# Sparse Learned Representations for Image Restoration

Julien Mairal - INRIA, WILLOW project

Francis Bach - INRIA
Michael Elad - The Technion
Jean Ponce - ENS/INRIA
Guillermo Sapiro - University of Minnesota
Andrew Zisserman - INRIA/Oxford University

Rennes, April 2009

## What this talk is about

- The dictionary learning paradigm.
- Various formulations for image and video processing.
- A fast online dictionary learning algorithm.
- Processing raw images from digital cameras.

1. Sparse representations for image denoising
2. Formulations for image and video processing
3. Online Dictionary Learning
4. Raw image processing

# Sparse representations for image denoising

$$\underbrace{\mathbf{y}}_{\text{measurements}} = \underbrace{\mathbf{x}_{orig}}_{\text{original image}} + \underbrace{\mathbf{w}}_{\text{white Gaussian noise}}$$

# Sparse representations for image denoising

Energy minimization problem - MAP estimation:

$$E(\mathbf{x}) = \underbrace{||\mathbf{y} - \mathbf{x}||_2^2}_{\text{relation to measurements}} + \underbrace{Pr(\mathbf{x})}_{\text{p}\textit{rior}}$$

# Sparse representations for image denoising

## Some classical priors

- Smoothness $\lambda||\mathbf{Lx}||_2^2$
- Total variation $\lambda||\nabla\mathbf{x}||_2^2$
- Wavelet sparsity $\lambda||\mathbf{Wx}||_1$
- ...

# Sparse representations for image denoising

## Sparsity and redundancy

$$Pr(\mathbf{x}) = \lambda ||\boldsymbol{\alpha}||_0 \text{ for } \mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$$

$$\underbrace{\begin{pmatrix} \\ \mathbf{x} \\ \\ \end{pmatrix}}_{\mathbf{x} \in \mathbb{R}^m} = \underbrace{\left(\begin{array}{c|c|c|c} & & & \\ \mathbf{d}_1 & \mathbf{d}_2 & \cdots & \mathbf{d}_k \\ & & & \end{array}\right)}_{\mathbf{D} \in \mathbb{R}^{m \times k}} \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix}}_{\boldsymbol{\alpha} \in \mathbb{R}^k, \text{sparse}}$$

# Sparse representations for image denoising

[Haar 1909], [Zweig, Morlet, Grossman ∼70s], [Meyer, Mallat, Daubechies, Coifman, Donoho, Candes ∼80s-today]...

## Which dictionary to choose?

- Wavelets
- Curvelets
- Wedgelets
- Bandlets
- ...lets

# Sparse representations for image denoising

[Fields & Olshausen '96], [MOD: Engan et. al '99],[Lewicki & Sejnowski '00],[K-SVD: Aharon, Elad & Bruckstein '04 '05],[FoE: Roth & Black '05],[Lee et al. '06],[Neural nets: Lecun, Hinton ~90s-today.]

## Learned dictionaries of patches

$$\min_{\boldsymbol{\alpha}_i, \mathbf{D} \in \mathcal{C}} \sum_i \underbrace{||\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i||_2^2}_{\text{reconstruction}} + \underbrace{\lambda \phi(\boldsymbol{\alpha}_i)}_{\text{sparsity}}$$

- $\phi(\boldsymbol{\alpha}) = ||\boldsymbol{\alpha}||_0$ ("$\ell_0$ pseudo-norm")
- $\phi(\boldsymbol{\alpha}) = ||\boldsymbol{\alpha}||_1$ ($\ell_1$ norm)

# Sparse representations for image denoising
## K-SVD: [Elad & Aharon ('06)]

### Key ideas for denoising

- Consider each patch of size $\sqrt{m} \times \sqrt{m}$ ($\sqrt{m} = 8$) in the image, including overlaps.
- learn the dictionary on the corrupted image.
- the Sparse Coding retrieve a sparse approximation of the *noisy* patches.
- Average the approximation of each patch to reconstruct the full image.

# Sparse representations for image denoising

## $\ell_0$ vs $\ell_1$

- Use $\ell_1$ for learning the dictionary!
- Use $\ell_0$ for reconstructing the image!

