Dictionary Learning from Phaseless Measurements

Andreas M. Tillmann (TUD) Yonina C. Eldar (Technion) Julien Mairal (INRIA)

2D Phase Retrieval

Task: For a linear function $\mathcal{F}: \mathbb{C}^{N_1 \times N_2} \to \mathbb{C}^{M_1 \times M_2}$, recover $\hat{X} \in \mathcal{X} \subseteq \mathbb{C}^{N_1 \times N_2}$ from (noisy) nonlinear measurements

$$Y := |\mathcal{F}(\hat{X})|^2 + N$$

Here: $\hat{X}$ is an image $\rightsquigarrow \mathcal{X} = [0,1]^{N_1 \times N_2}$

Known: Sparsity of $\hat{X}$ helpful for recovery, but usually, $\hat{X}$ is not sparse itself and dictionary for sparse representation is unknown a priori

Dictionary Learning

Modeling assumption: vector $\mathbb{C}^s \ni x \approx Da$ with a sparse for unknown D

Popular model: $\min_{D,a} \frac{1}{2}\|Da - x\|_2^2 + \lambda \|a\|_1$

$\rightsquigarrow$ A common solution approach: Alternating Minimization
(a-update: ISTA / gradient descent + soft-thresholding)

Variant: $\min_{D,a} \frac{1}{2}\|Da - x\|_2^2$ s.t. $\|a\|_0 \leq k$ (OMP for a-update)

DOLPHIn – DictiOnary Learning for PHase retrIeval

Phase Retrieval Model

Goal: Improve image reconstruction from noisy phaseless data by simultaneously learning a dictionary $D \in \mathbb{R}^{s \times n}$ to sparsely represent patches x^i of image X.
Phase-Retrieval Dictionary-Learning (DOLPHIn) Model:

$$\min_{X,D,A} \frac{1}{4}\|Y - |\mathcal{F}(X)|^2\|_F^2 + \frac{\mu}{2}\|\mathcal{E}(X) - DA\|_F^2 + \lambda \sum_{i=1}^{p} \|a^i\|_1 \quad \text{s.t.} \quad X \in \mathcal{X} := [0,1]^{N_1 \times N_2},\; D \in \mathcal{D} := \{\, D \in \mathbb{R}^{s \times n} : \|D_{\cdot j}\|_2 = 1 \;\forall j \,\}$$

DOLPHIn Algorithm

Input: initial image estimate $X_{(0)} \in [0,1]^{N_1 \times N_2}$, initial dictionary $D_{(0)} \in \mathcal{D} \subset \mathbb{R}^{s \times n}$, parameters $\mu, \lambda > 0$, iteration limits K_1, K_2

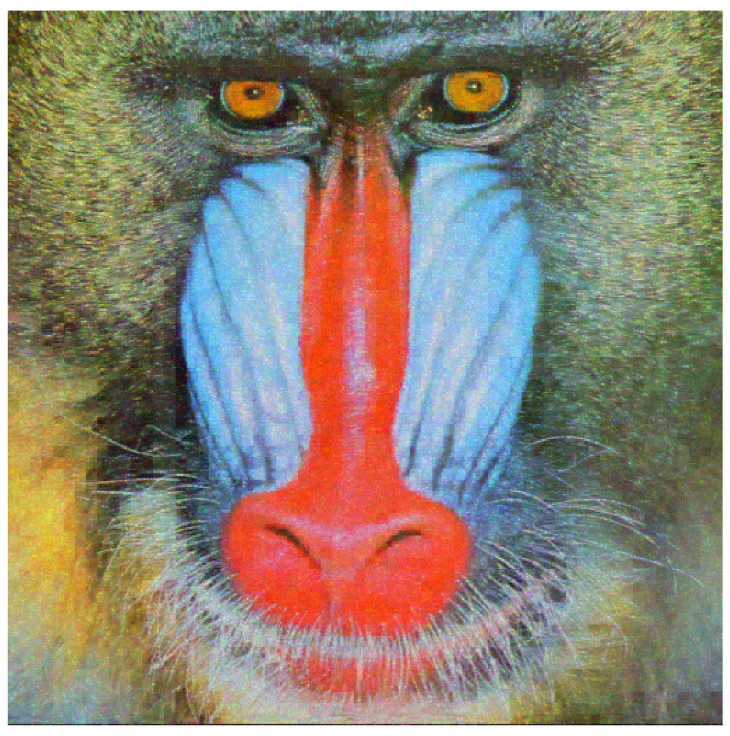
Output: Learned dictionary $D = D_{(K)}$, patch representations $(a^1, \dots, a^p) = A = A_{(K)}$, image reconstruction $X = X_{(K)}$

