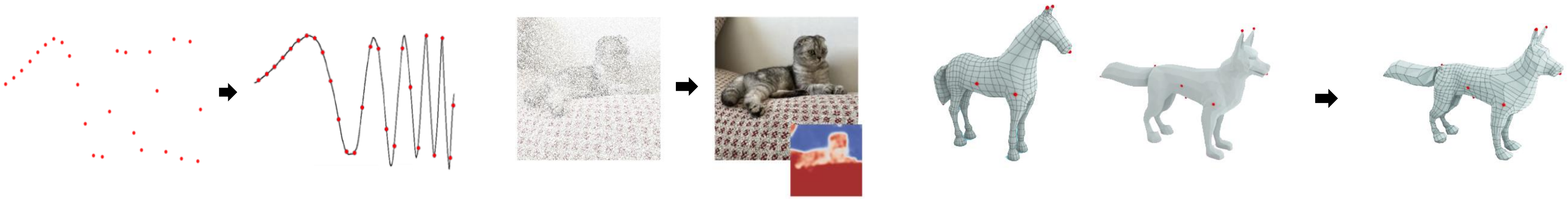
The Blavatnik School of Computer Science
The Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University

SAPE: Spatially-Adaptive Progressive Encoding for Neural Optimization

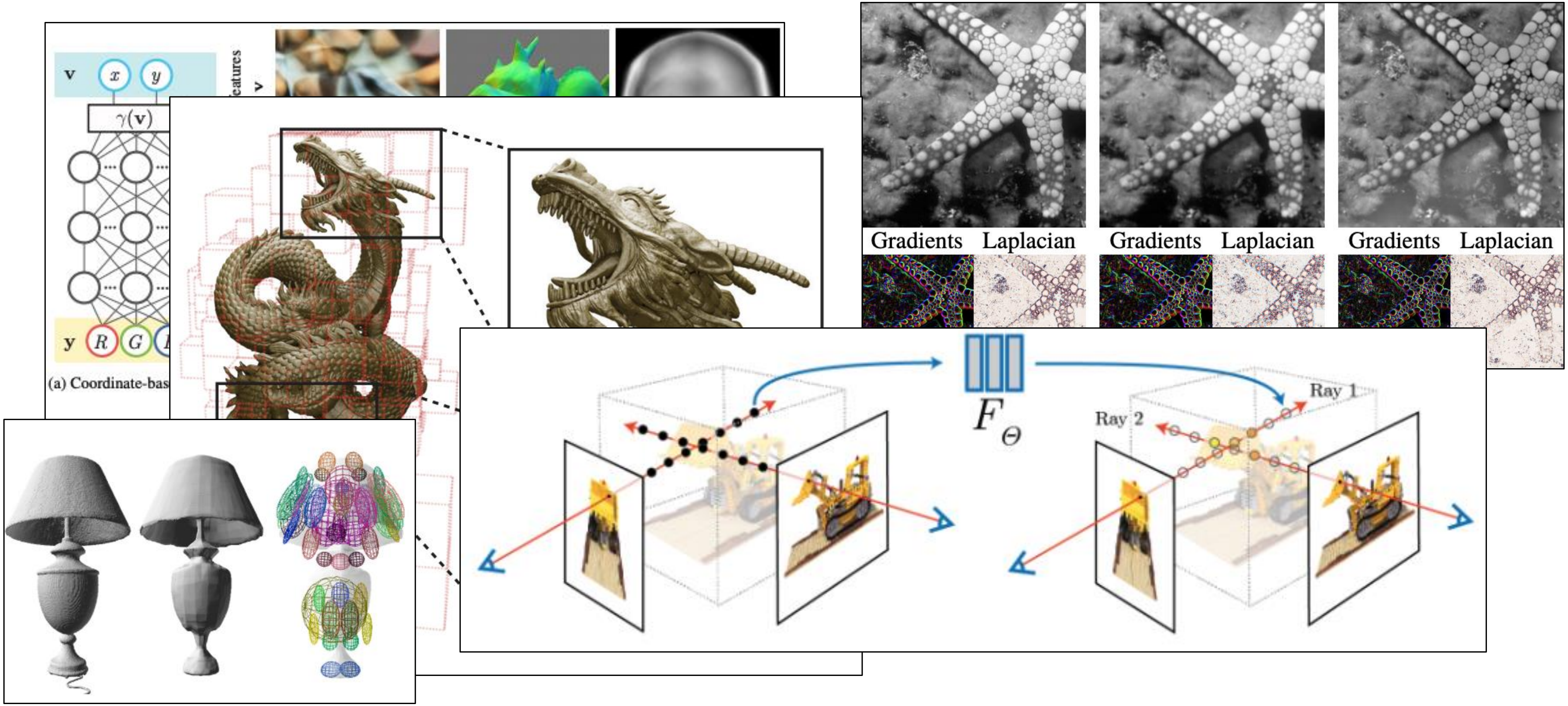
Amir Hertz[1], Or Perel[1],
Raja Giryes[1], Olga Sorkine-Hornung[2], Daniel Cohen-Or[1]

NeurIPS 2021



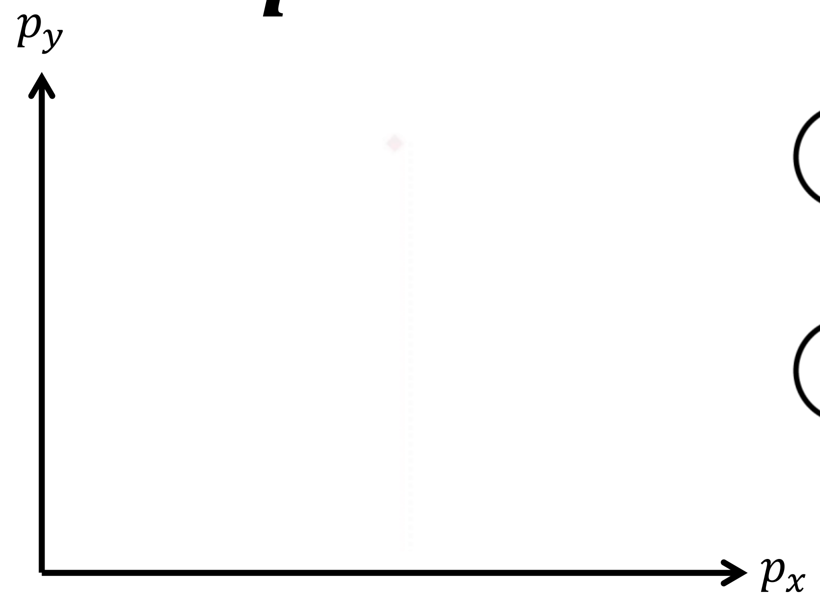
[1] *Tel Aviv University*
[2] *ETH Zurich, Switzerland*

