Large Inpainting of Face Images with Trainlets

Jeremias Sulam

joint work with Michael Elad

International Conference on Acoustics, Speech and Signal Processing

New Orleans, 5 - 9 March, 2017

Technion
Israel Institute of Technology

Supported by ERC Grant

no. 320649

Image Inpainting

Degradation model

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \mathbf{v}$$

Image Inpainting

Degradation model

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \mathbf{v}$$

Image Inpainting

Degradation model

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \mathbf{v}$$

- Low-level image restoration methods

Image Inpainting

Degradation model

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \mathbf{v}$$

- Low-level image restoration methods

Image Inpainting

Degradation model

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \mathbf{v}$$

- Low-level image restoration methods
- Diffusion/content propagation methods

Image Inpainting

Degradation model

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \mathbf{v}$$

- Low-level image restoration methods
- Diffusion/content propagation methods

Image Inpainting

Degradation model

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \mathbf{v}$$

- Low-level image restoration methods
- Diffusion/content propagation methods
- Intrinsic need of a **global model**

Learning High Dimensional Model

Difficulties

- Computational hard problem
- Curse of dimensionality

Learning High Dimensional Model

Difficulties

- Computational hard problem
- Curse of dimensionality

Related methods

- Manifold learning techniques
- Some global models
- Dictionary Learning

Learning High Dimensional Model

Difficulties

- Computational hard problem
- Curse of dimensionality

Related methods

- Manifold learning techniques
- Some global models
- Dictionary Learning

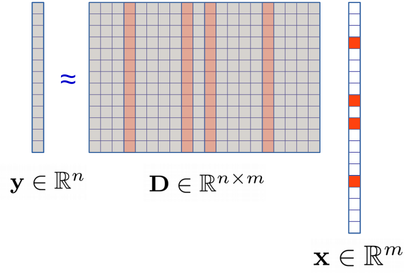
Our solution

- We employ a high dimensional dictionary learning method to learn a global model of face images
- Solve an inverse problem regularized with a sparse prior

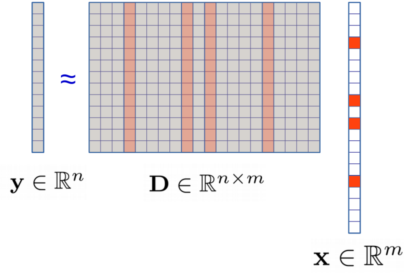
Contents

Sparse Representations

Sparse Representations



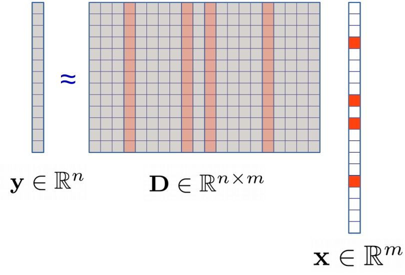
Sparse Representations



Sparse Coding

$$\min_{\mathbf{x}} \ ||\mathbf{x}||_0 \text{ subject to } ||\mathbf{y} - \mathbf{D}\mathbf{x}||_2^2 \leq \epsilon^2,$$

Sparse Representations

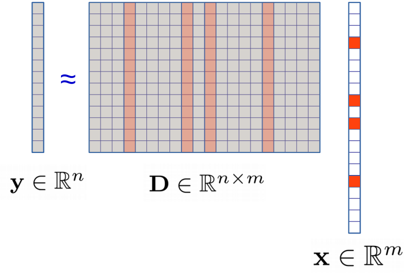


Sparse Coding

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ subject to } \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 \leq \epsilon^2,$$

- Greedy Pursuit (OMP, . . .)

Sparse Representations



Sparse Coding

$$\min_{\mathbf{x}} \ ||\mathbf{x}||_0 \text{ subject to } ||\mathbf{y} - \mathbf{D}\mathbf{x}||_2^2 \leq \epsilon^2,$$

- Greedy Pursuit (OMP, ...)
- Relaxation Methods

Sparse Representations

$$\min_{\mathbf{X}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{DX}\|_F^2 \qquad \text{subject to} \quad \|\mathbf{x}_i\|_0 \leq T$$

Sparse Representations

$$\min_{\mathbf{X}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{DX}\|_F^2 \qquad \text{subject to} \quad \|\mathbf{x}_i\|_0 \leq T$$