for $\ell = 0, 1, 2, \dots, K_1 + K_2$ **do**
- choose step size γ_ℓ^A and update $A_{(\ell+1)} \leftarrow \mathcal{S}_{\lambda \gamma_\ell^A / \mu}\Big(A_\ell - \gamma_\ell^A D_{(\ell)}^\top \big(D_{(\ell)} A_{(\ell)} - \mathcal{E}(X_{(\ell)})\big)\Big)$
- choose step size γ_ℓ^X and update $X_{(\ell+1)} \leftarrow \text{proj}_{\mathcal{X}}\Big(X_{(\ell)} - \gamma_\ell^X \Big(\Re\big(\mathcal{F}^*(\mathcal{F}(X) \odot (|\mathcal{F}(X)|^2 - Y))\big) + \mu R \odot \mathcal{R}(\mathcal{E}(X) - DA)\Big)\Big)$
- **if** $\ell < K_1$ **then**
  - keep $D_{(\ell+1)} \leftarrow D_{(\ell)}$
- **else**
  - perform one iteration of block-coordinate descent to obtain updated dictionary $D_{(\ell+1)}$

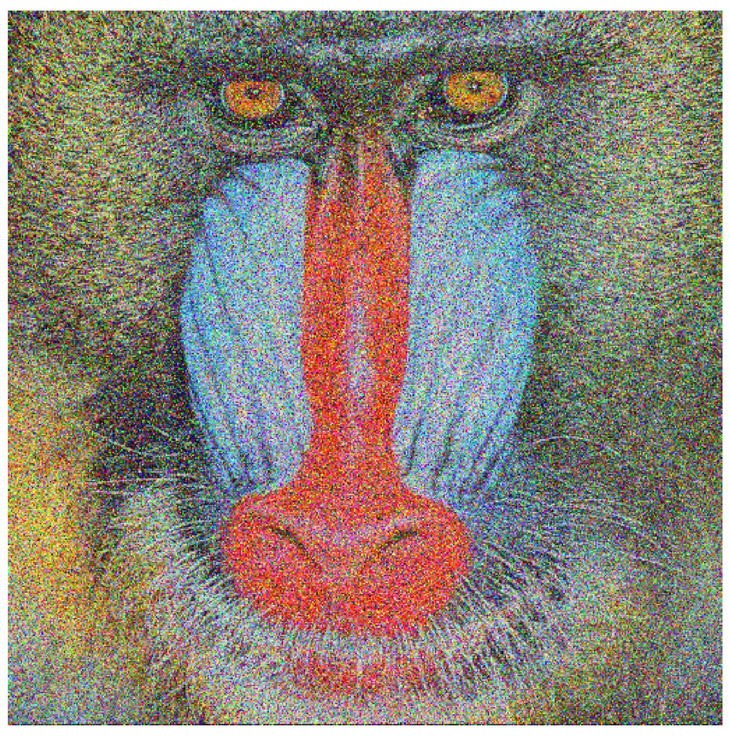
Notation: $\mathcal{E}(X) = (x^1, \dots, x^p)$ extracts patches from image, $\mathcal{R}$ reassembles image from patches, R: weight matrix for averaging pixel values (if patches overlap), $\mathcal{S}_\tau(Z) =:= \max\{0, |Z| - \tau\} \odot \text{sign}(Z)$ (soft-thresholding), $\text{proj}_{\mathcal{X}}(Z) := \max\{0, \min\{1, Z\}\}$

Convergence: Appropriate (Armijo line search) step size selection ensures convergence to a stationary point; mild further conditions give linear convergence rate.

