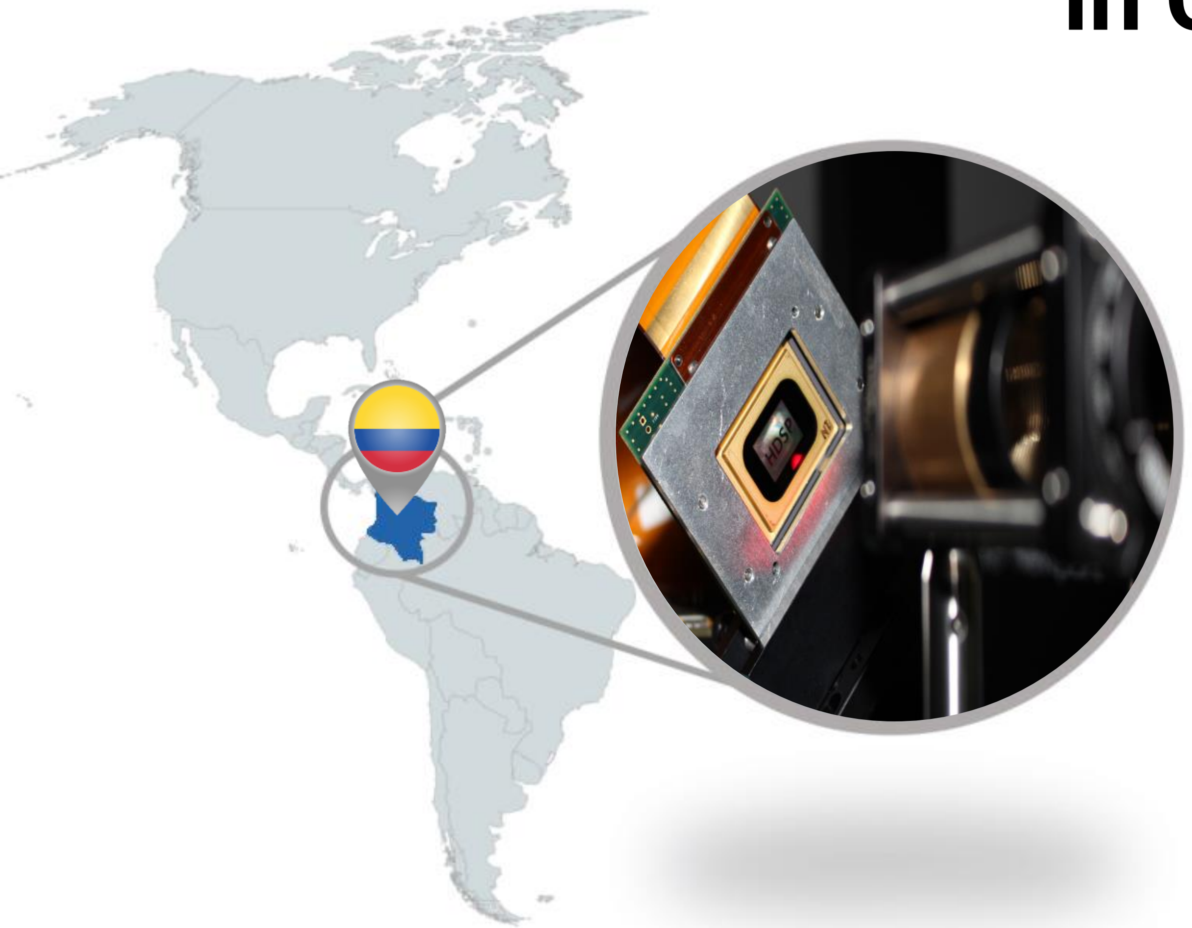
Data-Driven Optical Coding Optimization in Computational Imaging

HDSP



Henry Arguello, PhD
Universidad Industrial de Santander
Department of Computer Science
Bucaramanga, Colombia

www.hdspgroup.com
henarfu@uis.edu.co

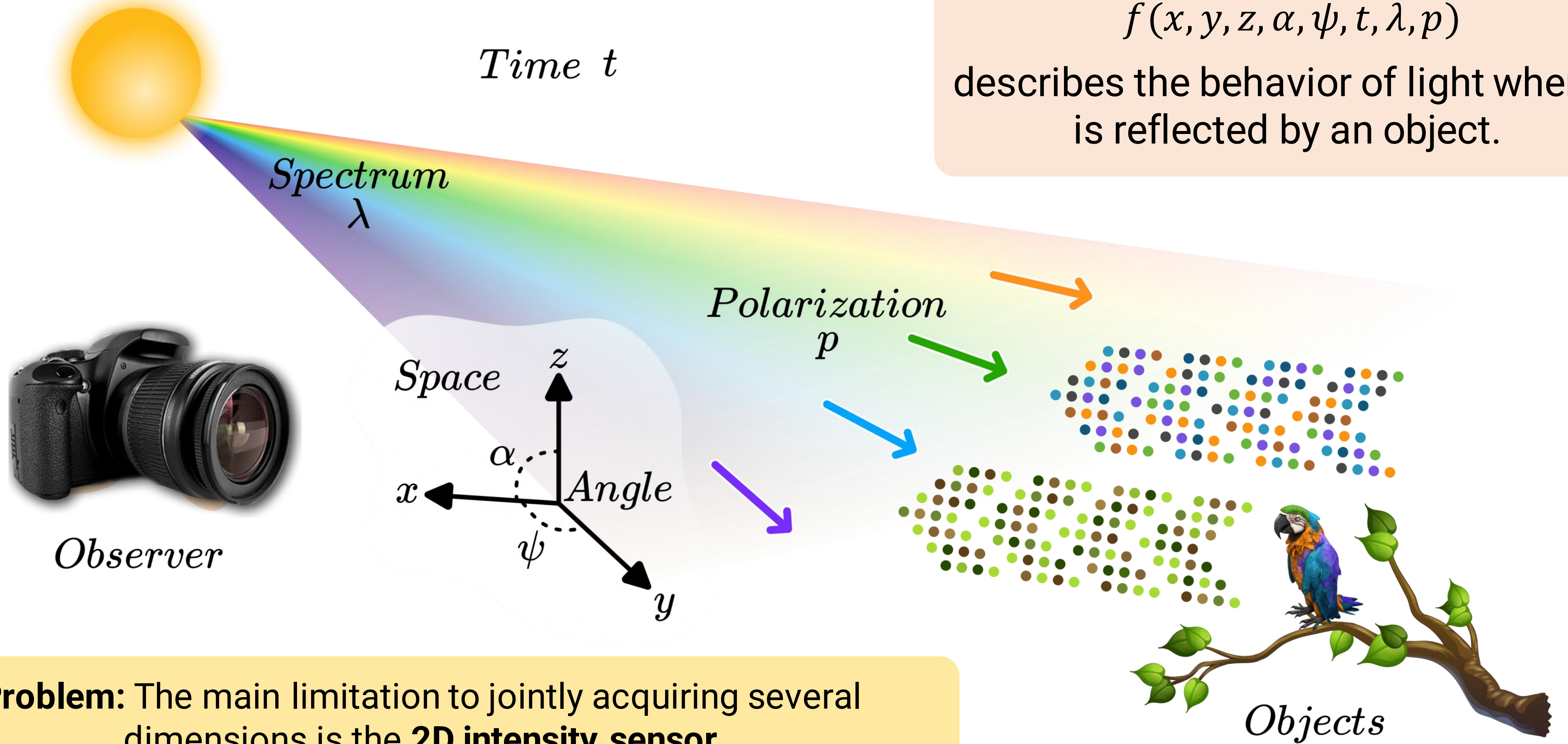
Universidad Industrial de Santander

Light

The plenoptic function[1]

$f(x, y, z, \alpha, \psi, t, \lambda, p)$

describes the behavior of light when it is reflected by an object.



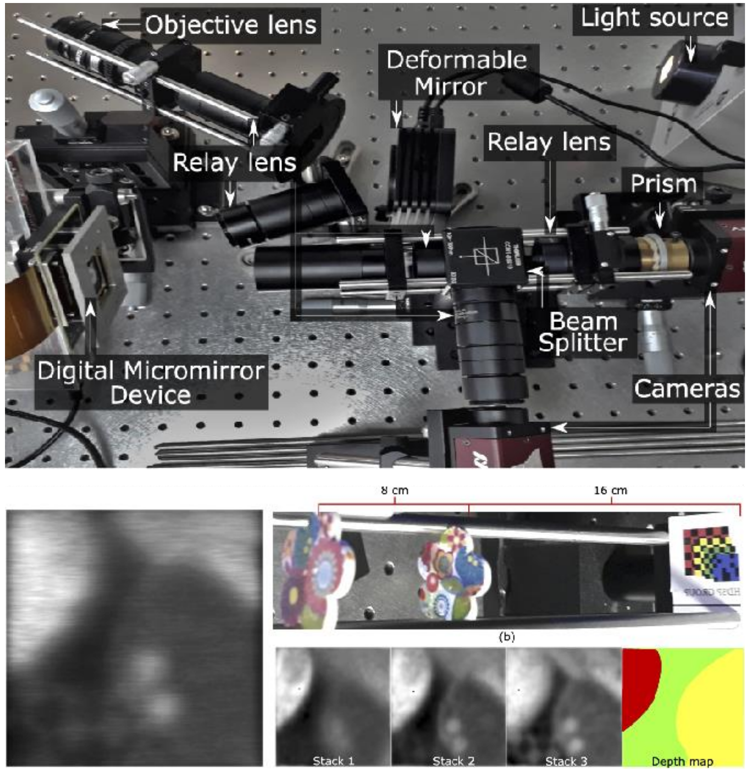
Problem: The main limitation to jointly acquiring several dimensions is the **2D intensity sensor**

1. Bergen, J. R., & Adelson, E. H. (1991). The plenoptic function and the elements of early vision. Computational models of visual processing.

Imaging Applications

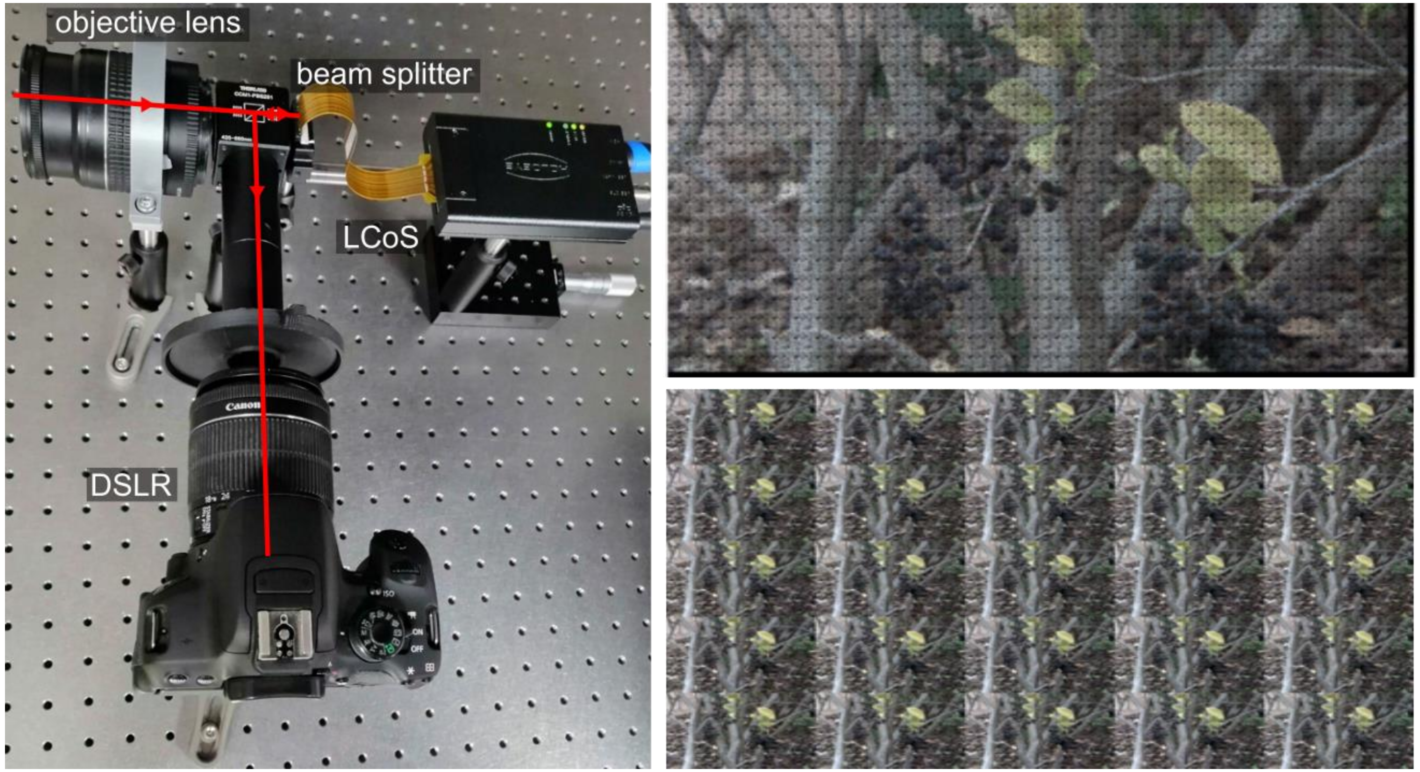
Images acquired with specialized cameras

Depth[1]



$$f(x,y,z)$$

Light Field[2]



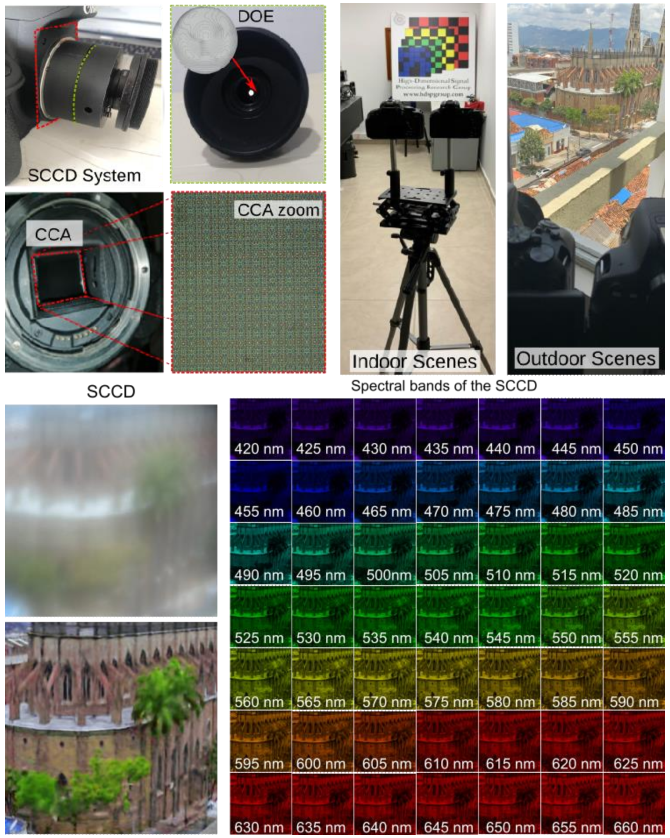
$$f(x,y,z,\alpha,\psi)$$

1. Marquez, *et al.* (2021). Snapshot compressive spectral depth imaging from coded aberrations. *Optics Express*.
2. Vargas, *et al.* (2021). Time-Multiplexed Coded Aperture Imaging: Learned Coded Aperture and Pixel Exposures for Compressive Imaging Systems. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

Imaging Applications

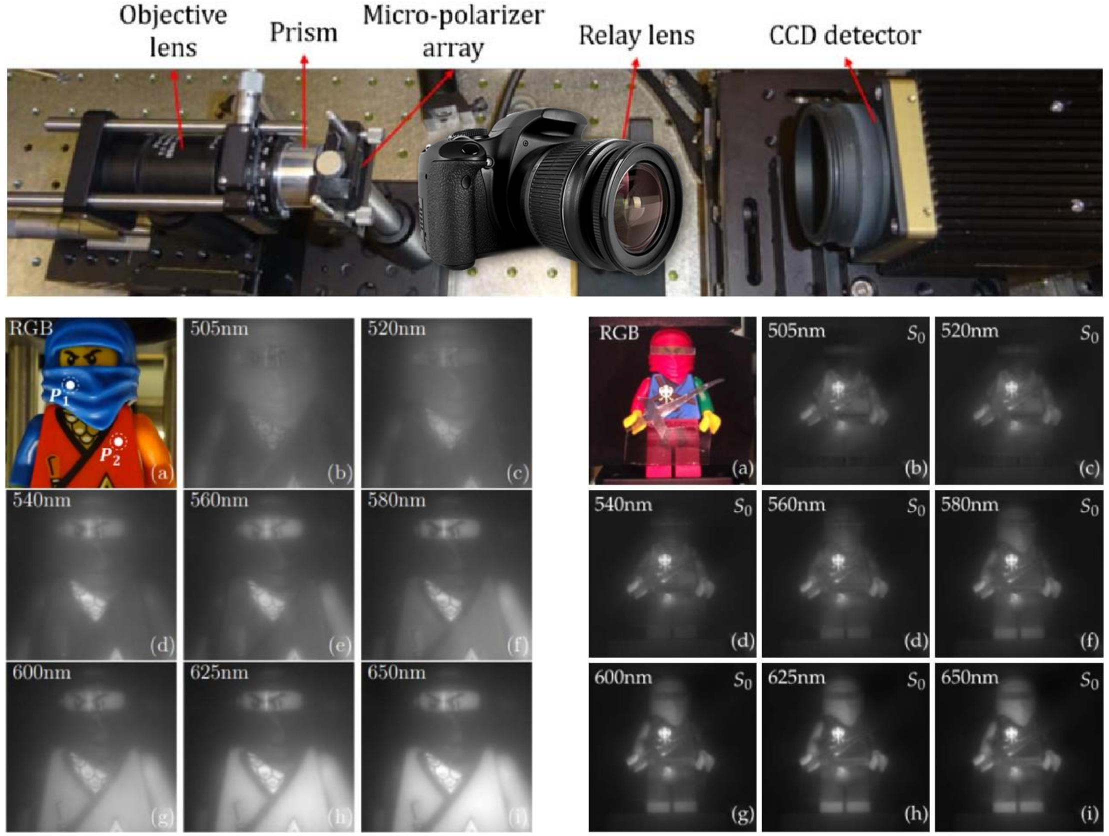
Images acquired with specialized cameras