The Choice of the Dictionary $\mathbf{D}$

- Transforms
  - Structured Matrices (Fast Algorithms)
  - Fair sparsification

Sparse Representations

$$\min_{\mathbf{X},\mathbf{D}} \quad \frac{1}{2}\|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 \qquad \text{subject to} \quad \|\mathbf{x}_i\|_0 \le T$$

The Choice of the Dictionary $\mathbf{D}$

- Transforms
  - Structured Matrices (Fast Algorithms)
  - Fair sparsification
- Learnt dictionaries
  - Unstructured matrices
  - Great sparsification

Sparse Representations

$$\min_{\mathbf{X},\mathbf{D}} \quad \frac{1}{2}\|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 \qquad \text{subject to} \quad \|\mathbf{x}_i\|_0 \leq T$$

The Choice of the Dictionary $\mathbf{D}$

- Transforms
  - Structured Matrices (Fast Algorithms)
  - Fair sparsification
- Learnt dictionaries
  - Unstructured matrices
  - Great sparsification

Dictionary Learning in Image Processing

- Many successful applications and results
  - Image denoising, inpainting, demosaicing
  - Image Compression
  - Reconstruction from incomplete measurements

Sparse Representations

$$\min_{\mathbf{X},\mathbf{D}} \quad \frac{1}{2}\|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 \qquad \text{subject to} \quad \|\mathbf{x}_i\|_0 \leq T$$

The Choice of the Dictionary $\mathbf{D}$

- Transforms
  - Structured Matrices (Fast Algorithms)
  - Fair sparsification
- Learnt dictionaries
  - Unstructured matrices
  - Great sparsification

Dictionary Learning in Image Processing

- Many successful applications and results
  - Image denoising, inpainting, demosaicing
  - Image Compression
  - Reconstruction from incomplete measurements
- Learning on small signal patches

Sparse Representations

$$\min_{\mathbf{X},\mathbf{D}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 \qquad \text{subject to} \quad \|\mathbf{x}_i\|_0 \leq T$$

The Choice of the Dictionary $\mathbf{D}$

- Transforms
  - Structured Matrices (Fast Algorithms)
  - Fair sparsification
- Learnt dictionaries
  - Unstructured matrices
  - Great sparsification

Dictionary Learning in Image Processing

- Many successful applications and results
  - Image denoising, inpainting, demosaicing
  - Image Compression
  - Reconstruction from incomplete measurements
- Learning on small signal patches
- Computational Constraints

Up-Scaling Dictionary Learning

Problem Formulation

$$\min_{\mathbf{A},\mathbf{X}} ||\mathbf{Y} - \boldsymbol{\Phi}\mathbf{A}\mathbf{X}||_F^2 \quad \text{subject to} \quad ||\mathbf{x}_i||_0 \leq p, \; ||\mathbf{a}_i||_0 \leq k$$

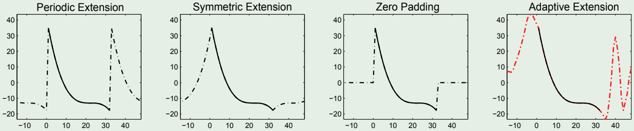
Up-Scaling Dictionary Learning

Problem Formulation

$$\min_{\mathbf{A},\mathbf{X}} ||\mathbf{Y} - \mathbf{\Phi A X}||_F^2 \quad \text{subject to} \quad ||\mathbf{x}_i||_0 \leq p, \ ||\mathbf{a}_i||_0 \leq k$$

Φ: Cropped Wavelets

The transform for signal $\mathbf{f}$ is defined in terms of a pursuit over a **convolutional and multi-scale dictionary**, providing sparsest wavelet representations by optimally (implicitly) extending the signal borders.



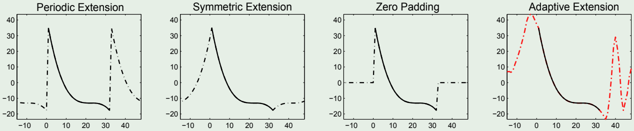
Up-Scaling Dictionary Learning

Problem Formulation

$$\min_{\mathbf{A},\mathbf{X}} ||\mathbf{Y} - \mathbf{\Phi A X}||_F^2 \quad \text{subject to} \quad ||\mathbf{x}_i||_0 \leq p,\ ||\mathbf{a}_i||_0 \leq k$$

Φ: Cropped Wavelets

The transform for signal $\mathbf{f}$ is defined in terms of a pursuit over a **convolutional and multi-scale dictionary**, providing sparsest wavelet representations by optimally (implicitly) extending the signal borders.