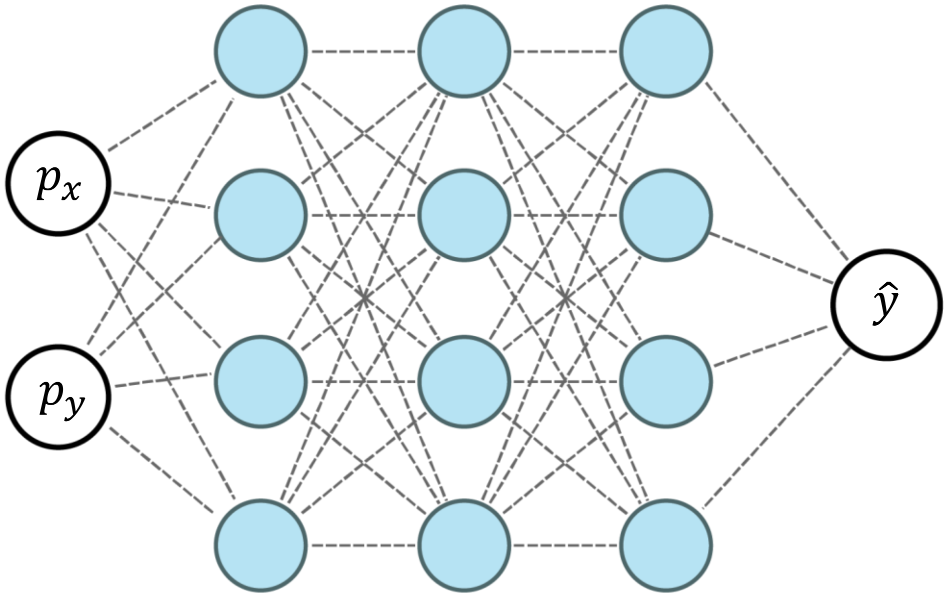
Neural Implicit Functions



Neural Implicit Functions

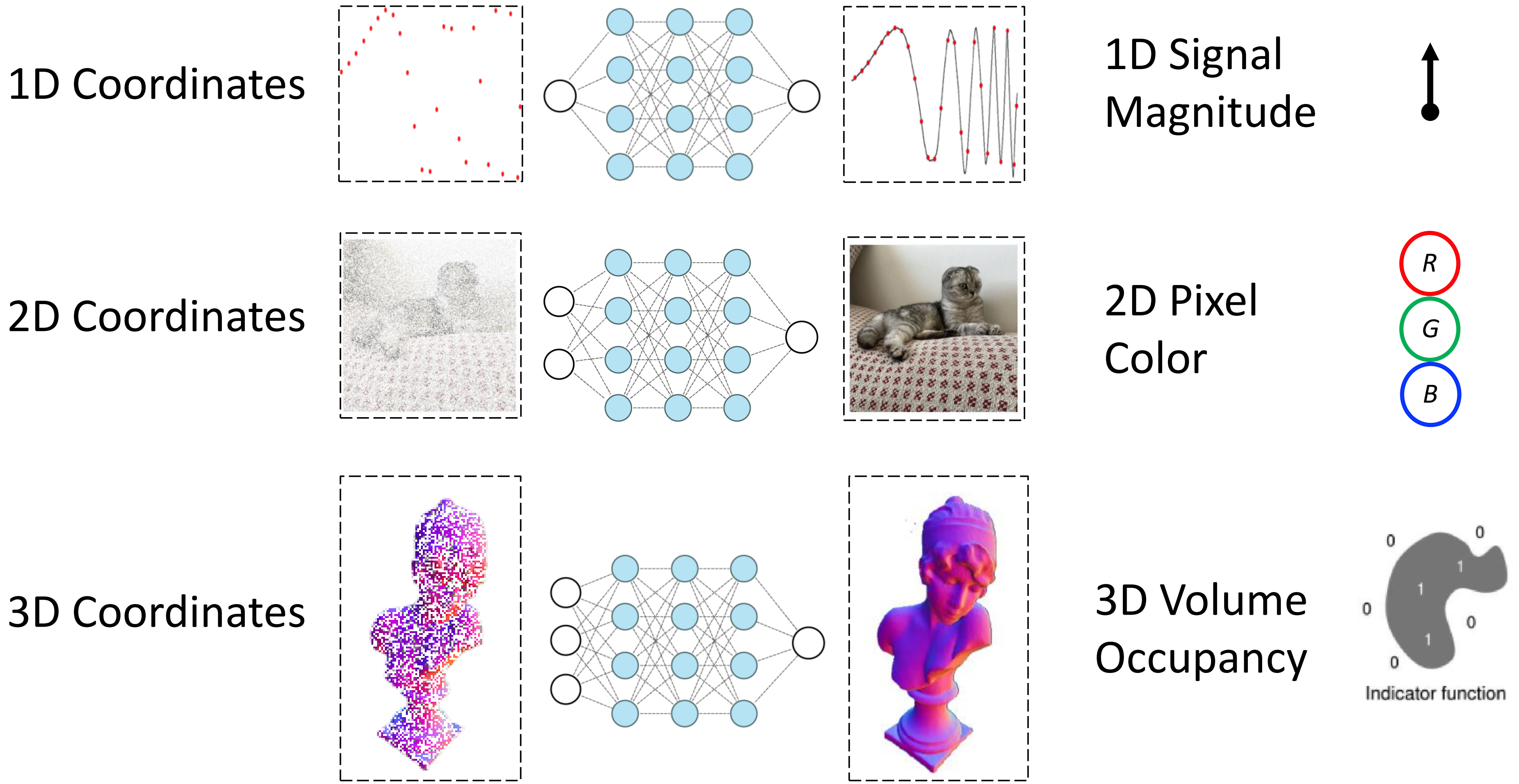
$\boldsymbol{p} \in \mathbb{R}^d$



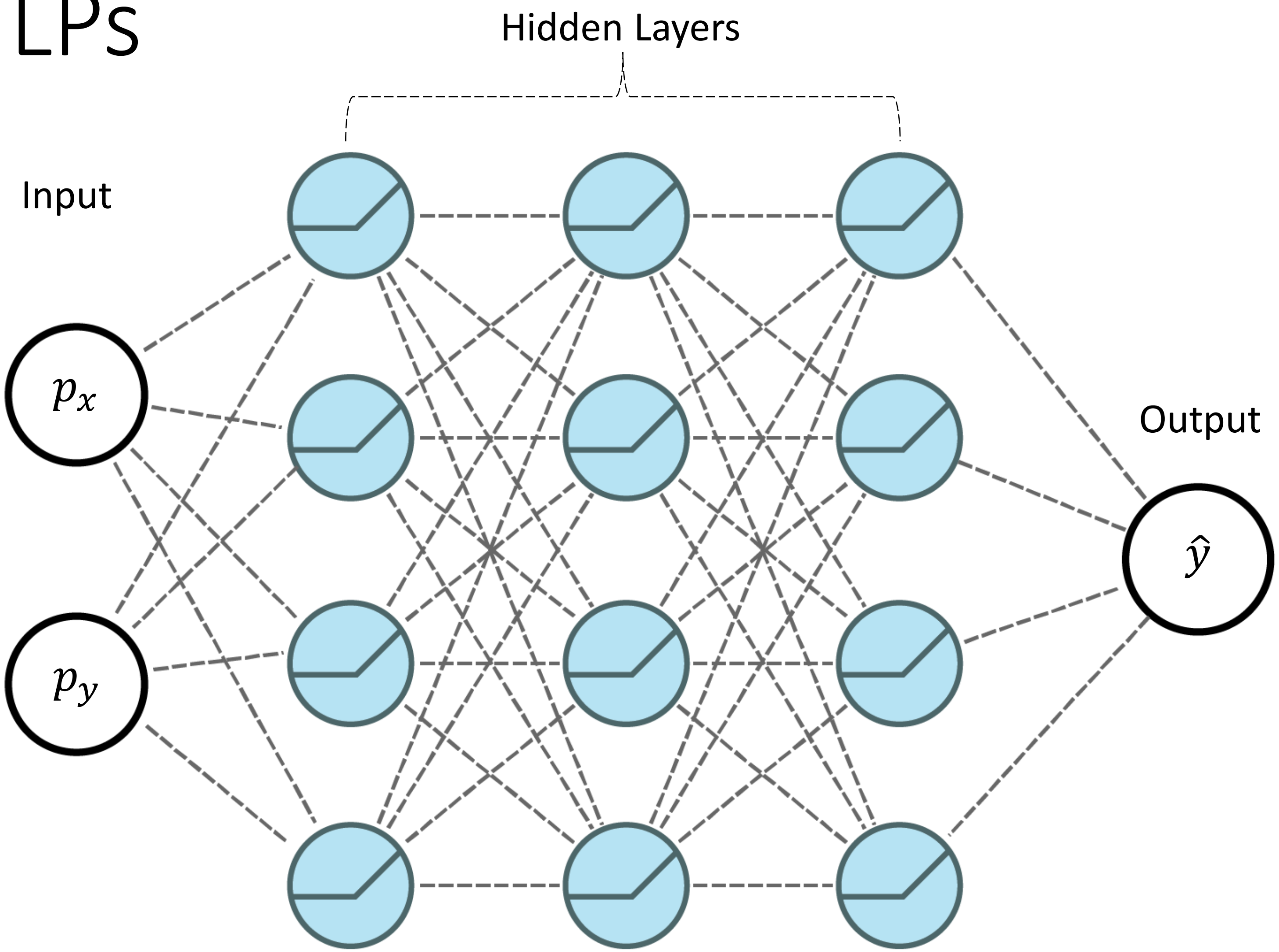


$\hat{\boldsymbol{y}}_{\boldsymbol{RGB}}$

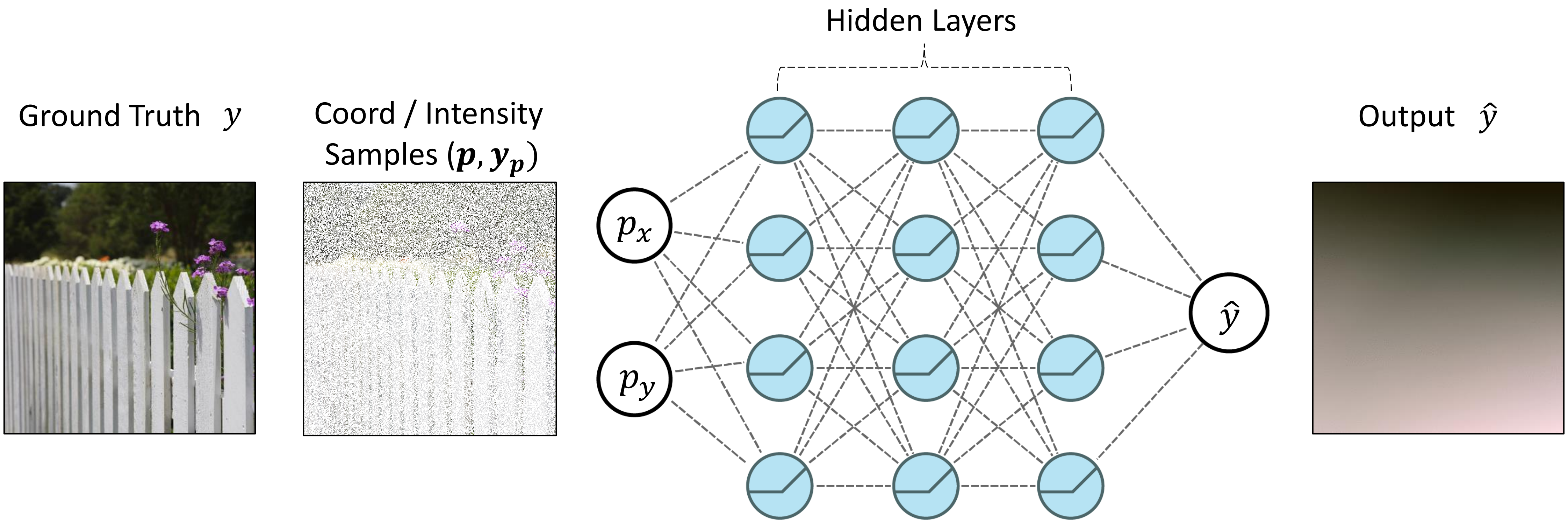
Our Setting


1D Coordinates
1D Signal Magnitude
2D Coordinates
2D Pixel Color
R
G
B
3D Coordinates
3D Volume Occupancy
Indicator function

ReLU MLPs

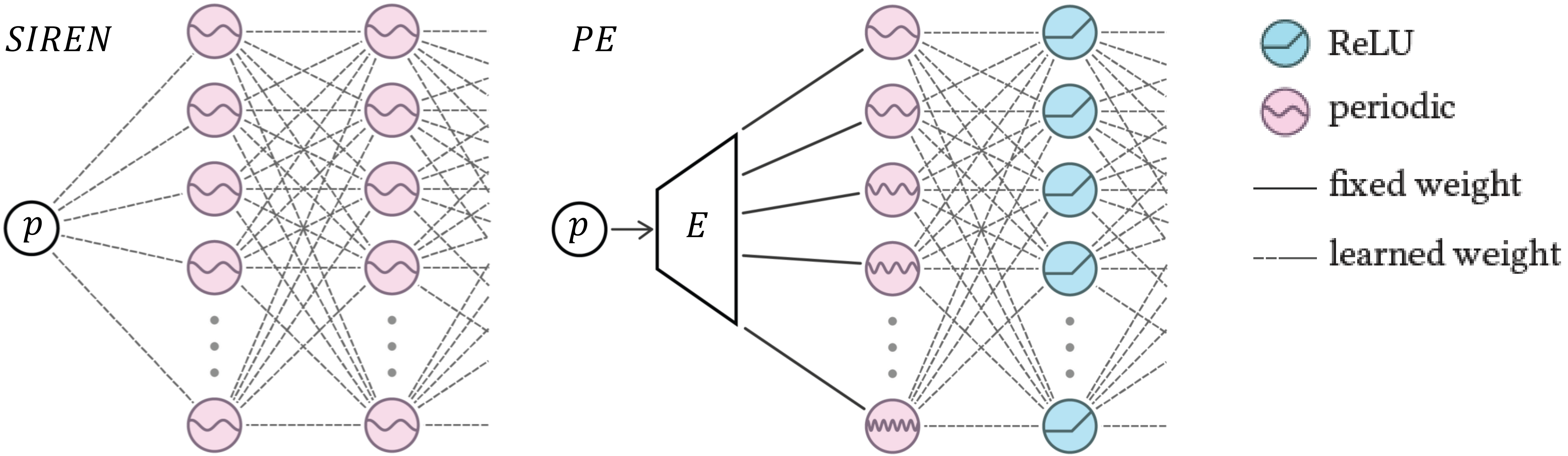


ReLU MLPs



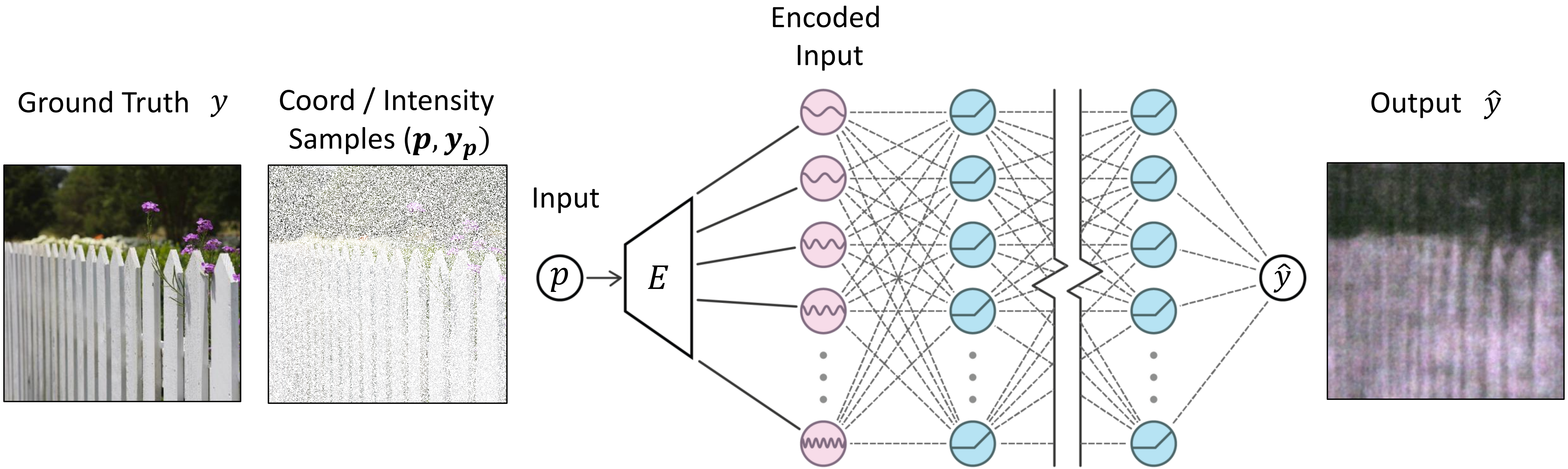
On the Spectral Bias of Neural Networks, Rahman et al. 2019