Spectral[1]



$$f(x, y, \lambda)$$

Polarization[2]

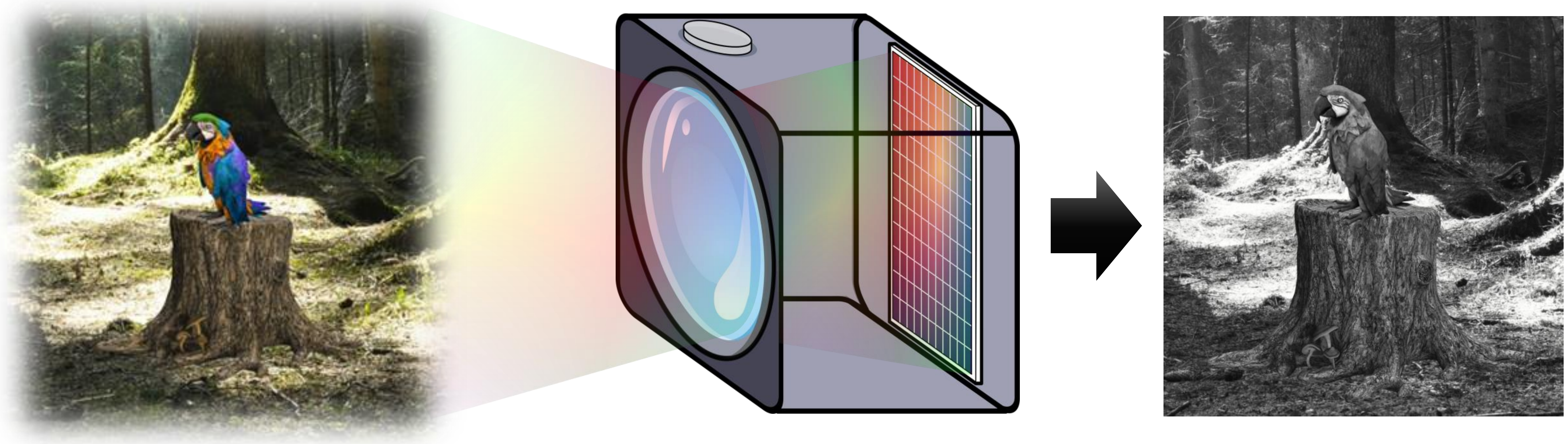


$$f(x, y, p)$$

1. Arguello, *et al*. (2021). Shift-variant color-coded diffractive spectral imaging system. *Optica*.
2. Fu, *et al*. (2015). Compressive spectral polarization imaging by a pixelized polarizer and colored patterned detector. *JOSA A*.

Imaging Applications

Scene | **Optical System** | **Image**



$f(x, y, z, \alpha, \psi, t, \lambda, p)$

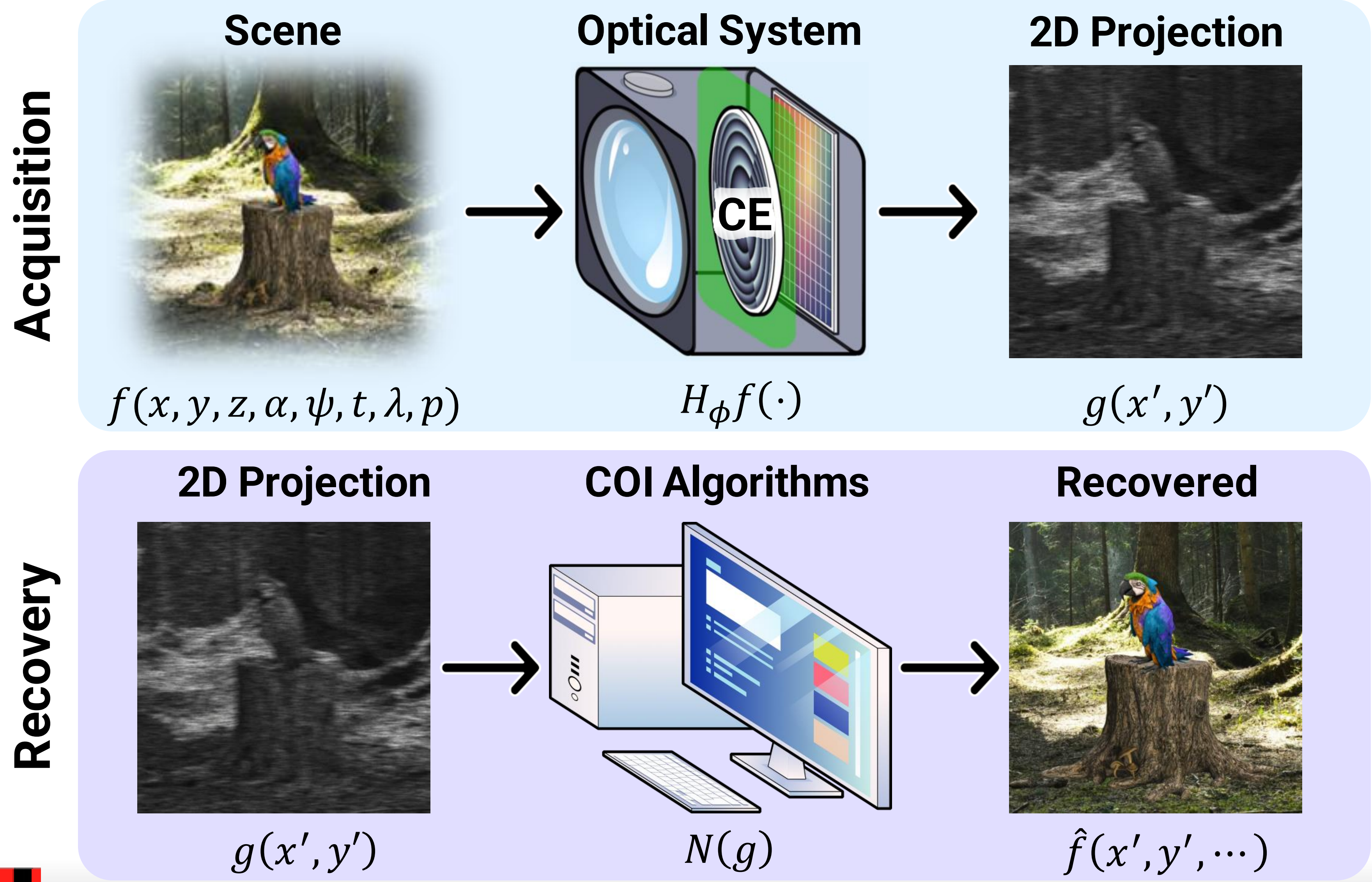
Lens

2D Sensor

$g(x', y') = \int_v f(\cdot) dv$

Current 2D sensors acquire the **photon flux of the incoming light**. Therefore, it is necessary **expensive** setup to obtain 3D or higher dimensional signals.

Computational Optical Imaging



CE
Coded Elements

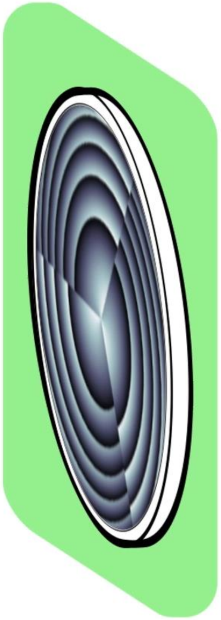
Low-cost acquisition can be done through specialized cameras and recovered through computational optical algorithms

COI
Computational Optical Imaging

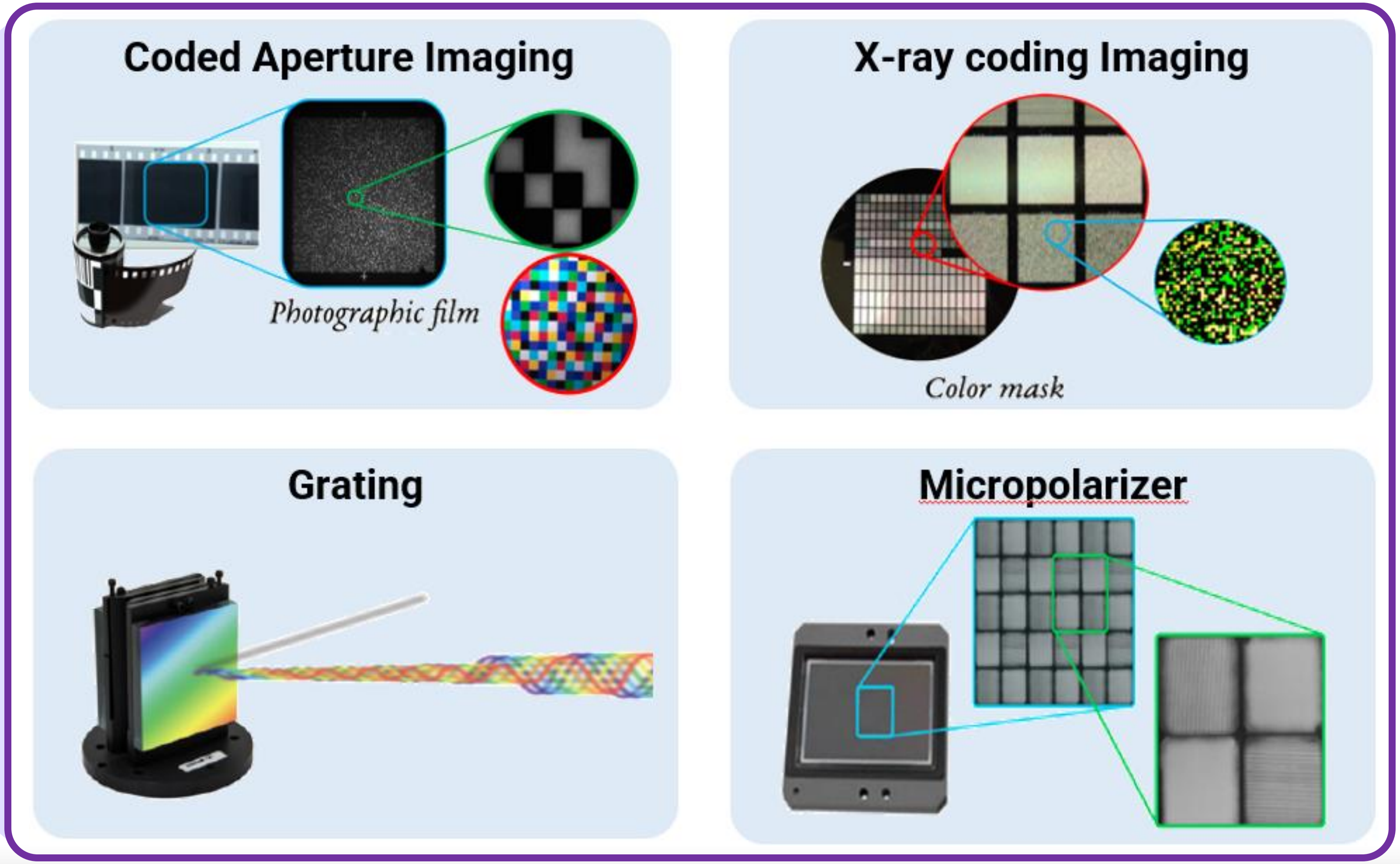
Computational Optical Imaging

Some Computational imaging tasks can be achieved thanks to the **Coded Elements:**

Coded Elements



ϕ_j

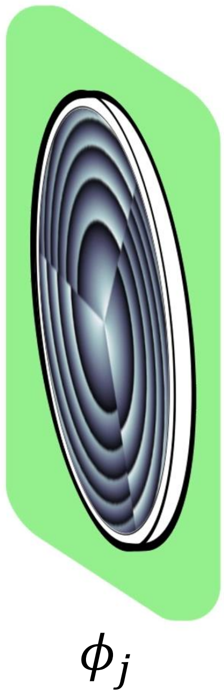


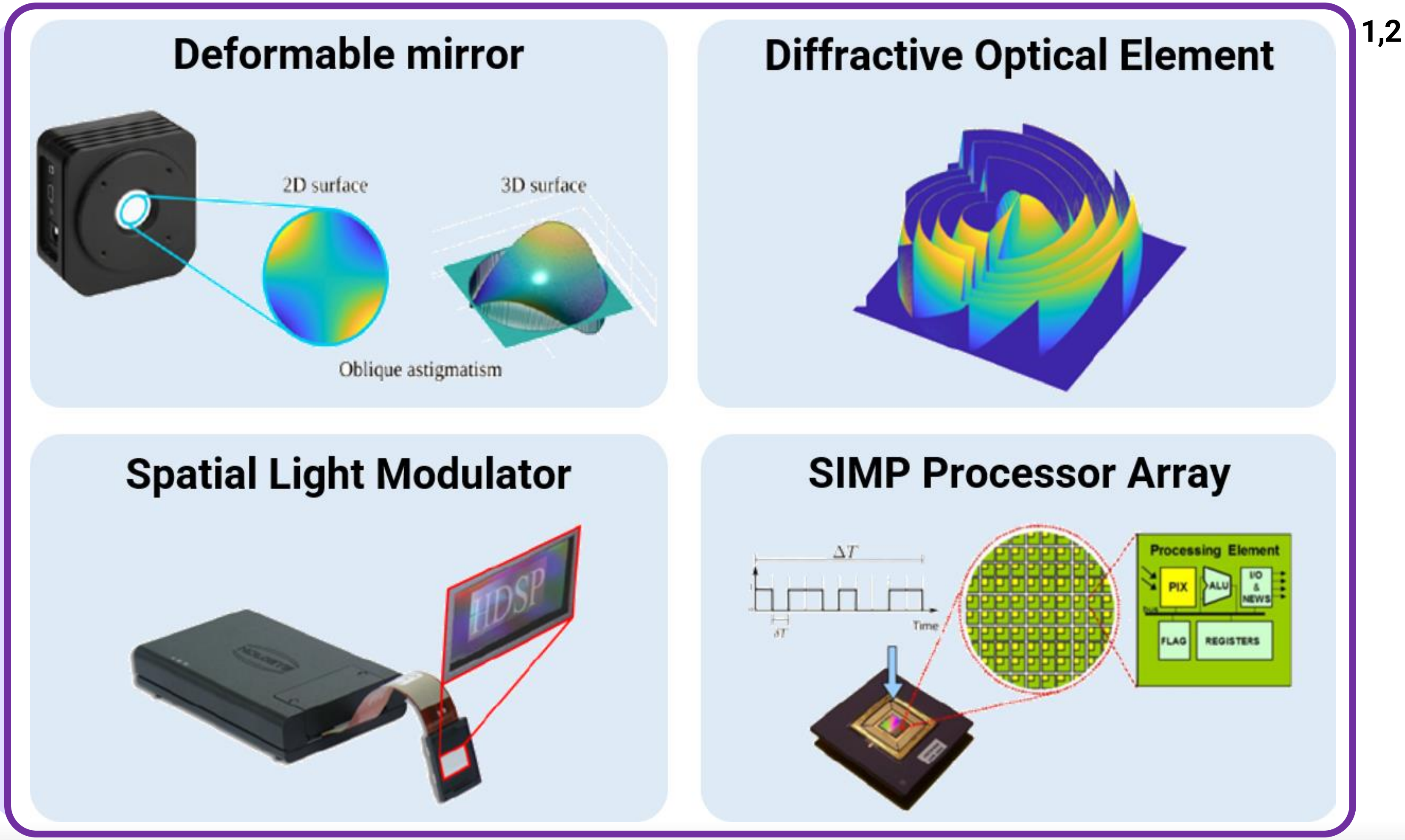
[1,2,3]

1. Rueda, et al. (2015) DMD-based implementation of patterned optical filter arrays for compressive spectral imaging. JOSA.
2. Rueda, et al. (2016) Compressive spectral testbed imaging system based on thin-film color-patterned filter arrays. Appl. Opt.
3. Pinilla, et al. (2018) "Coded diffraction system in X-ray crystallography using a boolean phase coded aperture approximation. Opt. Comm.

