Deep Precomputed Radiance Transfer for Deformable Objects

Yue Li
University of Pensylvania??
yueli.cg@gmail.com

Pablo Wiedemann
Edinburgh Napier University
p.wiedemann@napier.ac.uk

Kenny Mitchell
Edinburgh Napier University
k.mitchell2@napier.ac.uk

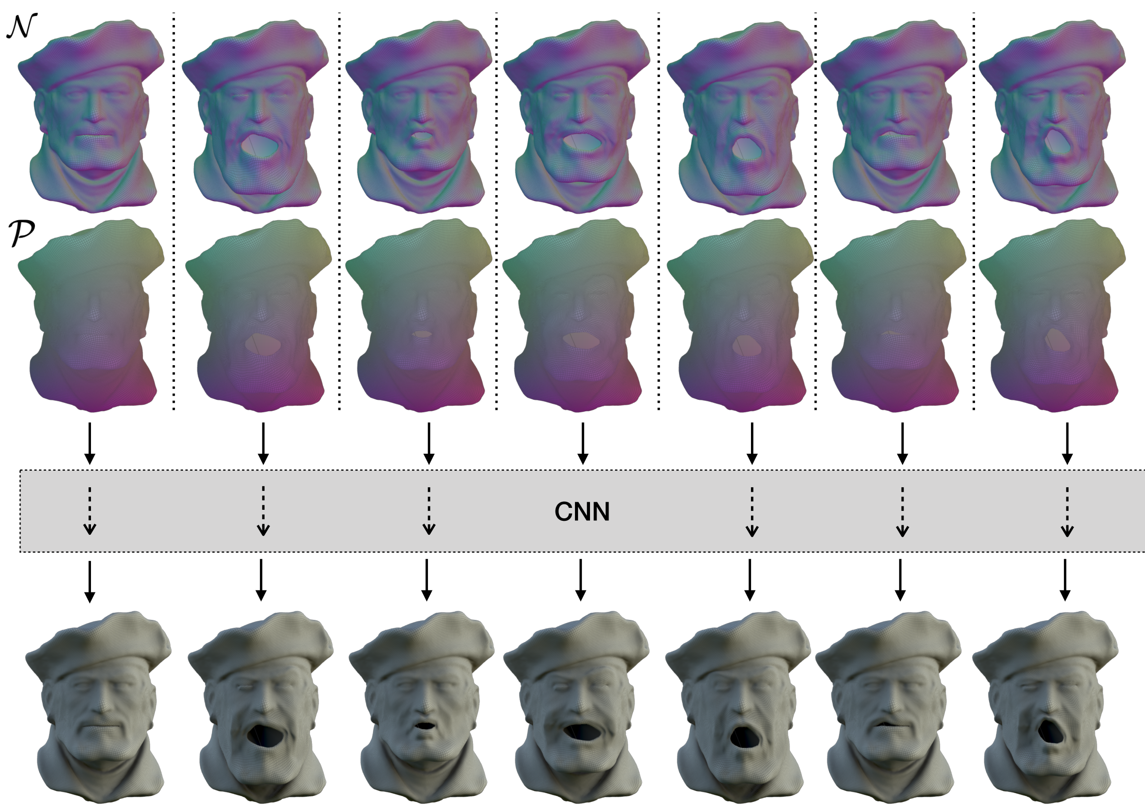

Figure 1: TODO

ABSTRACT

TODO:
(Written by Yue Li)
Traditional Precomputed Radiance Transfer methods require a lot of memory to save the precomputed data for real time rendering. For an animation sequence such data would be gigantic. We proposed a deep learning precomputed radiance transfer framework (DPRT) for deforming object, saving only the weights of the network, which lowers the memory cost in orders of magnitudes. Object is first parameterized via harmonic mapping and reconstructed to form geometry and normal image as the inputs of a carefully designed fully convolutional network.

