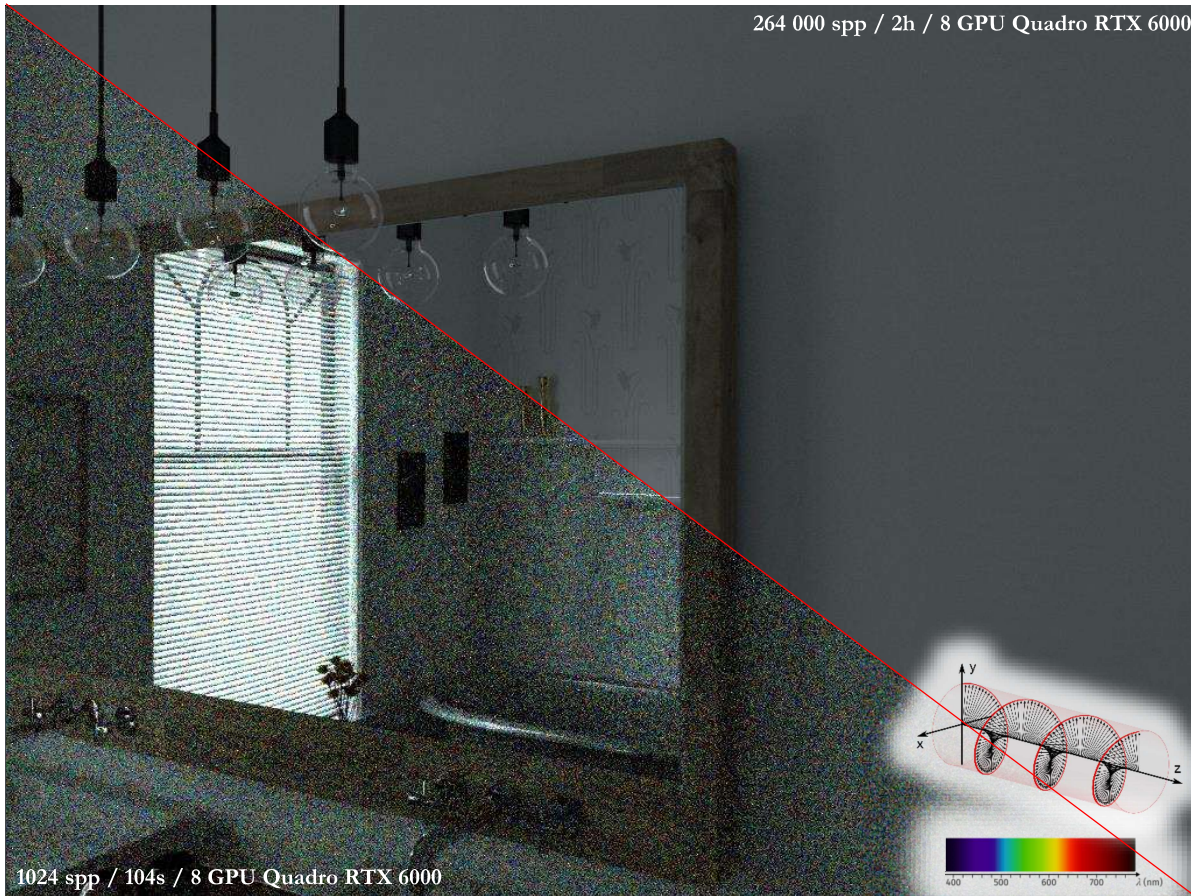
# GENERAL CONTEXT

- Predictive visualisation – in interactive time – of complex materials for industry (CA<sup>2</sup>O)
  - Coupling optical simulation and machine learning
    - How can rendering methods be combined with Deep Learning?
    - How can they be adapted for HPC architecture?





# MONTE CARLO RENDERING



1024 spp / 104s / 8 GPU Quadro RTX 6000

Rendering by UVR Predict Engine

18 octobre 2023

JCAD 2023

- Rendering equation

$$L_o(x, \omega) = L_e(\dots) + \int_{\Omega} f_r \cdot L_i(\dots) \cdot \cos \theta \, d\vec{\omega}_i$$

- Complex analytic resolution

- Recursive
- High dimension

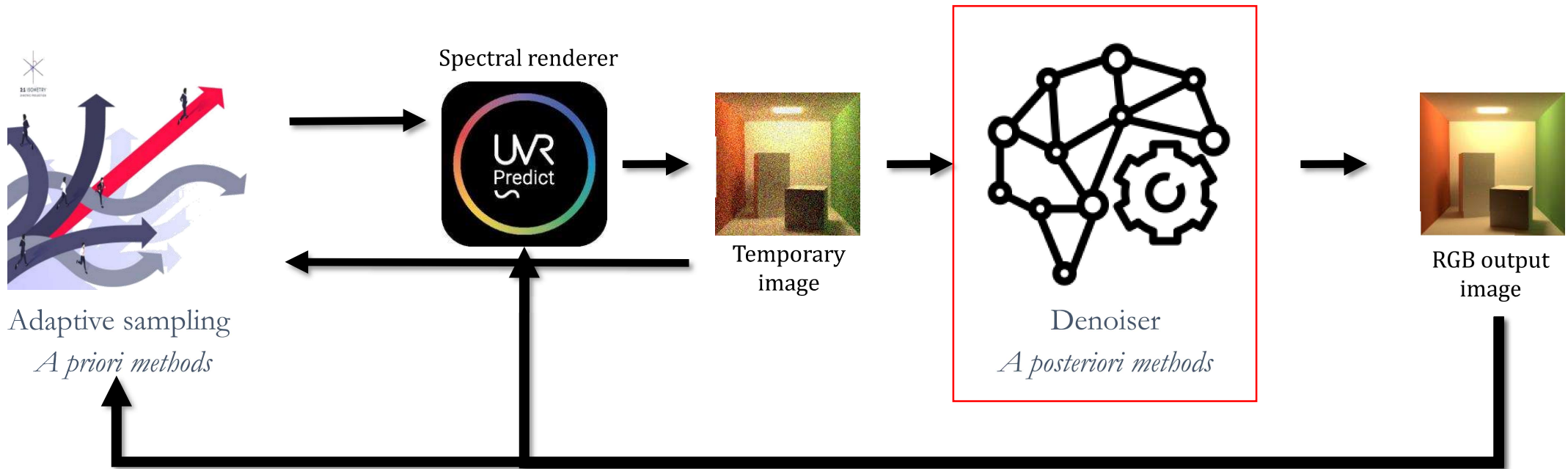
- Resolution based Monte Carlo approach

$$\langle F \rangle = \frac{1}{N} \left[ \frac{f(X)}{p(X)} \right] \approx \int f(x) dx$$

