Efficient Graphics Representation with Differentiable Indirection

ANONYMOUS AUTHOR(S)
SUBMISSION ID: PAPERS_000

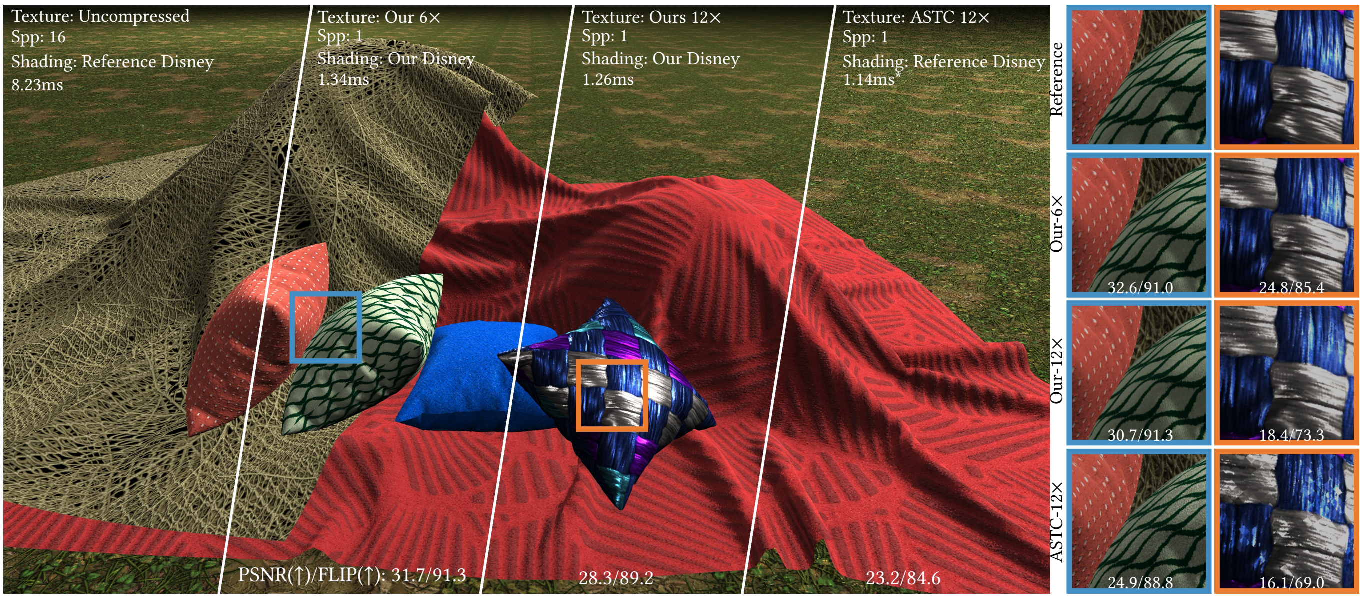

Fig. 1. Figure shows the use of *differentiable indirection* for texture compression/sampling and parametric shading. Our technique relies on few linearly interpolated memory lookups and applies to a wide range of tasks in the graphics pipeline including signed distance and radiance fields.

We introduce *differentiable indirection* – a novel neural primitive that employs differentiable multi-scale lookup tables as an effective substitute for traditional compute and data operations across the graphics pipeline. We demonstrate its flexibility on a number of graphics tasks, i.e., geometric and image representation, texture mapping, shading, and radiance field representation. In all cases, neural indirection seamlessly integrates into existing architectures, trains rapidly, and yields both versatile and efficient results.

CCS Concepts: • **Computing methodologies** → *Reflectance modeling*; *Non-photorealistic rendering*; **Image compression**; *Shape representations*; **Visibility**; *Machine learning*; *Rasterization.*

Additional Key Words and Phrases: Differentiable graphics representations

ACM Reference Format:
Anonymous Author(s). 2023. Efficient Graphics Representation with Differentiable Indirection. 1, 1 (June 2023), 10 pages. https://doi.org/a.bb/ccc.dddd

1 INTRODUCTION

Neural primitives are the fundamental building block of neural networks and used for a variety of purposes in graphics applications, such as appearance capture [34], shading [24], radiance caching [19], view-synthesis [17], and generative modeling [14]. Having *efficient* neural primitives is vital, due to their impact on latency, scalability, power, and training speed. Achieving high runtime performance with neural primitives is essential to the adoption of neural networks in real-time and low-power applications, such as AR/VR.

We introduce a simple neural primitive with exceptional runtime characteristics, featuring *low compute flops*, *minimal memory reads per query*, and *a compact parameter size.* Many networks rely primarily on multilayer perceptrons (MLP) due to their appeal as universal function approximators; however, MLP layers are often the most computationally expensive component of a network and often scale quadratically (both in flops and bytes transferred) with quality due to large matrix operations [25]. Conversely, combining memory grids with fixed function non-linearities such as Spherical Harmonics (SH) [9] or ReLUs [10] reduces compute and memory transfer but a large parameter cost. Our novel neural primitive – *differentiable indirection* – strikes a balance across these criteria and serves as a drop-in replacement for memory grid-based representations. It is compatible with any differentiable logic, such as MLPs or fixed function approaches, but significantly reduces or even eliminates reliance on MLPs. Notably, all of our examples are MLP-free, thereby eliminating the need for specialized hardware acceleration in real-time applications [7]. *Differentiable indirection* draws its expressive power solely from memory indirections and linear interpolation. This approach aligns well with the emerging computing paradigm of *compute in memory* [15, 31], which departs from traditional von Neumann model that MLPs are based on. We

XXXX-XXXX/2023/6-ART $15.00
https://doi.org/a.bb/ccc.dddd