# Sparse representations for image denoising

## K-SVD: [Elad & Aharon ('06)]

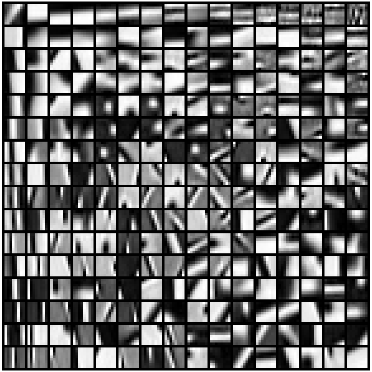
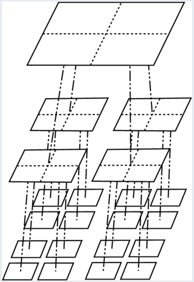


Figure: Dictionary trained on a noisy version of the image boat using K-SVD.

# A multiscale extension
[Mairal, Sapiro & Elad ('07)]

## The key changes

- A Quadtree for each patch
- One dictionary per scale
- multiscale decomposition of each patch



# A multiscale extension
## [Mairal, Sapiro & Elad ('07)]

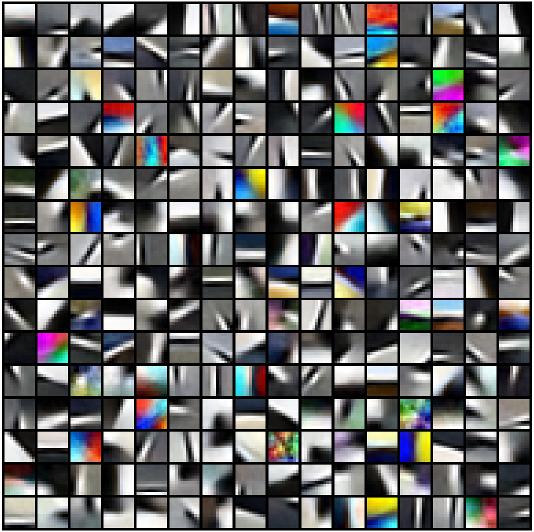
Figure: On the left: original image. In the middle, image corrupted ($\sigma = 15$). On the right, the result with 3 scales (PSNR=32.01dB)

# Preliminary results
## [Mairal, Bach, Ponce, Sapiro & Zisserman ('09)]

# Preliminary results
## [Mairal, Bach, Ponce, Sapiro & Zisserman ('09)]

# Sparse representations for image restoration

## Color denoising, [Mairal, Sapiro & Elad ('08)]



- Most of the atoms are grey!
- Color sparse approximations suffers from color artefacts.
- Average color should be taken into account during sparse approximation!

# Sparse representations for image restoration

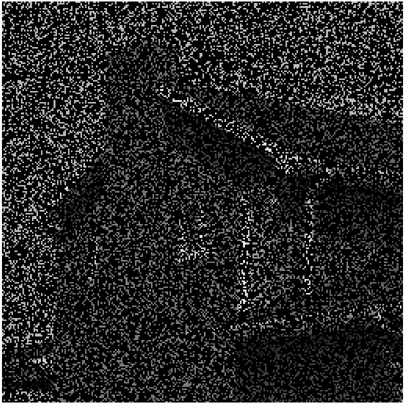
## Color denoising, [Mairal, Sapiro & Elad ('08)]

Figure: Denoising result for $\sigma = 25$ and 2 scales.

$$\min_{\boldsymbol{\alpha}_i, \mathbf{D} \in \mathcal{C}} \sum_i \underbrace{||\boldsymbol{\beta}_i(\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i)||_2^2}_{\text{reconstruction}} + \underbrace{\lambda \phi(\boldsymbol{\alpha}_i)}_{\text{sparsity}}$$

# Sparse representations for image restoration

## Inpainting, [Mairal, Sapiro & Elad ('08)]



Restored image on the right.

# Sparse representations for image restoration

## Inpainting, [Mairal, Sapiro & Elad ('08)]

# Sparse representations for image restoration

## Inpainting, [Mairal, Sapiro & Elad ('08)]

# Sparse representations for image restoration

## Inpainting, [Mairal, Sapiro & Elad ('08)]

# Sparse representations for image restoration
## Inpainting, [Mairal, Sapiro & Elad ('08)]

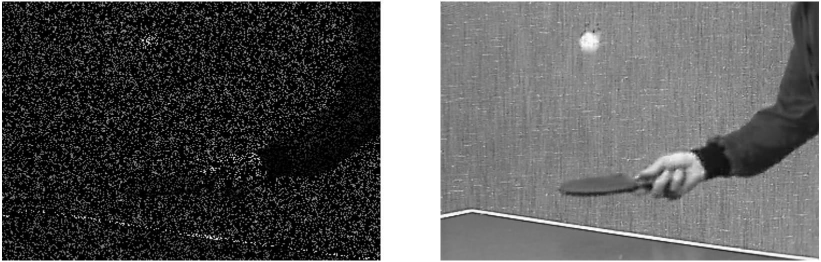


Figure: Inpainting results with two scales.

