Sparse Coding and Dictionary Learning for Image Analysis

Part III: Learning for the task

Francis Bach, Julien Mairal, Jean Ponce and Guillermo Sapiro

ICCV'09 tutorial, Kyoto, 28th September 2009

What this part is about

- Learning dictionaries with a discriminative cost function...
- ...and a few applications to computer vision applications.
- Compressed sensing with learned dictionaries and why you should not use random sensing matrices.

Learning dictionaries with a discriminative cost function

Idea:

Let us consider 2 sets S_-, S_+ of signals representing 2 different classes. Each set should admit a specific dictionary best adapted to its reconstruction.

Classification procedure for a signal $\mathbf{x} \in \mathbb{R}^n$:

$$\min(\mathbf{R}^\star(\mathbf{x}, \mathbf{D}_-), \mathbf{R}^\star(\mathbf{x}, \mathbf{D}_+))$$

where

$$\mathbf{R}^\star(\mathbf{x}, \mathbf{D}) = \min_{\boldsymbol{\alpha} \in \mathbb{R}^p} ||\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}||_2^2 \text{ s.t. } ||\boldsymbol{\alpha}||_0 \leq L.$$

"Reconstructive" training

$$\begin{cases} \min_{\mathbf{D}_-} \sum_{i \in S_-} \mathbf{R}^\star(\mathbf{x}_i, \mathbf{D}_-) \\ \min_{\mathbf{D}_+} \sum_{i \in S_+} \mathbf{R}^\star(\mathbf{x}_i, \mathbf{D}_+) \end{cases}$$

[Grosse et al., 2007], [Huang and Aviyente, 2006] (see also [Wright et al., 2009])

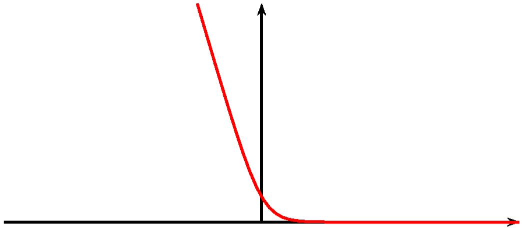
Learning dictionaries with a discriminative cost function

"Discriminative" training

[Mairal, Bach, Ponce, Sapiro, and Zisserman, 2008a]

$$\min_{\mathbf{D}_-,\mathbf{D}_+} \sum_i \mathcal{C}\Big(\lambda z_i \big(\mathbf{R}^\star(\mathbf{x}_i, \mathbf{D}_-) - \mathbf{R}^\star(\mathbf{x}_i, \mathbf{D}_+)\big)\Big),$$

where $z_i \in \{-1, +1\}$ is the label of $\mathbf{x}_i$.



Logistic regression function

Learning dictionaries with a discriminative cost function

Mixed approach

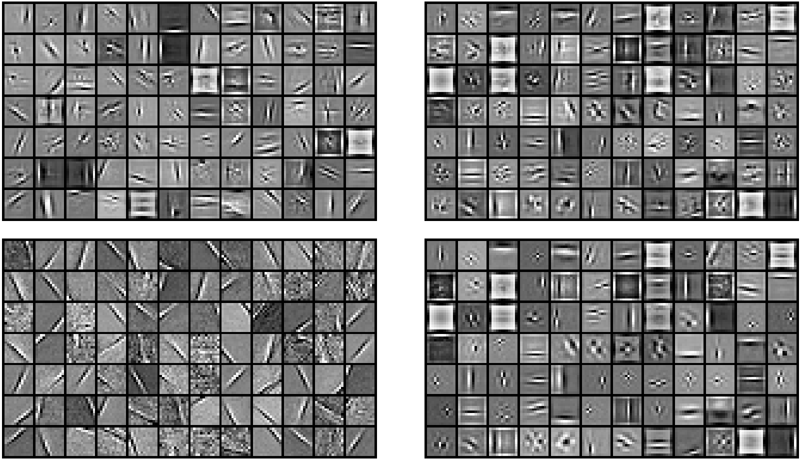
$$\min_{\mathbf{D}_-,\mathbf{D}_+} \sum_i \mathcal{C}\Big(\lambda z_i\big(\mathbf{R}^\star(\mathbf{x}_i,\mathbf{D}_-) - \mathbf{R}^\star(\mathbf{x}_i,\mathbf{D}_+)\big)\Big) + \mu \mathbf{R}^\star(\mathbf{x}_i,\mathbf{D}_{z_i}),$$

where $z_i \in \{-1,+1\}$ is the label of $\mathbf{x}_i$.