- Approximate the solution from a number of sample

# OUR GOAL

Converting a pipeline





# AVAILABLE DATA FROM SPECTRAL MONTE CARLO RENDERING

Spectral renderer



Noisy spectral input ( $n$  bins)

$$\begin{aligned}
 X &= \int_{\Lambda} X(\lambda)I(\lambda)d\lambda \\
 Y &= \int_{\Lambda} Y(\lambda)I(\lambda)d\lambda \\
 Z &= \int_{\Lambda} Z(\lambda)I(\lambda)d\lambda
 \end{aligned}$$

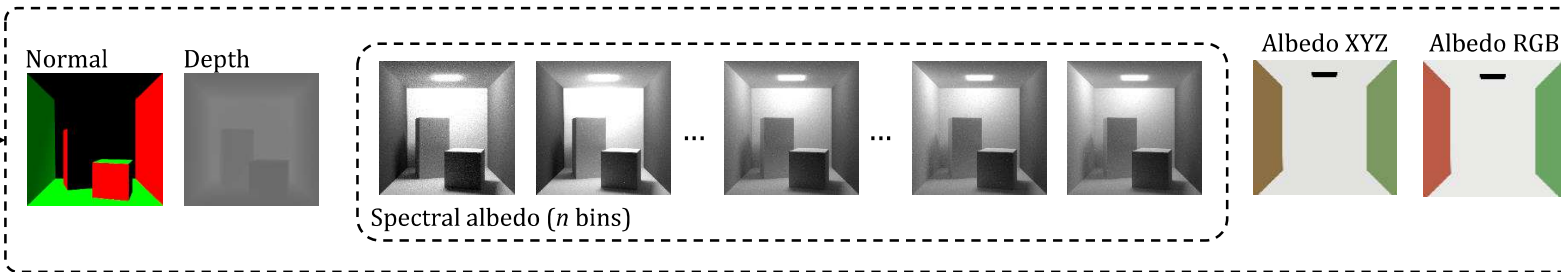
CIE-1931



XYZ output image

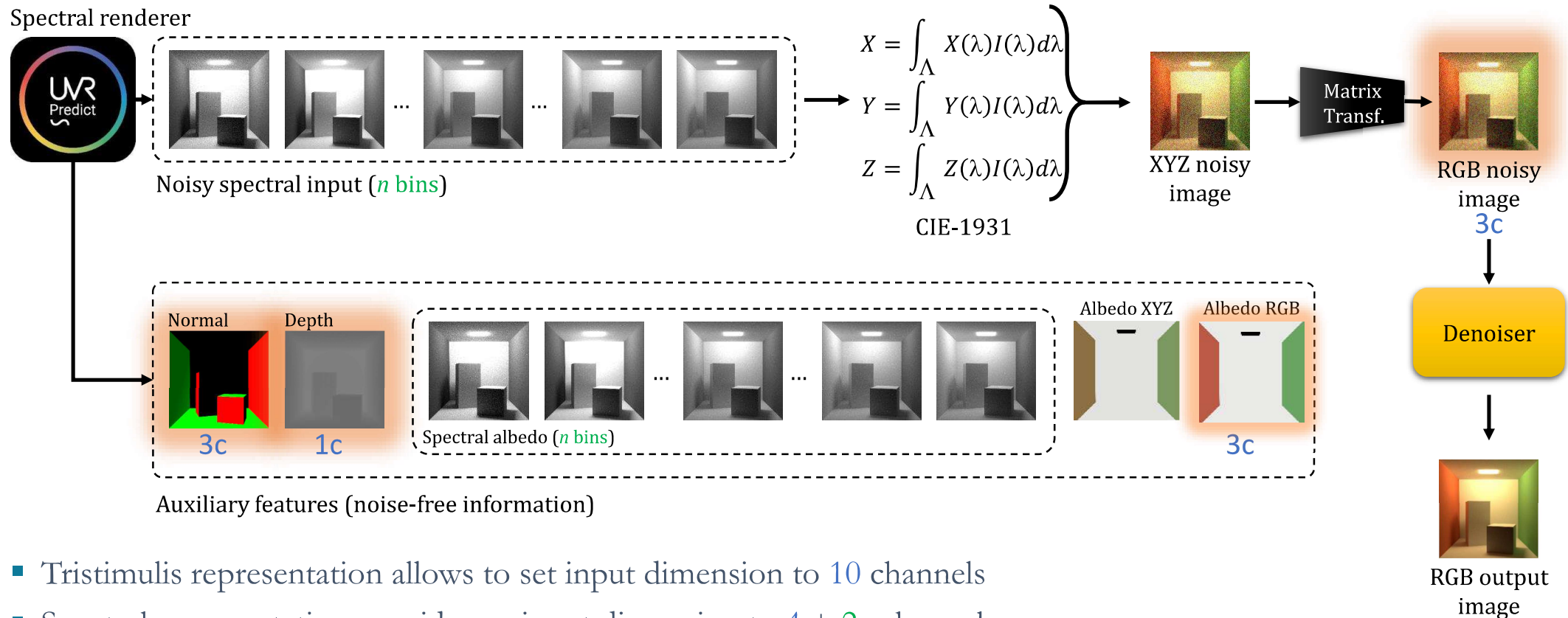


RGB output image



Auxiliary features (noise-free information)

