Sparse Representation for Computer Vision and Pattern Recognition

Yifei Han, Jixuan Fan, Yiqin Wang, Guanxing Lu

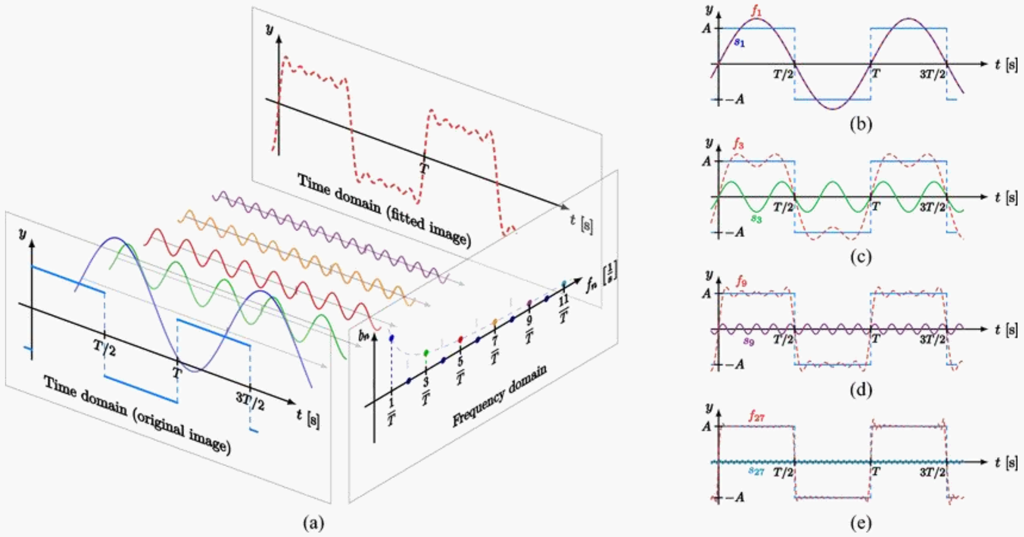
8/4/2024

Content

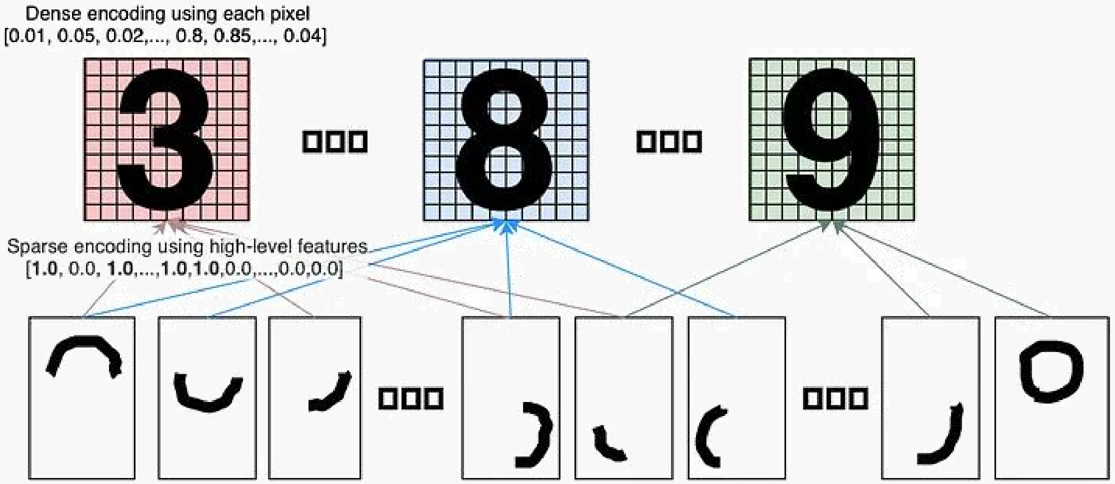
- Basic Concepts
- Task Definition
- Motivation
- Method
- Experiments
- Conclusions

Basic Concepts

- Sparse representation



Signal processing



Computer vision

Express data efficiently!