Computational Optical Imaging

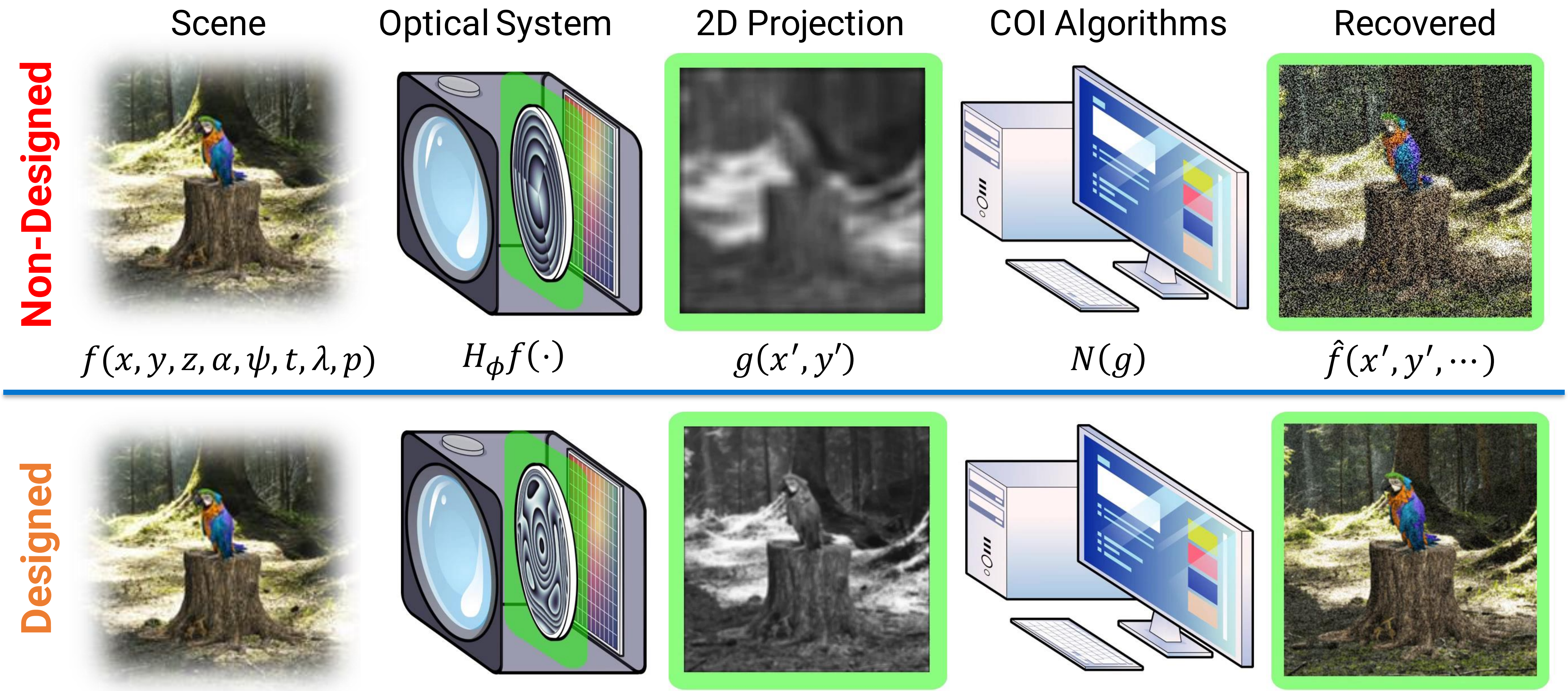
Some Computational imaging tasks can be achieved thanks to the **Coded Elements:**





1. Marquez, et al. (2019) Compressive spectral imaging via deformable mirror and colored-mosaic detector. Opt. Exp.
2. Bacca, et al. (2018) Single pixel compressive spectral polarization imaging using a movable micro-polarizer array. Rev. Fac. de Ing. Univ. de Ant.

Coded Elements



If the coded element is **designed**, the quality of the reconstruction **increases**.

Designed Coded Elements

Without data

$$\mathcal{R}(\mathbf{H}_{\phi_1}) = \|\mathbf{H}_{\phi_1}^{\top}\mathbf{H}_{\phi_1} - \mathbf{I}_n\|_2^2$$

Depends on its own CE

Algorithm

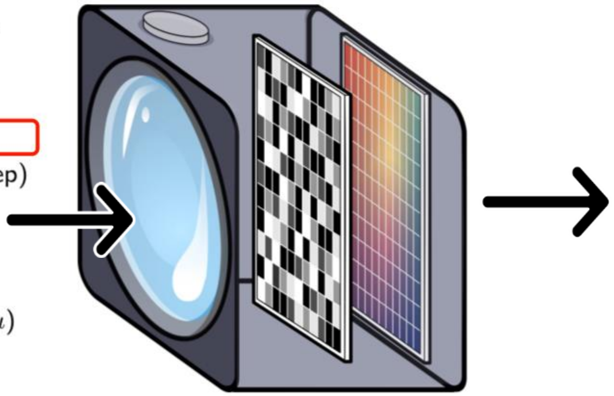
Step 0 Constants $\gamma, \gamma_1 \in (0,1)$, T maximum number of iterations, and $\mu_0 > 0$.

Step 2 (Main loop)

For $i = 0 : T-1$ do
$\mathbf{H}_{i+1} = \mathbf{H}_i - \alpha_i \partial g_{\mu_i}(\mathbf{H}_i)$ (update step)
if $\|\partial g_{\mu_i}(\mathbf{H}_{i+1})\|_F \geq \gamma \mu_i$ then
$\mu_{i+1} = \mu_i$
else
$\mu_{i+1} = \gamma_1 \mu_i$ (Key step: decreasing μ)
end if
end for

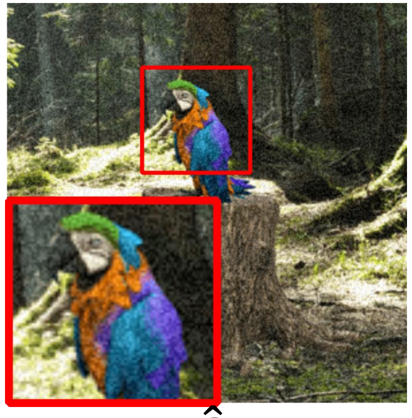
Step 3 (Return) $\mathbf{H}_T$

Optical System



ϕ_1

Recovery



$\hat{\mathbf{f}}_j$

Data Driven

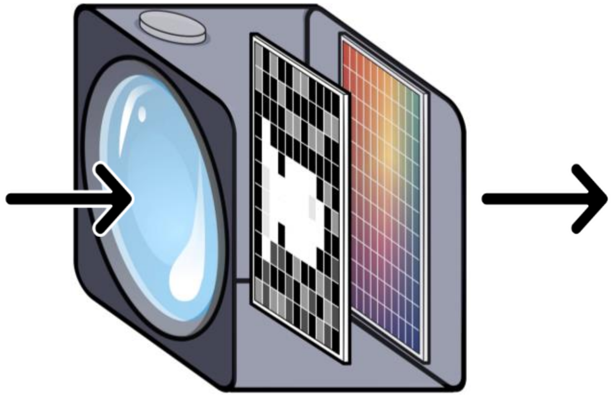
$$\mathcal{R}(\mathbf{H}_{\phi_2}\mathbf{f}_k) = \|\mathbf{f}_k - \mathbf{H}_{\phi_2}^{\top}\mathbf{H}_{\phi_2}\mathbf{f}_k\|_2^2$$

Depends on data

Dataset

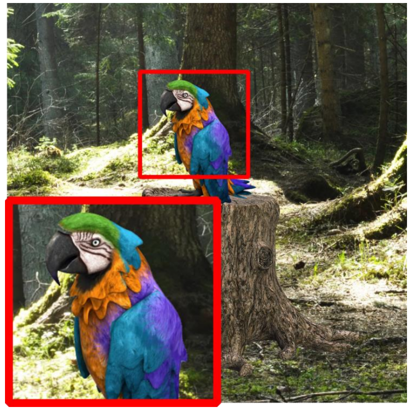
$\mathbf{f}_k$

Optical System



ϕ_2

Recovery



$\hat{\mathbf{f}}_j$

Non-Designed

Designed

Random

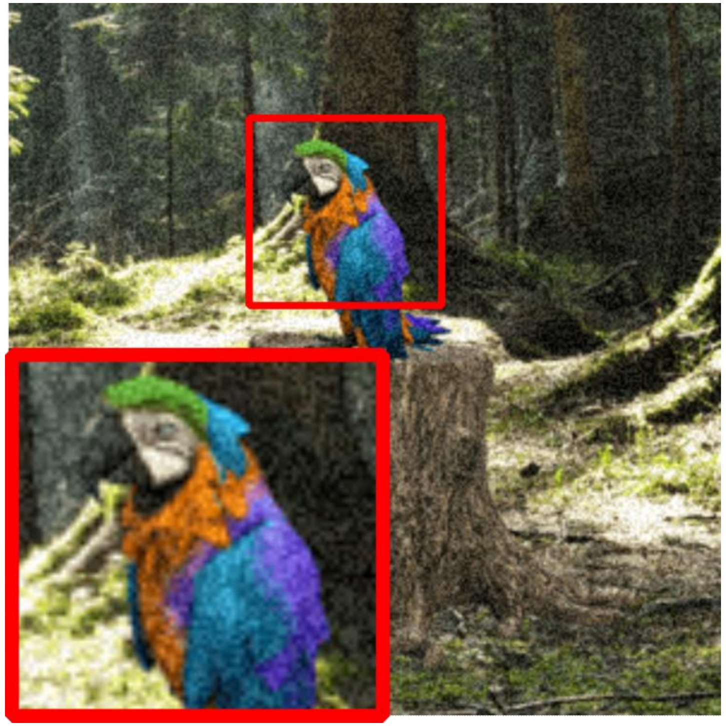
$\mathbf{H}_{\phi_{ij}} = Bernoulli(0,5)$

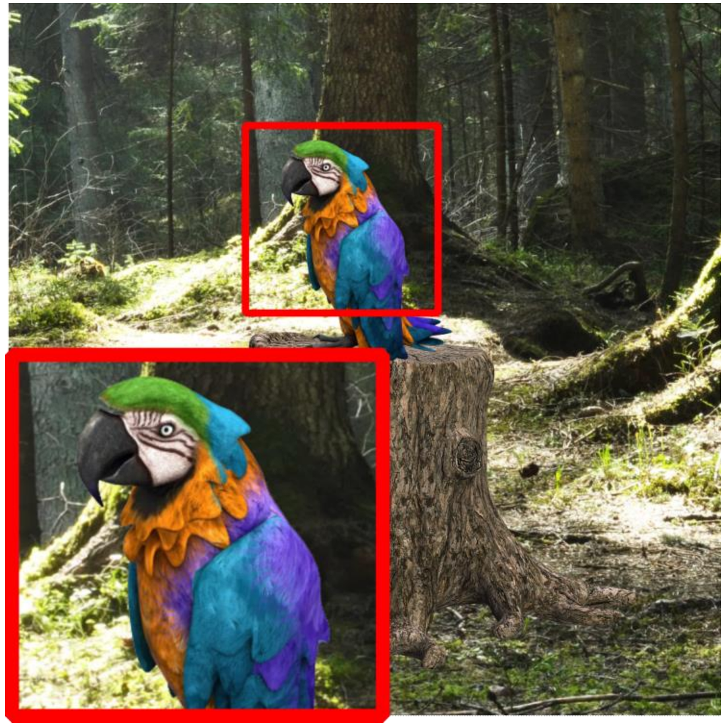
Without Data

$\mathcal{R}(\mathbf{H}_\phi) = \|\mathbf{H}_\phi^\top \mathbf{H}_\phi - \mathbf{I}_n\|_2^2$

Data-Driven

$\mathcal{R}(\mathbf{H}_\phi \mathbf{f}_k) = \|\mathbf{f}_k - \mathbf{H}_\phi^\top \mathbf{H}_\phi \mathbf{f}_k\|_2^2$





Random pattern produces low reconstruction performance

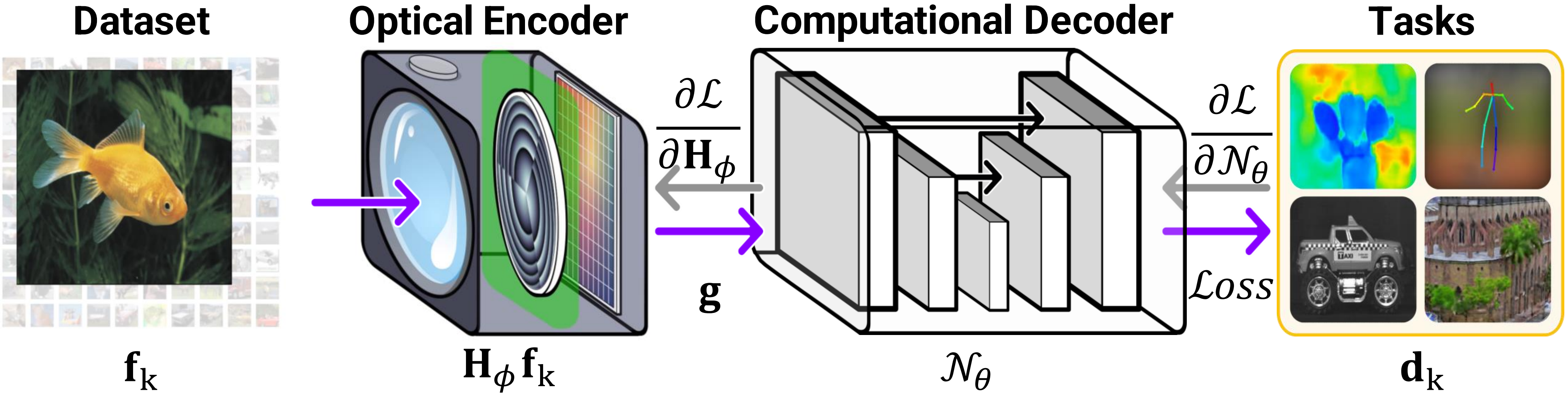
Without data, the design improves random patterns

Data-driven designed pattern provides the **best** reconstruction performance

Data-Driven Optical Coding Optimization

End-to-end approach

The optical system has to be modeled as a layer



$$\{\phi^*, \theta^*\} = \underset{\phi,\theta}{\arg\min}\, \mathrm{E}\left[\mathcal{L}_{task}\left(\mathcal{N}_\theta\left(\mathbf{H}_\phi \mathbf{f}_\mathrm{k}\right), \mathbf{d}_k\right) + \rho\mathcal{R}(\phi) + \tau\mathcal{R}\left(\mathbf{H}_\phi \mathbf{f}\right)\right]$$

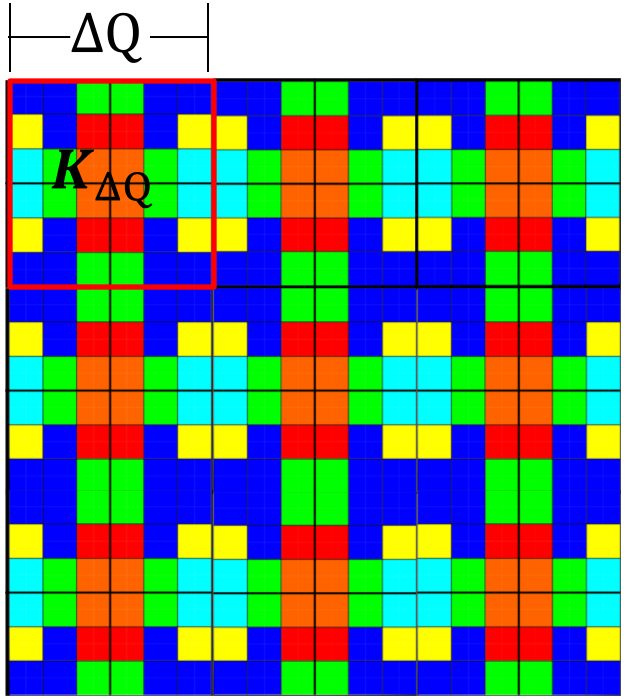
- $\mathcal{L}_{task}(\cdot)$ is the **loss-function** of a determined COI task, $\mathbf{d}_\mathrm{k}$ is the desired output of the training image $\mathbf{f}_\mathrm{k}$
- $\mathcal{R}(\cdot)$ represents the **physical constraints** in the optical encoder ϕ.
- The **tasks** can be depth, privacy, super resolution, spectral imaging, among as.