CCS CONCEPTS

• **Computing methodologies → Rendering**; **Ray tracing**;

KEYWORDS

ray tracing, global illumination, octrees, quadtrees

ACM Reference Format:
Yue Li, Pablo Wiedemann, and Kenny Mitchell. 2018. Deep Precomputed Radiance Transfer for Deformable Objects. In *Proceedings of Interactive 3D Graphics and Games.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/8888888.7777777

Interactive 3D Graphics and Games, 2019, Montreal

ACM ISBN 978-1-4503-1234-5/17/07.
https://doi.org/10.1145/8888888.7777777

1 INTRODUCTION

Rendering photo-realistic appearances entails solving the *rendering equation* for each point on an objects surface. This computation can be extremely demanding, especially considering global illumination effects where the problem becomes highly recursive.
Precomputed Radiance Transfer (PRT) is a technique addressed to overcome this computational overkill, simplifying the rendering equation but still enabling high-quality renderings for complex illuminations. The quintessence is to perform a single pre-computation step of the light-transport information and only evaluate the equation at runtime.
Classic PRT algorithms function well for static scenes; however, these are destined to fail eyeing dynamic and/or interactive environments, in which considered objects undergo significant deformations. Responsible is a term in the rendering equation called the *transfer function* which is fully dependent on the shape of the surface. That is, any object deformation implies a re-computation of the *transfer function*, requiring expensive ray-casting. Hence, using classic PRT to render deformable objects would involve pre-computing large amounts of data, leading to immense storage consumptions. On top of that, the costly and memory consuming pre-computation of these *transfer functions*, presumes knowledge of all future deformations of the regarded object. Nevertheless, dynamic or interactive scenes may require on-the-fly adaptive, previously unknown, object deformations. Examples of such are: interactive physically based deformations for cloth or soft-bodies [find references] ; or more recent developements in the field of automatic character animations involving on-the-fly pose adaptation [for instance, Holden paper, deepmotion,etc...].

We propose a Deep Learning framework addressed to overcome the limitations of traditional PRT algorithms described above. In particular, we replace expensive ray-casting algorithms by a deep Convolutional Neural Network (CNN) that, for a given deformation, infers the corresponding set of SH - coefficients that represent the *transfer function*. Thus, regardless of the number of deformations our method maintains a constant size (fixed storage consumption). Moreover, due to the inherent generalisation capabilities of DNN's, DPRT is able to accurately predict appearances of previously unknown shapes. We call this approach *Deep Precomputed Radiance Transfer* (DPRT).

Finding an appropriate representation of shape, or manifold like, data to use in a CNN framework is a challenging task due to the non-Euclidean nature of the domain in which the data is defined on. Here, basic operations, such as the convolution, are not well defined being a major impediment for Deep Learning (DL) to fully flourish in this particular field. Nonetheless, more recently some authors have started to address the paradigm of DL on non-Euclidean data proposing a variety of approaches [Bronstein et al. 2016; Maron et al. 2017; Masci et al. 2015] [http://geometricdeeplearning.com/]. In particular, we propose learning on *geometry images*, a parametrisation proposed by [Gu et al. 2002] and further explored within the DL context by [**?**].

The main contributions of our approach are:

- enabling arbitrary and adaptive deformations,
- while maintaining a compact representation.

2 RELATED WORK

Precomputed Radiance Transfer (PRT)

PRT was first proposed by [Sloan et al. 2002] to address low-frequency global illumination effects on objects for real-time applications. This technique exploits the limitation of static objects by making a single pre-computation step of the Transfer Function, allowing fast computations at runtime.