Positional Encodings



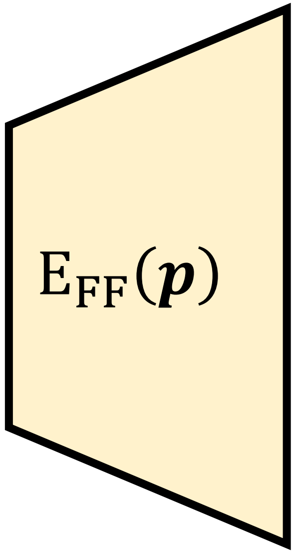
NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall et al. 2020
Implicit Neural Representations with Periodic Activation Functions, Sitzmann et al. 2020

Positional Encoding



Positional Encoding

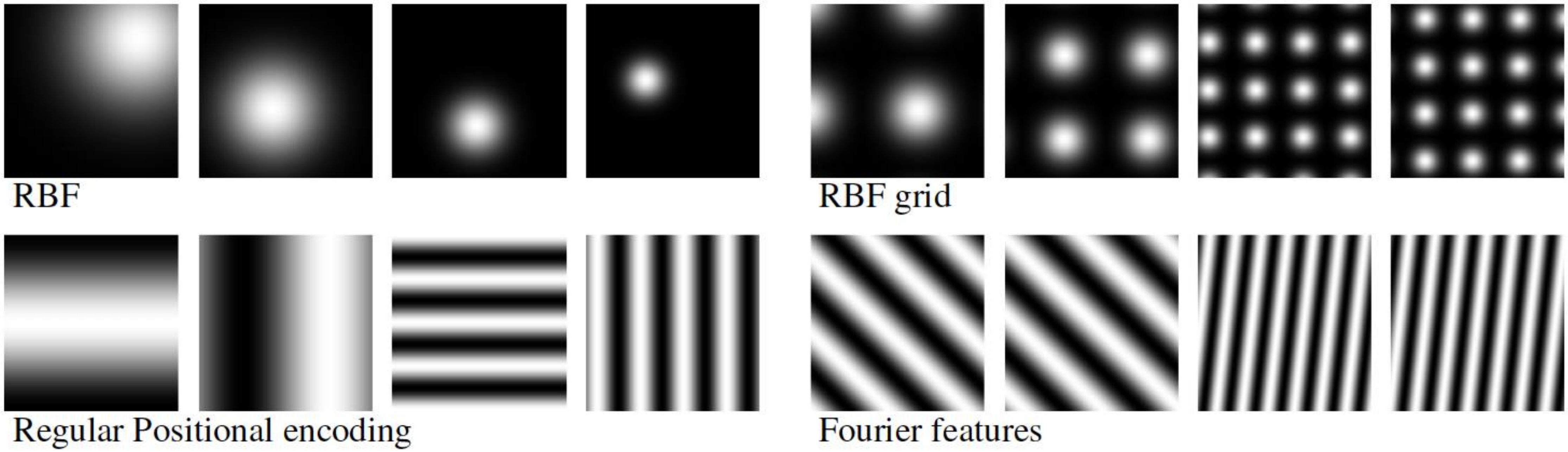
$\boldsymbol{p} \in \mathbb{R}^d \longrightarrow$



$$\longrightarrow \begin{bmatrix} \cos(2\pi \boldsymbol{b}_1^\top \boldsymbol{p}) \\ \sin(2\pi \boldsymbol{b}_1^\top \boldsymbol{p}) \\ \vdots \\ \cos(2\pi \boldsymbol{b}_n^\top \boldsymbol{p}) \\ \sin(2\pi \boldsymbol{b}_n^\top \boldsymbol{p}) \end{bmatrix}$$

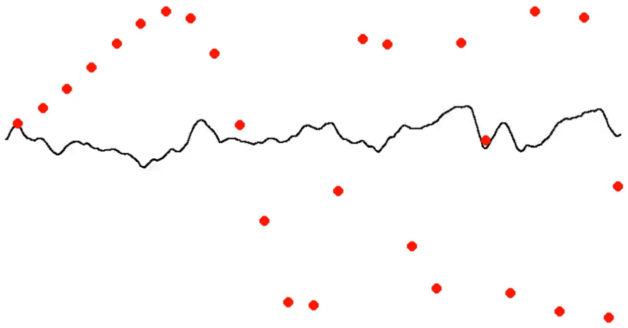
Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains, Tancik et al. 2020

Positional Encodings

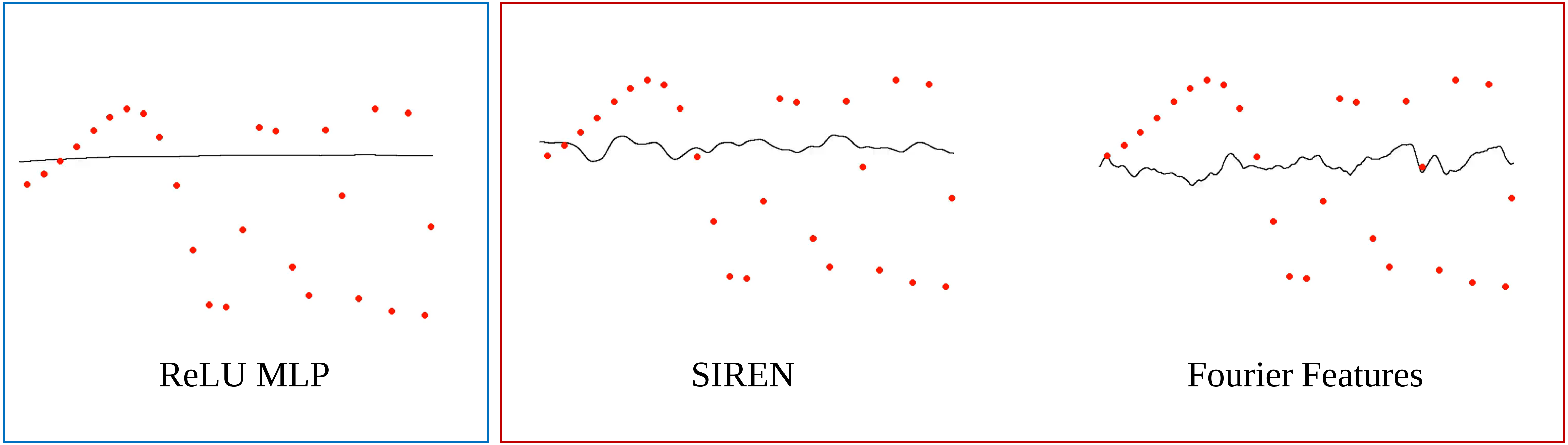


What's the problem here?