Keys of the optimization framework

- Alternation of sparse coding and dictionary updates (not online yet).
- Continuation path with decreasing values of μ.
- OMP to address the NP-hard sparse coding problem...
- ...or LARS when using ℓ_1.
- Use softmax instead of logistic regression for $N > 2$ classes.

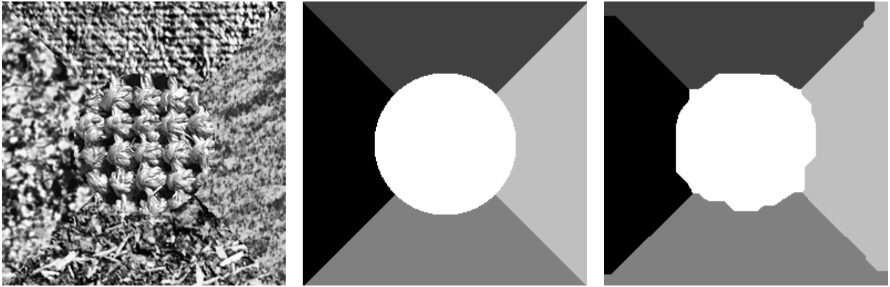
Learning dictionaries with a discriminative cost function
Examples of dictionaries



Top: reconstructive, Bottom: discriminative, Left: Background, Right: Bicycle

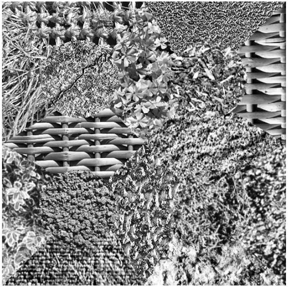
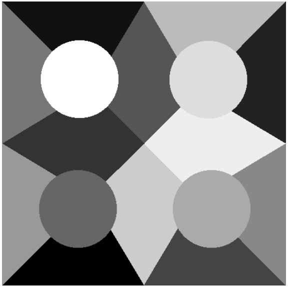
Learning dictionaries with a discriminative cost function

Texture segmentation

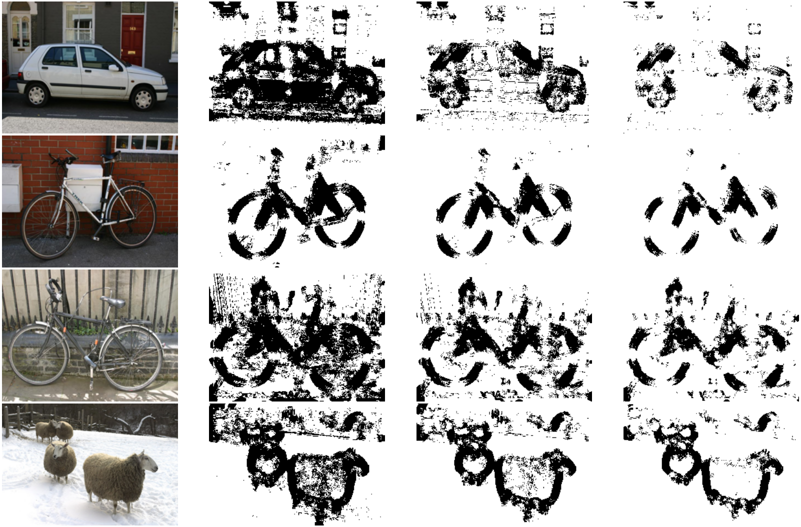


Learning dictionaries with a discriminative cost function

Texture segmentation

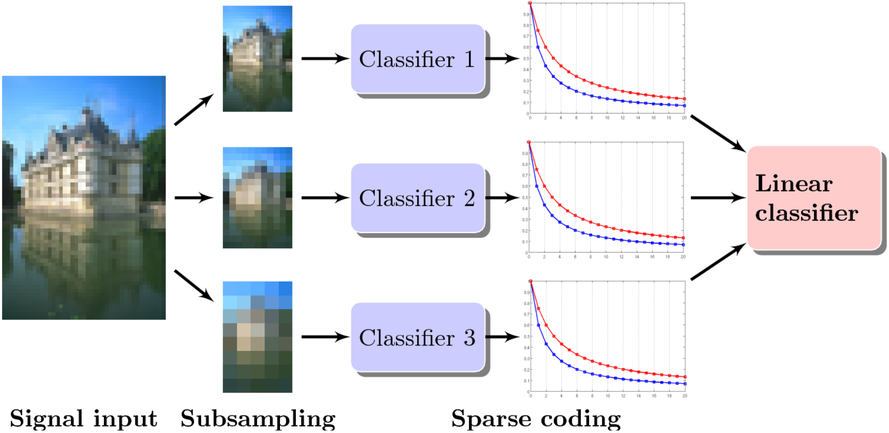




Learning dictionaries with a discriminative cost function
Pixelwise classification