PRT for dynamic or deformable objects would require pre-computing the Transfer Function for each conceivable pose, resulting in data sets that increase in size proportionally to the number of poses; hence, rapidly becoming unwieldy for such applications.
Our aim is to extend traditional PRT to arbitrary deformable geometries while preserving a rather manageable and compact storage consumption.
To our knowledge, literature that regards PRT for deformable objects is, on one hand, relatively narrow and, on the other hand, mostly very limited and concise with respect to their proposed solutions.
One extension of PRT was introduced by [Sloan et al. 2005] to enable transfer of local illumination effects, such as bumps and wrinkles, to arbitrary deformations. Nevertheless, this method cannot account for shadowing effects that arise from global shape deformations, such as the cast shadow from a limb to the trunk from an articulated figure.
Other approaches [Dachsbacher et al. 2007; Dong et al. 2007] circumvent the pre-computation problem by proposing an alternative algorithm to efficiently compute an approximation of the Visibility Function (implicit in T) at near real-time frame rates. (However, ...)
Data-based approaches, in principle aim to reduce the dimensionality of the problem, and thus the storage consumption, by exploiting the information of the dataset:
A data-based compression scheme of precomputed radiance transfer matrices is presented in [Feng et al. 2007]. Precomputed transfer matrices of surface samples, deformed by *skinning*, are clustered and compressed, such that de-compression and interpolation can be performed efficiently.
An appearance model, that approximates PRT lighting, is presented in [James and Fatahalian 2003]. The model is based on a reduced state space of deformable shapes that allows only very limited kind of poses/shapes.
Similarly, [Schneider et al. 2017] suggest a linear self-shadowing model to predict the coefficients of the Transfer Function from shape parameters of Morphable Models (MoMo) [Blanz and Vetter 1999]. Their proposed model show good results while operating within the reduced shape space of MoMo; nevertheless, our aim is to provide a more generic PRT-model enabling good approximations for more general arbitrary deformations. To that end, we rather propose a non-linear model with well known strong generalisation properties namely Deep Neural Networks [LeCun et al. 2015]. To the extend of our knowledge, our work is the first to tackle the problem of PRT for deformable objects.

Nevertheless, Deep Learning (DL) has been used for appearance predictions before, mostly focusing on learning illumination effects from screen-space data. For instance, [Nalbach et al. 2017] and also [Thomas and Forbes 2017] learn on image data gathered from the shading buffers to predict illumination effects in screen space. However, this approach does not leverage the underlying structure of the geometry (....)
Alternatively, we propose learning from geometric data, in particular our aim is to apply a fully Convolutional Neural Network (CNN), due to its remarkable classification properties [Karpathy et al. 2014; Krizhevsky et al. 2012], on surfaces data. However, basic operations such as the convolution are not well-defined on Surfaces, hence making the problem rather challenging.
One approach is to circumvent this difficulty by representing the surface data as a probability distribution on a 3D grid and apply volumetric CNN's [Wu et al. 2015]. However, this extrinsic representation has many shortcomings when applied to deformable geometries: They are very sensitive to deformations are computationally expensive and, equally to the screen-space strategies, do not exploit the intrinsic structure of the geometry... Deep Surface Light Fields [Chen et al. 2018]

Learning on surfaces

Geometry image:

- is utilized to preserve 3D mesh as 2D image. The regular shape of these images could be used in different areas of graphics reseaches and applications [Gu et al. 2002].
- To get rid of classic geometry image artefacts [Praun and Hoppe 2003; Sinha et al. 2016]

3 METHOD

Precomputer Radiance Transfer (PRT).

In PRT the integrands of the *rendering equation*

$$L(\boldsymbol{x}, \boldsymbol{\omega}_0) = \int_{\Omega} L_{\epsilon}(\boldsymbol{x}, \boldsymbol{\omega}_i) T(\boldsymbol{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_0, \boldsymbol{N}) \, d\boldsymbol{\omega}_i, \quad (1)$$

are split into two terms:

1. The *lighting function*: $L_{\epsilon}(\boldsymbol{x}, \boldsymbol{\omega}_i)$, describing all incoming radiance over the hemisphere around the surface point $\boldsymbol{x}$,
2. and the *transfer function* :

$$T(\boldsymbol{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_0, \boldsymbol{N}) = f(\boldsymbol{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_0) G(\boldsymbol{x}, \boldsymbol{\omega}_i, \boldsymbol{N})$$

describing the surface reflectance properties f (BRDF) and the geometric information G surrounding $\boldsymbol{x}$. Where $\boldsymbol{N}$ is the normal vector of the surface at $\boldsymbol{x}$.

Both functions L_{ϵ} and T are projected onto a suitable set of orthonormal basis functions, in our case *Spherical Harmonics* (SH), for faster evaluation.
For n number of SH bands and l_i, t_i being the i-th SH coefficient of L and T respectively, the rendering equation 1 reduces to [for more PRT SH, see citations]:

$$L(\boldsymbol{x}, \boldsymbol{\omega}_0) \approx \sum_{i}^{n^2} l_i \cdot t_i$$

Goal. The objective of our CNN is to directly predict the coefficients t_i's, for a given shape, skipping the costly ray-casting computations of G.