FFN $\sigma = 5$ FFN $\sigma = 25$



ReLU MLP vs. Positional Encoding

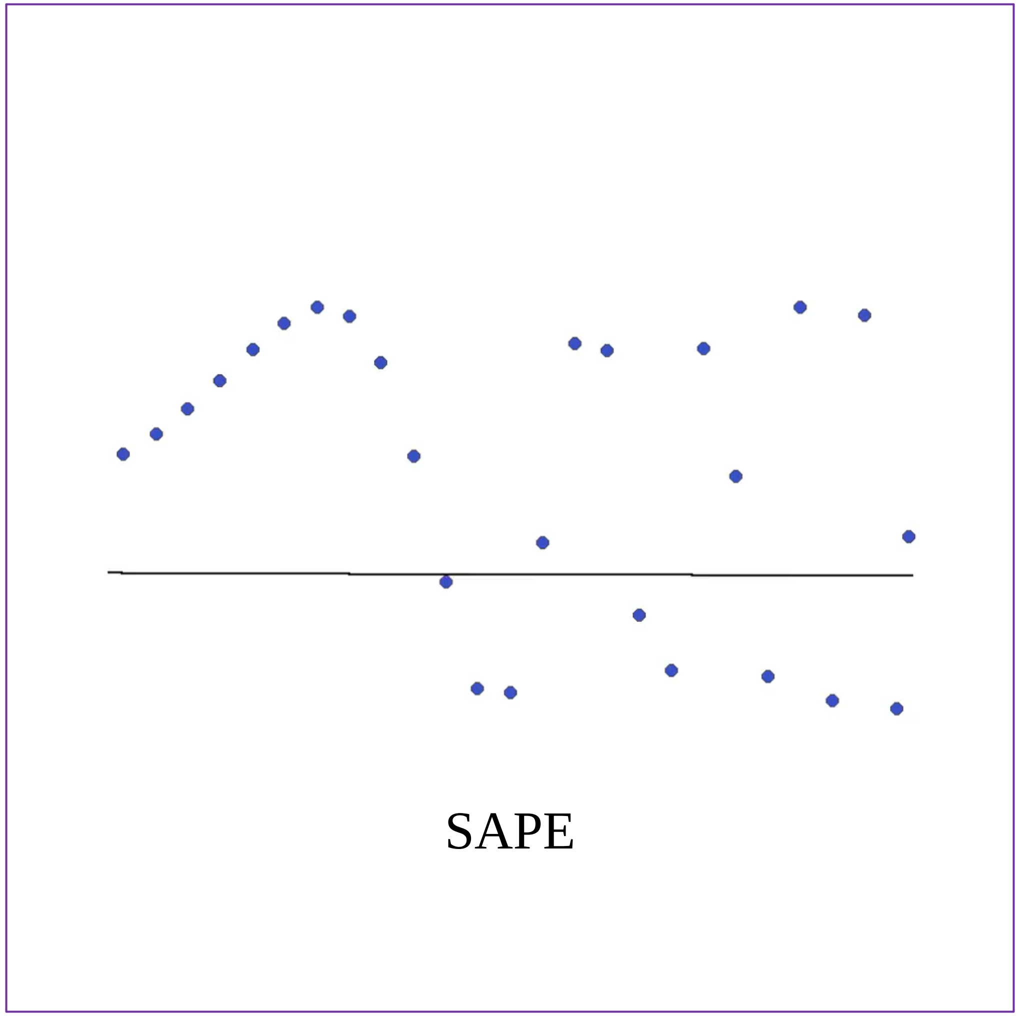


Spectral Bias
Low Frequencies
Globally Stable

Positional Encoding
High Frequencies
Fits Delicate Details

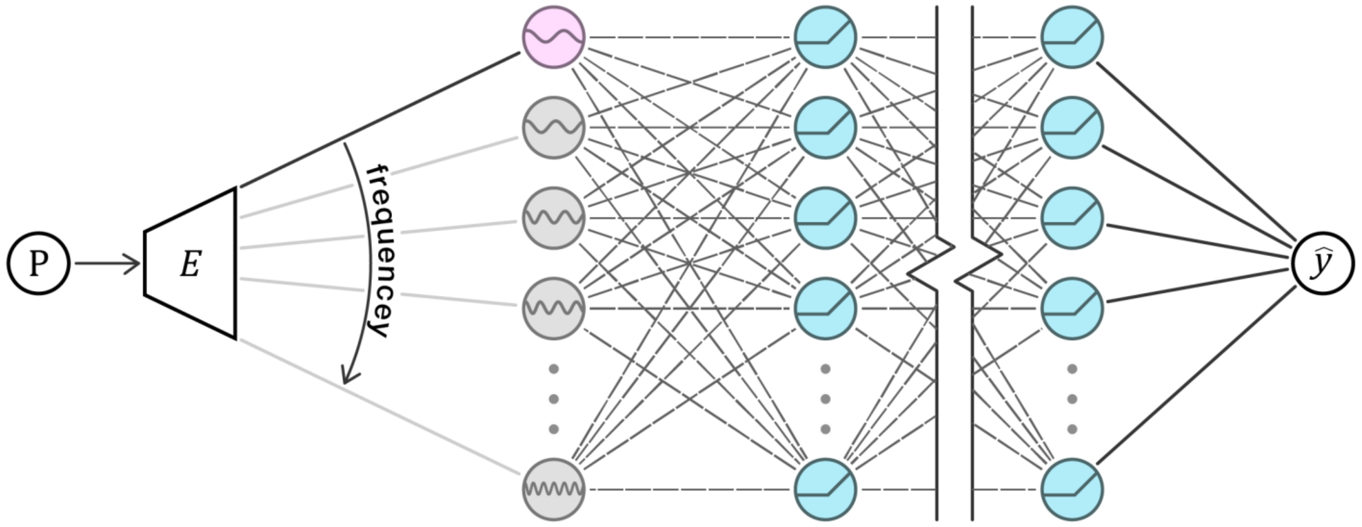
SAPE: The Best of Both Worlds

Spectral Bias
Low Frequencies
Globally Stable

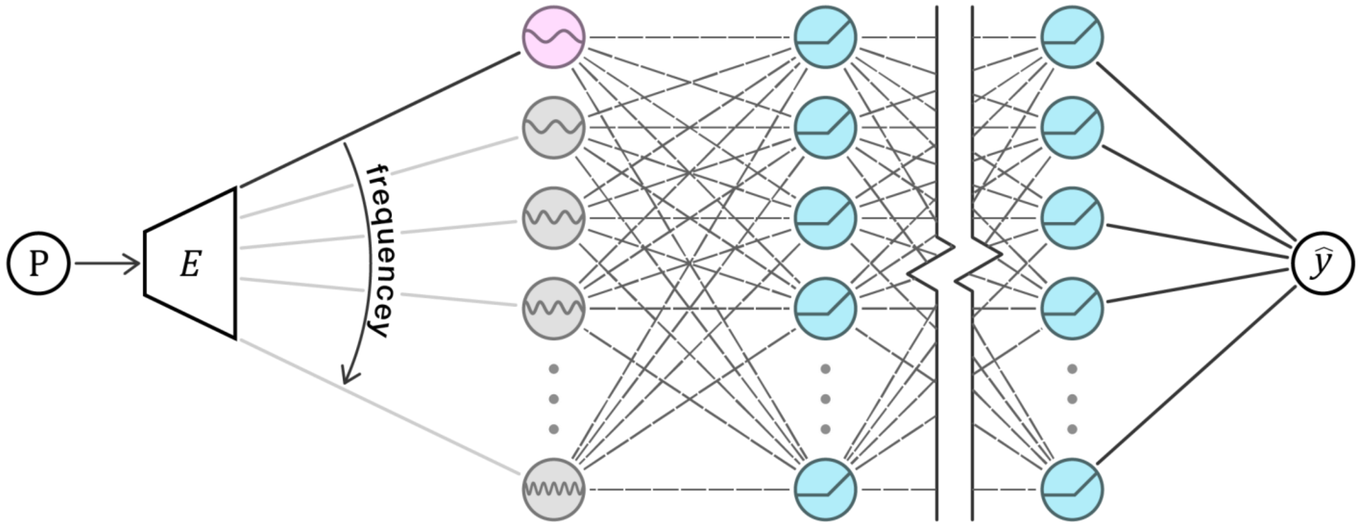


Positional Encoding
High Frequencies
Fits Delicate Details

