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

Supported by ERC Grant
no. 320649
Image Inpainting

Degradation model

\[ y = Mz + v \]
Image Inpainting

Degradation model

\[ y = Mz + v \]
Image Inpainting

Degradation model

\[ y = Mz + v \]

- Low-level image restoration methods
Degradation model

\[ y = Mz + v \]

- Low-level image restoration methods
Image Inpainting

Degradation model

\[ y = Mz + v \]

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

\[ y = Mz + v \]

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

\[ y = Mz + v \]

- 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
Background

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 by solving an inverse problem regularized with a sparse prior.
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

1 Background

2 Learning the Model

3 Inpainting algorithm

4 Results

5 Conclusion
Learning the Model

Sparse Representations

$$\min_x \|x\|_0 \text{ subject to } \|y - Dx\|_2^2 \leq \epsilon^2,$$

Greedy Pursuit (OMP, ...)

Relaxation Methods
Sparse Representations

\[ y \in \mathbb{R}^n \quad D \in \mathbb{R}^{n \times m} \quad x \in \mathbb{R}^m \]
Sparse Representations

\[ y \in \mathbb{R}^n \quad D \in \mathbb{R}^{n \times m} \quad x \in \mathbb{R}^m \]

Sparse Coding

\[
\min_{x} \|x\|_0 \quad \text{subject to} \quad \|y - Dx\|_2^2 \leq \epsilon^2,
\]
Sparse Representations

\[ y \in \mathbb{R}^n \quad D \in \mathbb{R}^{n \times m} \quad x \in \mathbb{R}^m \]

Sparse Coding

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

- Greedy Pursuit (OMP, ...)

Learning the Model
Face Images Inpainting with Trainlets
March 7, 2017 5 / 17
Sparse Representations

\[ y \in \mathbb{R}^n \quad \text{D \in \mathbb{R}^{n \times m}} \quad x \in \mathbb{R}^m \]

Sparse Coding

\[
\min_{x} \|x\|_0 \quad \text{subject to} \quad \|y - Dx\|_2^2 \leq \epsilon^2,
\]

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

\[
\min_X \quad \frac{1}{2} \|Y - DX\|_F^2, \quad \text{subject to} \quad \|x_i\|_0 \leq T
\]
Sparse Representations

\[ \min_X \frac{1}{2} \| Y - DX \|_F^2 \quad \text{subject to} \quad \| x_i \|_0 \leq T \]

The Choice of the Dictionary \( D \)

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

\[ \min_{X,D} \frac{1}{2} \|Y - DX\|_F^2, \quad \text{subject to} \quad \|x_i\|_0 \leq T \]

The Choice of the Dictionary \(D\)

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

- **Learnt dictionaries**
  - Unstructured matrices
  - Great sparsification
Sparse Representations

\[ \min_{X,D} \frac{1}{2} \| Y - DX \|_F^2 \quad \text{subject to} \quad \| x_i \|_0 \leq T \]

The Choice of the Dictionary \( D \)

- **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

\[
\min_{X,D} \quad \frac{1}{2} \|Y - DX\|_F^2, \quad \text{subject to} \quad \|x_i\|_0 \leq T
\]

The Choice of the Dictionary \( D \)
- **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

$$\min_{X,D} \frac{1}{2} \|Y - DX\|_F^2, \quad \text{subject to } \|x_i\|_0 \leq T$$

### The Choice of the Dictionary $D$

- **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

\[
\min_{A, X} \|Y - \Phi AX\|^2_F \quad \text{subject to} \quad \|x_i\|_0 \leq p, \ |a_i|_0 \leq k
\]
Up-Scaling Dictionary Learning

Problem Formulation

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

\( \Phi \): Cropped Wavelets

The transform for signal \( f \) 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.
Learning the Model

Up-Scaling Dictionary Learning

Problem Formulation

\[
\min_{A,X} \| Y - \Phi A X \|_F^2 \quad \text{subject to} \quad \| x_i \|_0 \leq p, \quad \| a_i \|_0 \leq k
\]