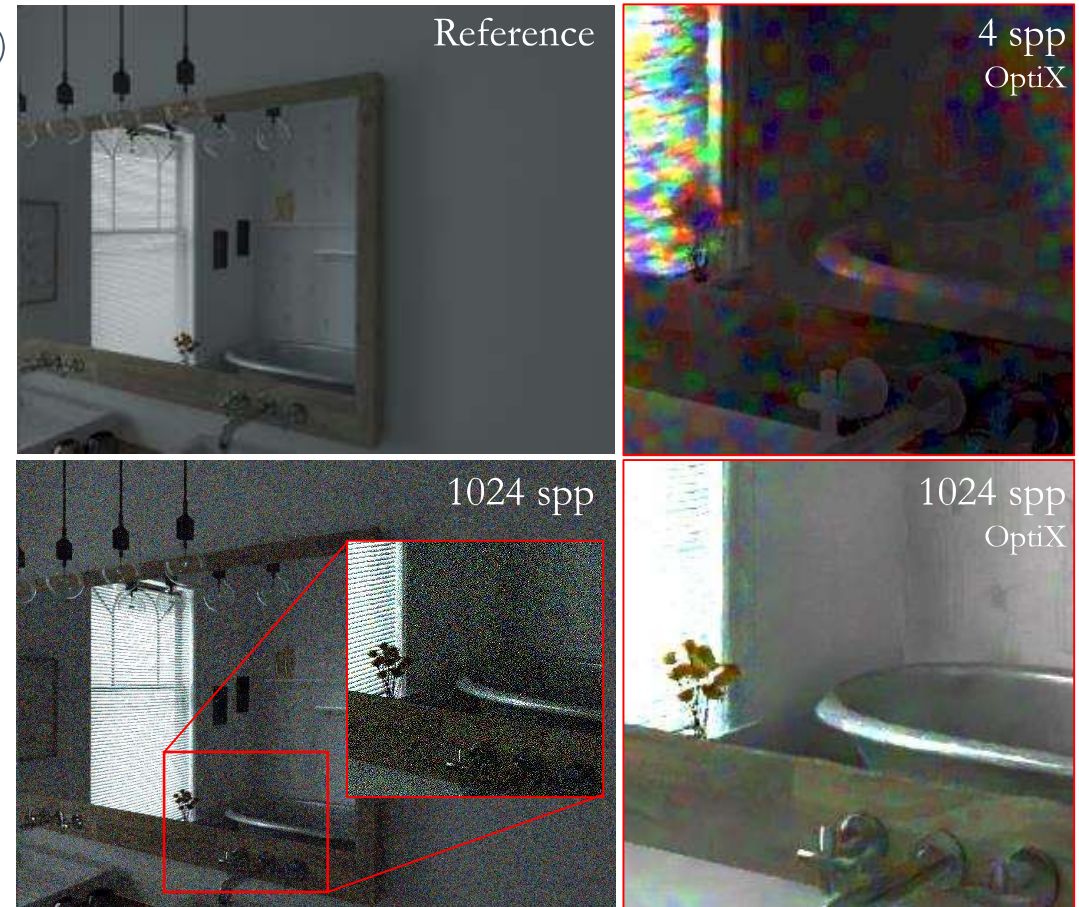
# STATE OF THE ART - DENOISING



- Tristimulus representation allows to set input dimension to 10 channels
- Spectral representation provides an input dimension to  $4 + 2n$  channels

# STATE OF THE ART – OFF-THE-SHELL DENOISER

- Denoise the tristimulus representation (RGB, XYZ...)
  - NVIDIA OptiX denoiser
  - Intel Open Image Denoiser (IOID)
- Apply denoiser with a well sample rate
- Limits
  - Compress all spectrum information into 3 dimensions
  - Bring chromatic aberration
- Questions
  - Has off-the-shell denoiser train on spectral rendering?
  - Do spectral denoisers exist?





# STATE OF THE ART – MONTE CARLO DENOISING

## Interactive Reconstruction of Monte Carlo Image Sequences using a Recurrent Denoising Autoencoder

CHAKRAVARTY R. ALLA CHAITANYA, NVIDIA, University of Montreal and McGill University  
 ANTON S. KAPLANIAN, NVIDIA  
 CHRISTOPH SCHIED, NVIDIA and Karlsruhe Institute of Technology  
 MARCO SALVI, NVIDIA  
 AARON LEFJOHN, NVIDIA  
 DEREK NONOVICZEZHAI, McGill University  
 TIAGO ALLA, NVIDIA

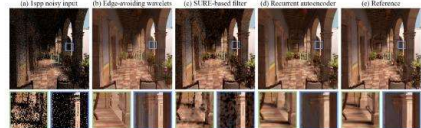


Fig. 1. Left to right (a) noisy image generated using path-traced global illumination with one indirect scene reflection and 1 sample/pixel. (b) edge-weighting weights that [Zhou et al., 2017] use as priors. (c) SI-ME based filter [Li et al., 2021]. (d) Rec. Autoencoder. (e) Reference path-traced image with 4096 samples/pixel.

We describe a machine learning technique for reconstructing image sequences rendered using Monte Carlo methods. Our primary focus is on reconstruction of global illumination with extremely low sampling budgets at interactive rates. Motivated by recent advances in image denoising with deep convolutional networks, we propose a variant of these networks tailored to the characteristics of Monte Carlo image sequences using a Recurrent Denoising Autoencoder. ACM Trans. Comput. Sci., 42(4):19, Oct 2021, 12 pages.