Online Learning

- Faster convergence
- Training on millions of examples

OSDL in practice

$$\min_{\mathbf{X},\mathbf{A}} \underbrace{\frac{1}{2}||\mathbf{Y}-\mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2}_{f(\mathbf{X},\mathbf{A})} \quad \text{s.t.} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 \leq k & \forall j \end{cases} .$$

OSDL in practice

$$\min_{\mathbf{X},\mathbf{A}} \underbrace{\frac{1}{2}||\mathbf{Y}-\mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2}_{f(\mathbf{X},\mathbf{A})} \quad \text{s.t.} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 \leq k & \forall j \end{cases}.$$

Data: Training samples $\{\mathbf{y}_i\}$, base-dictionary $\mathbf{\Phi}$, initial sparse matrix $\mathbf{A}^0$

for $t = 1, \ldots, T$ **do**

  Draw a mini-batch $\mathbf{Y}_t$ at random ;

OSDL in practice

$$\underbrace{\min_{\mathbf{X},\mathbf{A}} \frac{1}{2}||\mathbf{Y}-\mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2}_{f(\mathbf{X},\mathbf{A})} \quad \text{s.t.} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 \leq k & \forall j \end{cases}.$$

Data: Training samples $\{\mathbf{y}_i\}$, base-dictionary $\mathbf{\Phi}$, initial sparse matrix $\mathbf{A}^0$

for $t = 1, \ldots, T$ **do**

Draw a mini-batch $\mathbf{Y}_t$ at random ;

$\mathbf{X}_t \leftarrow$ Sparse Code $(\mathbf{Y}_t, \mathbf{\Phi}, \mathbf{A}^t, \mathbf{G}^t)$;

OSDL in practice

$$\min_{\mathbf{X},\mathbf{A}} \underbrace{\frac{1}{2}||\mathbf{Y}-\mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2}_{f(\mathbf{X},\mathbf{A})} \quad \text{s.t.} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 \leq k & \forall j \end{cases}.$$

Data: Training samples $\{\mathbf{y}_i\}$, base-dictionary $\mathbf{\Phi}$, initial sparse matrix $\mathbf{A}^0$

for $t = 1, \ldots, T$ **do**

Draw a mini-batch $\mathbf{Y}_t$ at random ;

$\mathbf{X}_t \leftarrow$ Sparse Code $(\mathbf{Y}_t, \mathbf{\Phi}, \mathbf{A}^t, \mathbf{G}^t)$;

$\mathbf{A}_{\mathcal{S}}^{t+1} = \mathcal{P}_k\left[\mathbf{A}_{\mathcal{S}}^t - \eta^t \nabla f(\mathbf{A}_{\mathcal{S}}^t)\right]$;

OSDL in practice

$$\underbrace{\min_{\mathbf{X},\mathbf{A}} \frac{1}{2}||\mathbf{Y} - \mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2}_{f(\mathbf{X},\mathbf{A})} \quad \text{s.t.} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 \leq k & \forall j \end{cases}.$$

Data: Training samples $\{\mathbf{y}_i\}$, base-dictionary $\mathbf{\Phi}$, initial sparse matrix $\mathbf{A}^0$

for $t = 1, \ldots, T$ **do**

- Draw a mini-batch $\mathbf{Y}_t$ at random ;
- $\mathbf{X}_t \leftarrow$ Sparse Code $(\mathbf{Y}_t, \mathbf{\Phi}, \mathbf{A}^t, \mathbf{G}^t)$;
- $\mathbf{A}_{\mathcal{S}}^{t+1} = \mathcal{P}_k \left[\mathbf{A}_{\mathcal{S}}^t - \eta^t \nabla f(\mathbf{A}_{\mathcal{S}}^t)\right]$;
- Update columns and rows of $\mathbf{G}$ by $\left(\mathbf{A}^{t+1}\right)^T \mathbf{G}_{\mathbf{\Phi}} \mathbf{A}_{\mathcal{S}}^{t+1}$;