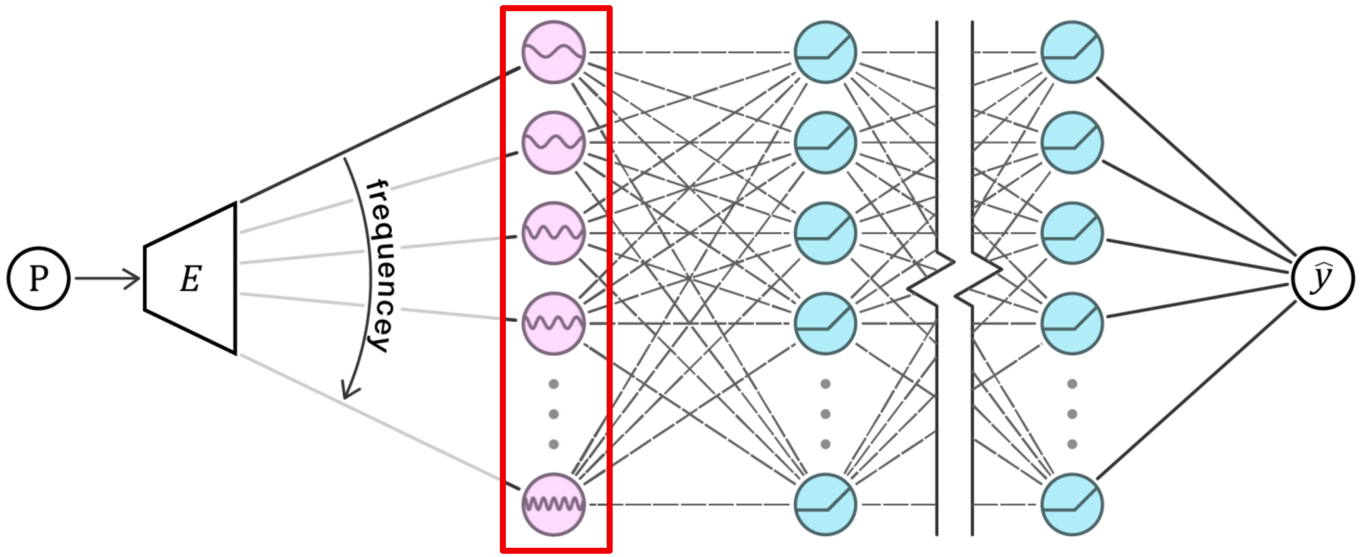
SAPE



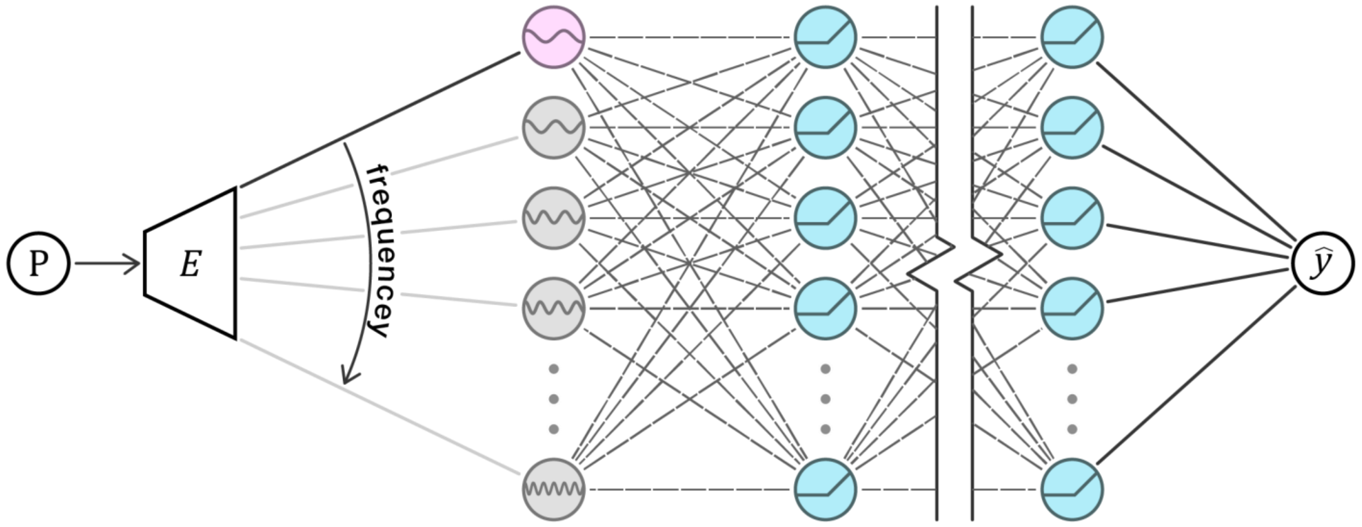
Progressive Encoding



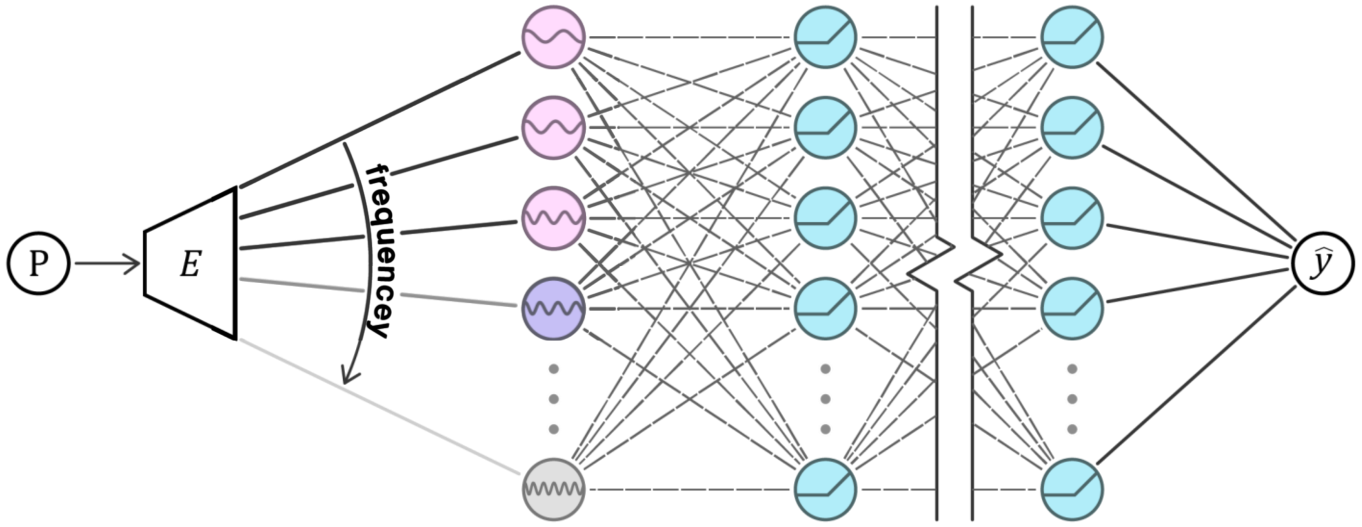
Progressive Encoding



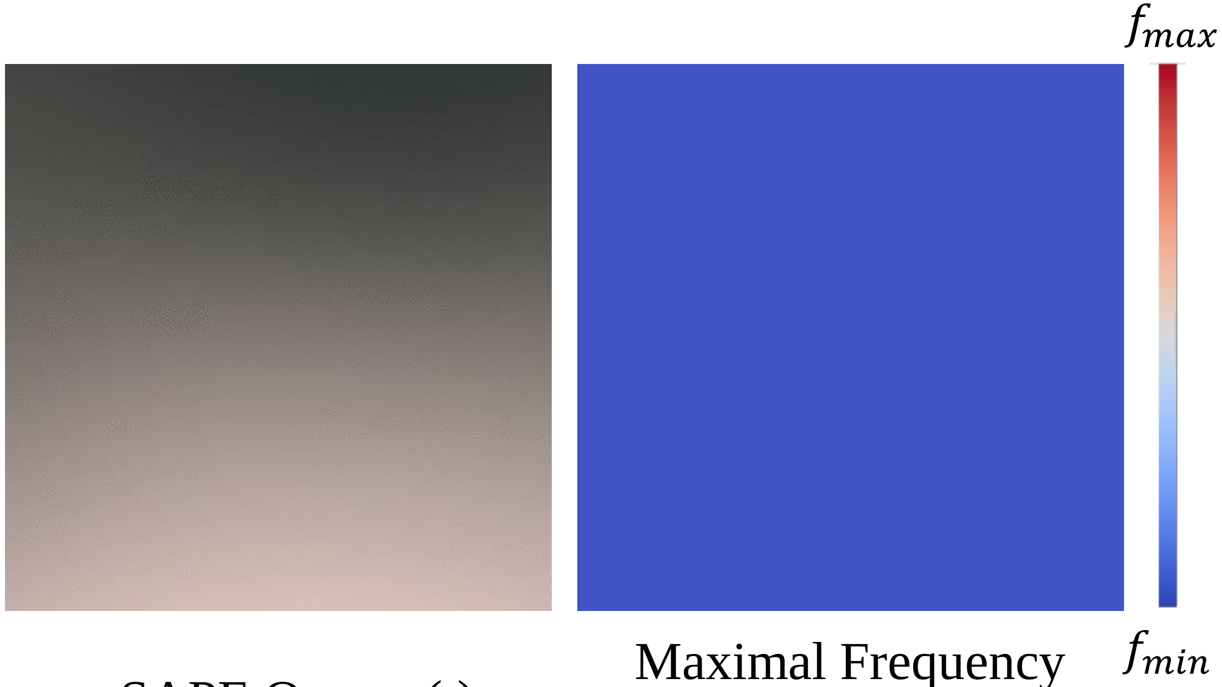
Progressive Encoding



Progressive Encoding



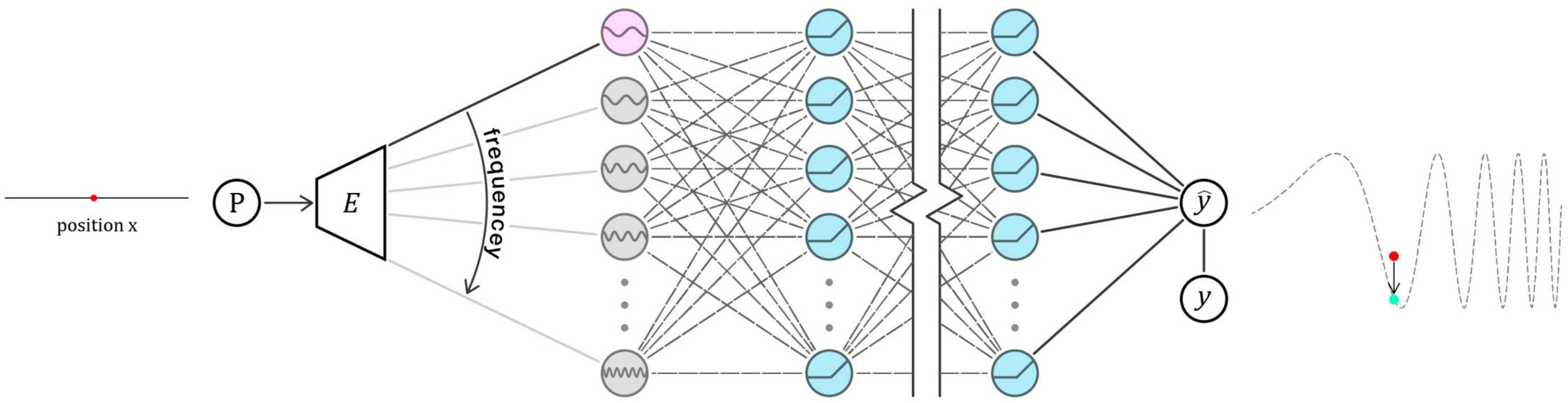
Spatially Adaptive



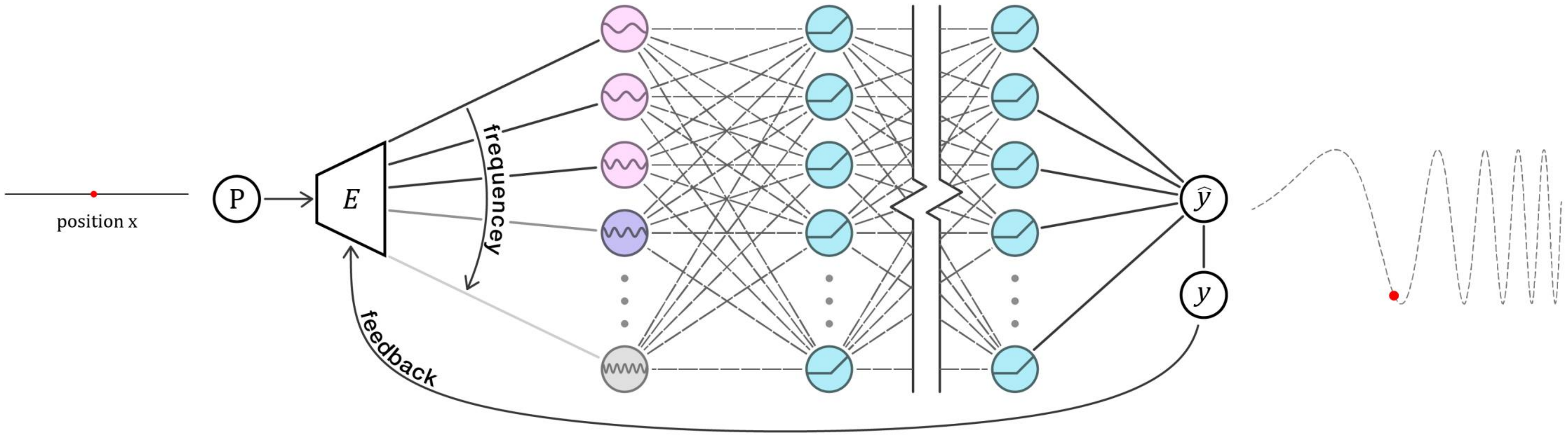
SAPE Output (t)

Maximal Frequency Used Per Spatial Location (t)