ACM Reference Format:  
 Chaitanya R. Alla Chaitanya, Anton S. Kaplanian, Christoph Schied, Marco Salvi, Aaron Lejohn, Derek Nonoviczezhai, and Tiago Alla, 2021. Interactive Reconstruction of Monte Carlo Image Sequences using a Recurrent Denoising Autoencoder. ACM Trans. Comput. Sci., 42(4):19, Oct 2021, 12 pages.  
 DOI: <https://doi.org/10.1145/3472935.3473011>

1 INTRODUCTION  
 Ray and path tracing have recently emerged as the rendering algorithms of choice for visual effects [Baker et al., 2013]. This has encouraged the development of filtering and denoising techniques designed to reduce the noise inherent in Monte Carlo rendering [Zirker et al., 2015], but the focus on this quality results provides knowledge to thousands of samples per pixel prior to filtering.

Meanwhile, advances have recently emerged towards physically-based shading from more empirical models [Liu et al., 2015], but much of the practical increase in realism from this transition hinges on the possibility of sampling light transport paths more flexibly than rasterization allows. Unfortunately, even the fastest ray tracers can only trace a few rays per pixel at 1080p and 60Hz. While this number doubles every few years, the trend is (at least partially)

ACM Trans. Comput. Sci., Vol. 42, No. 4, Article . Publication date: July 2021.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted by ACM, provided that the copies are not made for commercial distribution, are not sold, and are not used for advertising or promotional purposes. This work is copyright © 2021 by ACM. All rights reserved. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

## Kernel-Predicting Convolutional Networks for Denoising Monte Carlo Renderings

STEVE BAKO<sup>1</sup>, University of California, Santa Barbara  
 THIAS VOGELS<sup>2</sup>, ETH Zurich & Disney Research  
 BRIAN MCWILLIAMS<sup>3</sup>, Disney Research  
 MARK MEYER<sup>4</sup>, Pixar Animation Studios  
 JAN NOVÁK<sup>5</sup>, Disney Research  
 ALEX HARVILL<sup>6</sup>, Pixar Animation Studios  
 PRADEEP SEN<sup>7</sup>, University of California, Santa Barbara  
 TONY DEUSE<sup>8</sup>, Pixar Animation Studios  
 FABRICE ROUSSELLE<sup>9</sup>, Disney Research



Fig. 1. We introduce a deep learning approach for denoising Monte Carlo-rendered images that produces high-quality results suitable for production. We train a convolutional neural network to learn the complex relationship between noisy and reference data across a large set of frames with varying distributed effects from the film *Toy Story 4* (see left). The trained network can then be applied to denoise new images from other films with significantly different style and content, such as *Cars 3* (right), with production-quality results.

Regression-based algorithms have shown to be good at denoising Monte Carlo (MC) renderings by leveraging the invariance by gradient of the feature buffers. However, while state-of-the-art models in handle complex scenes, there is always a trade-off between the quality of the denoising and the ability to denoise. To address this problem, we propose a novel, supervised learning approach that allows for denoising based on more complex and generalizable features. We propose a deep convolutional neural network (CNN) denoising framework. In our convolutional net framework, the CNN directly predicts the final denoised pixel value as a highly non-linear combination of the input features. In a second approach, we introduce a novel, lower-level prediction network which uses the CNN to estimate the local weighting kernel used to compute each denoised pixel from its neighbors. We train and evaluate our system on two scenes.

ACM Reference Format:  
 Steve Bako, Thias Vogels, Brian McWilliams, Mark Meyer, Jan Novák, Alex Harvill, Pradeep Sen, Tony Deuse, and Fabrice Roussele, 2021. Kernel-Predicting Convolutional Networks for Denoising Monte Carlo Renderings. ACM Trans. Comp. Sci., 42(4):19, Oct 2021, 12 pages.  
 DOI: <https://doi.org/10.1145/3472935.3473018>

1 INTRODUCTION  
 In recent years, physically-based image synthesis has become widespread in feature animation and visual effects [Baker et al., 2015].

ACM Trans. Comput. Sci., Vol. 42, No. 4, Article . Publication date: July 2021.

[Chaitanya et al. 2017]

[Bako et al. 2017]

[Yang et al. 2019]

[Back et al. 2020]

Direct prediction  
 ~1-4 spp\*  
 Loss details

Kernel prediction  
 ~64-128 spp\*  
 Details reconstructed

Direct prediction  
 ~1-4 spp\*  
 Reduce artifact induce by auxiliary features  
 Less of diffuse details lost

Kernel prediction  
 independent of spp\*  
 Improve diffuse details reconstruction

spp\* = 1 sample in RGB

18 octobre 2023

JCAD 2023

11

Yang X, Wang D, He W et al. DEEM: A deep dual-encoder network for denoising Monte Carlo rendering. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 34(5): 1123-1135, Sept. 2019. DOI: [10.1007/s11390-019-1064-2](https://doi.org/10.1007/s11390-019-1064-2)

## DEMC: A Deep Dual-Encoder Network for Denoising Monte Carlo Rendering

Xin Yang<sup>1</sup>, Dawei Wang<sup>2</sup>, Wenhao He<sup>3</sup>, Li-Jing Zhao<sup>4</sup>, Bao-Cai Yin<sup>1</sup>, Qiang Zhang<sup>1</sup>, Xiao-Peng Wu<sup>5</sup> and Hengbo Fu<sup>6</sup>

<sup>1</sup>Department of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China  
<sup>2</sup>Department of Computer Science, The University of Hong Kong, Pokfulam, Hong Kong, China  
<sup>3</sup>Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong, China  
<sup>4</sup>School of Creative Media, City University of Hong Kong, Kowloon, Hong Kong, China

E-mail: xiyang@outlook.com, dawei@hkust.hk, wenhao@cs.cuhk.hk, 7859714@mail.dlut.edu.cn  
 [f.fu, zhang, wu]@hkust.hk, fuhs@cs.cuhk.hk, fuhs@cityu.edu.hk

Received January 15, 2019; revised May 28, 2019.

