# **Encoding Lightfields with Coordinate-based** Neural Networks Samuel Cole, Computer Engineering Dr. Jonathan Ventura, Computer Science

### **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]

				_		
0 0 0 0 0 0 0 0	1 1 1 1 1 1 1	2 2 2 2 2 2 2 2 2	3 3 3 3 3 3 3 3	444	5 5 5 5 5 5 5 5	
6 6 6 6 6 6	7 7 7 7 7 7 7 7 7	8 8 8 8 8	9999 999 999	10 10 10 10 10 10 10 10	57 57 57 57 57 57 11 17	
12 12 12 12 12 12 12 12	13 13 13 13 13 13 13 13	14 14 14 14 14 14	15 15 15 15 15 15 15 15	16 16 16 16 16 16 16 16	17 17 17 17 17 17 17 17 17	
18 18 18 18 18 18 18 18	19 19 19 19 19 19 19 19 19	20 20 20 20 20 20 30 <u>20</u>	21 21 21 21 21 21 21 21 21	22 22 22 22 22 22 22 22 22 22 22	20 20 20 28 28 23 38 <mark>23</mark>	
24 24 24 24 24 24 26 24	25 25 25 25 25 25 26 <b>26</b>	26 26 26 26 26 26 26 26 26	27 27 27 27 27 27 22 27 27	28 28 28 28 28 28 28 28 28	20 20 20 20 20 20 20 <b>20</b>	
30 30 30 30 30 30 30 <mark>30</mark>	313131 313131 31 <mark>31</mark> 31	32 32 32 32 32 32 32 32 32	33 33 33 33 33 33 35 <mark>33</mark>	34 34 34 34 34 34 34 34	35 35 35 35 35 35 35 <mark>35</mark>	

Micro images

### Sub-apertures



Transforming Micro Images to sub-aperture images [1]



A Lightfield with closeup of sub-aperture images

### 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.

# **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  $\rightarrow$  3D Coordinate
  - Output Vector  $\rightarrow$  RGB Pixel Value



RGB Image as a matrix [2]

- On more complex media i.e. 5D Voxel grid, a network can provide high compression
  - 5D Voxel Grid (+10Gb)  $\rightarrow$  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]





RGB Coordinate-based Neural Net [3]

### **Incomplete sampling with** MAML and CBN







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





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



Number of Iterations

Number of Iterations

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