Learning dictionaries with a discriminative cost function

Multiscale scheme

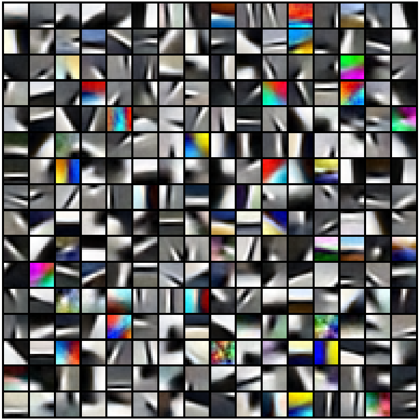


Learning dictionaries with a discriminative cost function

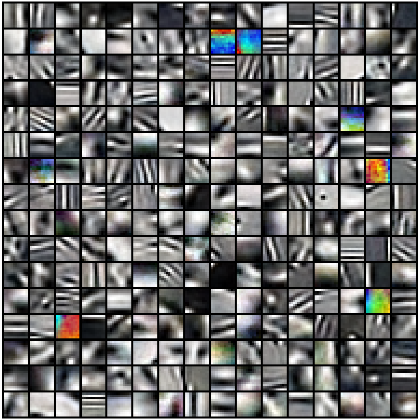
weakly-supervised pixel classification

Application to edge detection and classification

[Mairal, Leordeanu, Bach, Hebert, and Ponce, 2008b]

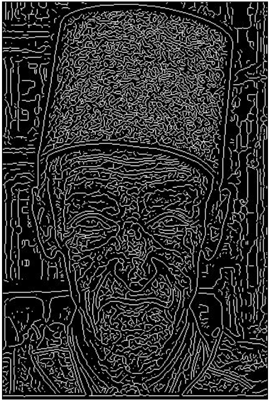


Good edges



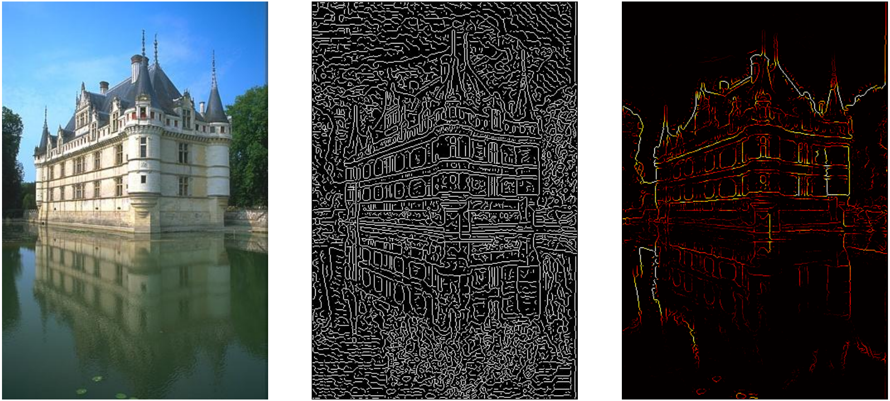
Bad edges

Application to edge detection and classification
Berkeley segmentation benchmark


Raw edge detection on the right

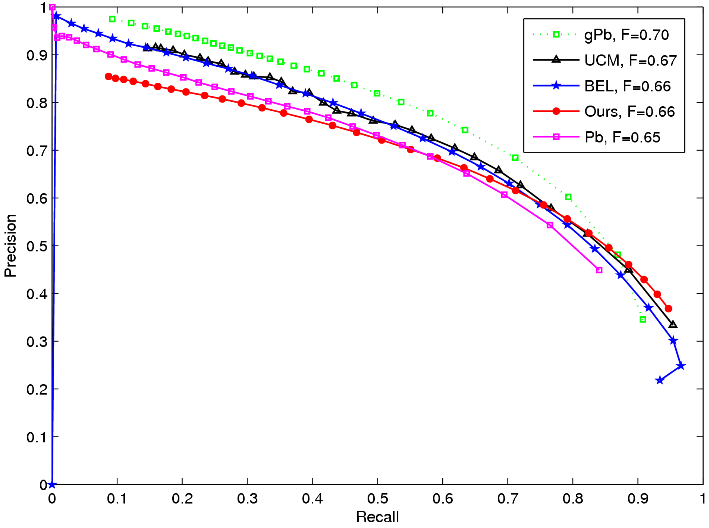
Application to edge detection and classification
Berkeley segmentation benchmark