Limitations

- ➢ **Physical constraints** reduce the degrees of freedom of the CE
- ➢ Fewer encoder layers compared with decoder layers produce **gradient vanishing**

Coding Element Parameterization

Colored coded aperture



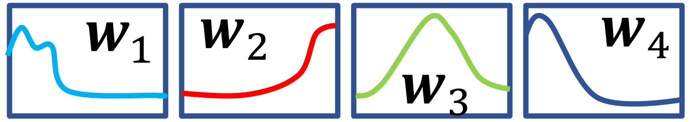
Spatial

$$K_{\Delta Q} = \mathbf{J}_{\Delta Q} K_{\Delta Q} \mathbf{J}_{\Delta Q} \text{ (Symmetry)}$$

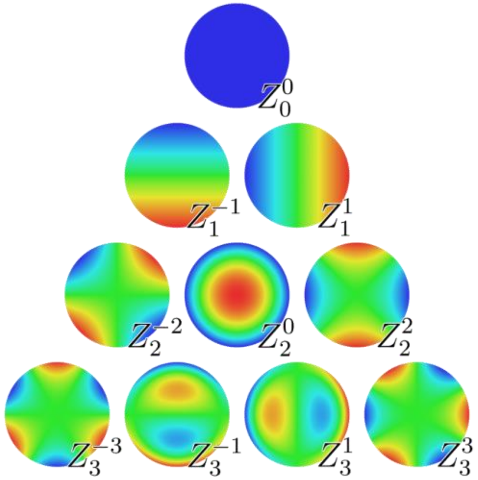
$$\phi = \mathbf{1} \otimes K_{\Delta Q} \text{ (Periodicity)}$$

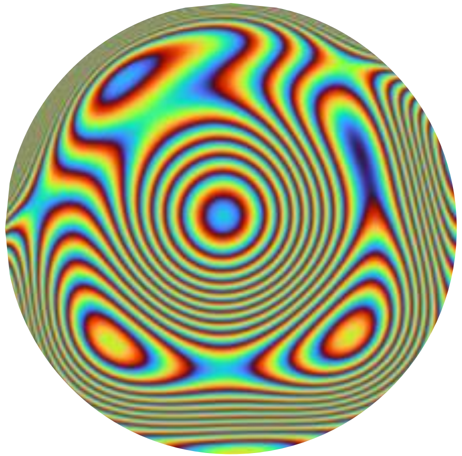
Spectral

$$K_{\Delta Q} = \sum_n \boldsymbol{\beta}_n \boldsymbol{w}_n$$



Diffractive optical element





$$\phi = \sum_n \boldsymbol{\beta}_n Z_n$$

- Instead of learning directly ϕ, the **trainable parameters** are $\boldsymbol{\beta}$
- Parameterizations allow **reducing** the number of trainable parameters and addressing **implementation constraints**

1. End-to-End Regularization Strategy

$$\{\phi^*, \theta^*\} = \underset{\phi,\theta}{\arg\min}\, \mathrm{E}\left[\mathcal{L}_{task}\left(\mathcal{N}_\theta\left(\mathbf{H}_\phi \mathbf{f}\right), \mathbf{f}\right) + \rho\mathcal{R}(\phi)\right]$$

Regularization **addresses physical constraints** for optimizing the Optical Design:

$$\phi^{i+1} = \phi^i - \alpha\left(\frac{\partial\mathcal{L}_{task}\left(\mathcal{N}_\theta\left(\mathbf{H}_\phi \mathbf{f}\right), \mathbf{f}\right)}{\partial\phi} + \rho\frac{\partial\mathcal{R}(\phi)}{\partial\phi}\right)$$

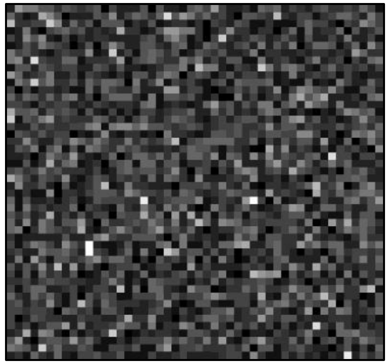
1. Address physical constraints

$\mathcal{R}(\phi)$

ϕ

1. $\frac{1}{n}\sum_{l=1}^{n}(\phi_l)^2(\phi_l - 1)^2$
2. $\sum_{j=1}^{S}\sqrt{\sum_{l=1}^{n}\left(\phi_l^j\right)^2}$
3. $\left(\frac{\sum_{l=1}^{n}\phi_l}{n} - T_r\right)^2$

Real values



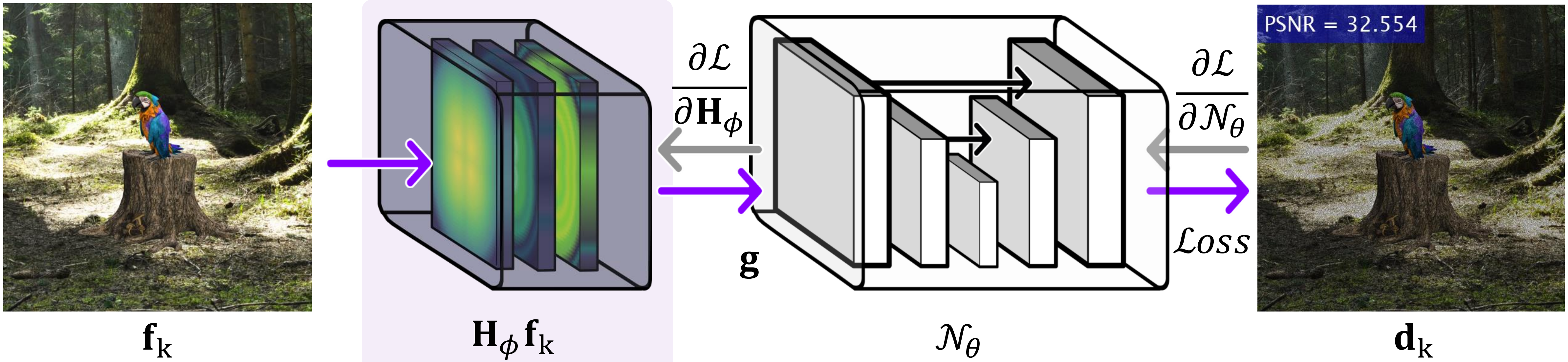
Hardly feasible

Binary values

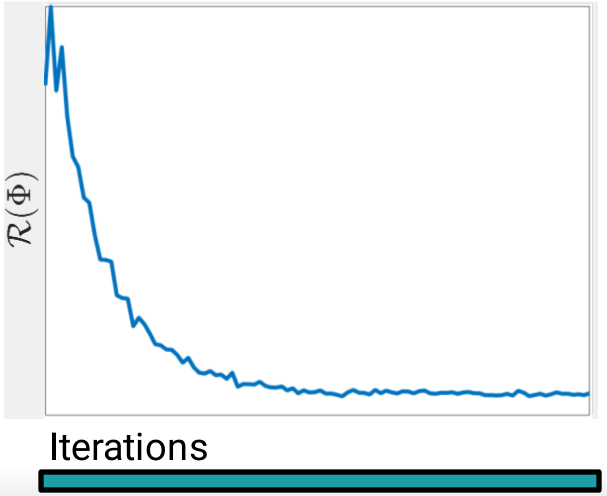
Physically feasible

1. Regularization to address physical constraints

Dataset | **Optical Encoder** | **Computational Decoder** | **Recovered**



Binarization

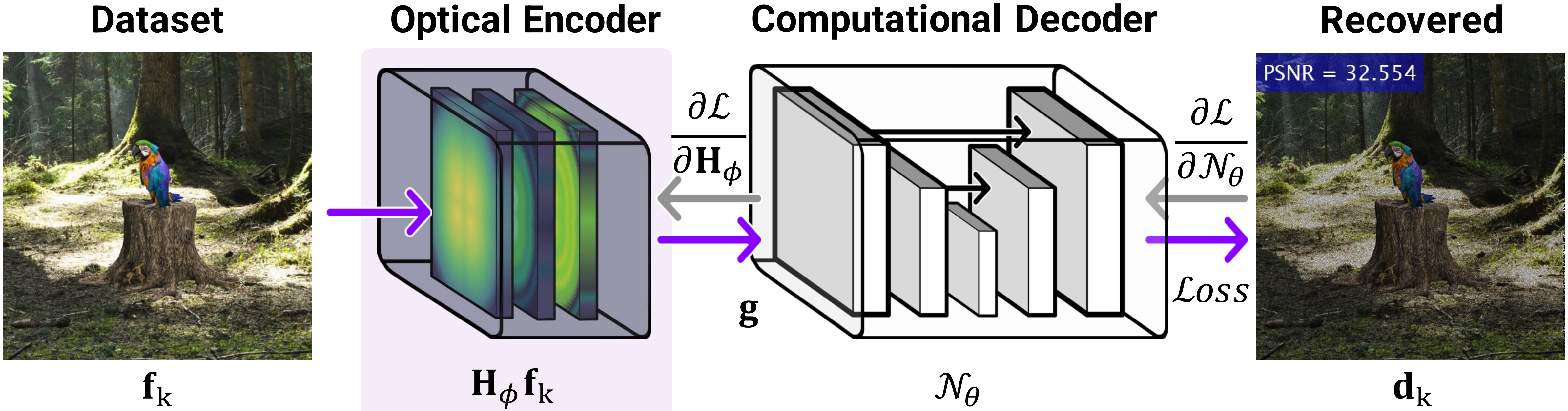


$$\mathcal{L}oss = \mathcal{L}_{task} + \mathcal{R}(\boldsymbol{\phi})$$

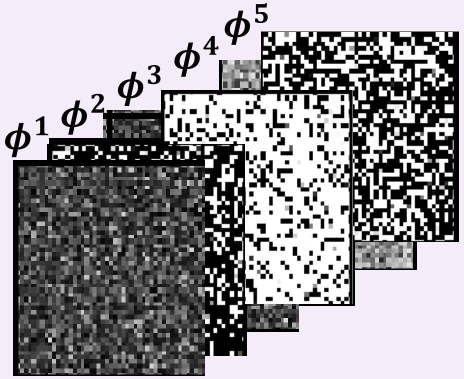
$$\mathcal{R}(\boldsymbol{\phi}) = \frac{1}{n}\sum_{l=1}^{n}(\boldsymbol{\phi}_l)^2(\boldsymbol{\phi}_l - 1)^2$$

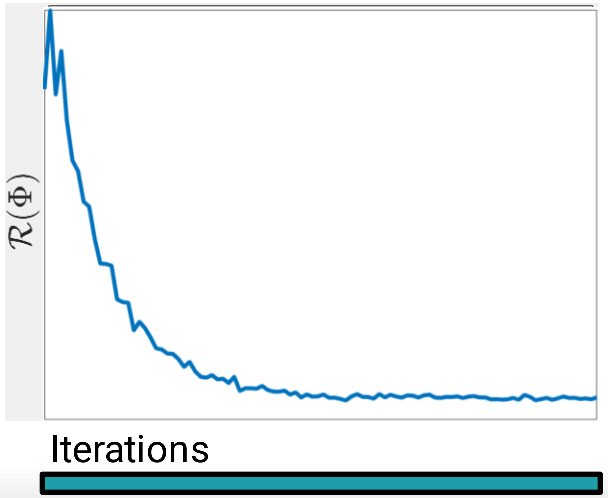
$$\mathcal{R}(\boldsymbol{\phi}) = \frac{1}{n}\sum_{l=1}^{n}(\sin(\boldsymbol{\phi}_l))^2(\sin(\boldsymbol{\phi}_l) - 1)^2$$

1. Regularization to address physical constraints



Minimum number of shots



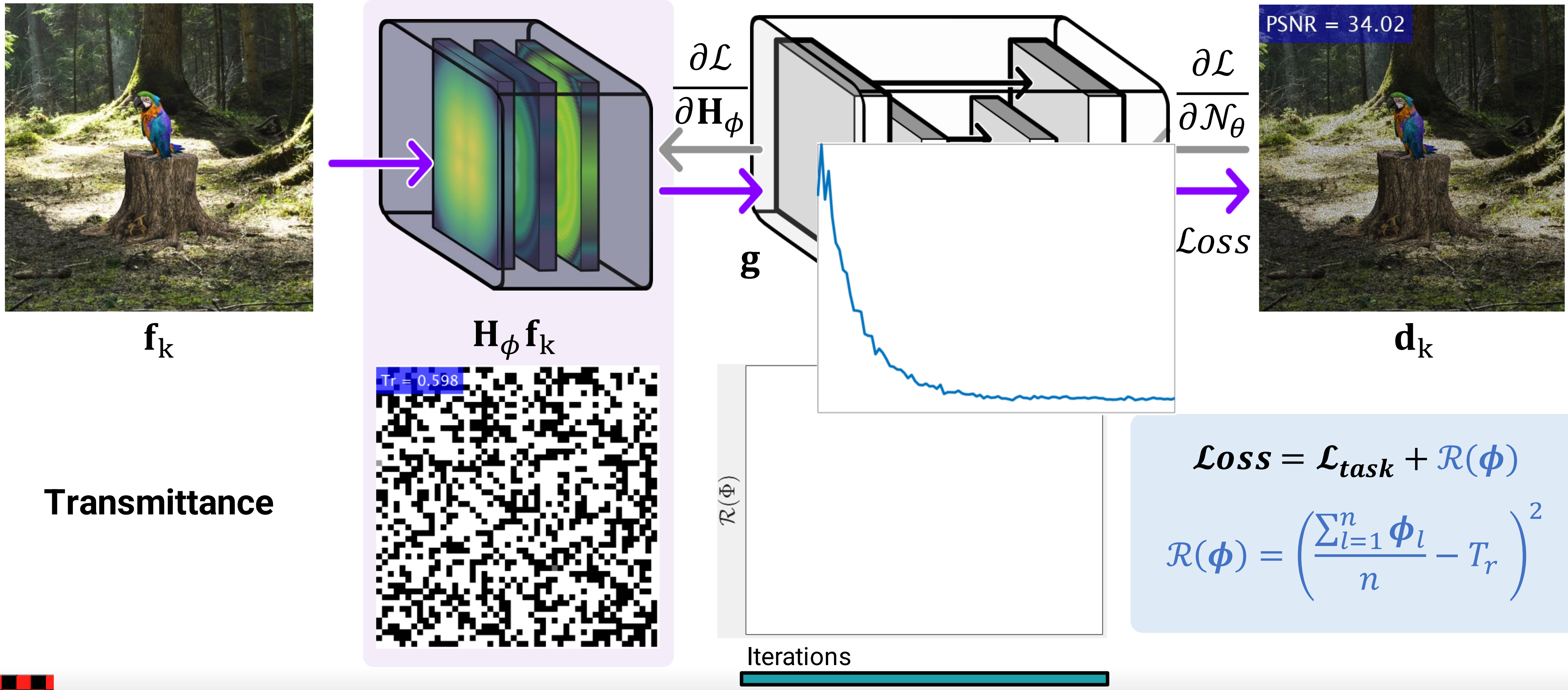


$$\mathcal{L}oss = \mathcal{L}_{task} + \mathcal{R}(\phi)$$

$$\mathcal{R}(\phi) = \sum_{s=1}^{L} \|\Phi^s\|_2$$

Bacca, J., Gelvez-Barrera, T., & Arguello, H. (2021). Deep coded aperture design: An end-to-end approach for computational imaging tasks. IEEE Transactions on Computational Imaging

1. Regularization to address physical constraints



$$\mathcal{L}oss = \mathcal{L}_{task} + \mathcal{R}(\boldsymbol{\phi})$$

$$\mathcal{R}(\boldsymbol{\phi}) = \left(\frac{\sum_{l=1}^{n} \boldsymbol{\phi}_l}{n} - T_r\right)^2$$

2. End-to-End Regularization Strategy

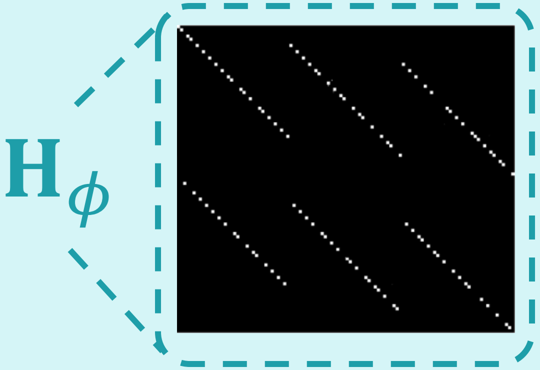
$$\{\phi^*, \theta^*\} = \arg\min_{\phi,\theta} \mathrm{E}\left[\mathcal{L}_{task}\left(\mathcal{N}_\theta(\mathbf{H}_\phi \mathbf{f}), \mathbf{f}\right) + \rho\mathcal{R}(\phi) + \tau\mathcal{R}(\mathbf{H}_\phi)\right]$$

Regularization **improves performance** by inducing properties in the Optical Design:

$$\phi^{i+1} = \phi^i - \alpha\left(\frac{\partial\mathcal{L}_{task}\left(\mathcal{N}_\theta(\mathbf{H}_\phi \mathbf{f}), \mathbf{f}\right)}{\partial\phi} + \rho\frac{\partial\mathcal{R}(\phi)}{\partial\phi} + \tau\frac{\partial\mathcal{R}(\mathbf{H}_\phi)}{\partial\phi}\right)$$