\( \Phi \): Cropped Wavelets

The transform for signal \( f \) 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

\[
\min_{X,A} \frac{1}{2}\|Y - \Phi AX\|_F^2 \quad \text{s.t.} \quad \begin{cases} 
\|x_i\|_0 \leq p & \forall i \\
\|a_j\|_0 \leq k & \forall j 
\end{cases}
\]
OSDL in practice

$$\min_{X,A} \frac{1}{2} \|Y - \Phi AX\|^2_F \quad \text{s.t.} \quad \begin{cases} \|x_i\|_0 \leq p & \forall i \\ \|a_j\|_0 \leq k & \forall j \end{cases}$$

**Data:** Training samples \( \{y_i\} \), base-dictionary \( \Phi \), initial sparse matrix \( A^0 \)

**for** \( t = 1, \ldots, T \) **do**

Draw a mini-batch \( Y_t \) at random ;
OSDL in practice

\[
\min_{X,A} \frac{1}{2} \| Y - \Phi A X \|_F^2 \quad \text{s.t.} \quad \begin{cases} 
\| x_i \|_0 \leq p & \forall i \\
\| a_j \|_0 \leq k & \forall j
\end{cases}
\]

**Data:** Training samples \( \{ y_i \} \), base-dictionary \( \Phi \), initial sparse matrix \( A^0 \)

**for** \( t = 1, \ldots, T \) **do**

- Draw a mini-batch \( Y_t \) at random;
- \( X_t \leftarrow \text{Sparse Code} \ (Y_t, \Phi, A^t, G^t) \);
Learning the Model

**OSDL in practice**

\[
\min_{X,A} \frac{1}{2} \|Y - \Phi AX\|_F^2 \quad \text{s.t.} \quad \begin{cases} \|x_i\|_0 \leq p & \forall i \\ \|a_j\|_0 \leq k & \forall j \end{cases}.
\]

**Data:** Training samples \( \{y_i\} \), base-dictionary \( \Phi \), initial sparse matrix \( A^0 \)

**for** \( t = 1, \ldots, T \)** do**

- Draw a mini-batch \( Y_t \) at random ;
- \( X_t \leftarrow \text{Sparse Code} \left( Y_t, \Phi, A^t, G^t \right) \);
- \( A_{t+1} = P_k \left[ A_t - \eta^t \nabla f(A_t) \right] \);

Incorporate a momentum variable

Analytical step-size

Replace repeated/unused atoms
OSDL in practice

\[
\min_{X,A} \frac{1}{2}\|Y - \Phi AX\|_F^2 \quad \text{s.t.} \quad \begin{cases} 
\|x_i\|_0 \leq p & \forall i \\
\|a_j\|_0 \leq k & \forall j
\end{cases}
\]

Data: Training samples \(\{y_i\}\), base-dictionary \(\Phi\), initial sparse matrix \(A^0\)

for \(t = 1, \ldots, T\) do

- Draw a mini-batch \(Y_t\) at random;
- \(X_t \leftarrow \text{Sparse Code} (Y_t, \Phi, A^t, G^t)\);
- \(A_{t+1}^S = \mathcal{P}_k [A_t^S - \eta^t \nabla f(A_t^S)]\);
- Update columns and rows of \(G\) by \((A^{t+1})^T G \Phi A^{t+1}_S\);
OSDL in practice

\[
\min_{X,A} \frac{1}{2} \|Y - \Phi AX\|_F^2 \quad \text{s.t.} \quad \begin{cases} 
\|x_i\|_0 \leq p \quad \forall i \\
\|a_j\|_0 \leq k \quad \forall j
\end{cases}
\]

Data: Training samples \(\{y_i\}\), base-dictionary \(\Phi\), initial sparse matrix \(A^0\)

for \(t = 1, \ldots, T\) do

- Draw a mini-batch \(Y_t\) at random;
- \(X_t \leftarrow \text{Sparse Code (}Y_t, \Phi, A^t, G^t\) ;
- \(A^{t+1}_S = P_k [A^t_S - \eta^t \nabla f(A^t_S)] ;
- \text{Update columns and rows of } G \text{ by } (A^{t+1})^T G \Phi A^{t+1}_S ;

end

Result: Sparse Dictionary \(A\)
OSDL in practice

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

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

**for** \( t = 1, \ldots, T \) **do**

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

**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.
- \( \Phi : \text{Db4 cropped-wavelets} \ (r = 1.37, 1D) \).
Learning the Model

