Memory-efficient Learning for Large-scale Computational Imaging

Michael Kellman*, Jon Tamir, Emrah Bostan, Michael Lustig & Laura Waller
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley
*kellman@berkeley.edu

Overview

- Learning for large-scale computational imaging systems is limited due to the memory capacity of modern graphical processing units.
- We propose a method that enables learning for large-scale computational imaging systems using constant memory.

Introduction

Conventional Image Reconstruction:

\[ x^* = \arg \min_x \sum_{k=1}^K \| y_k - A_k(x) \|^2 + \mathcal{P}(x) \]

The system’s image prior and experimental design can be learned to improve overall performance. Unroll the iterations of the reconstruction’s optimizer to form the layers of a network [1]. Optimizers often alternate between applying data and prior updates to form the network:

1st \[ x^{(0)} \] → 2nd \[ x^{(1)} \] → 3rd \[ x^{(2)} \] → ... → Nth \[ x^{(N)} \]

Gradients are computed for the network using automatic differentiation [2]. For large-scale computational imaging systems, this will exceed the GPU’s memory capacity.

Memory-efficient Learning

Rather than storing the whole graph for auto-differentiation, smaller graphs can be formed for each layer one at a time in reverse order by recalculating intermediate variables using each layer’s inverse.

- Recalculate input to layer using layer’s inverse
- Recompute layer’s auto-differentiation graph
- Backpropagate gradients through layer

Results: Multi-channel MRI

- Multi-channel MRI reduces scan time by undersampling and relying on image priors for reconstruction.
- Learning image priors for 3D multi-channel MRI [3] is restricted by memory available on GPU.

MR image reconstruction [3]:

\[ x^{(k+1)} = \text{Conjugate Gradient} \] → Denoiser (learnable) \[ x^{(k)} \]

Our method enables learning at practical scales for this system, ordinarily requiring >40GB using only 10GB.

Results: Super Resolution Microscopy

- Fourier Ptychography performs super resolution on an LED array microscope, however, has poor time resolution.
- Learn the illumination design to compress the information into fewer measurements [4].

Fourier Ptychography

Our method enables learning at practical scales for this system, ordinarily requiring 500GB using only 3GB.

References


* M. Kellman is supported in part by the National Science Foundation (NSF) under grant DGE 1106400.