**Abstract** In this paper, we present DEMC, a deep dual-encoder network to remove Monte Carlo noise efficiently while preserving details. Denoising Monte Carlo rendering is different from natural image denoising since uncorrelated pixel features (feature buffers) can be extracted in the rendering stage. Most of them are noise-free and can provide sufficient details for image reconstruction. However, these feature buffers also contain redundant information. Hence, the main challenge of this topic is how to extract useful information and reconstruct clean images. To address this problem, we propose a novel network structure, dual-encoder network with a feature buffer sub-network, to fuse feature buffers flexibly, then encode the fused feature buffers using a noisy image simultaneously, and finally reconstruct a clean image by a decoder network. Compared with the state-of-the-art methods, our model is more robust on a wide range of scenes, and is able to generate satisfactory results in a significantly faster way.

**Keywords** Monte Carlo rendering, Monte Carlo denoising, neural network

### 1 Introduction

Producing a photorealistic image from 3D models needs complex computations at every pixel of the image. For example, a ray tracing algorithm requires complex complex integral, over all the ray paths between light sources and every point on image sensors. Monte Carlo (MC) ray tracing<sup>1</sup> introduces a method to approximate this complex integral by tracing light paths in a multi-dimensional space, in order to obtain an estimated value of the integral expression. Although Monte Carlo rendering has been widely accepted by many movie production studios, it suffers from noise pollution, which can only be mitigated by increasing

the number of samples exponentially, making the synthesis of a noise-free and photo-realistic image very time-consuming. However, some industry applications, such as real-time game rendering, virtual augmented reality, require rendering high-quality images in a faster way. Recently, a variety of methods<sup>2,3,4</sup> for accelerating Monte Carlo rendering have been proposed. The core idea of these methods is to render a noisy image with a few samples per pixel (SPP) flexibly, and then use denoising algorithms to reconstruct a perceptually indistinguishable image from the noisy image and auxiliary feature buffers. Here, the auxiliary feature buffers are inexpensive by-products generated in the rendering stage, which contain geometry and texture information.

Regular Paper  
 Received by JCVI 2019.  
 This work was supported in part by the National Natural Science Foundation of China under Grant No. 91781404, U1911403, 61672166, 61472166, and 61771201. The National Key Research and Development Program of China under Grant No. 2016YFB0301006, the Open Project Program of the State Key Laboratory of CAD&GP of Zhejiang University of China under Grant No. A1901, and the Open Research Fund of Beijing Key Laboratory of Big Data Technology for Public Safety Project under Grant No. FJ2019-0101P.  
 \*Corresponding Author.  
 ©2019 Springer Science + Business Media, LLC & Science Press, China.

## Deep Combiner for Independent and Correlated Pixel Estimates

JONGHEE BACK, Gwangju Institute of Science and Technology, South Korea  
 BINH-SUN HUA, VnA Research, Vietnam and Vrije University, Vietnam  
 TOSHIYA HACHISUKA, The University of Tokyo, Japan  
 ROCHANG MOON, Georgia Institute of Science and Technology, South Korea

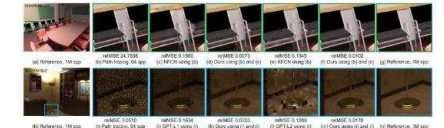


Fig. 1. Our framework allows us to combine two different types of images, independent pixel estimates (i.e., path-traced images) and correlated pixel estimates (i.e., denoised images), and reduce remaining errors (red marks) or systematic errors of the existing methods such as Nonlocal-weighted Fast Color Regression (NWR) [Baker et al., 2015], Kernel-Predicting Convolutional Networks (KPCN) [Bako et al., 2017], and Gradient-domain Path Tracing with L1 and L2 reconstruction (GPL) and GPL2 [Battar et al., 2015]. The numbers are the relative mean square error (RMSE) [Battar et al., 2015].

ACM Reference Format:  
 Jonghee Back, Binh-Sun Hua, Toshiya Hachisuka and Rochang Moon, 2020. Deep Combiner for Independent and Correlated Pixel Estimates. ACM Trans. Comput. Sci., 42(4):19, October 2020, 12 pages. <https://doi.org/10.1145/3472935.3473017>

1 INTRODUCTION  
 Monte Carlo (MC) rendering [Kajiya 1986] has been recognized as a powerful tool for light transport simulation, which has been widely adopted in production rendering recently [Pfeiffer 2018]. MC rendering can simulate a variety of lighting effects by randomly sampling light paths and averaging their contributions at every pixel. Pixel estimates in MC rendering are typically independent of each other. The main problem is that to get noisy pixel estimates, which also stems from its random nature. In general, a large number of samples (often a considerable amount of computation time) are needed to reduce such noise to an unacceptable level.

A popular class of noise reduction methods in MC rendering is image-space denoising [Oosterlinck et al., 2009; Sun and Durkin 2012]. Its main advantage is that it can handle different types of random noise generated by complex lighting effects without suffering from the complexity of light transport. Learning-based denoising [Bako et al., 2017; Chaitanya et al., 2017; Ghahramani et al., 2019; Xu et al., 2019] has achieved an impressive level of noise reduction recently. The denoising process typically introduces correlation among pixels.