2. Improve performance

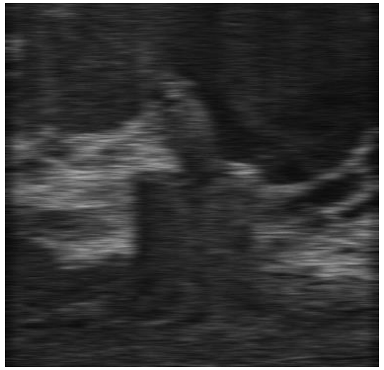
$\mathcal{R}(\mathbf{H}_\phi)$



$\mathbf{G}_\phi = \mathbf{H}_\phi^T\mathbf{H}_\phi$

1. $\|\mathbf{f}_k - \mathbf{G}_\phi\mathbf{f}_k\|_2^2$
2. $\|\mathbf{f}_k - (\mathbf{G}_\phi + \gamma\mathbf{I})^{-1}\mathbf{G}_\phi\mathbf{f}_k\|_2^2$
3. $\sum_j\sum_i\left(\frac{\|\mathbf{H}_\phi(\mathbf{f}_i - \mathbf{f}_j)\|_2}{\|\mathbf{f}_i - \mathbf{f}_j\|_2} - 1\right)^2$
4. $D_{KL}\left(q_\phi(\mathbf{H}_\phi\mathbf{f}_k|\mathbf{f}_k)\,\|\,p(\mathbf{H}_\phi\mathbf{f}_k)\right)$

Non-Optimized



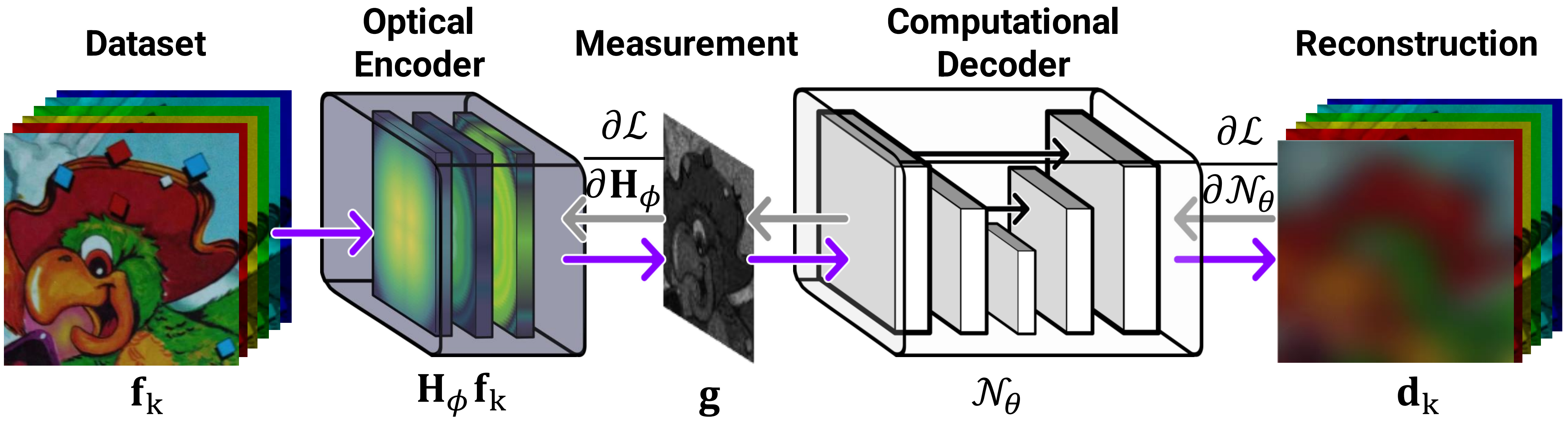
Bad conditioned

Optimized

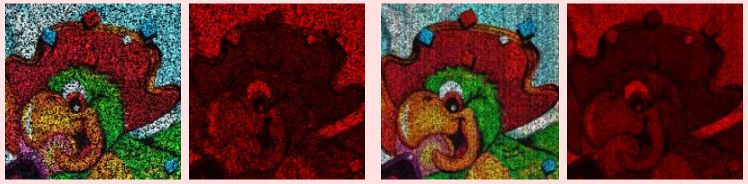
Well conditioned

2. Regularization to improve the performance

Data-driven conditionality $\mathcal{R}(\mathbf{H}_\phi \mathbf{f}_\mathrm{k}) = \|\mathbf{f}_\mathrm{k} - \mathbf{H}_\phi^T \mathbf{H}_\phi \mathbf{f}_\mathrm{k}\|_2^2$



Without regularization ✗



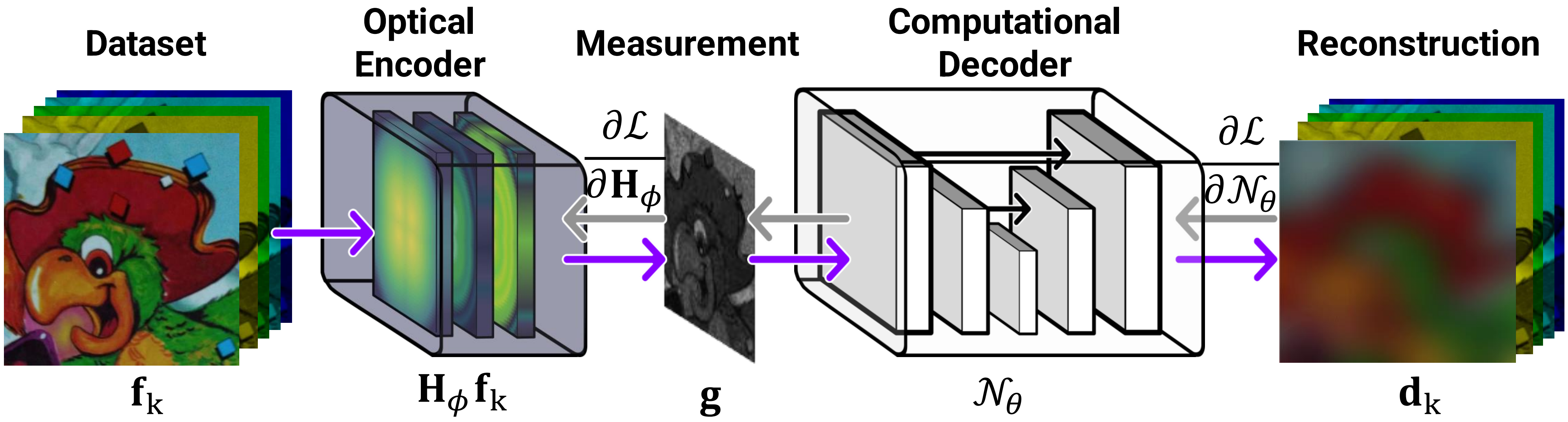
$\mathbf{G}_\phi \mathbf{f}_\mathrm{k}$ $\mathcal{N}_\theta(\mathbf{H}_\phi \mathbf{f}_\mathrm{k})$

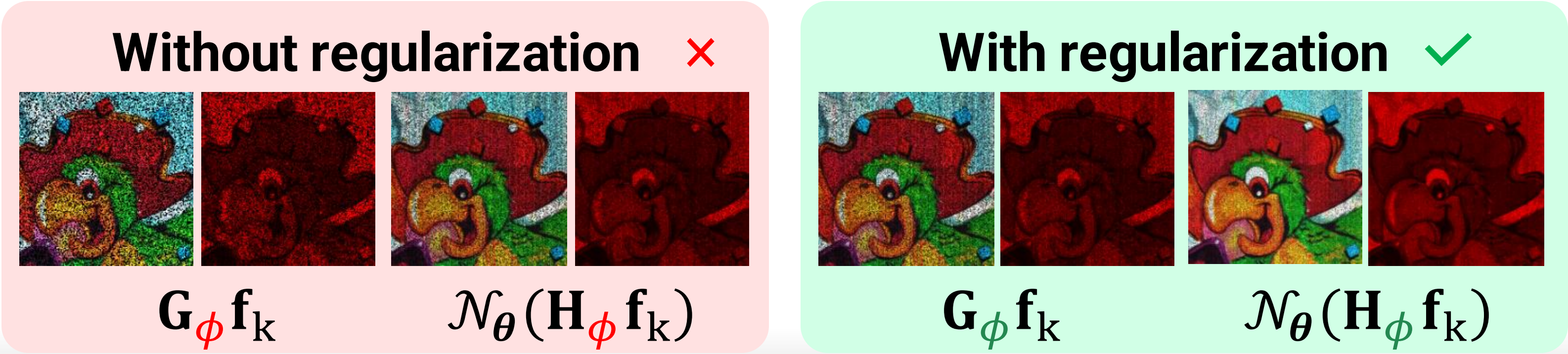
With regularization ✓

$\mathbf{G}_\phi \mathbf{f}_\mathrm{k}$ $\mathcal{N}_\theta(\mathbf{H}_\phi \mathbf{f}_\mathrm{k})$

2. Regularization to improve the performance

Inversion regularizer $\|\mathcal{R}(\mathbf{H}_\phi \mathbf{f}_k) = \|\|\, \mathbf{f}_k - \left(\mathbf{G}_\phi + \gamma \mathbf{I}\right)^{-1} \mathbf{G}_\phi \mathbf{f}_k \,\|_2^2$

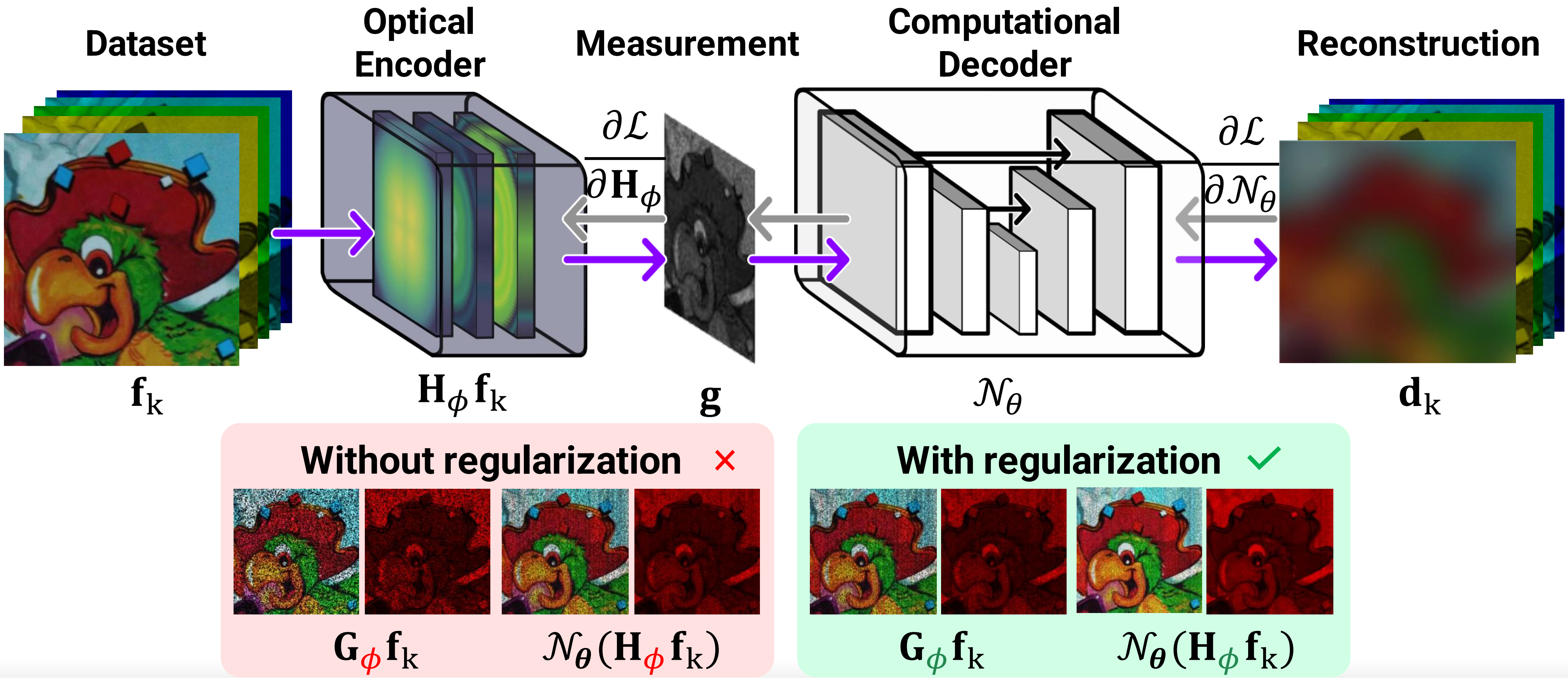




2. Regularization to improve the performance

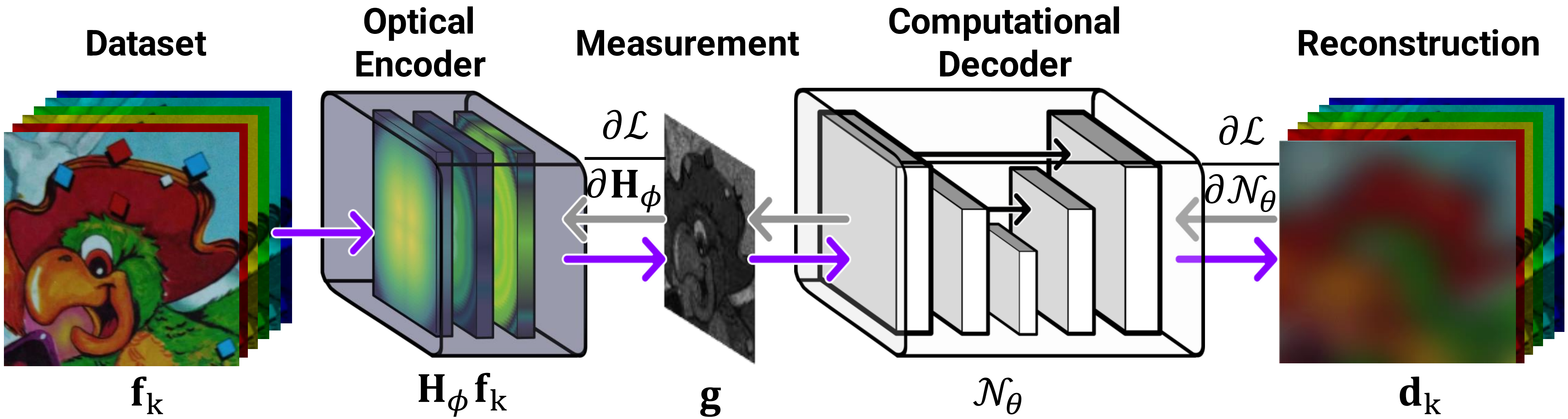
Restricted Isometry Property

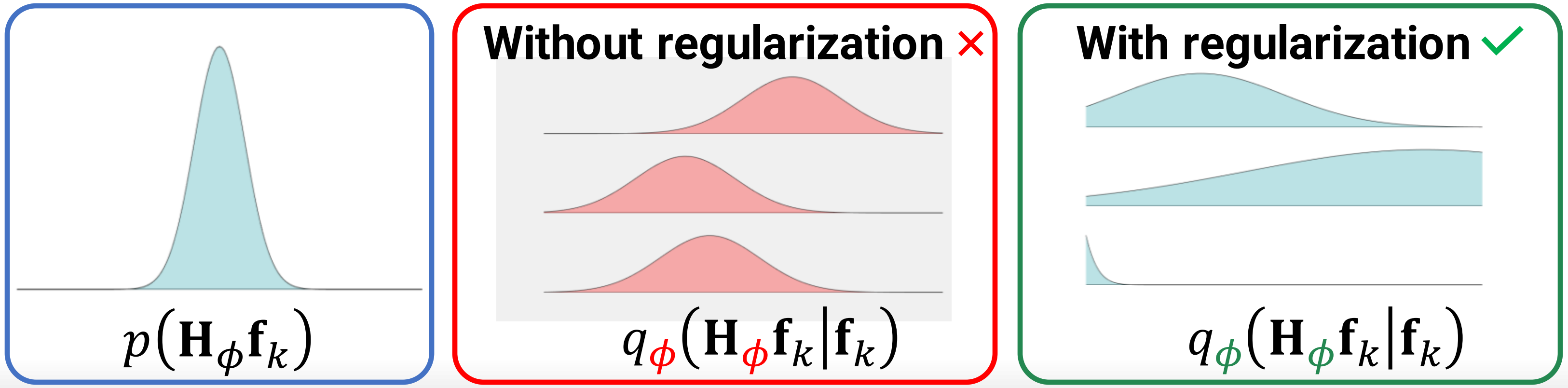
$$\mathcal{R}(\mathbf{H}_\phi \mathbf{f}) = \sum_j \sum_i \left(\frac{\|\mathbf{H}_\phi(\mathbf{f}_i - \mathbf{f}_j)\|_2}{\|\mathbf{f}_i - \mathbf{f}_j\|_2} - 1 \right)^2$$