Raw edge detection on the right

Application to edge detection and classification

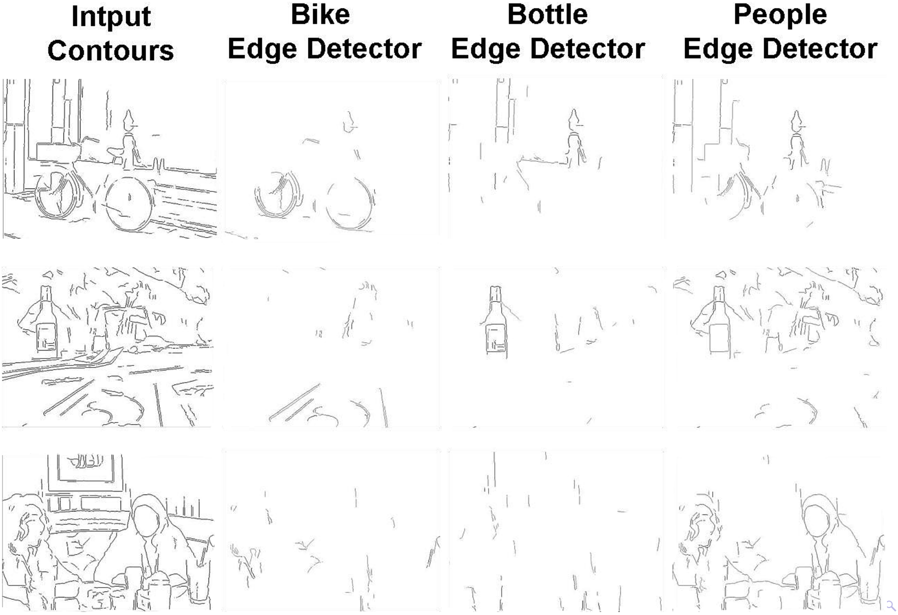
Berkeley segmentation benchmark



Application to edge detection and classification
Contour-based classifier: [Leordeanu, Hebert, and Sukthankar, 2007]

Is there a bike, a motorbike, a car or a person on this image?

Application to edge detection and classification



Application to edge detection and classification
Performance gain due to the prefiltering

Ours + [Leordeanu '07]	[Leordeanu '07]	[Winn '05]
96.8%	89.4%	76.9%

Recognition rates for the same experiment as [Winn et al., 2005] on VOC 2005.

Category	Ours+[Leordeanu '07]	[Leordeanu '07]
Aeroplane	71.9%	61.9%
Boat	67.1%	56.4%
Cat	82.6%	53.4%
Cow	68.7%	59.2%
Horse	76.0%	67%
Motorbike	80.6%	73.6%
Sheep	72.9%	58.4%
Tvmonitor	87.7%	83.8%
Average	75.9%	64.2 %

Recognition performance at equal error rate for 8 classes on a subset of images from Pascal 07.

A partial conclusion on discriminative learned dictionaries

- The learning of sparse representations should be discriminative for recognition tasks.
- Discriminative sparse representations are well adapted to edge analysis.
- Local prefiltering of edges dramatically helps contours-based classifiers.
- promising, but still a lot of work to do...

References I

R. Grosse, R. Raina, H. Kwong, and A. Y. Ng. Shift-invariant sparse coding for audio classification. In *Proceedings of the Twenty-third Conference on Uncertainty in Artificial Intelligence*, 2007.

K. Huang and S. Aviyente. Sparse representation for signal classification. In *Advances in Neural Information Processing Systems*, Vancouver, Canada, December 2006.

M. Leordeanu, M. Hebert, and R. Sukthankar. Beyond local appearance: Category recognition from pairwise interactions of simple features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.

J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Discriminative learned dictionaries for local image analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008a.

J. Mairal, M. Leordeanu, F. Bach, M. Hebert, and J. Ponce. Discriminative sparse image models for class-specific edge detection and image interpretation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2008b.

J. Winn, A. Criminisi, and T. Minka. Object categorization by learned universal visual dictionary. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2005.

References II

J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 210–227, 2009.