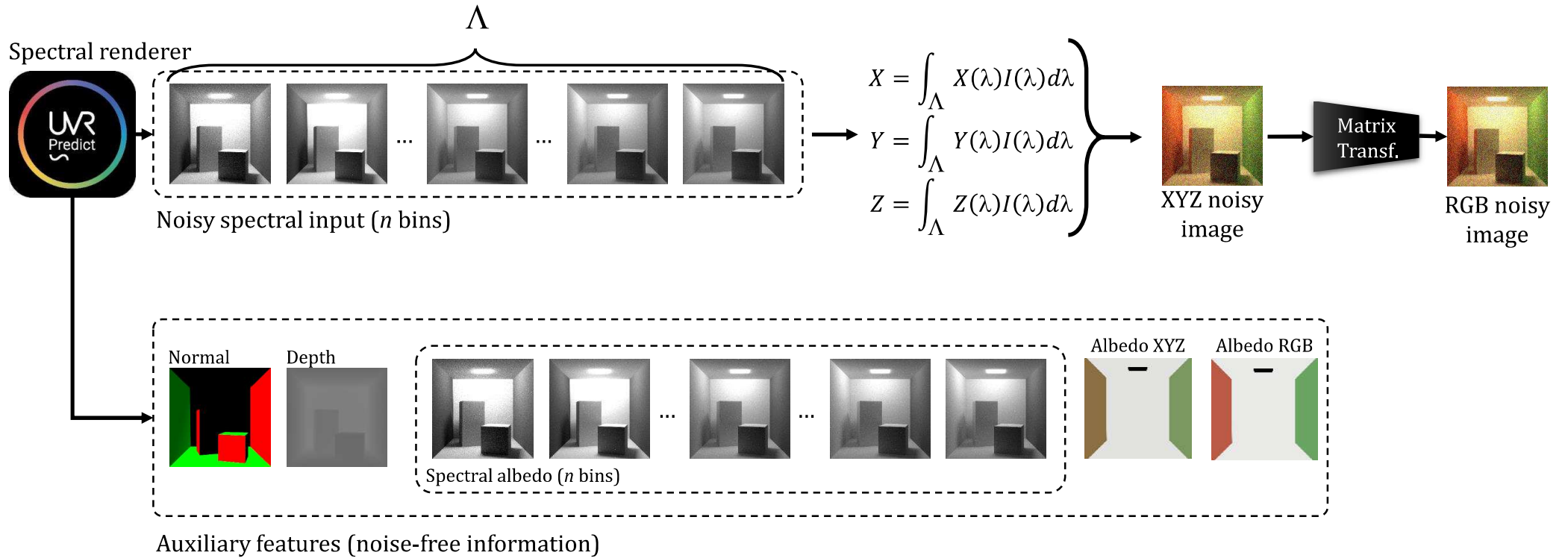
ACM Trans. Comput. Sci., Vol. 42, No. 4, Article . Publication date: December 2020.

ACM Reference Format:  
 Rochang Moon, Jonghee Back, Gwangju Institute of Science and Technology, South Korea, Binh-Sun Hua, VnA Research, Vietnam and Vrije University, Vietnam, Toshiya Hachisuka, The University of Tokyo, Japan, Rochang Moon, Georgia Institute of Science and Technology, South Korea.

© 2020 Copyright held by the owner/author(s). Publication rights reserved to ACM.  
 This is the author's version of the work as it appeared in the journal, with the final version of the work as it appeared in the journal. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