Trainlets for Data Approximation

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

![Graph showing PSNR against coefficients for different methods]

- Sparse Dictionary (OSDL)
- Separable Cropped Wavelets
- Wavelets
- PCA
Training the Model

- Several face-images dataset ≈ 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: $100 \times 100$
- 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: $100 \times 100$
- No pre-processing (alignment, coherent scaling, etc)

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

1 Background

2 Learning the Model

3 Inpainting algorithm

4 Results

5 Conclusion
### Problem Formulation

\[
\min_x \|x\|_0 \quad \text{subject to} \quad \|y - MDx\|_2 \leq \epsilon.
\]
Inpainting algorithm

Large Image Inpainting

Problem Formulation

\[
\min_x \|x\|_0 \ \text{subject to} \ \|y - MDx\|_2 \leq \epsilon.
\]

Relaxed to:

\[
\min_x \|y - MDx\|_2 + \lambda \|x\|_1,
\]
Problem Formulation

\[
\min_x \|x\|_0 \text{ subject to } \|y - MDx\|_2 \leq \epsilon.
\]

Relaxed to:

\[
\min_x \|y - MDx\|_2 + \lambda \|x\|_1,
\]

- Effect of regularization

\[
\lambda \rightarrow
\]
Results

Inpainting Results
Inpainting Results

Masked Image  Patch Propagation  PCA  SEDIL  Trainlets  Original Image
Inpainting Results

Masked Image | Patch Propagation | PCA | SEDIL | Trainlets | Original Image
---|---|---|---|---|---
![Masked Image](image1.png) | ![Patch Propagation](image2.png) | ![PCA](image3.png) | ![SEDIL](image4.png) | ![Trainlets](image5.png) | ![Original Image](image6.png)
Inpainting Results

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

[Images of inpainting results for different techniques applied to face images]
## Inpainting Results

<table>
<thead>
<tr>
<th>Masked Image</th>
<th>Patch Propagation</th>
<th>PCA</th>
<th>SEDIL</th>
<th>Trainlets</th>
<th>Original Image</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Masked Image" /></td>
<td><img src="image2" alt="Patch Propagation" /></td>
<td><img src="image3" alt="PCA" /></td>
<td><img src="image4" alt="SEDIL" /></td>
<td><img src="image5" alt="Trainlets" /></td>
<td><img src="image6" alt="Original Image" /></td>
</tr>
<tr>
<td><img src="image7" alt="Masked Image" /></td>
<td><img src="image8" alt="Patch Propagation" /></td>
<td><img src="image9" alt="PCA" /></td>
<td><img src="image10" alt="SEDIL" /></td>
<td><img src="image11" alt="Trainlets" /></td>
<td><img src="image12" alt="Original Image" /></td>
</tr>
<tr>
<td><img src="image13" alt="Masked Image" /></td>
<td><img src="image14" alt="Patch Propagation" /></td>
<td><img src="image15" alt="PCA" /></td>
<td><img src="image16" alt="SEDIL" /></td>
<td><img src="image17" alt="Trainlets" /></td>
<td><img src="image18" alt="Original Image" /></td>
</tr>
</tbody>
</table>

J. Sulam

Face Images Inpainting with Trainlets

March 7, 2017
Inpainting Results

Masked Image  Patch Propagation  PCA  SEDIL  Trainlets  Original Image
Inpainting Results

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

- Masked Image
- Patch Propagation
- PCA
- SEDIL
- Trainlets
- Original Image
Inpainting Results

Masked Image  Patch Propagation  PCA  SEDIL  Trainlets  Original Image

J. Sulam
Face Images Inpainting with Trainlets
March 7, 2017
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