Task Definition

☐ Datasets

Extended Yale B Face (Yale-B)

2,414 frontal-face images

Columbia University Image Library (COIL-20)

72 views of 20 objects each

MNIST

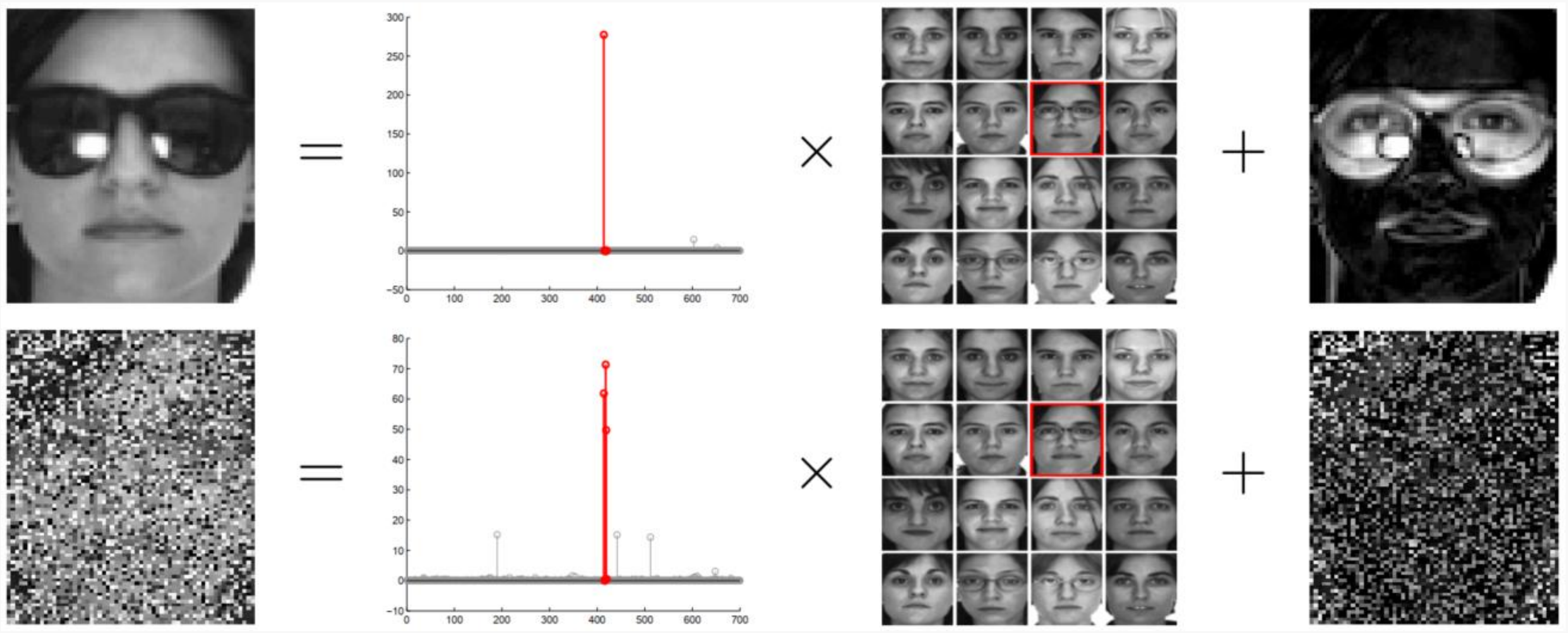
60,000 handwritten numbers

Face Recognition

◻ Face Recognition as Sparse Representation

Potentially occluded

Corrupted



Test image

Training images

Sparse errors

Representing the test signal as a sparse linear combination of the training signals

Face Recognition

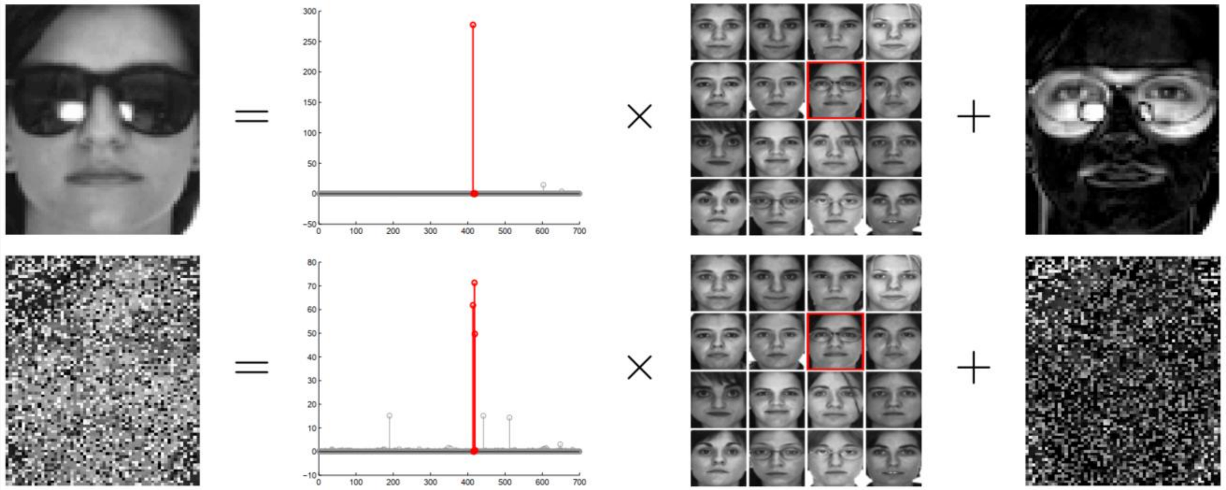
Face Recognition as Sparse Representation