Spatially Adaptive



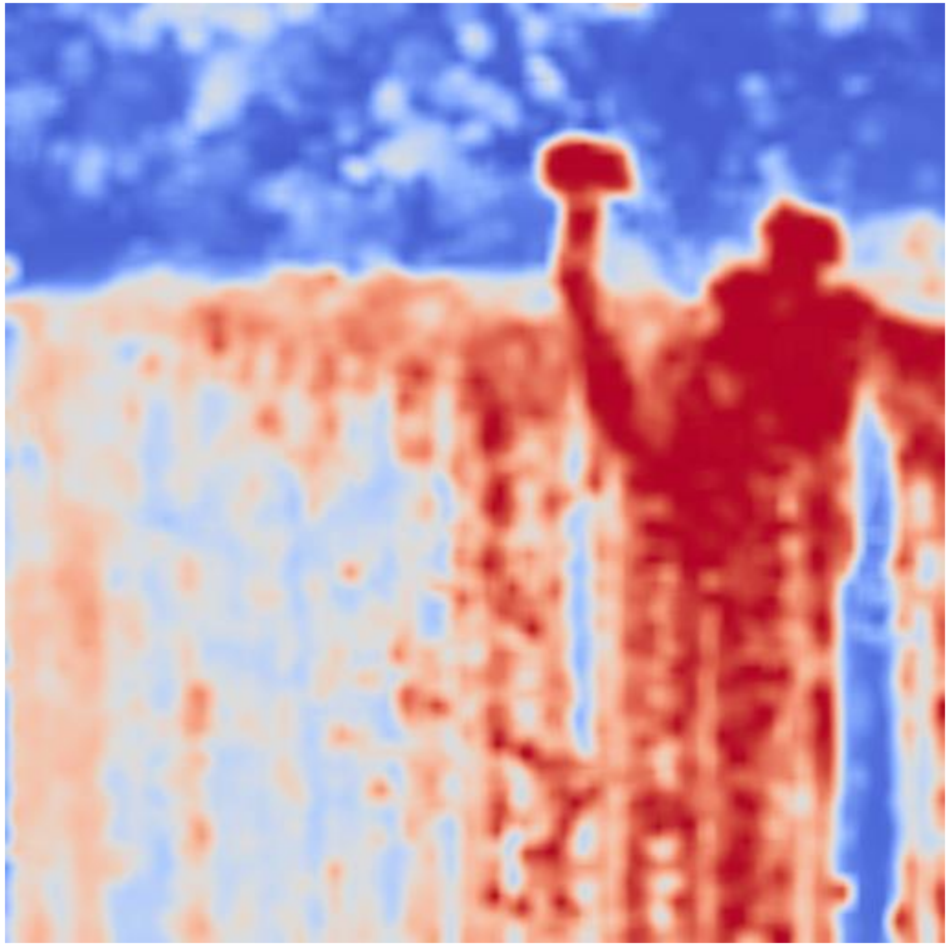
Spatially Adaptive

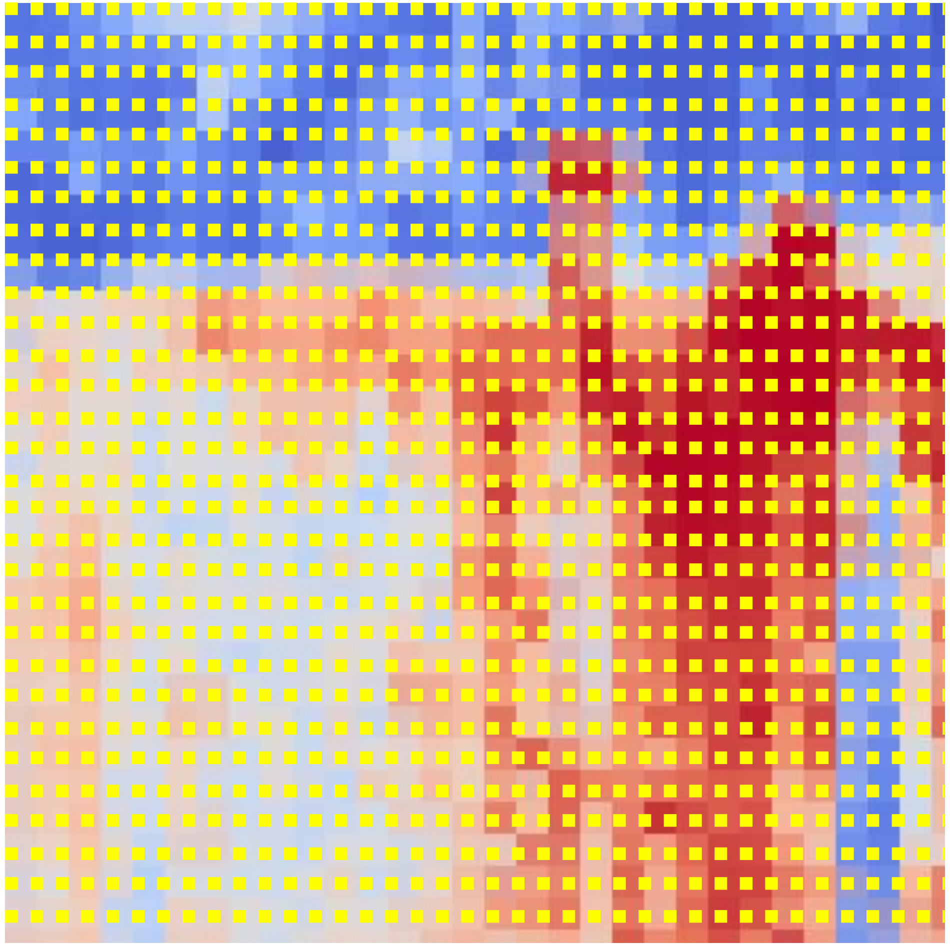


Spatially Adaptive

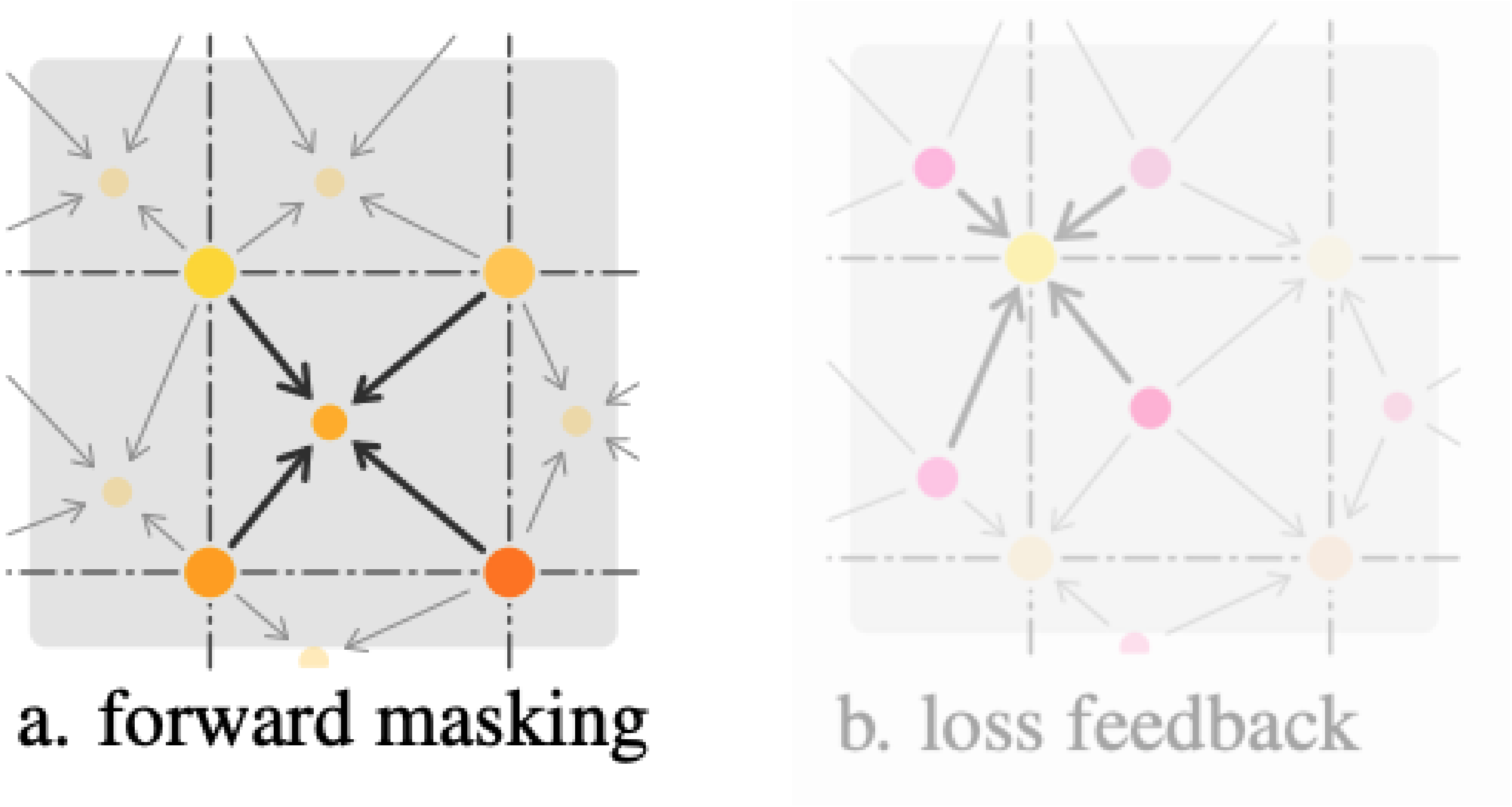
Figure 10: Non-spatial (left) vs. spatial encoding (middle). SAPE fit high frequencies while maintaining a smooth background using a different spatial encoding (right).

Adaptive Grid



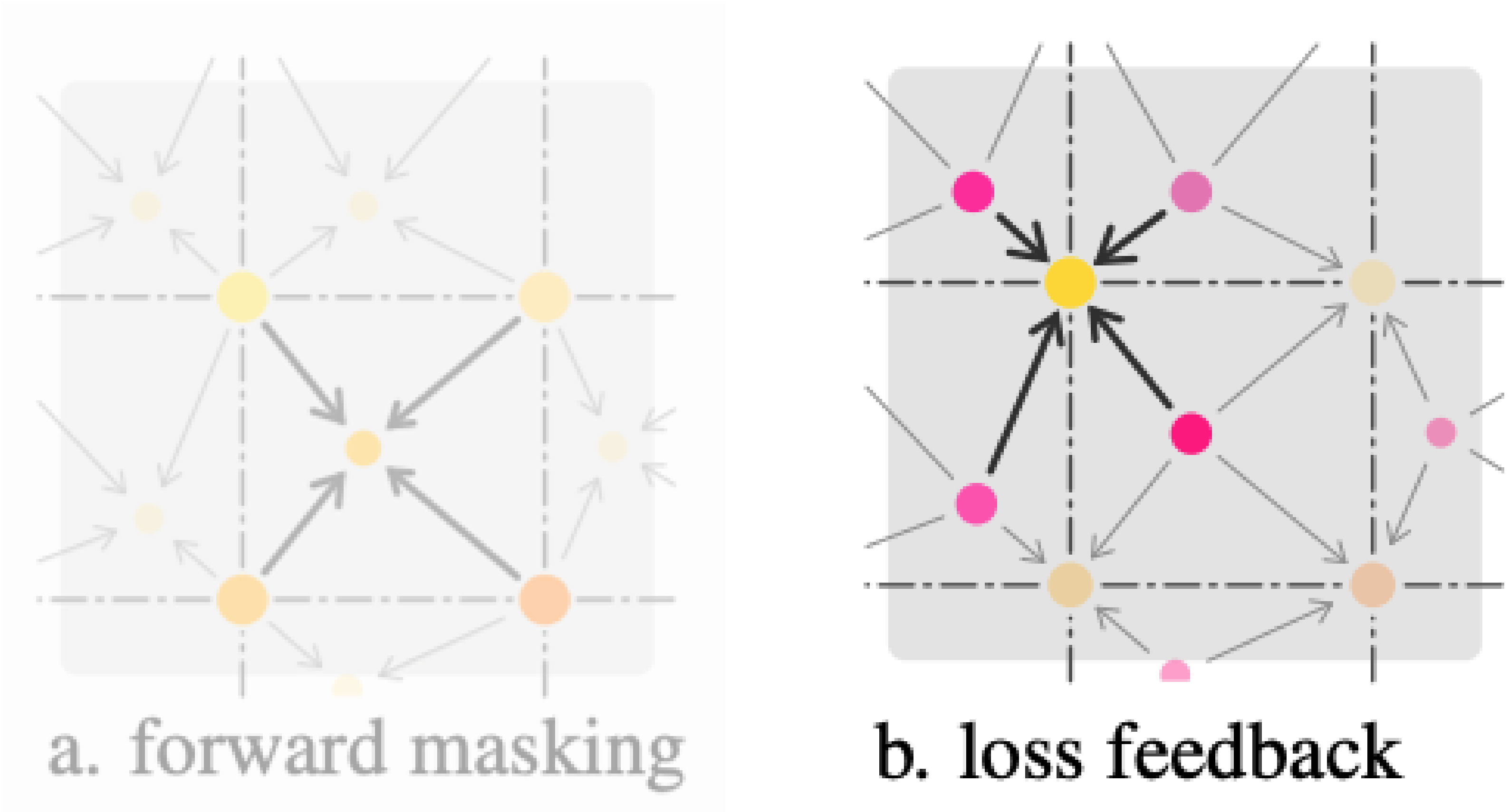


Adaptive Grid



$$\boldsymbol{\alpha}(t, \mathbf{p}) = \sum_{\mathbf{u} \in \mathcal{N}_G(\mathbf{p})} w_{p,\mathbf{u}} \boldsymbol{\alpha}(t, \mathbf{u})$$

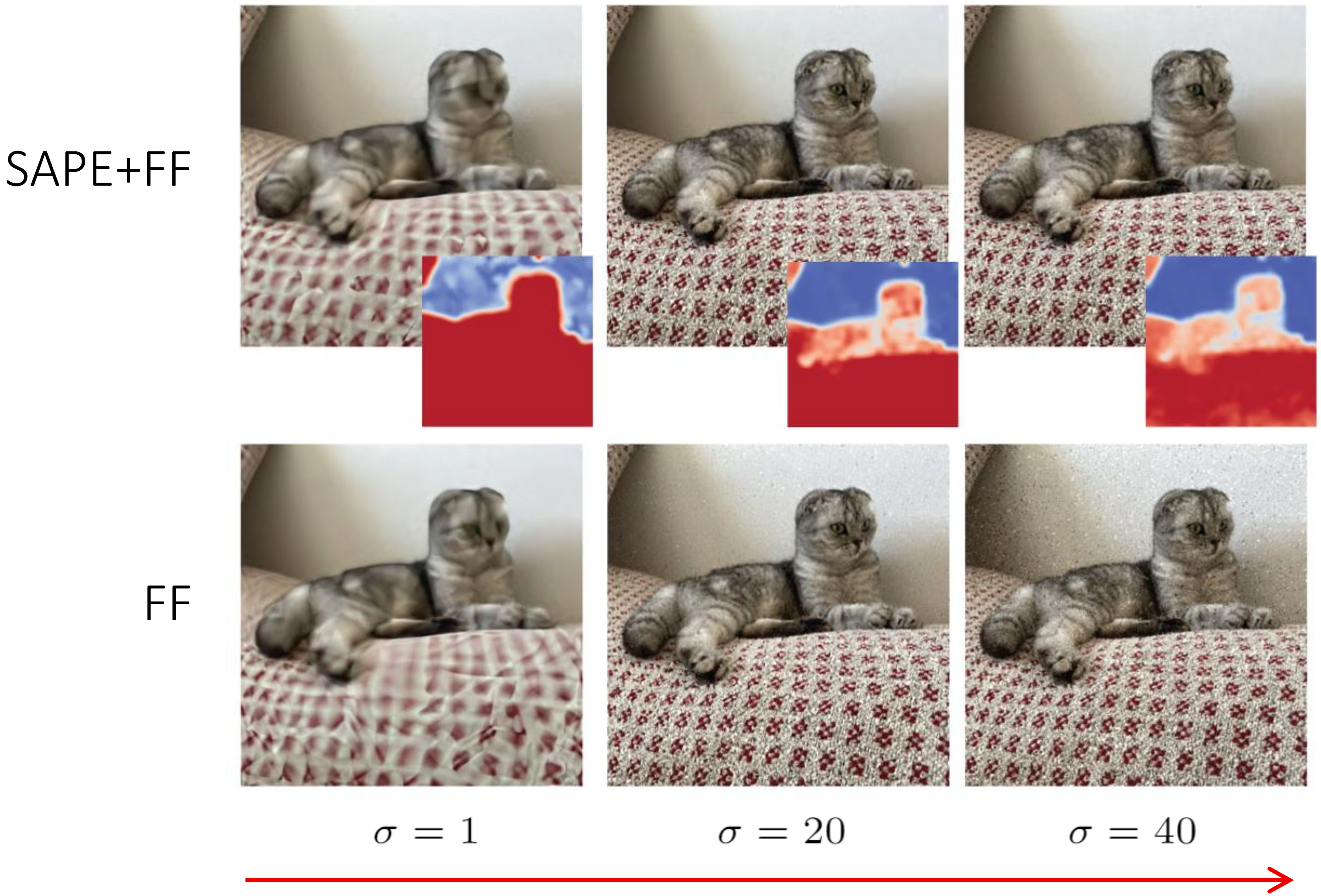
Adaptive Grid



$$\mathcal{L}(t,\mathbf{u}) = \frac{1}{\sum_{\mathbf{p}} w_{\mathbf{p},\mathbf{u}}} \sum_{\mathbf{p}} w_{\mathbf{p},\mathbf{u}} \mathcal{L}(t,\mathbf{p}).$$