2. Regularization to improve the performance

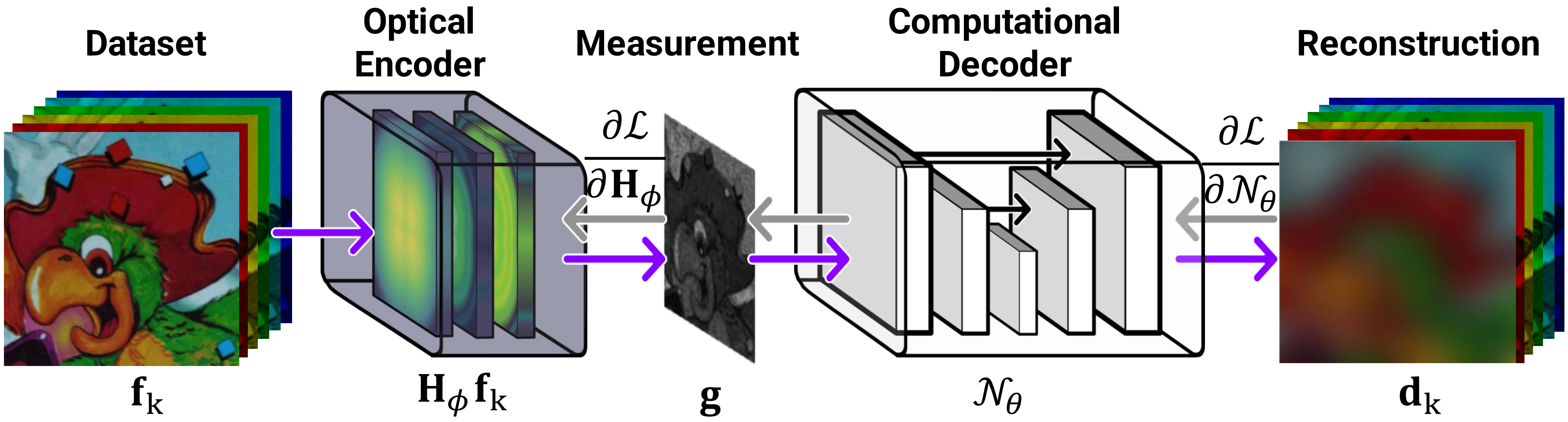
Distribution regularization $\mathcal{R}(\mathbf{H}_\phi \mathbf{f}) = \mathrm{D}_{KL}\left(q_\phi(\mathbf{H}_\phi \mathbf{f}|\mathbf{f}) \,\|\, p(\mathbf{H}_\phi \mathbf{f})\right)$





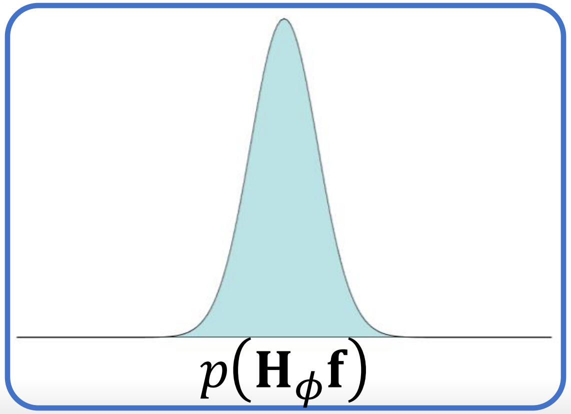
2. Regularization to improve the performance

Distribution regularization $\mathcal{R}(\mathbf{H}_\phi \mathbf{f}) = \mathrm{D}_{KL}\left(q_\phi(\mathbf{H}_\phi \mathbf{f} | \mathbf{f}) \,\|\, p(\mathbf{H}_\phi \mathbf{f})\right)$



What is the best prior distribution for the regularization?

Depending on the dataset and the computational task, it must be chosen the optimal distribution of the measurements.



This regularization can be used in any computational task

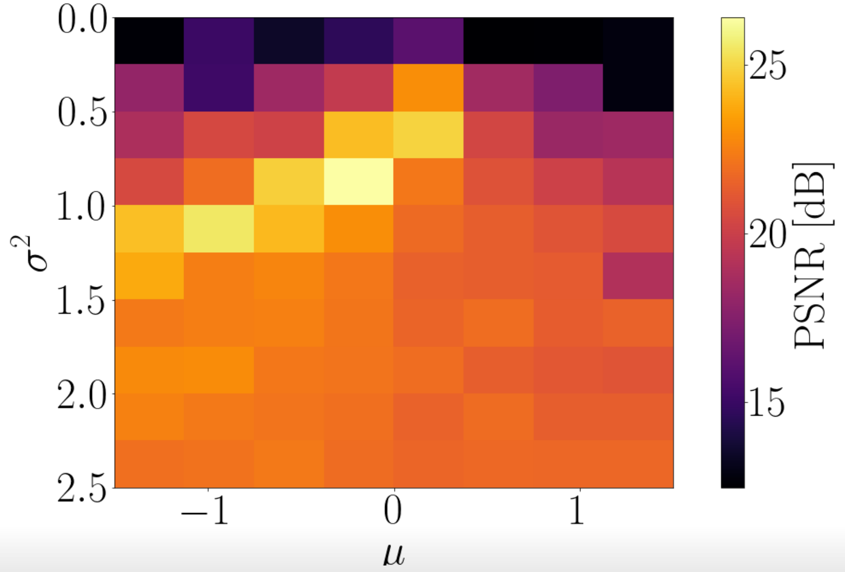
2. Regularization to improve the performance

Distribution regularization

When the prior distribution is Gaussian, *i.e.* $p(\mathbf{H}_\phi \mathbf{g}) = \mathcal{N}(\mu, \sigma^2)$ what is the best configuration of μ and σ^2 for a given task?

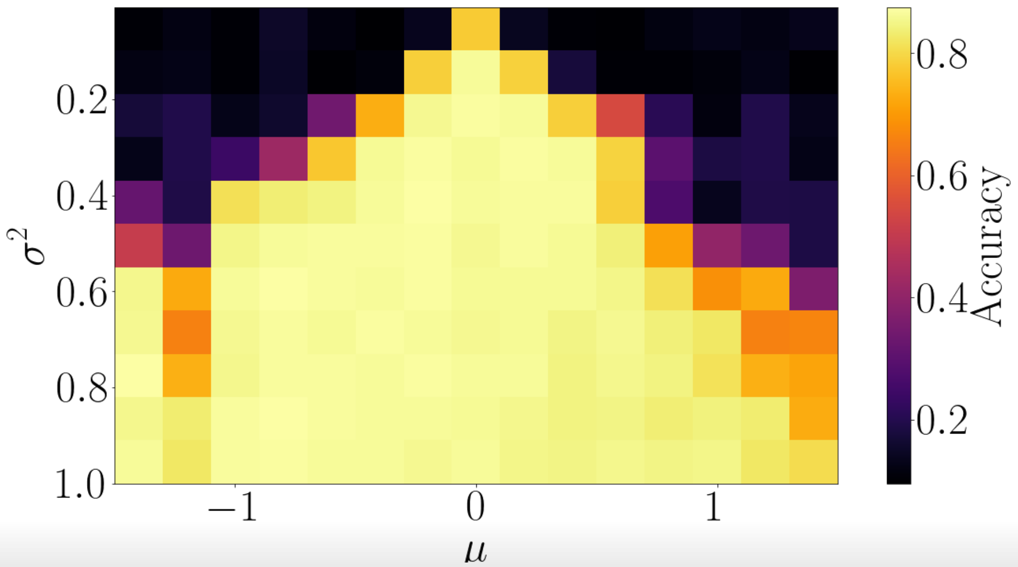
Recovery Task

More concentrated measurements allows better reconstruction



Classification Task

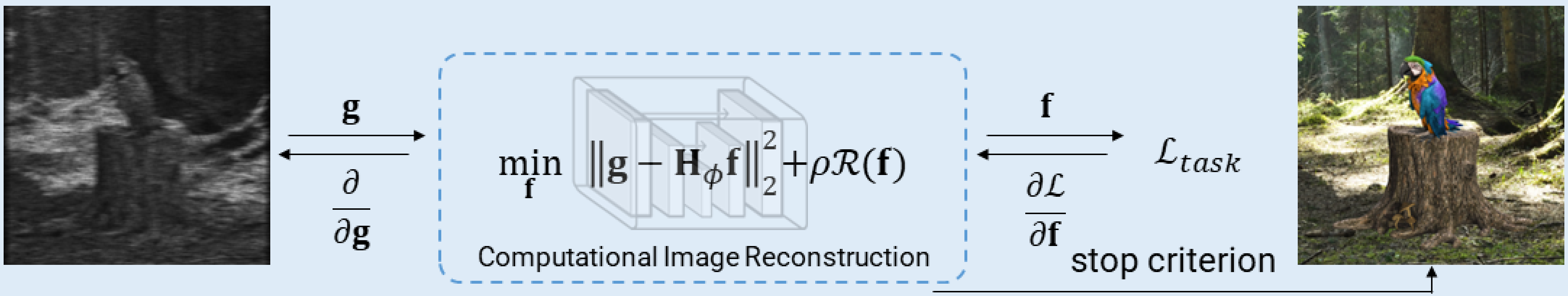
More separated measurements makes easier the classification



Computational Decoder

Optimization Algorithms

Differentiable based[1]



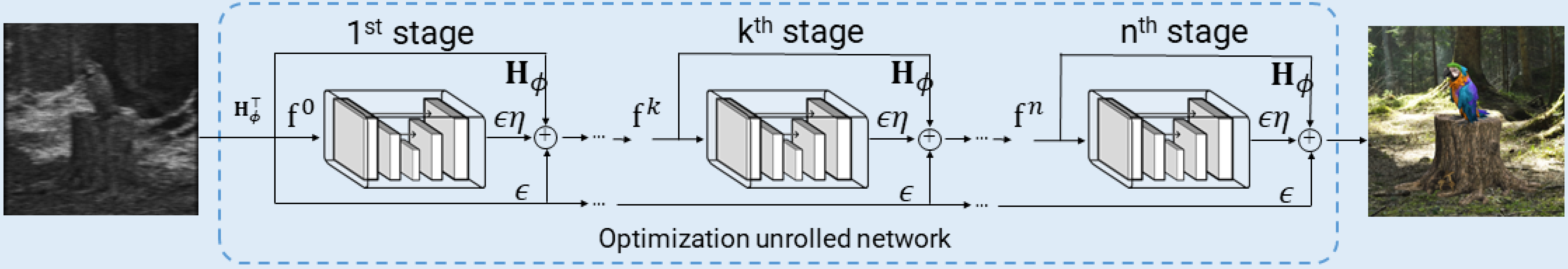
- **Fully- differentiable** optimization model for image recovery.
- Jointly optimization of **optical parameters** and image processing **algorithm parameters**.

1. Sitzmann, V. et al. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. ACM Transactions on Graphics

Computational Decoder

Deep Learning Decoder

Unrolling Optimization based[1]



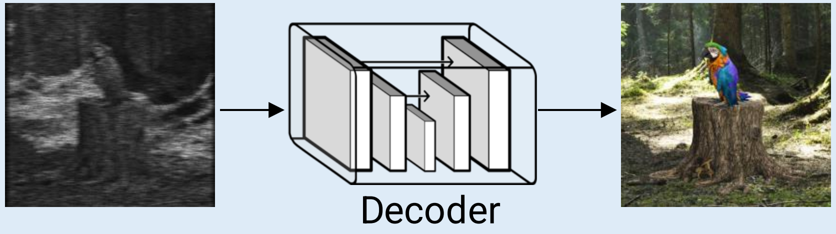
- Advantages of **differentiable-based** approach.
- Unrolled network **boosts the reconstruction speed** by freezing the parameters of iteration[2].
- Each stage aims to solve an iteration equation, which makes the network **explainable**[3].

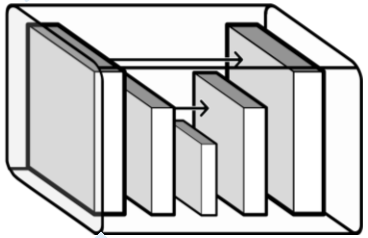
1. Monga, V et al. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. IEEE Signal Processing Magazine.
2. Huang, L. et al. Spectral imaging with deep learning. Light: Science & Applications.
3. Monroy, B. et al. JR2net: A Joint Non-Linear Representation and Recovery Network for Compressive Spectral Imaging. arXiv preprint.

Computational Decoder

Deep Learning Decoder

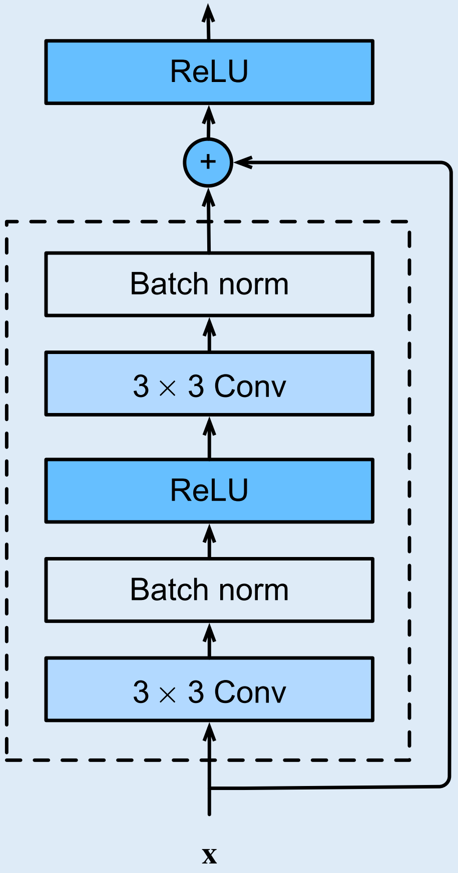
Black-box based[1]





- **Non-linear inverse mapping** from measurements to image recovery
- **Several architectures** for vast compressive imaging applications.

ResNet

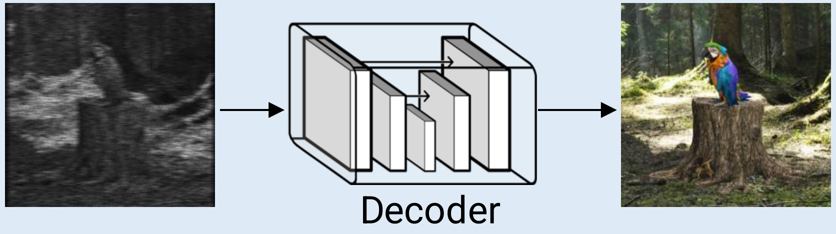


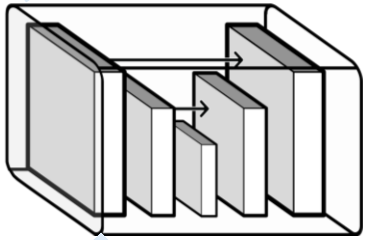
1. Zhang, T. et al . Hyperspectral image reconstruction using deep external and internal learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision.

Computational Decoder

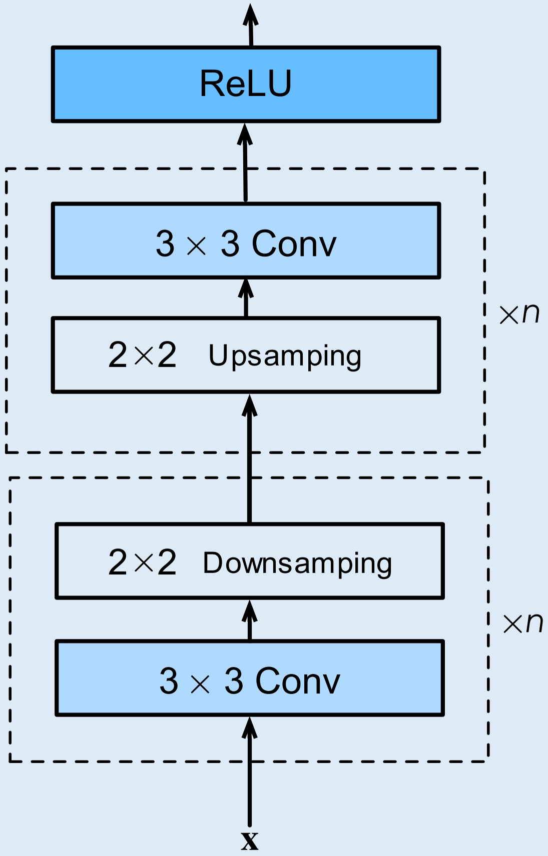
Deep Learning Decoder

Black-box based[1]





Autoencoder



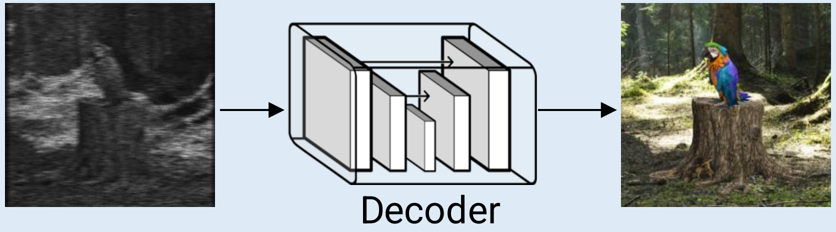
- **Non-linear inverse mapping** from measurements to image recovery
- **Several architectures** for vast compressive imaging applications.