Problem. To this end, the network has to be fed with shape information.

3.1 Object Reconstution

We chose a parametric approach called *geometry Iimage* which translates the problem into a uniform grid on which standard CNNs can be applied [Gu et al. 2002; Sinha et al. 2016]. The parametrisation we chose is a harmonic mapping and the geometric data we transform into *geomtry images* are vertex positions and normals (lets call them postion and normal image respectively) (see Figure 2 picturing input and output of network)

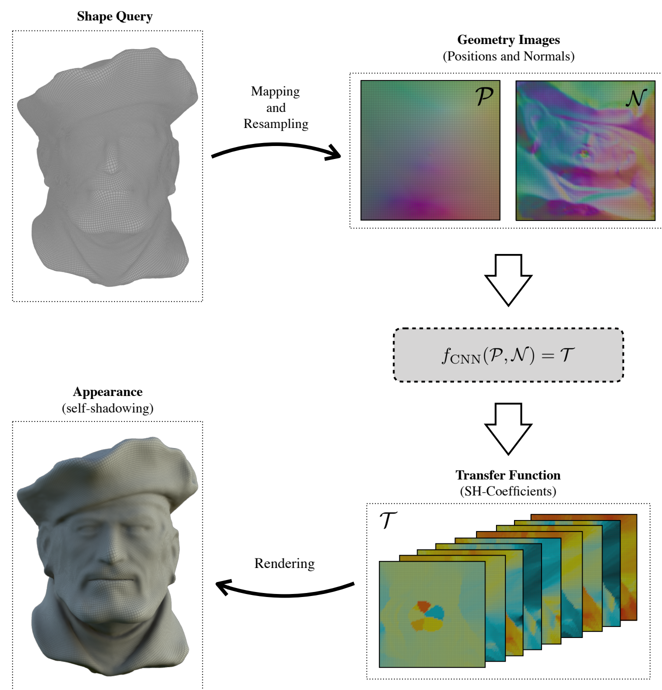

Figure 2: Method Overview

3.2 DPRT

3.2.1 Data. We synthesise our training data by generating animations (such as cloth or faces, see example section) of deforming objects of 500 frames of duration. For each frame, a set

$$\mathcal{D} = \{\mathcal{P}, \mathcal{N}, \mathcal{T}\}$$

of position and normal images ($\mathcal{P} = \{P_x, P_y, P_z\}$ and $\mathcal{N} = \{N_x, N_y, N_z\}$) are created, where

$$P_i, N_i \in \mathbb{R}^{M \times M}, \quad i \in \{x, y, z\}$$

Besides, the a set of precomputed SH-Coefficients of the Transfer Function $\mathcal{T} = \{T_0, T_1, T_2, T_3..., T_{n^2}\}$ for the ground truth data...

3.2.2 *Network.*

4 RESULTS AND COMPARISONS

4.1 DPRT on Diffuse Object

4.2 DPRT on Glossy Object

5 FUTURE WORK

- Use stretch-minimizing parametrization of paper [Gu et al. 2002] instead of current harmonic mapping.
  → paper claims: reconstruction is more uniformly distributed after re-sampling on image space.
  (see slide 21 in: http://hhoppe.com/gim.ppt)
- extend parametrisation to surfaces with higher genus

REFERENCES

Volker Blanz and Thomas Vetter. 1999. A Morphable Model for the Synthesis of 3D Faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '99)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 187–194. https://doi.org/10.1145/311535.311556

Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. 2016. Geometric deep learning: going beyond Euclidean data. *CoRR* abs/1611.08097 (2016). arXiv:1611.08097 http://arxiv.org/abs/1611.08097

Anpei Chen, Minye Wu, Yingliang Zhang, Nianyi Li, Jie Lu, Shenghua Gao, and Jingyi Yu. 2018. Deep Surface Light Fields. *Proc. ACM Comput. Graph. Interact. Tech.* 1, 1, Article 14 (July 2018), 17 pages. https://doi.org/10.1145/3203192

Carsten Dachsbacher, Marc Stamminger, George Drettakis, and Frédo Durand. 2007. Implicit Visibility and Antiradiance for Interactive Global Illumination. *ACM Trans. Graph.* 26, 3, Article 61 (July 2007). https://doi.org/10.1145/1276377.1276453