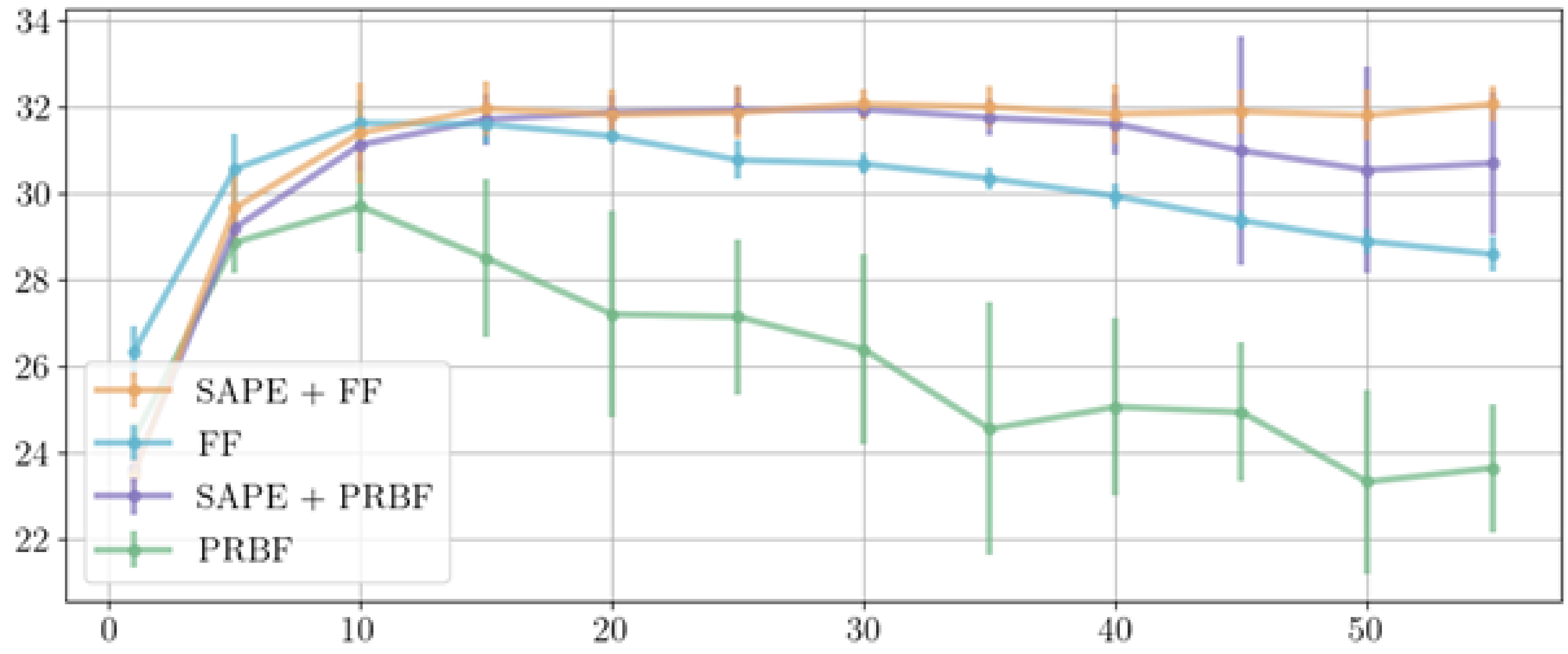
Results

Robustness to the Frequency Range



Robustness to the Frequency Range

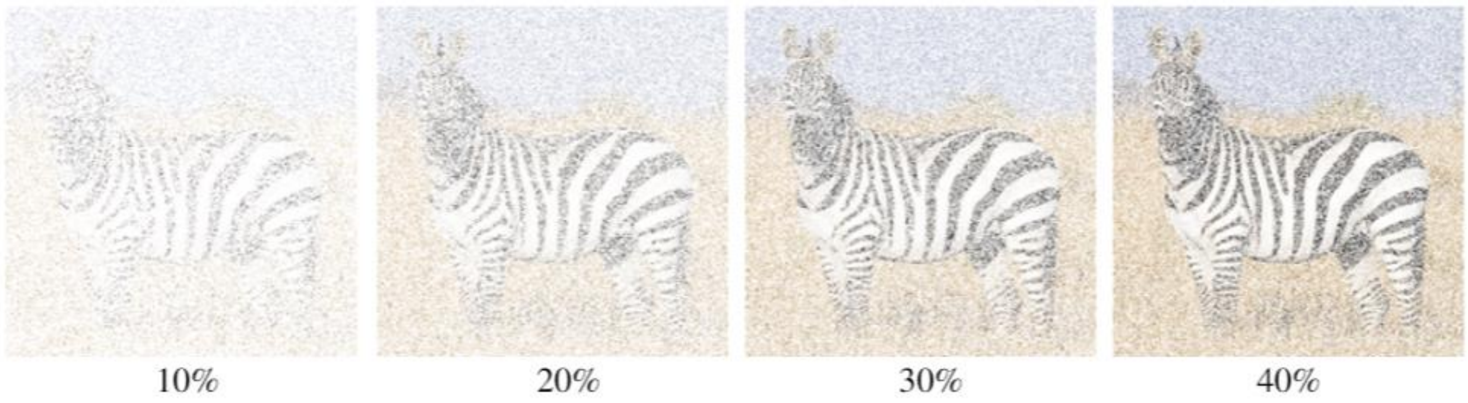
PSNR



Standard deviation of the frequencies

Robustness to Sample Size

Sampled pixels

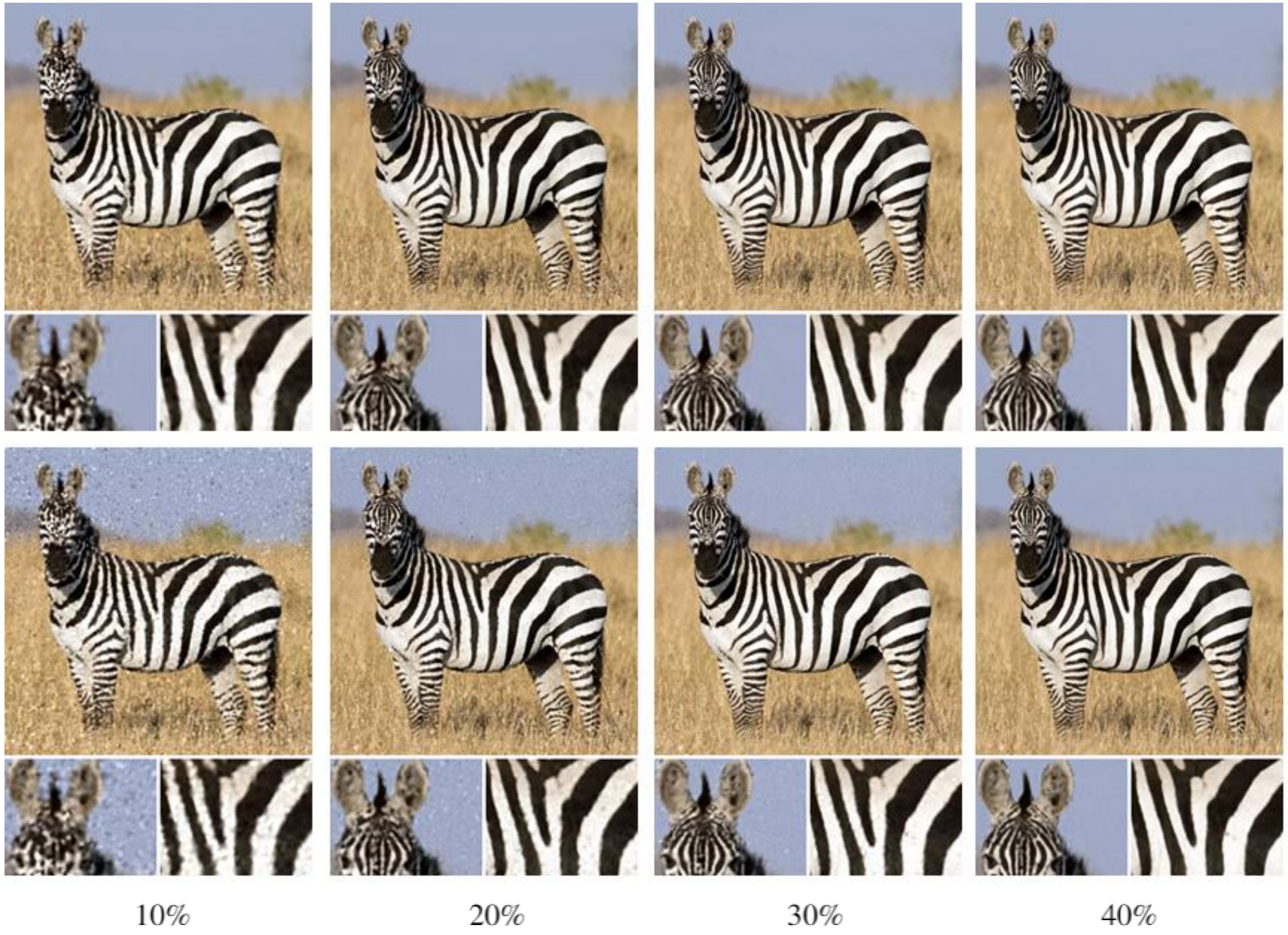


Sample size

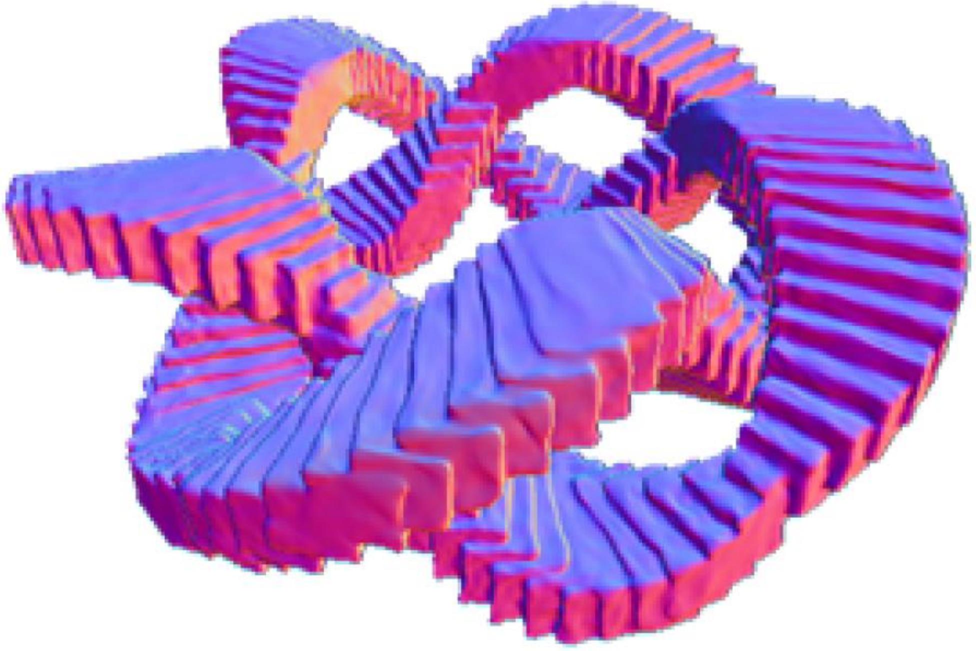
Robustness to Sample Rate

SAPE+FF

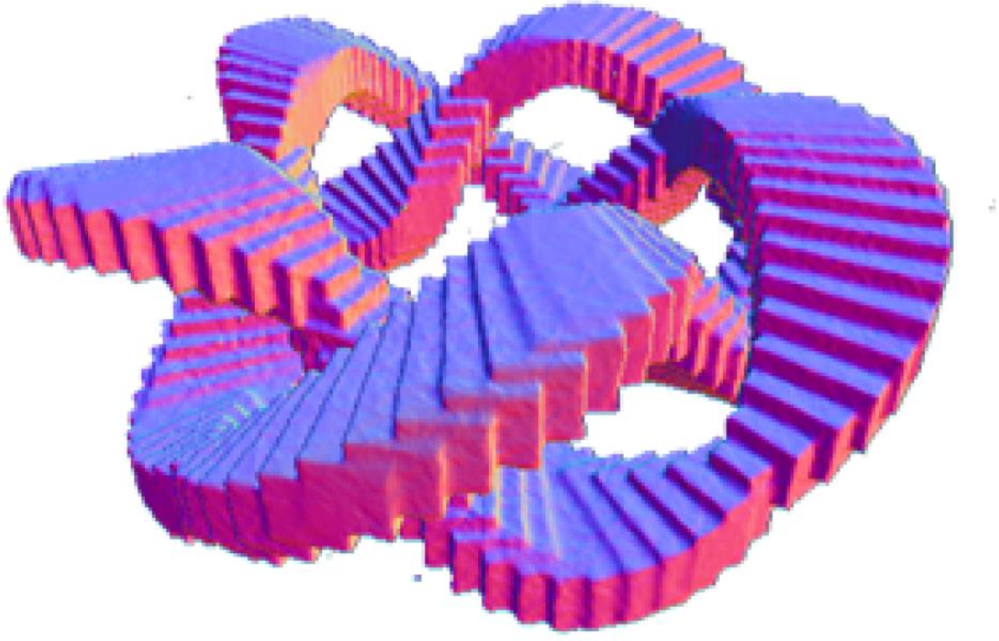
FF



Results: 3D Occupancy



Siren

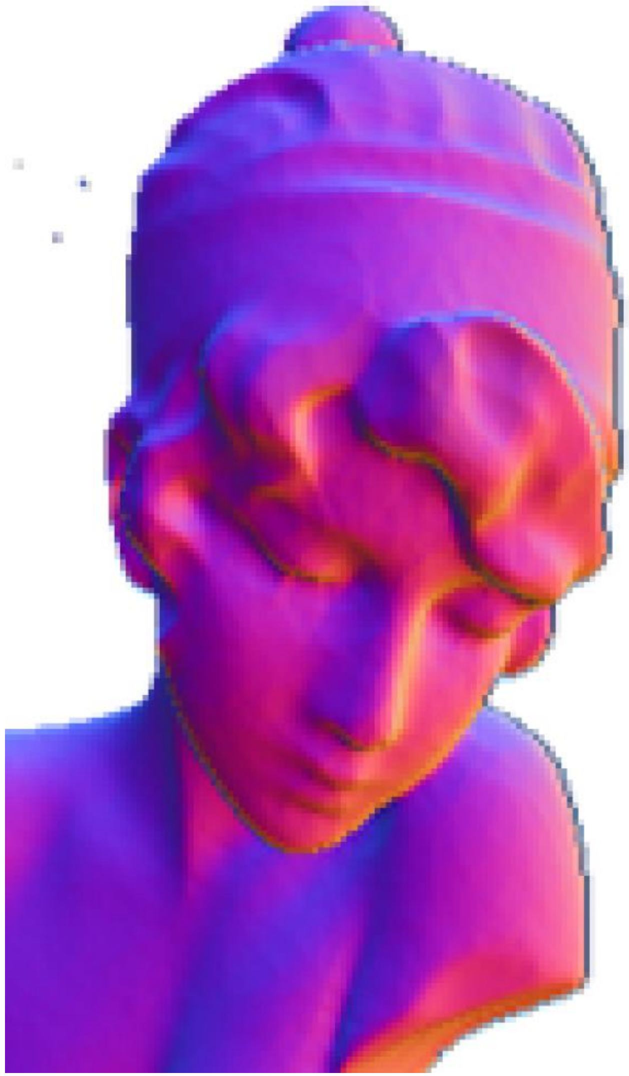