OSDL in practice

$$\min_{\mathbf{X},\mathbf{A}} \underbrace{\frac{1}{2}||\mathbf{Y}-\mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2}_{f(\mathbf{X},\mathbf{A})} \quad \text{s.t.} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 \leq k & \forall j \end{cases} .$$

Data: Training samples $\{\mathbf{y}_i\}$, base-dictionary $\mathbf{\Phi}$, initial sparse matrix $\mathbf{A}^0$

for $t = 1, \ldots, T$ **do**

Draw a mini-batch $\mathbf{Y}_t$ at random ;

$\mathbf{X}_t \leftarrow$ Sparse Code $(\mathbf{Y}_t, \mathbf{\Phi}, \mathbf{A}^t, \mathbf{G}^t)$;

$\mathbf{A}_{\mathcal{S}}^{t+1} = \mathcal{P}_k \left[\mathbf{A}_{\mathcal{S}}^t - \eta^t \nabla f(\mathbf{A}_{\mathcal{S}}^t)\right]$;

Update columns and rows of $\mathbf{G}$ by $\left(\mathbf{A}^{t+1}\right)^T \mathbf{G}_{\mathbf{\Phi}} \mathbf{A}_{\mathcal{S}}^{t+1}$;

end

Result: Sparse Dictionary $\mathbf{A}$

OSDL in practice

$$\min_{\mathbf{X},\mathbf{A}} \underbrace{\frac{1}{2}||\mathbf{Y} - \mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2}_{f(\mathbf{X},\mathbf{A})} \quad \text{s.t.} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 \leq k & \forall j \end{cases}.$$

Data: Training samples $\{\mathbf{y}_i\}$, base-dictionary $\mathbf{\Phi}$, initial sparse matrix $\mathbf{A}^0$

for $t = 1, \ldots, T$ **do**

- Draw a mini-batch $\mathbf{Y}_t$ at random ;
- $\mathbf{X}_t \leftarrow$ Sparse Code $(\mathbf{Y}_t, \mathbf{\Phi}, \mathbf{A}^t, \mathbf{G}^t)$;
- $\mathbf{A}_{\mathcal{S}}^{t+1} = \mathcal{P}_k \left[\mathbf{A}_{\mathcal{S}}^t - \eta^t \nabla f(\mathbf{A}_{\mathcal{S}}^t)\right]$;
- Update columns and rows of $\mathbf{G}$ by $\left(\mathbf{A}^{t+1}\right)^T \mathbf{G}_{\mathbf{\Phi}} \mathbf{A}_{\mathcal{S}}^{t+1}$;

end

Result: Sparse Dictionary $\mathbf{A}$

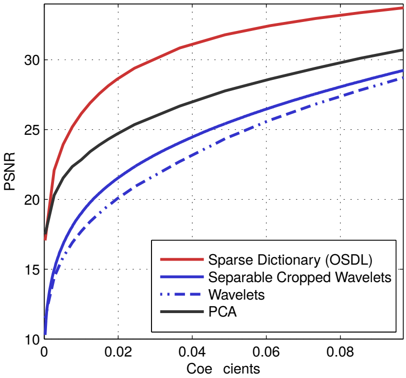
- Incorporate a momentum variable
- Analytical step-size
- Replace repeated/unused atoms

Trainlets for Data Approximation

- 64×64 Images.
- $\approx$ 12K training examples.
- Non-Redundant Dictionary ($\approx$ 4K atoms).
- Atoms Sparsity: 300.
- $\mathbf{\Phi}$: Db4 cropped-wavelets ($r = 1.37$, 1D) .

Trainlets for Data Approximation

- 64×64 Images.
- ≈ 12K training examples.
- Non-Redundant Dictionary (≈ 4K atoms).
- Atoms Sparsity: 300.
- $\mathbf{\Phi}$: Db4 cropped-wavelets ($r = 1.37$, 1D) .



Training the Model

- Several face-images dataset $\approx 19K$ images
- Images size: 100×100
- No pre-processing (alignment, coherent scaling, etc)

Training the Model

- Several face-images dataset $\approx 19K$ images
- Images size: 100×100
- No pre-processing (alignment, coherent scaling, etc)

- Cropped Daubechies Wavelets (4 v.m.)
- $6,000$ atoms, each $1,000$ nnz (6% sparse!)