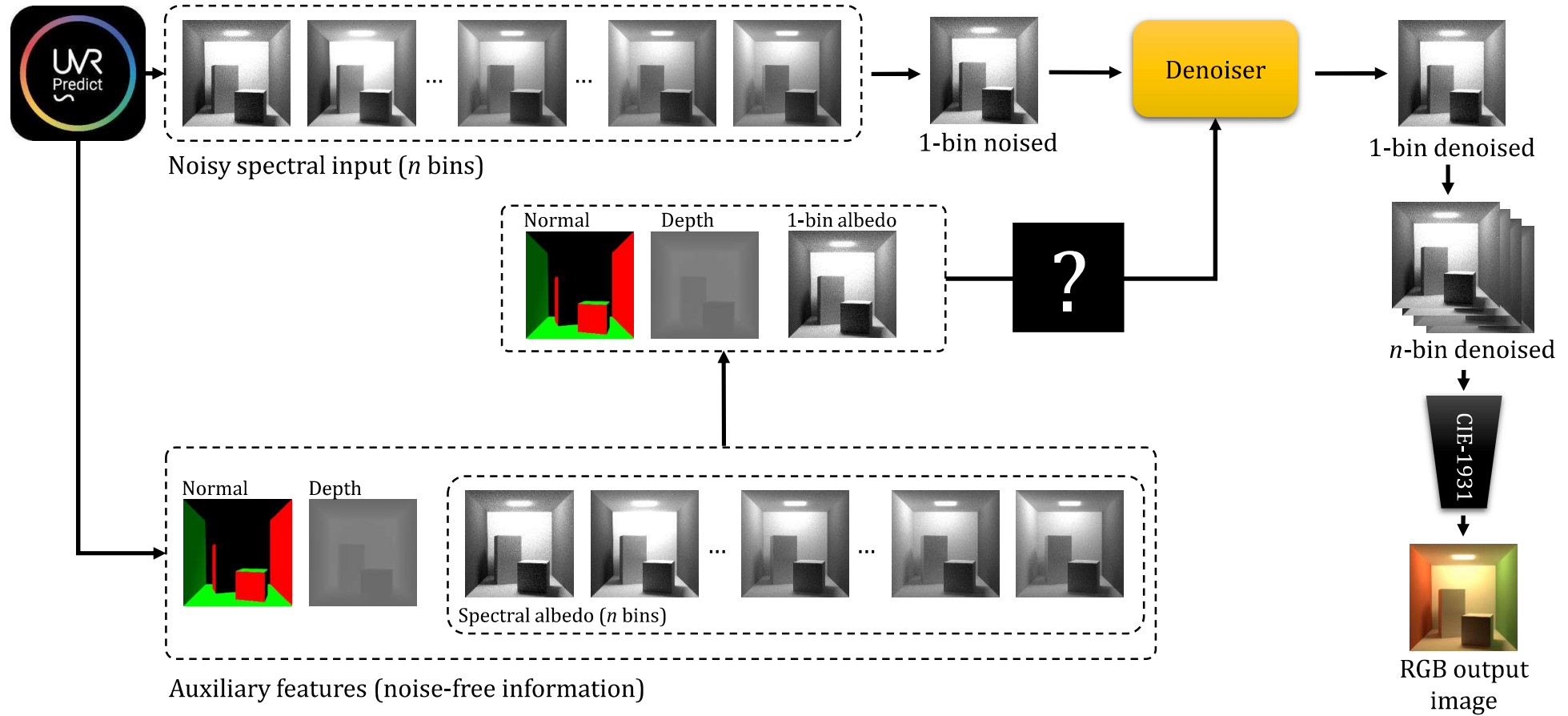


# DATA MANAGING



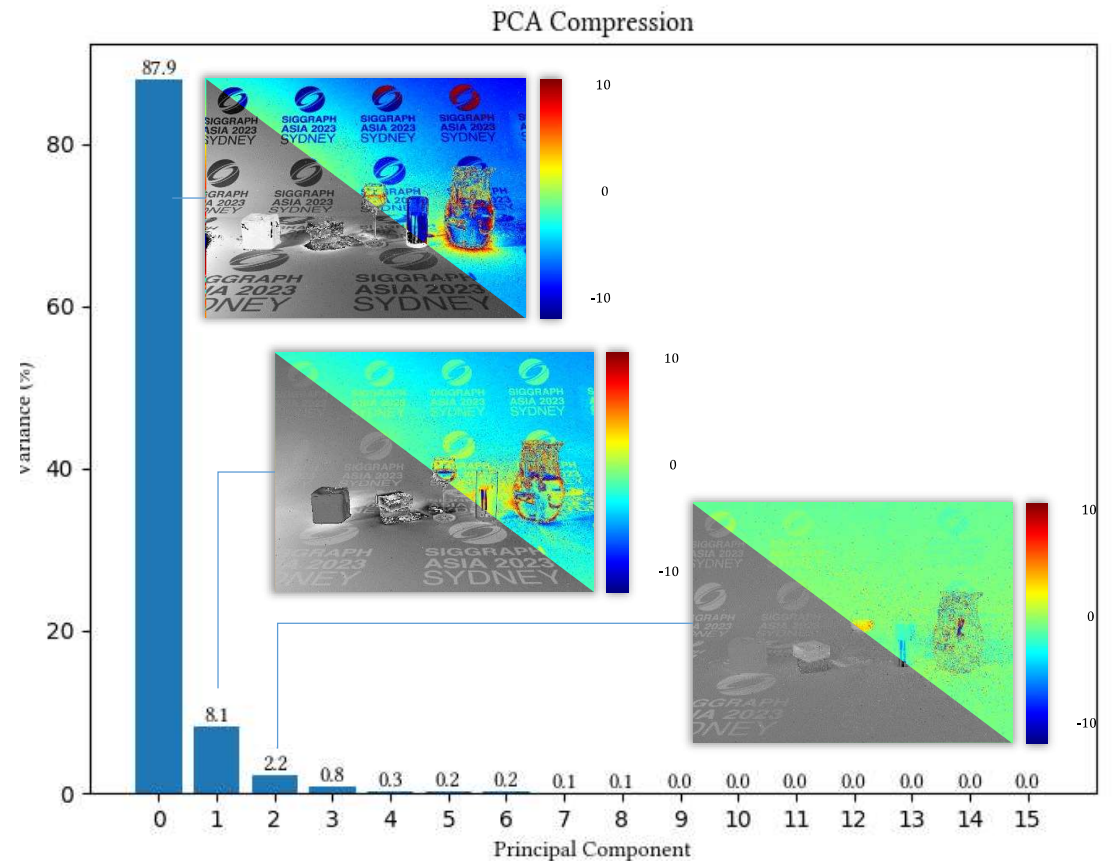
# DATA MANAGING

Spectral renderer

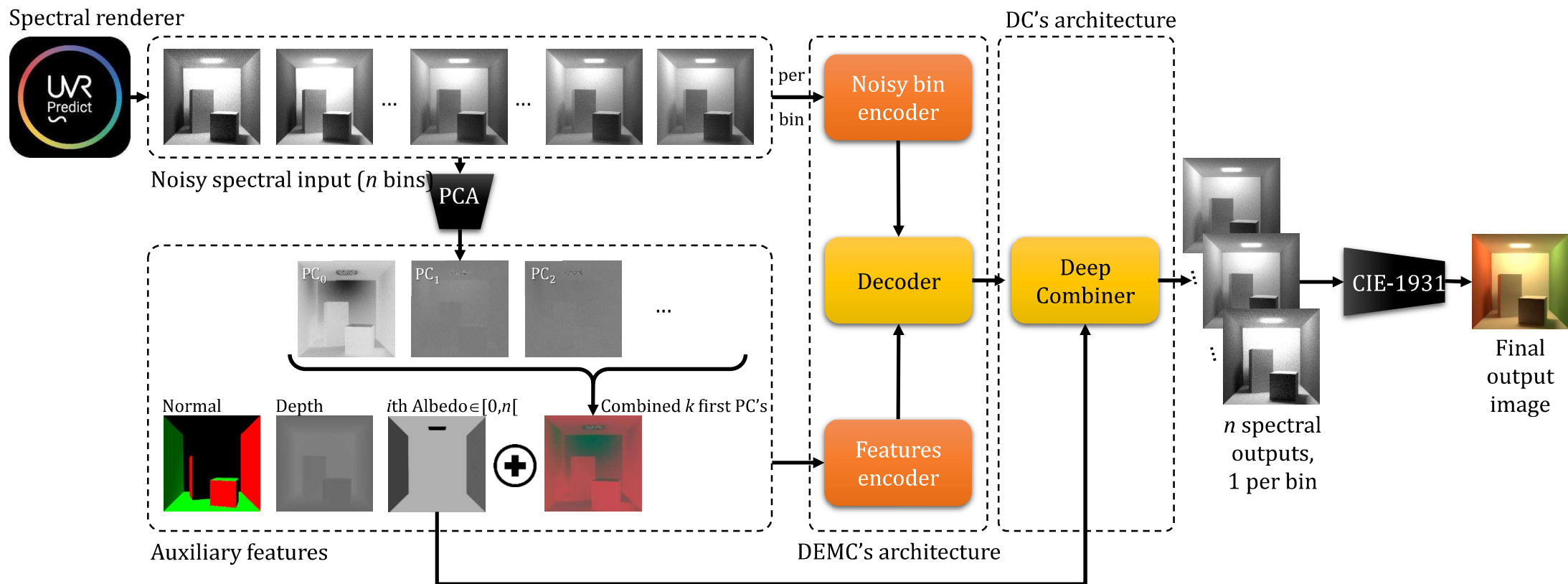


# SPECTRAL COMPRESSION

- Aim to provide full light spectrum information
  - With fixed set input dimension
- Tri-chromatic representation
  - Set to 3 dimensions
  - Provide a displayable information
  - Change the nature of data representation
- PCA
  - No truncate spectral information
  - Compress without data lost
  - The 3 first PC represent 98.2% of initial information