SAPE + FF

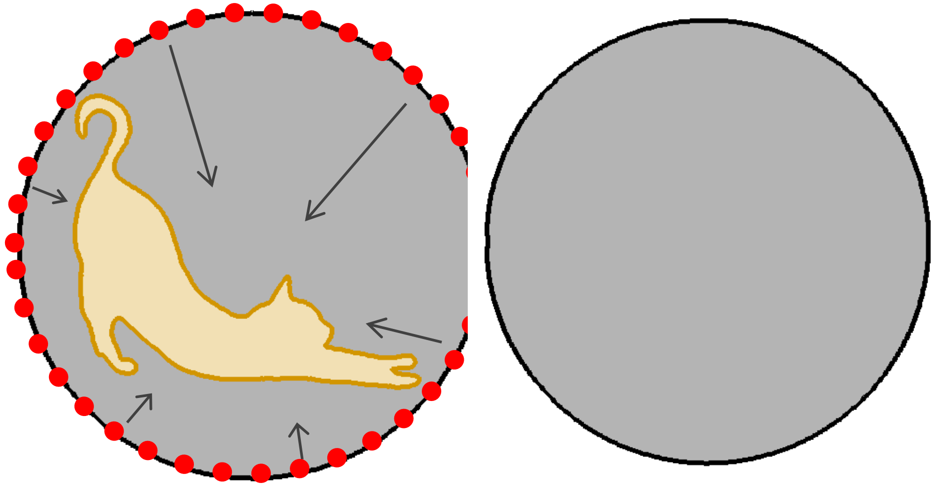
Results: 3D Occupancy

Siren

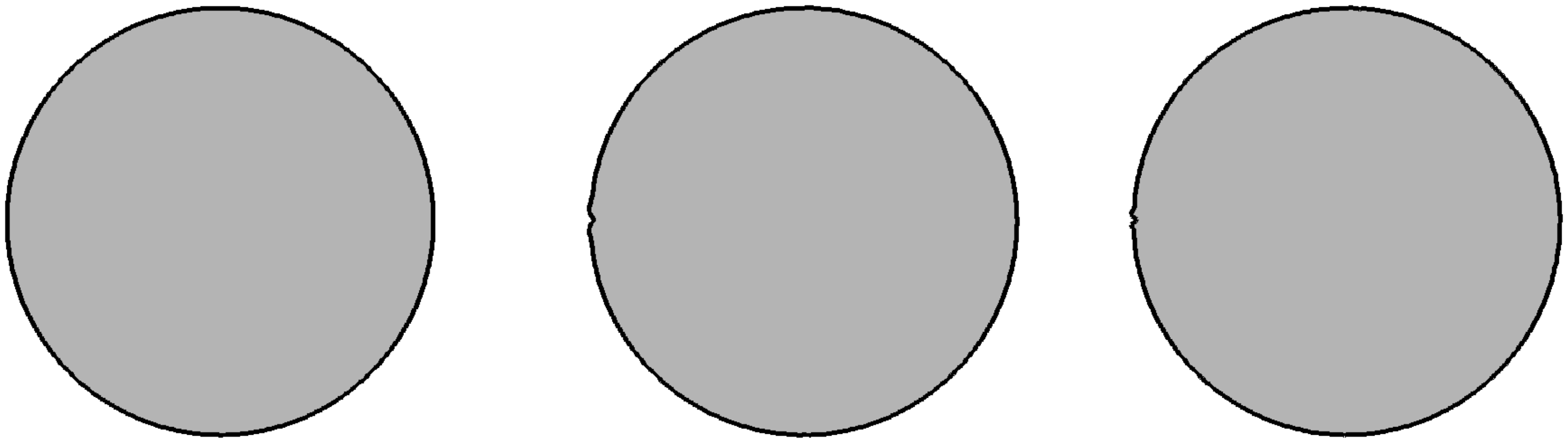
SAPE + FF



Results: 2D Silhouettes Deformation

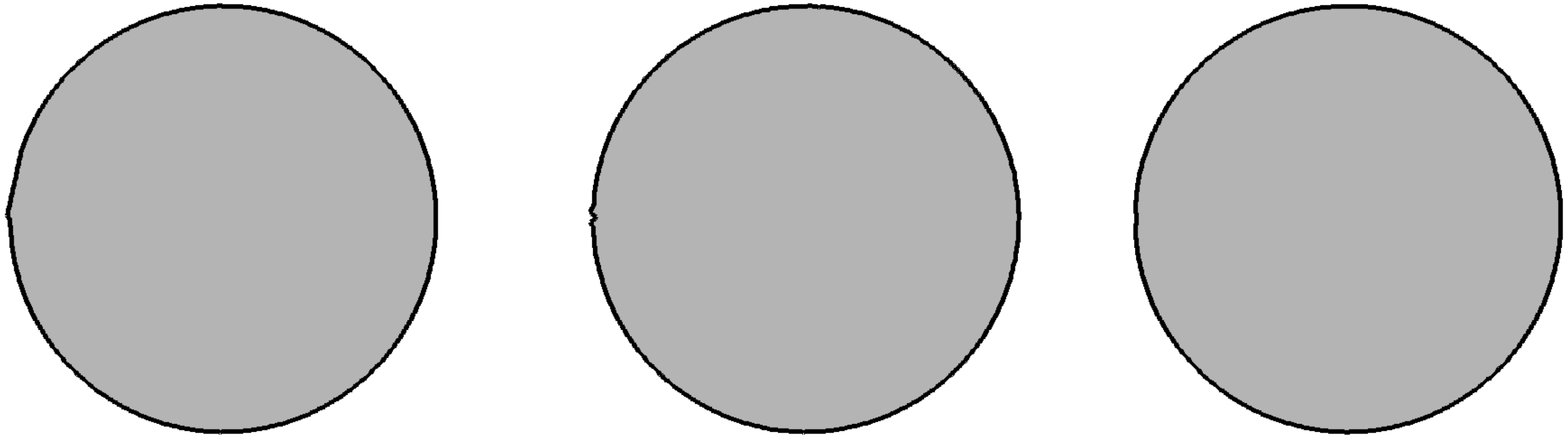


Results: 2D Silhouettes Deformation



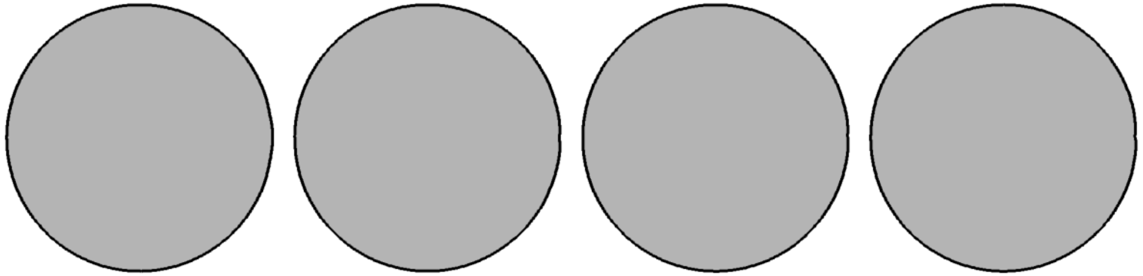
FF $\sigma = 1$ FF $\sigma = 3$ FF $\sigma = 5$

Results: 2D Silhouettes Deformation



No Encoding | Fourier Features | SAPE

Thank you! ☺



Visit our project page: amirhertz.github.io/sape