# Sparse representations for image restoration
## Inpainting, [Mairal, Sapiro & Elad ('08)]

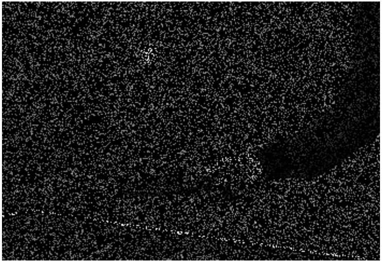
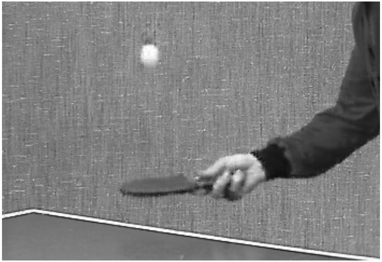
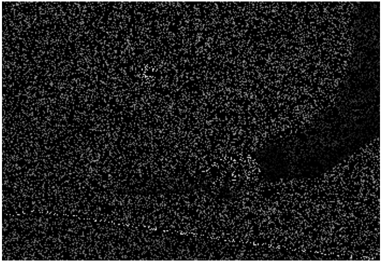



Figure: Inpainting results with two scales.

# Sparse representations for image restoration
## Inpainting, [Mairal, Sapiro & Elad ('08)]



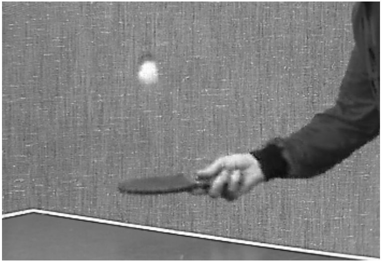


Figure: Inpainting results with two scales.

# Sparse representations for image restoration
## Inpainting, [Mairal, Sapiro & Elad ('08)]

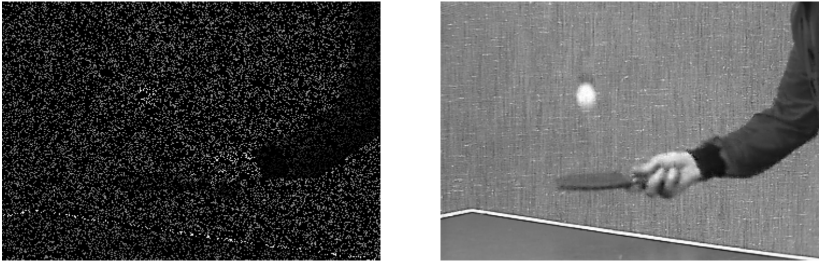
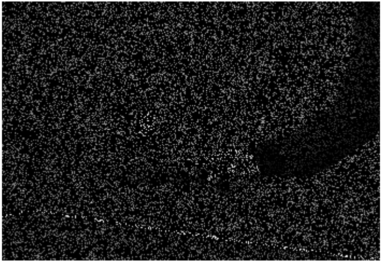


Figure: Inpainting results with two scales.

# Sparse representations for image restoration

## Inpainting, [Mairal, Sapiro & Elad ('08)]



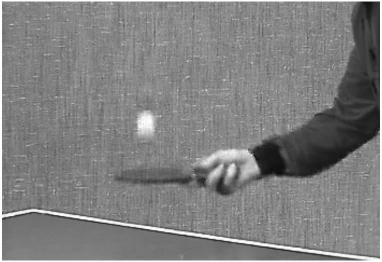


Figure: Inpainting results with two scales.