Representing the test signal as a sparse linear combination of the training signals

$$\boldsymbol{D} \doteq [\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_c] = \left[\boldsymbol{d}_{1,1}, \boldsymbol{d}_{1,2}, \ldots, \boldsymbol{d}_{k,N_k}\right]. \quad (1)$$

$$\boldsymbol{x} = \boldsymbol{D}\alpha_0 \in \mathbb{R}^m \quad (2)$$

$$\alpha_0 = [0, \cdots, 0, \alpha_i^T, 0, \ldots, 0]^T \in \mathbb{R}^N$$



Consider occlude and corrupt：

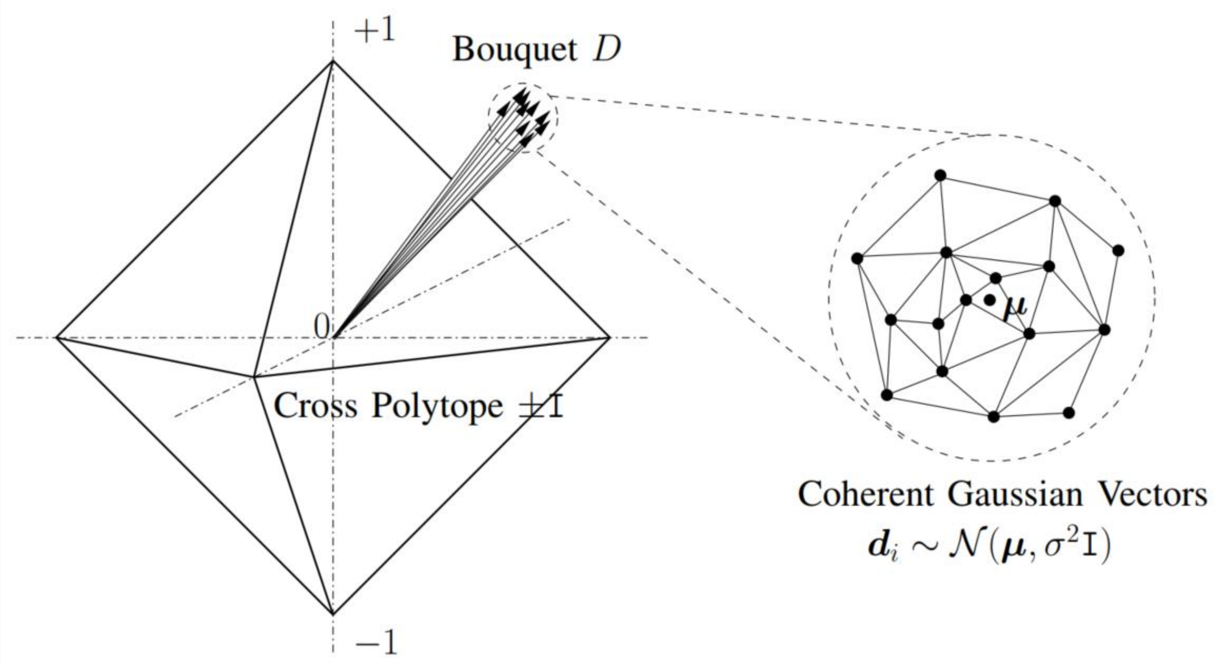
$$\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{e}_0 = \boldsymbol{D}\alpha_0 + \boldsymbol{e}_0 \quad (3)$$

$$(\alpha_0, \boldsymbol{e}_0) = \arg\min \|\alpha\|_0 + \|\boldsymbol{e}\|_0 \quad \text{subject to} \quad \boldsymbol{x} = \boldsymbol{D}\alpha + \boldsymbol{e}. \quad (4)$$

$$\min \|\alpha\|_1 + \|\boldsymbol{e}\|_1 \quad \text{subject to} \quad \boldsymbol{x} = \boldsymbol{D}\alpha + \boldsymbol{e} \quad (5)$$

Face Recognition

The “cross-and-bouquet” model



The cross–and–bouquet polytope is spanned by vertices from both the bouquet and the cross