# OUR CONTRIBUTION



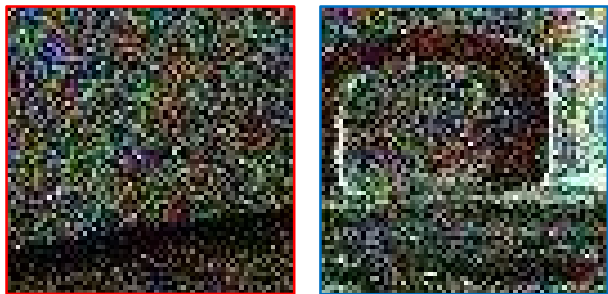


# TRAINING INFORMATION

- Loss function: SMAPE
- Dataset
  - 5500 image peers (4554 for training, 946 for testing)
  - 23 scenes (22 points of views)
  - Resolution of image's crop 128×128
- Training parameters (for each network)
  - Epochs : 5000
  - Learning rate :  $10^{-4}$
  - Optimizer : Adam
- Training time: ~3 days
- 4 GPU Nvidia Tesla P100 (16 Go VRAM)

# RESULTS

Noisy image



Reference image



Denoised image with our method

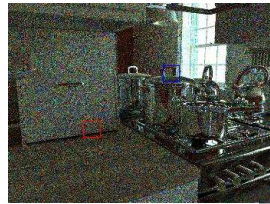


SPP: 1024  
Time: 89 s on 1 GPU  
RelMSE: 0.70

1 M  
≈ 36 h on 4 GPUs  
GT

1024  
89.16 s on 1 GPU  
0.0313

# RESULTS

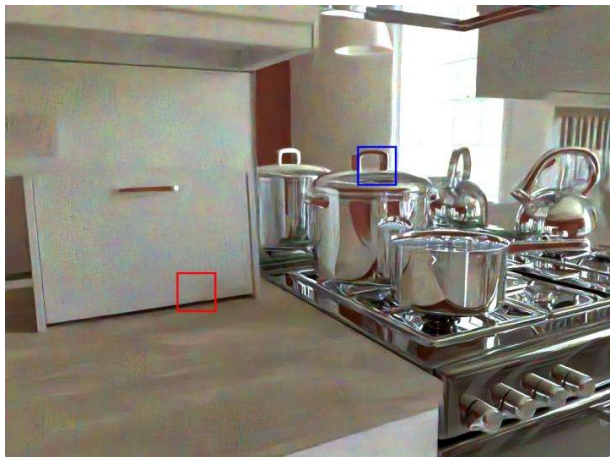


Noisy image  
1024 SPP



Reference image  
1M SPP

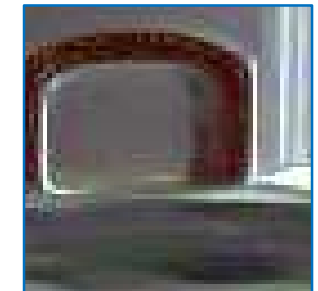
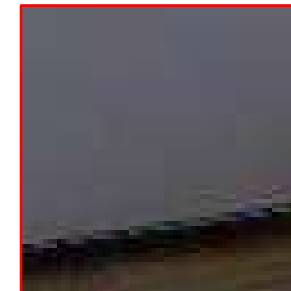
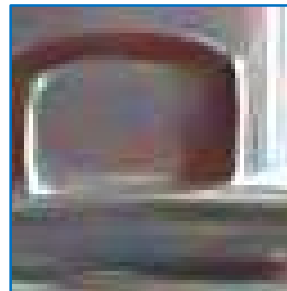
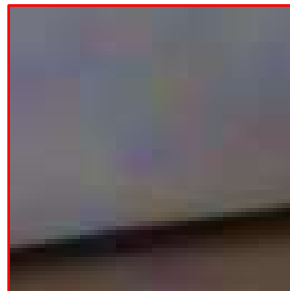
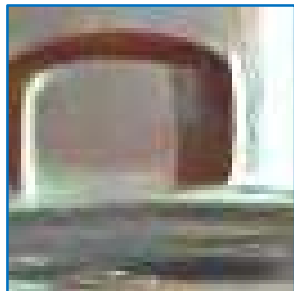
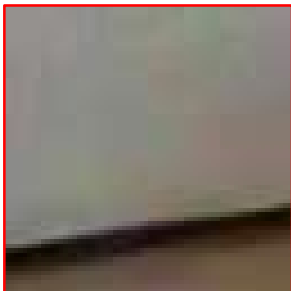
Intel



OptiX



Ours



ReIMSE: 0.176

0,205

0.0313

# CONCLUSION & FUTURE WORKS

- Contributions
  - First spectral denoiser
    - based on the spectral bins processing
    - Tailoring input, auxiliary and output features to favorize spectral information
  - Out-perform off-the-shell denoiser (with RelMSE measure)
  - Submitted to Eurographics 2024
- Future works
  - Improve border reconstruction
  - Improve albedo computation to reduce artifacts



Thanks for your attention!

LICHS  
LRC  
DIGIT

JCAD 2023