Training the Model

- Several face-images dataset $\approx 19K$ images
- Images size: 100×100
- No pre-processing (alignment, coherent scaling, etc)

- Cropped Daubechies Wavelets (4 v.m.)
- $6,000$ atoms, each $1,000$ nnz (6% sparse!)

Contents

Large Image Inpainting

Problem Formulation

$$\min_{\mathbf{x}} \ \|\mathbf{x}\|_0 \text{ subject to } \|\mathbf{y} - \mathbf{MDx}\|_2 \leq \epsilon.$$

Large Image Inpainting

Problem Formulation

$$\min_{\mathbf{x}} \ \|\mathbf{x}\|_0 \text{ subject to } \|\mathbf{y} - \mathbf{MDx}\|_2 \leq \epsilon.$$

Relaxed to:

$$\min_{\mathbf{x}} \ \|\mathbf{y} - \mathbf{MDx}\|_2 + \lambda\|\mathbf{x}\|_1,$$

Large Image Inpainting

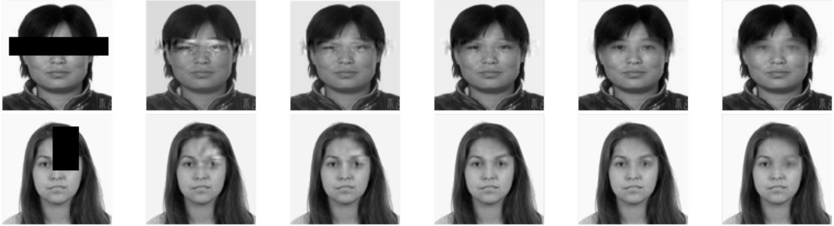
Problem Formulation

$$\min_{\mathbf{x}} \ \|\mathbf{x}\|_0 \text{ subject to } \|\mathbf{y} - \mathbf{MDx}\|_2 \leq \epsilon.$$

Relaxed to:

$$\min_{\mathbf{x}} \ \|\mathbf{y} - \mathbf{MDx}\|_2 + \lambda\|\mathbf{x}\|_1,$$

