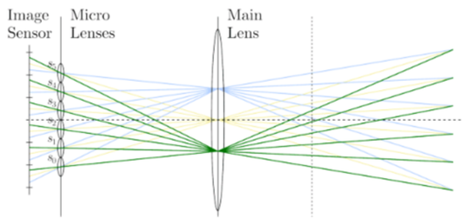
Encoding Lightfields with Coordinate-based Neural Networks

Samuel Cole, Computer Engineering
Dr. Jonathan Ventura, Computer Science

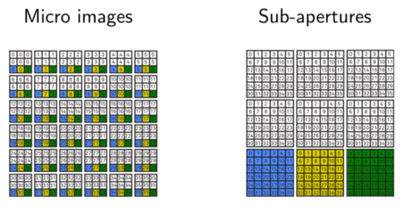
CAL POLY

Neural Lightfields

- Lightfields are a 4D structure consisting of sub aperture images arranged in a sphere around the viewer
- A Plenoptic camera, which simulates an array of individual cameras, is used to capture lightfields
- They contain data capable of recreating 3D scenes, say for virtual reality
- Neural Lightfields are created using a trained neural netowrk



Model of a Plenoptic Camera used to capture a Lightfield [1]

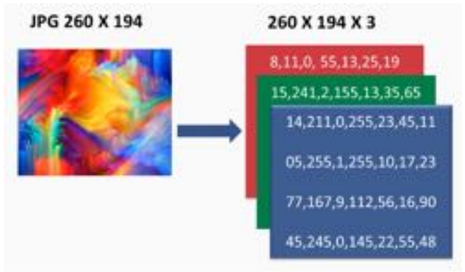


Transforming Micro Images to sub-aperture images [1]

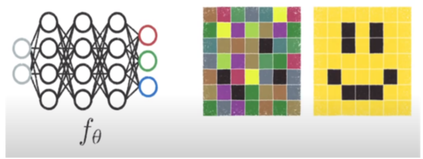
A Lightfield with closeup of sub-aperture images

Coordinate-based Neural Networks

- Coordinate Based Neural Networks are used in Computer Vision tasks to encode and compress N-Dimensional structures
- Using a simple RGB image as an example:
  - Input Vector → 3D Coordinate
  - Output Vector → RGB Pixel Value



RGB Image as a matrix [2]

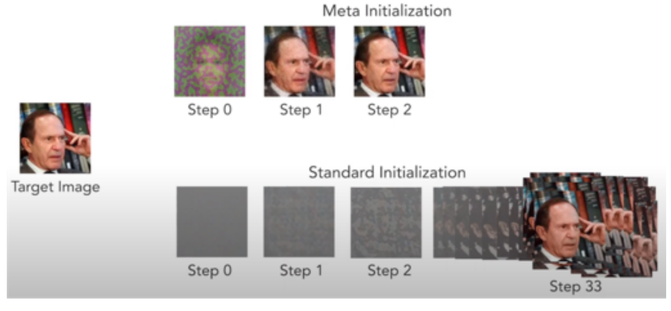


RGB Coordinate-based Neural Net [3]

- On more complex media i.e. 5D Voxel grid, a network can provide high compression
  - 5D Voxel Grid (+10Gb) → Trained weights (~5Mb)
- Model Weights are trained to encode a single N-Dimensional structure
  - Every new example must be trained individually
  - This is slow and not ideal for real time applications

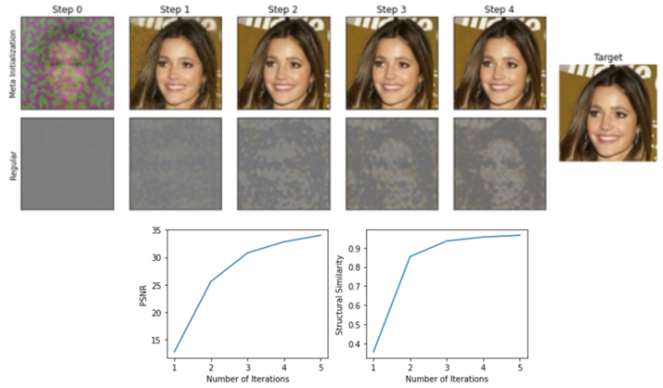
Model Agnostic Meta Learning (MAML)

- Normally, starting weights for training a model are random
- If the targets of training are similar images for CBN, we can find starting weights that make training converge faster

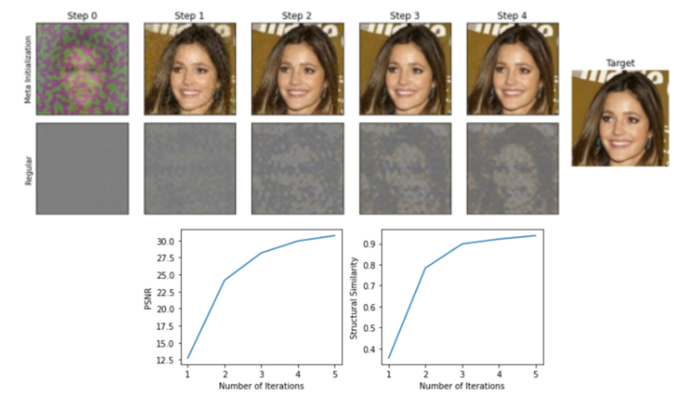


Training with MAML vs Standard Initialization [3]

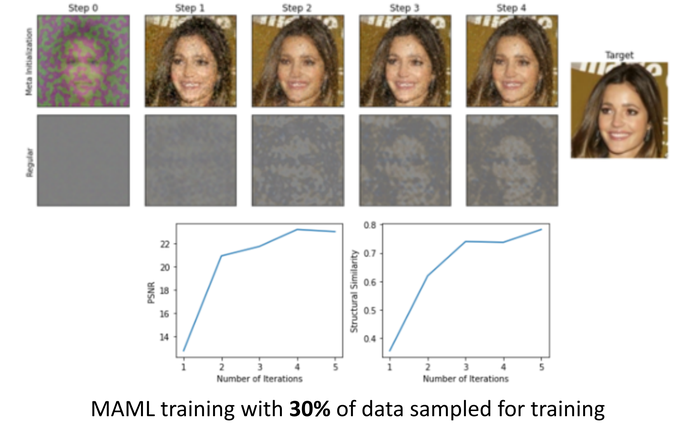
Incomplete sampling with MAML and CBN



Baseline MAML training with **100%** of data sampled for training



MAML training with **70%** of data sampled for training



MAML training with **30%** of data sampled for training

REFERENCES

1 C. Hahne, "The Standard Plenoptic camera," *The Plenoptic Camera aka Light Field Camera*.
2 "Convolution on RGB images," *Packt subscription*, 2020.
3 *Learned Initializations for Optimizing Coordinate-Based Neural Representations*. YouTube, 2020.