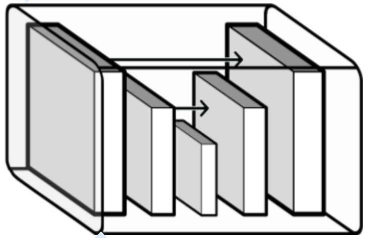
1. Zhang, T. et al . Hyperspectral image reconstruction using deep external and internal learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision.

Computational Decoder

Deep Learning Decoder

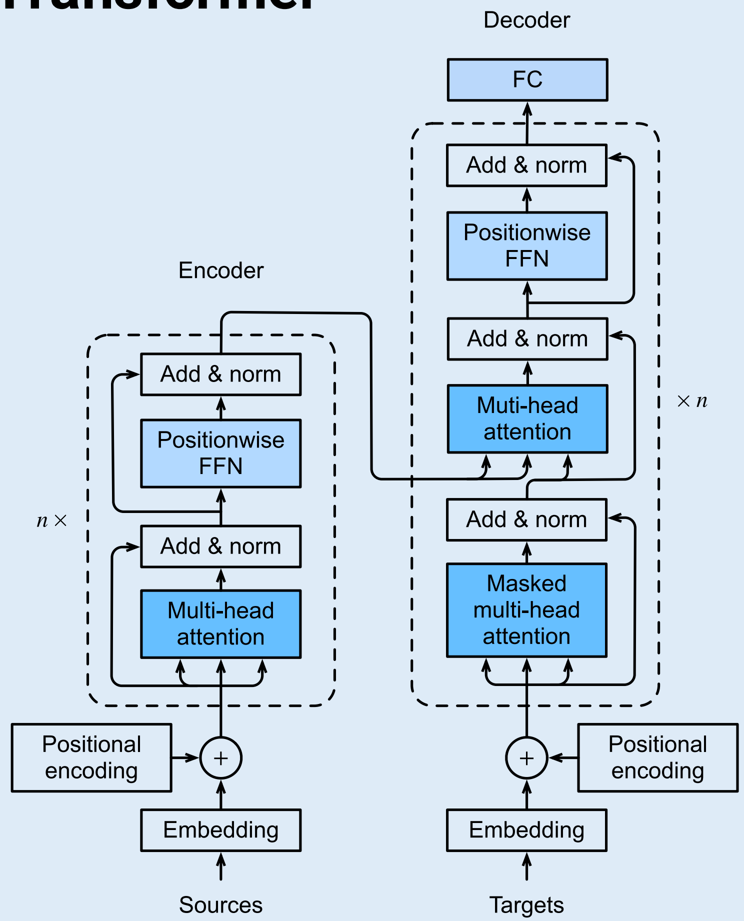
Black-box based[1]





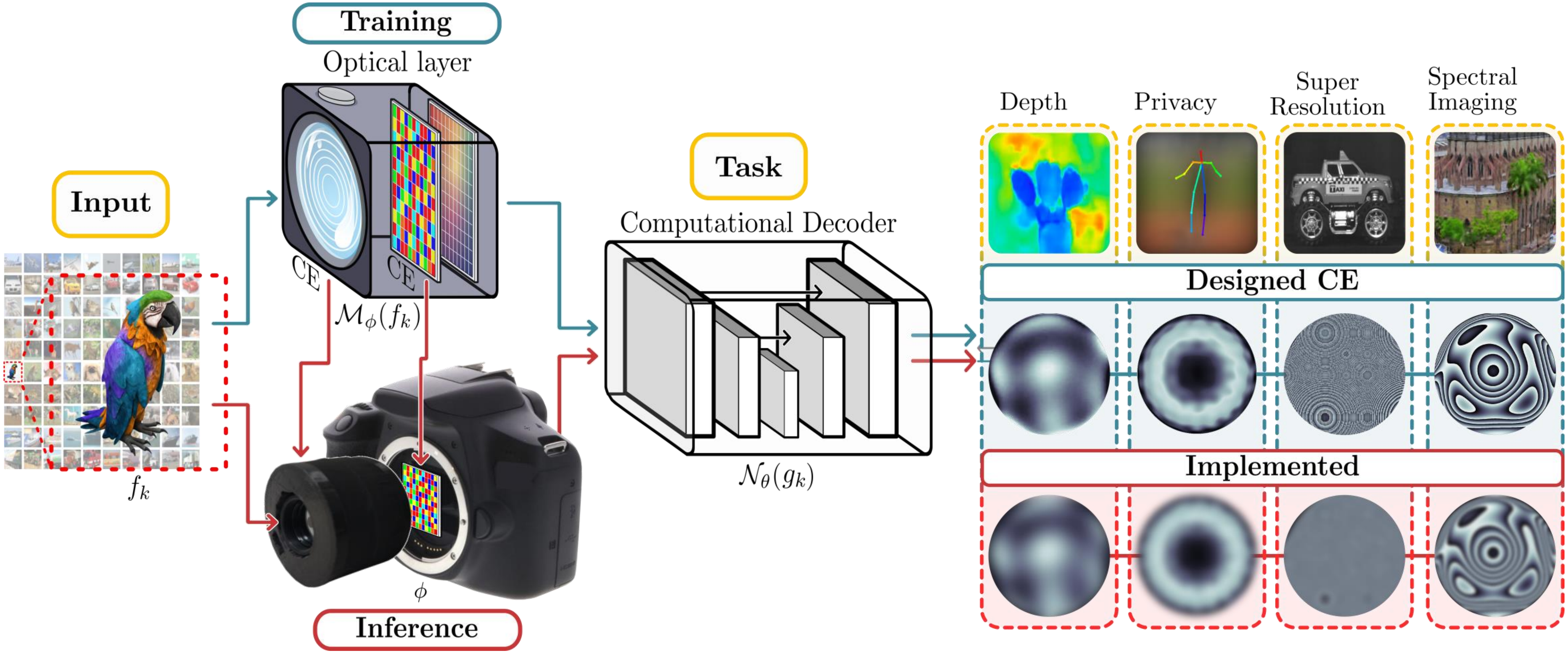
- **Non-linear inverse mapping** from measurements to image recovery
- **Several architectures** for vast compressive imaging applications.

Transformer



Implementations and Applications

Implementation and Fabrication of Coding Elements

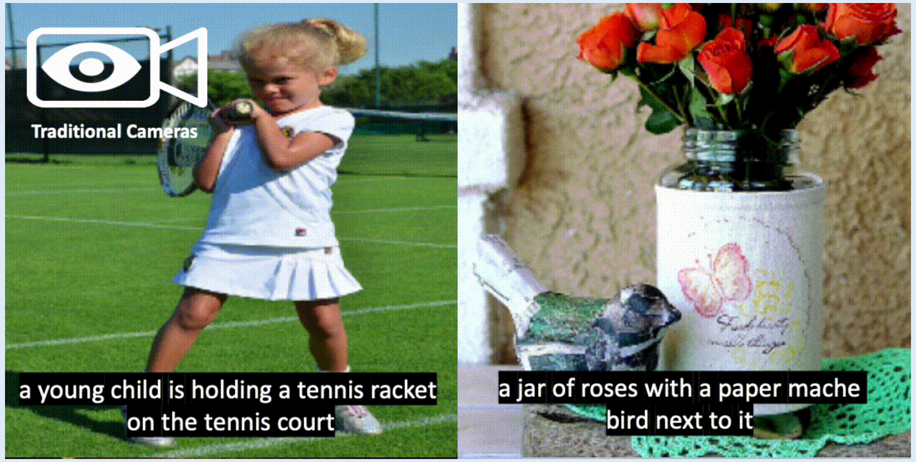


Once the CE is designed, it is fabricated and implemented in real setups.

Applications

Privacy: Pose Estimation[1]

Privacy: Scene Captioning[2]



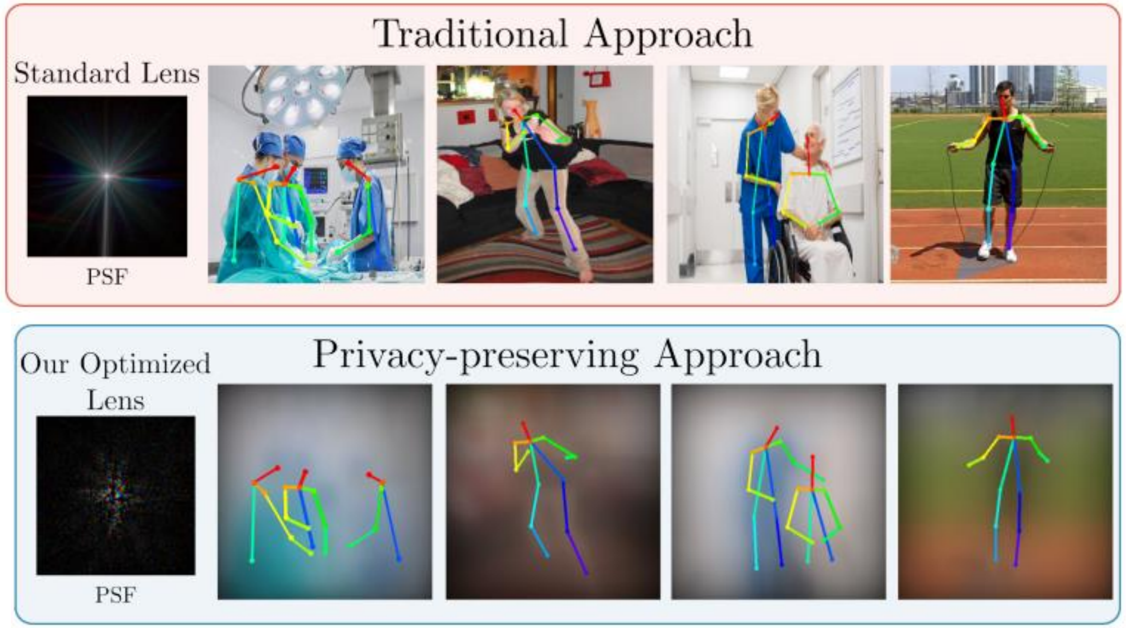
Compressive Sensing: Refractive[3]

Random TMCA

Compressive Sensing: Diffractive[4]

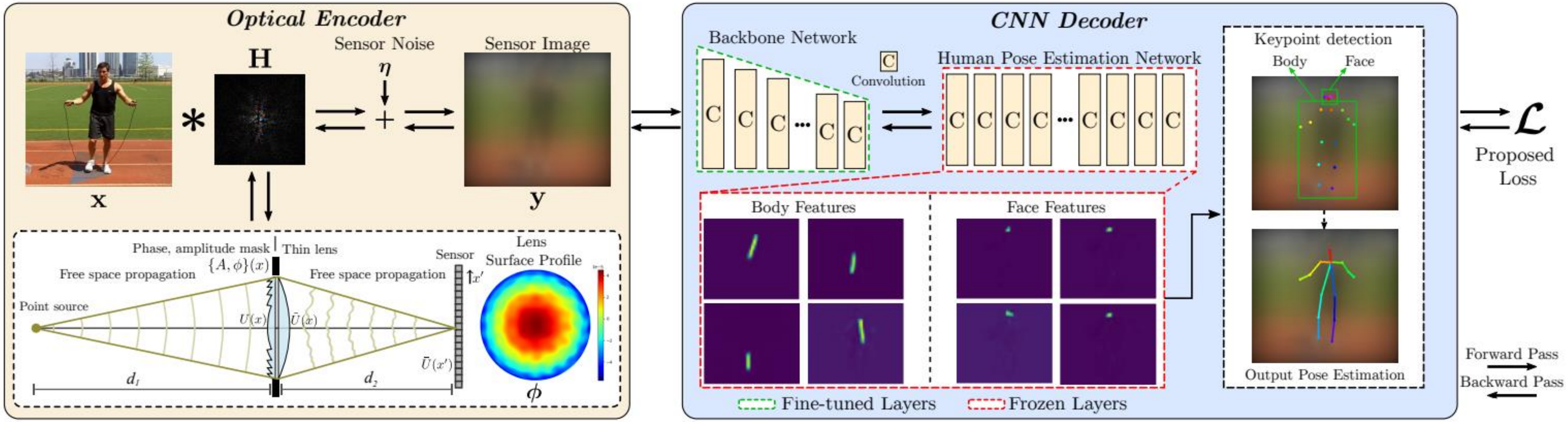
1. Hinojosa, *et al*. (2021). Learning Privacy-preserving Optics for Human Pose Estimation. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*.
2. Arguello, *et al*. (2021). Shift-variant color-coded diffractive spectral imaging system. *Optica*.
3. Vargas, *et al*. (2021). Time-Multiplexed Coded Aperture Imaging: Learned Coded Aperture and Pixel Exposures for Compressive Imaging Systems. In *Proceedings of the IEEE/CVF ICCV*.
4. Hinojosa, *et al*. (2021). Learning Privacy-preserving Optics for Human Pose Estimation. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*.

Privacy-Preserving: Pose Estimation



The goal is to **estimate the pose** of the people in the scene while maintaining **privacy**.

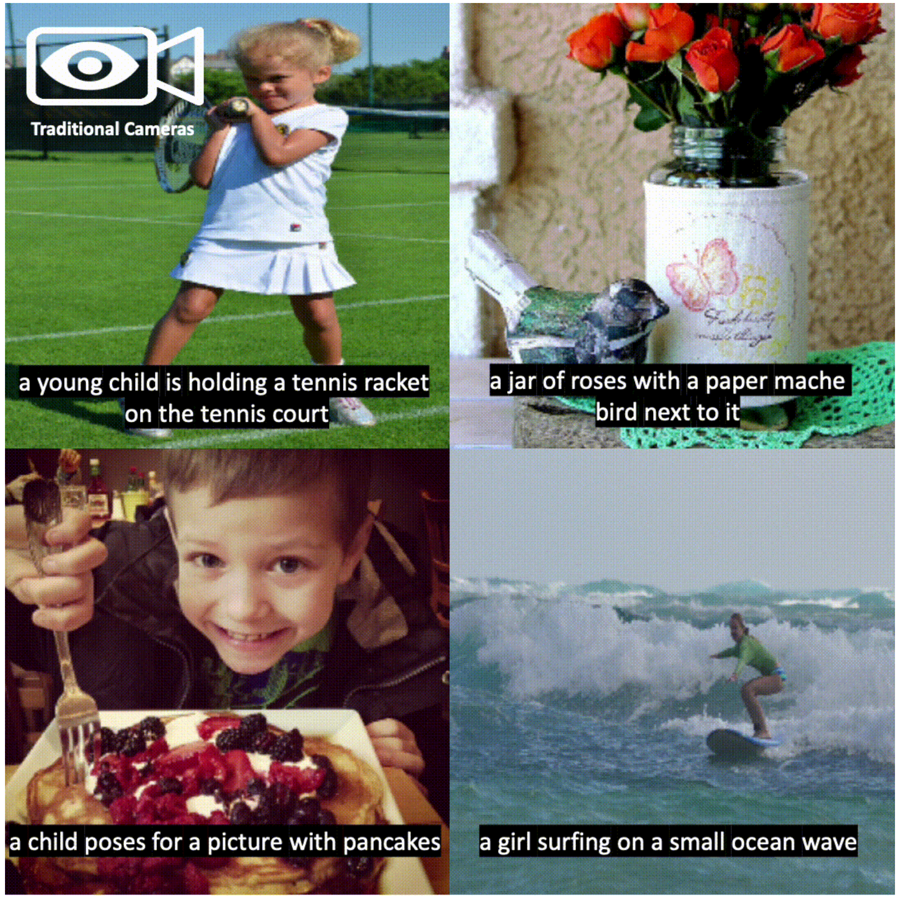
Privacy-Preserving: Pose Estimation

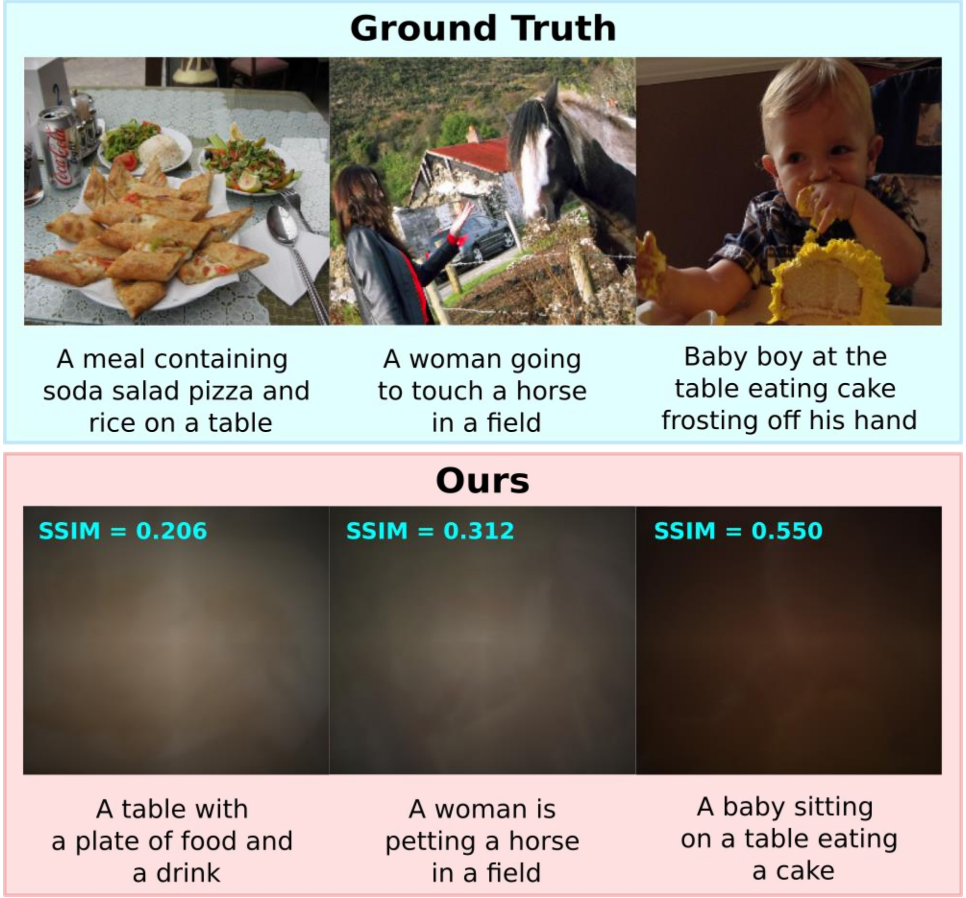


The goal is to **estimate the pose** of the people in the scene while maintaining **privacy**.