Z. Dong, J. Kautz, C. Theobalt, and H. Seidel. 2007. Interactive Global Illumination Using Implicit Visibility. In *15th Pacific Conference on Computer Graphics and Applications (PG'07)*. 77–86. https://doi.org/10.1109/PG.2007.37

Wei-Wen Feng, Liang Peng, Yuntao Jia, and Yizhou Yu. 2007. Large-scale Data Management for PRT-based Real-time Rendering of Dynamically Skinned Models. In *Proceedings of the 18th Eurographics Conference on Rendering Techniques (EGSR'07)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 23–34. https://doi.org/10.2312/EGWR/EGSR07/023-034

Xianfeng Gu, Steven J Gortler, and Hugues Hoppe. 2002. Geometry images. *ACM Transactions on Graphics (TOG)* 21, 3 (2002), 355–361.

Doug L. James and Kayvon Fatahalian. 2003. Precomputing Interactive Dynamic Deformable Scenes. *ACM Trans. Graph.* 22, 3 (July 2003), 879–887. https://doi.org/10.1145/882262.882359

A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei. 2014. Large-Scale Video Classification with Convolutional Neural Networks. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 1725–1732. https://doi.org/10.1109/CVPR.2014.223

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet Classification with Deep Convolutional Neural Networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12)*. Curran Associates Inc., USA, 1097–1105. http://dl.acm.org/citation.cfm?id=2999134.2999257

Yann LeCun, Yoshua Bengio, and Geoffrey E. Hinton. 2015. Deep learning. *Nature* 521, 7553 (2015), 436–444. https://doi.org/10.1038/nature14539

Haggai Maron, Meirav Galun, Noam Aigerman, Miri Trope, Nadav Dym, Ersin Yumer, Vladimir G. Kim, and Yaron Lipman. 2017. Convolutional Neural Networks on Surfaces via Seamless Toric Covers. *ACM Trans. Graph.* 36, 4, Article 71 (July 2017), 10 pages. https://doi.org/10.1145/3072959.3073616

Jonathan Masci, Davide Boscaini, Michael M. Bronstein, and Pierre Vandergheynst. 2015. ShapeNet: Convolutional Neural Networks on Non-Euclidean Manifolds. *CoRR* abs/1501.06297 (2015).

Oliver Nalbach, Elena Arabadzhiyska, Dushyant Mehta, Hans-Peter Seidel, and Tobias Ritschel. 2017. Deep Shading: Convolutional Neural Networks for Screen-Space Shading. 36, 4 (2017).

Emil Praun and Hugues Hoppe. 2003. Spherical Parametrization and Remeshing. *ACM Trans. Graph.* 22, 3 (July 2003), 340–349. https://doi.org/10.1145/882262.882274

A. Schneider, S. SchÃűnborn, B. Egger, L. Frobeen, and T. Vetter. 2017. Efficient Global Illumination for Morphable Models. In *2017 IEEE International Conference on Computer Vision (ICCV)*. 3885–3893. https://doi.org/10.1109/ICCV.2017.417

Ayan Sinha, Jing Bai, and Karthik Ramani. 2016. Deep learning 3D shape surfaces using geometry images. In *European Conference on Computer Vision*. Springer, 223–240.

Peter-Pike Sloan, Jan Kautz, and John Snyder. 2002. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In *ACM Transactions on Graphics (TOG)*, Vol. 21. ACM, 527–536.

Peter-Pike Sloan, Ben Luna, and John Snyder. 2005. Local, Deformable Precomputed Radiance Transfer. ACM, 1216–1224. https://www.microsoft.com/en-us/research/publication/local-deformable-precomputed-radiance-transfer/

Manu Mathew Thomas and Angus Graeme Forbes. 2017. Deep Illumination: Approximating Dynamic Global Illumination with Generative Adversarial Network. *CoRR* abs/1710.09834 (2017). arXiv:1710.09834 http://arxiv.org/abs/1710.09834

Zhirong Wu, S. Song, A. Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and J. Xiao. 2015. 3D ShapeNets: A deep representation for volumetric shapes. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 1912–1920. https://doi.org/10.1109/CVPR.2015.7298801