# Sparse representations for image restoration

## Color Video denoising, [Mairal, Sapiro & Elad ('08)]

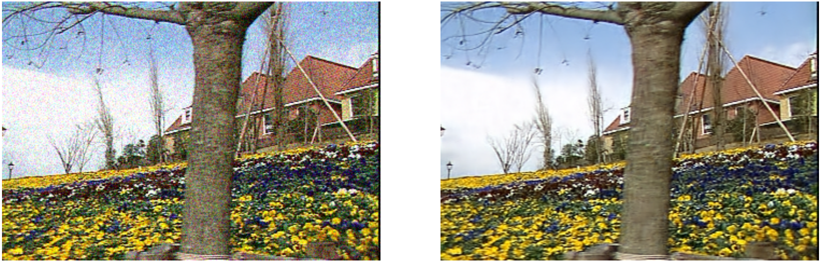


Figure: Denoising results with two scales. $\sigma = 25$

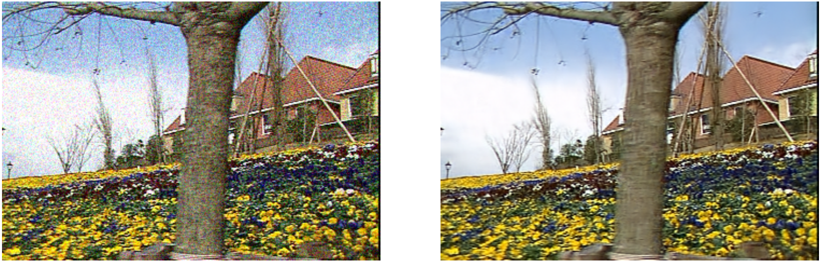


Figure: Denoising results with two scales. $\sigma = 25$

# Sparse representations for image restoration

## Color Video denoising, [Mairal, Sapiro & Elad ('08)]

Figure: Denoising results with two scales. $\sigma = 25$

# Sparse representations for image restoration

## Color Video denoising, [Mairal, Sapiro & Elad ('08)]

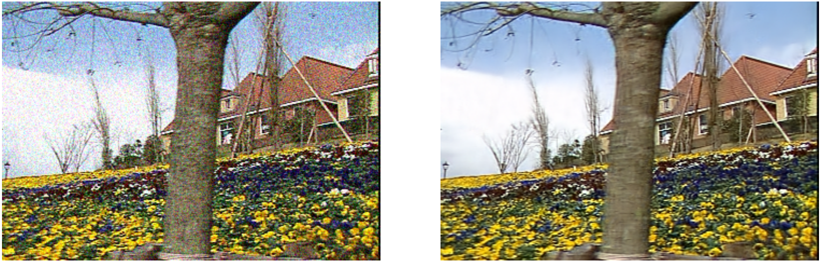


Figure: Denoising results with two scales. $\sigma = 25$

# Sparse representations for image restoration

## Color Video denoising, [Mairal, Sapiro & Elad ('08)]

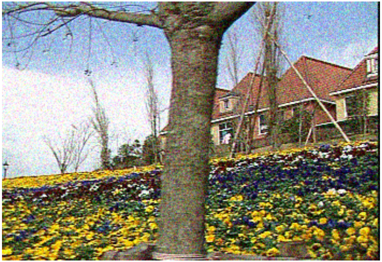


Figure: Denoising results with two scales. $\sigma = 25$

# Online Dictionary Learning
[Mairal, Bach, Ponce & Sapiro ('09)]

## Classical formulation for dictionary learning

$$\min_{\mathbf{D}\in\mathcal{C}} f_n(\mathbf{D}) = \min_{\mathbf{D}\in\mathcal{C}} \frac{1}{n}\sum_{i=1}^{n} l(\mathbf{x}_i, \mathbf{D}),$$

where

$$l(\mathbf{x}, \mathbf{D}) = \min_{\boldsymbol{\alpha}\in\mathbb{R}^k} \frac{1}{2}||\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i||_2^2 + \lambda||\boldsymbol{\alpha}||_1.$$

# Online Dictionary Learning

[Mairal, Bach, Ponce & Sapiro ('09)]

## Which formulation are we interested in?

$$\min_{\mathbf{D} \in \mathcal{C}} f(\mathbf{D}) = \mathbb{E}_{\mathbf{x}}[l(\mathbf{x}, \mathbf{D})] = \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=1}^{n} l(\mathbf{x}, \mathbf{D}).$$

# Online Dictionary Learning
[Mairal, Bach, Ponce & Sapiro ('09)]

## Online learning can

- handle potentially infinite datasets.
- be dramatically faster than batch algorithms.
- adapt to dynamic training sets.

# Online Dictionary Learning

[Mairal, Bach, Ponce & Sapiro ('09)]

Initialization of $\mathbf{D}_0$,

**for** $t = 1$ to $T$ **do**

  Draw $\mathbf{x}_t$.