Numerical Examples



$\mathcal{R}(DA) \approx X_{\text{DOLPHIn}}$



X_{WF}

Reconstructions by DOLPHIn and Wirtinger Flow from 2 ternary coded diffraction patterns, corrupted by noise N such that $\text{SNR}(Y, |\mathcal{F}(\hat{X})|^2) = 15\,\text{dB}$ (512×512 RGB image, 8×8 patches, $(\mu, \lambda)/m_Y = (0.05, 0.003)$, $K_1 = 25$, $K_2 = 50$). X_{DOLPHIn}: PSNR 22.31 dB, SSIM 0.61; avg. $\|a^i\|_0$: 17.27; X_{WF}: PSNR 12.66 dB, SSIM 0.20; $t_{\text{DOLPHIn}} \approx 119\,\text{s}$, $t_{\text{WF}} \approx 44\,\text{s}$.

		256 × 256 instances					512 × 512 instances				
$\mathcal{F}$ type	reconstr.	$(\mu, \lambda)/m_Y$	time	PSNR	SSIM	$\varnothing \|a^i\|_0$	$(\mu, \lambda)/m_Y$	time	PSNR	SSIM	$\varnothing \|a^i\|_0$
$G\hat{X}$	X_{DOLPHIn}	(0.5,0.105)	9.28	24.59	0.5673		(0.5,0.105)	48.29	23.41	0.6530	
	$\mathcal{R}(DA)$			23.06	0.6618	3.81			22.67	0.6802	6.40
	X_{WF}		4.75	18.83	0.2840	–		28.94	18.80	0.3770	–
$G\hat{X}G^*$	X_{DOLPHIn}	(0.5,0.210)	33.79	22.71	0.4146		(0.5,0.210)	190.82	22.56	0.5281	
	$\mathcal{R}(DA)$			23.70	0.7321	7.51			23.42	0.7654	11.51
	X_{WF}		32.75	22.71	0.4147	–		192.66	22.57	0.5281	–
$G\hat{X}H^*$	X_{DOLPHIn}	(0.5,0.210)	33.81	22.58	0.4102		(0.5,0.210)	190.54	22.47	0.5241	
	$\mathcal{R}(DA)$			23.63	0.7249	7.61			23.41	0.7647	11.68
	X_{WF}		32.90	22.57	0.4098	–		196.85	22.47	0.5240	–
CDP	X_{DOLPHIn}	(0.05,0.003)	6.75	27.19	0.7414		(0.05,0.003)	27.52	27.34	0.7820	
	$\mathcal{R}(DA)$			26.61	0.7650	8.04			26.33	0.7663	11.50
	X_{WF}		1.45	12.81	0.1093	–		6.63	12.98	0.1537	–

Test results for m_Y Gaussian-type and coded diffraction pattern (CDP) measurements. Reported are mean values (geom. mean for CPU times) per measurement type, obtained from three instances with random $X_{(0)}$ and noise for each of three 256×256 and five 512×512 images, w.r.t. the reconstructions from DOLPHIn (X_{DOLPHIn} and $\mathcal{R}(DA)$) with parameters (μ, λ) and (real-valued, $[0,1]$-constrained) Wirtinger Flow (X_{WF}), resp. ($K_1 = 25$, $K_2 = 50$, 8×8 patches (no overlap)).

CPU times [s], PSNR [dB]. Gauss-type measurements: $G: 4N_1 \times N_2$, noise-SNR 10, CDP: 2 masks, noise-SNR 20.

The work of J. Mairal was funded by the French National Research Agency [Macaron project, ANR-14-CE23-0003-01]. The work of Y. Eldar was funded by the European Union's Horizon 2020 research and innovation programme under grant agreement ERC-BNYQ, and by the Israel Science Foundation under Grant no. 335/14.