- Effect of regularization



$$\lambda \rightarrow$$

Contents

Inpainting Results

Inpainting Results

Masked Image

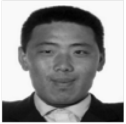
Patch Propagation

PCA

SEDIL

Trainlets

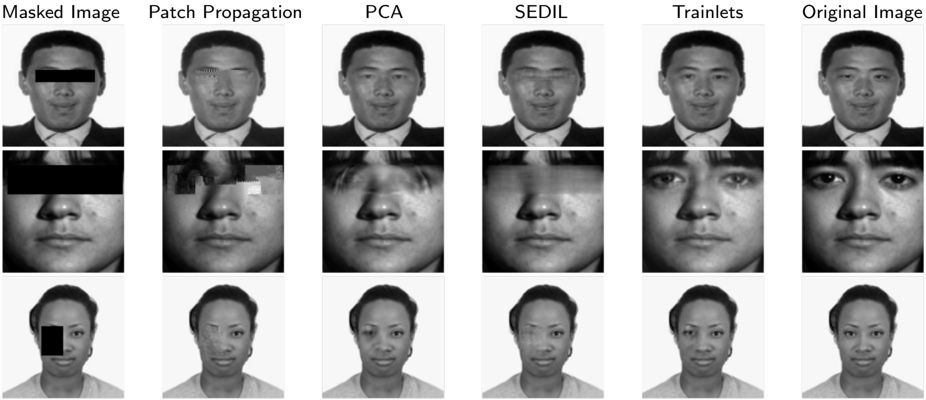
Original Image



Inpainting Results

Masked Image | Patch Propagation | PCA | SEDIL | Trainlets | Original Image

Inpainting Results



Inpainting Results

Masked Image | Patch Propagation | PCA | SEDIL | Trainlets | Original Image

Inpainting Results

Masked Image

Patch Propagation

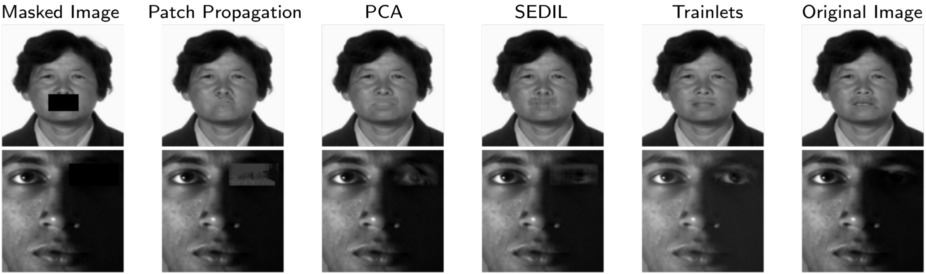
PCA

SEDIL

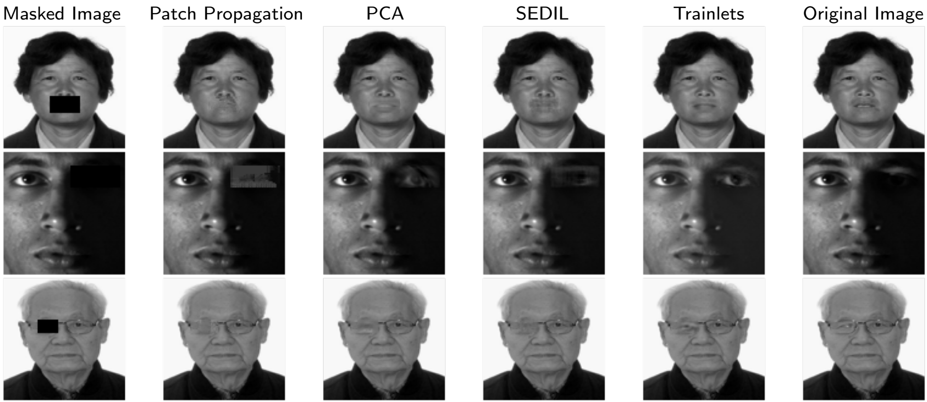
Trainlets

Original Image

Inpainting Results

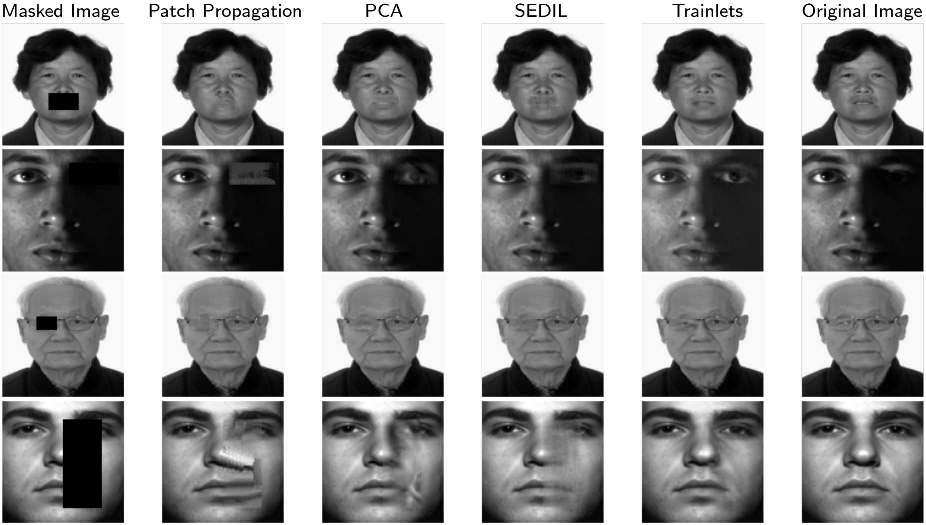


Inpainting Results



Inpainting Results

Inpainting Results



Contents

Concluding Remarks

- We exploit the representation power of Trainlets to learn a global model
- Very simple problem formulation
- No extra algorithmic manipulation are needed
- Plausible reconstructions – while different from the original images

- Larger dataset would boost the model
- Other type of inverse problems?

Concluding Remarks

- We exploit the representation power of Trainlets to learn a global model
- Very simple problem formulation
- No extra algorithmic manipulation are needed
- Plausible reconstructions – while different from the original images

- Larger dataset would boost the model
- Other type of inverse problems?

Questions?

Code and model available at `jsulam.cswp.cs.technion.ac.il`