  Sparse coding: compute using LARS

$$\boldsymbol{\alpha}_t \leftarrow \underset{\boldsymbol{\alpha} \in \mathbb{R}^k}{\arg\min} \frac{1}{2} ||\mathbf{x}_t - \mathbf{D}_{t-1}\boldsymbol{\alpha}||_2^2 + \lambda ||\boldsymbol{\alpha}||_1.$$

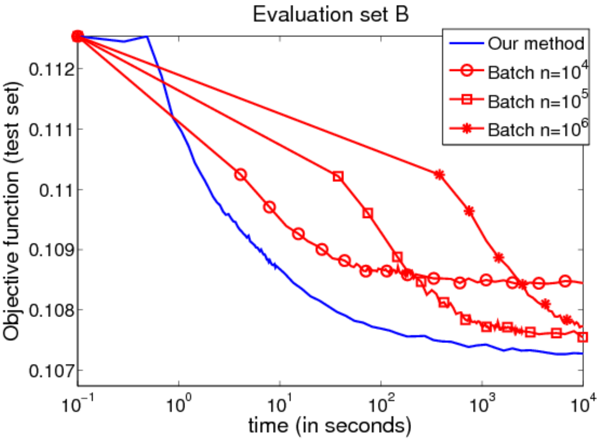
  Compute $\mathbf{D}_t$ using $\mathbf{D}_{t-1}$ as warm restart,

$$\mathbf{D}_t \leftarrow \underset{\mathbf{D} \in \mathcal{C}}{\arg\min} \frac{1}{t} \sum_{i=1}^{t} \frac{1}{2} ||\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i||_2^2 + \lambda ||\boldsymbol{\alpha}_i||_1.$$

**end for**

**return** $\mathbf{D}_T$ (learned dictionary).

# Online Dictionary Learning
## [Mairal, Bach, Ponce & Sapiro ('09)]




$m = 12 \times 12$ color patches, $k = 512$.

# Online Dictionary Learning

## Inpainting a 12Mpixels photograph

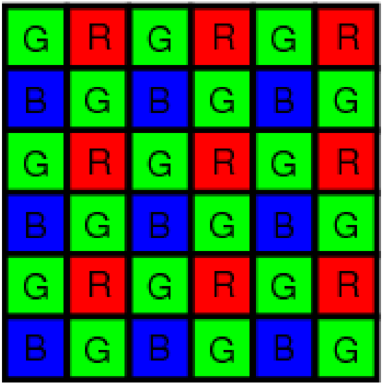
# Online Dictionary Learning
## [Mairal, Bach, Ponce & Sapiro ('09)]

### A few simple extensions

- sparse dictionaries.
- non-negative matrix factorization.
- sparse PCA.

# Raw image processing

What is a raw image?



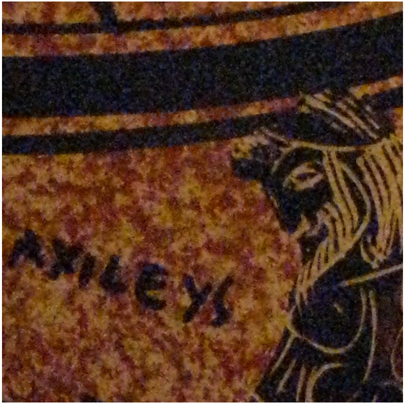

$+$ noise

# Raw image processing

## The raw image processing pipeline

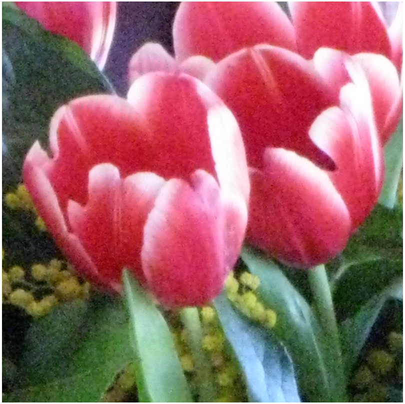
1. Denoising of the mosaick.
2. Demosaicking.
3. Color conversion to sRGB.

# Preliminary results
[Mairal, Bach, Ponce, Sapiro & Zisserman ('09)]



# Preliminary results
## [Mairal, Bach, Ponce, Sapiro & Zisserman ('09)]

# Preliminary results
## [Mairal, Bach, Ponce, Sapiro & Zisserman ('09)]



# Preliminary results
## [Mairal, Bach, Ponce, Sapiro & Zisserman ('09)]

# Conclusion

## Learned sparse represtentations

- can adapt to various type of data.
- lead to state-of-the-art results for several tasks.
- are computationally cheap thanks to online learning.