Hinojosa, *et al*. (2021). Learning Privacy-preserving Optics for Human Pose Estimation. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*.

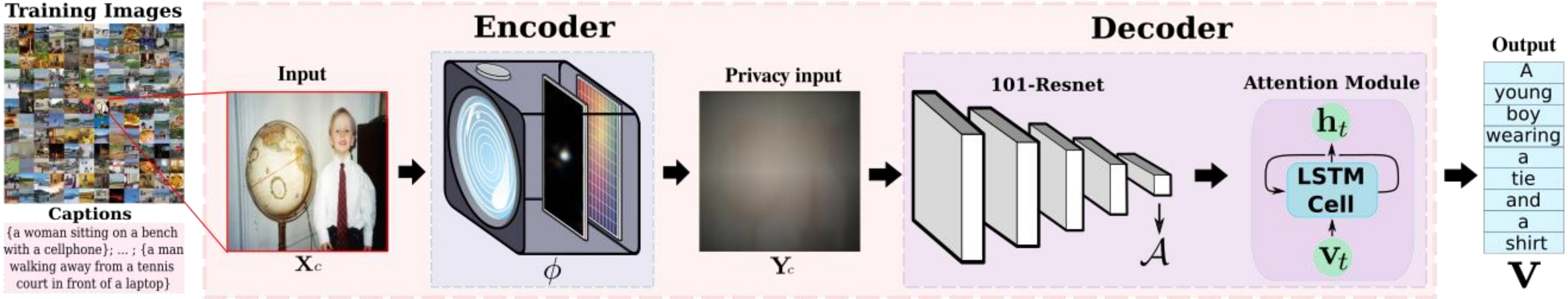
Privacy-Preserving: Scene Captioning

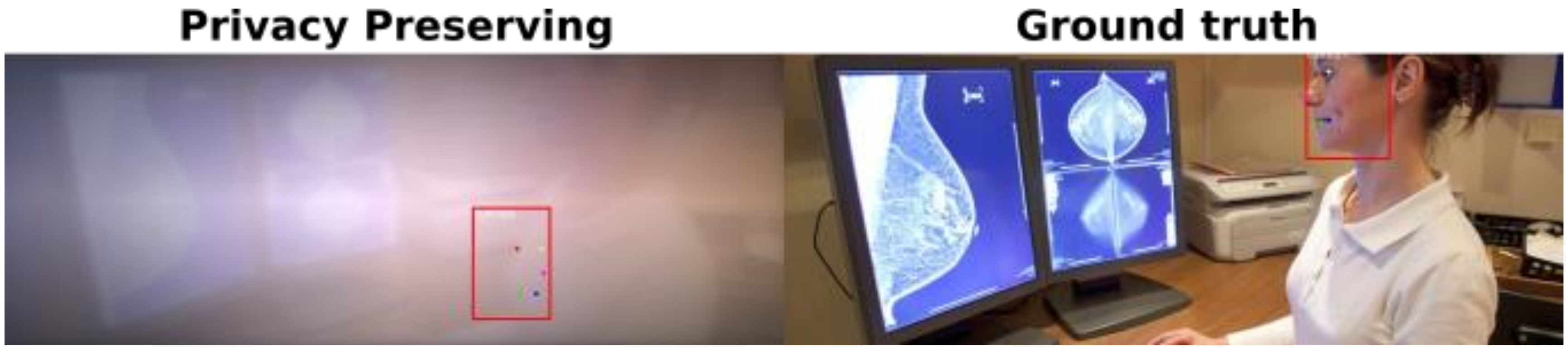




The goal is to **preserve privacy** while performing the **image captioning task**.

Privacy-Preserving: Scene Captioning



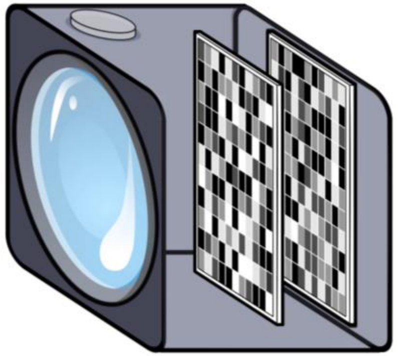


The goal is to **preserve privacy** while performing the **image captioning task.**

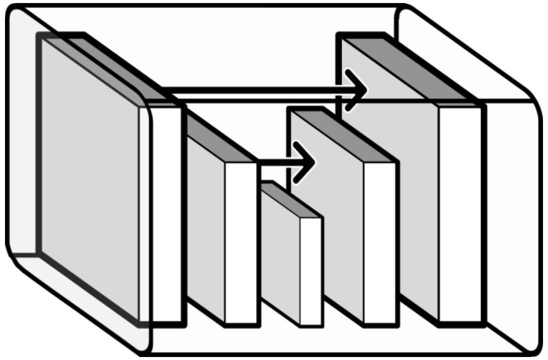
Compressive Sensing: Reffractive Imaging

Time Multiplexed Coded Aperture (TMCA)

Optical Encoder



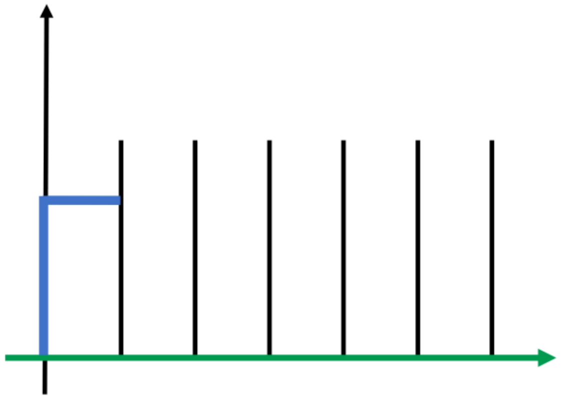
Computational Decoder



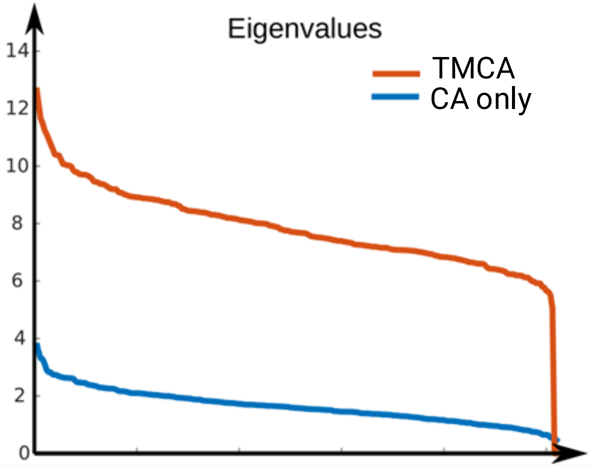
Traditional

TMCA

Recovery



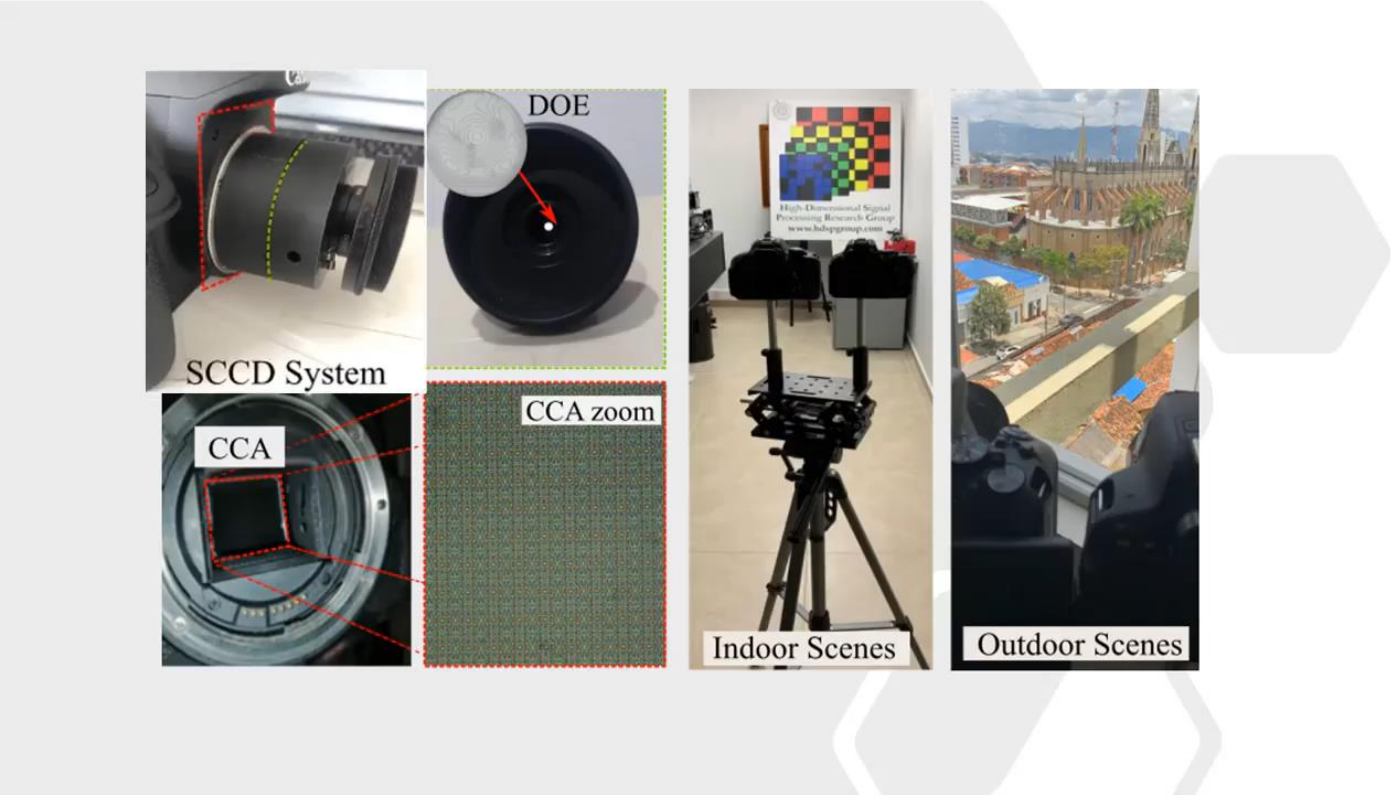
$$Shutter\ at\ pixel_{x,y}(t)$$



TMCA improves the **conditioning** of sensing matrices

Vargas, E., Martel, J. N., Wetzstein, G., & Arguello, H. (2021). Time-Multiplexed Coded Aperture Imaging: Learned Coded Aperture and Pixel Exposures for Compressive Imaging Systems. In *Proceedings of the IEEE/CVF International Conference on Computer Vision.*

Compressive Sensing: Diffractive Imaging



Non-Data Driven Designed

Data Driven Optical Design

Conclusions

To take away: The optical design can be addressed by parameterization and regularization in an E2E approach.

$$\{\phi^*, \theta^*\} = \arg\min_{\phi,\theta} \mathrm{E}\left[\mathcal{L}_{task} + \rho\mathcal{R}(\phi) + \tau\mathcal{R}(\mathbf{H}_\phi)\right]$$

1. Address physical constraints

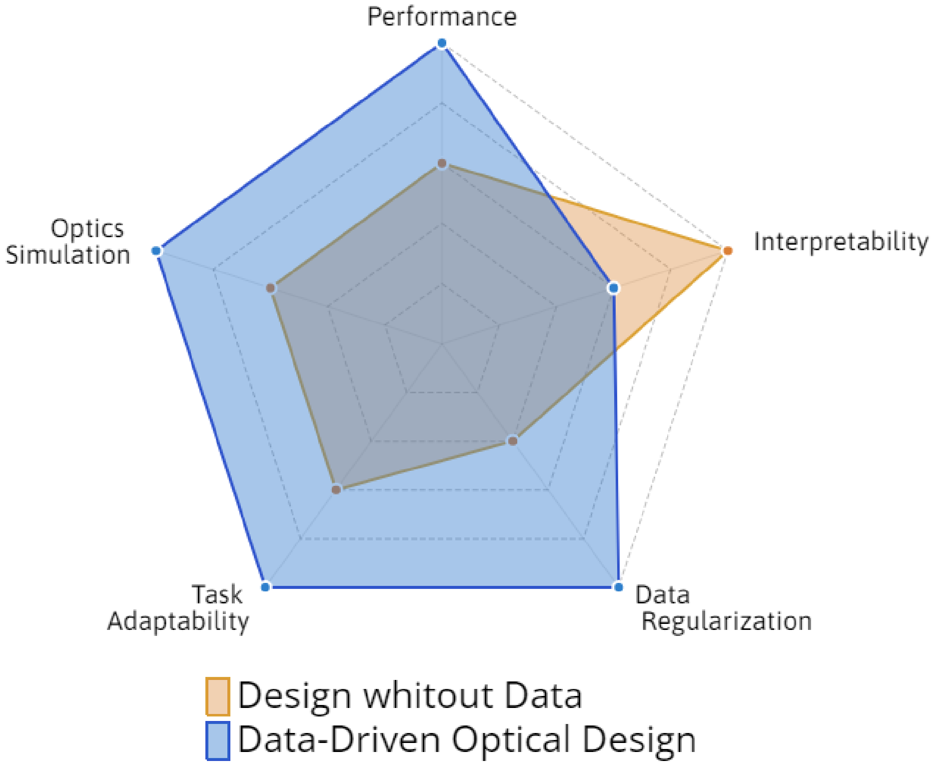
$\mathcal{R}(\phi)$

1. $\frac{1}{n}\sum_{l=1}^{n}(\phi_l)^2(\phi_l - 1)^2$
2. $\sum_{j=1}^{S}\sqrt{\sum_{l=1}^{n}(\phi_l^j)^2}$
3. $\left(\frac{\sum_{l=1}^{n}\phi_l}{n} - T_r\right)^2$

2. Improve Performance

$\mathcal{R}(\mathbf{H}_\phi)$ $\quad \mathbf{G}_\phi = \mathbf{H}_\phi^\top \mathbf{H}_\phi$

1. $\|\mathbf{f}_k - \mathbf{G}_\phi \mathbf{f}_k\|_2^2$
2. $\|\mathbf{f}_k - (\mathbf{G}_\phi + \gamma \mathbf{I})^{-1}\mathbf{G}_\phi \mathbf{f}_k\|_2^2$
3. $\sum_j \sum_i \left(\frac{\|\mathbf{H}_\phi(\mathbf{f}_i - \mathbf{f}_j)\|_2}{\|\mathbf{f}_i - \mathbf{f}_j\|_2} - 1\right)^2$
4. $D_{KL}\left(q_\phi(\mathbf{H}_\phi \mathbf{f}_k | \mathbf{f}_k) \,\|\, p(\mathbf{H}_\phi \mathbf{f}_k)\right)$



1. Sitzmann, et al. (2018). End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. ACM Transactions on Graphics (TOG).
2. Lin, X., et al. (2018). All-optical machine learning using diffractive deep neural networks. Science.
3. Shi, J., et al. (2021). Multiple-view D 2 NNs array: realizing robust 3D object recognition. Optics Letters.

WACCI

2nd International Workshop on Adaptive, Compressive and Computational Imaging

SPONSORED BY

OPTICA | Formerly OSA

Universidad Industrial de Santander UIS

ANH AGENCIA NACIONAL DE HIDROCARBUROS

U.S. AIR FORCE

U.S.ARMY

Ministerio de Ciencia, Tecnología, Conocimiento e Innovación — Gobierno de Chile

El conocimiento es de todos — Minciencias

IEEE

Bucaramanga, Colombia

November 8-10, 2022

For more information, please contact us at the email
wacci2022@uis.edu.co

Thank you!

www.hdspgroup.com
henarfu@uis.edu.co

Please contact me if you want to visit our group!

Professors

Postdocs

Visiting students