Face Recognition

- The "cross-and-bouquet" model

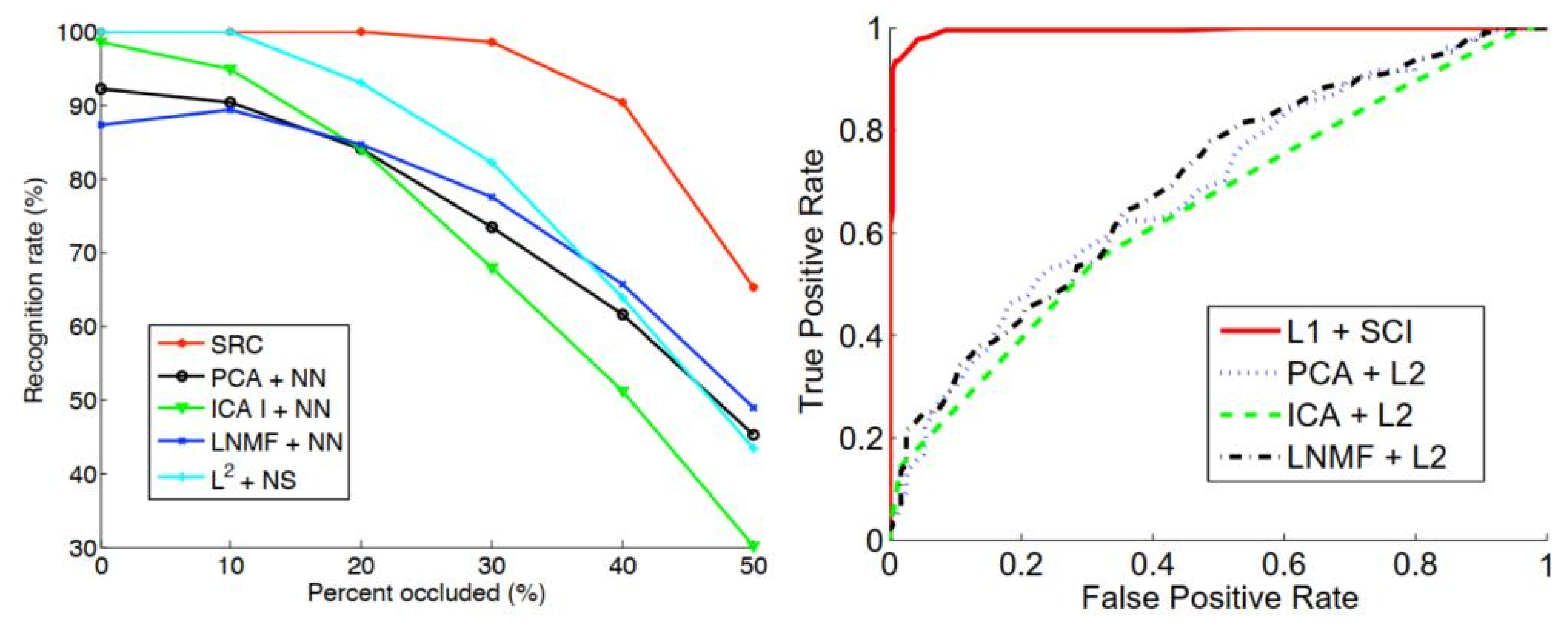
$$\begin{gathered} \boldsymbol{D} = [\boldsymbol{d}_1 \ldots \boldsymbol{d}_N] \in \mathbb{R}^{m \times N}, \quad \boldsymbol{d}_i \sim_{iid} \mathcal{N}\left(\boldsymbol{\mu}, \frac{\nu^2}{m} \mathtt{I}_m\right), \\ \|\boldsymbol{\mu}\|_2 = 1, \qquad \|\boldsymbol{\mu}\|_\infty \leq C_\mu m^{-1/2}. \end{gathered} \tag{6}$$

$$\begin{aligned} (\boldsymbol{\alpha}_0, \boldsymbol{e}_0) \quad = \quad & \arg\min \; \|\boldsymbol{\alpha}\|_1 + \|\boldsymbol{e}\|_1 \\ & \quad \text{subject to } \; \boldsymbol{D}\boldsymbol{\alpha} + \boldsymbol{e} = \boldsymbol{D}\boldsymbol{\alpha}_0 + \boldsymbol{e}_0, \end{aligned} \tag{8}$$

清華大学 i-VisionGroup

Face Recognition

- Dense Error Correction by L1 -Minimization



Data: Extended Yale B Face

Receiver Operating Characteristic (ROC) for validation with 30% occlusion

Conclusion: The sparse representation significantly outperforms the competitors

Experiments

Sparse Modeling for Image Reconstruction

Classification error rate (%)

Cluster #	ℓ^1-graph	G-g	LE-g	LLE-g	PCA+Km
USPS : 7	**0.962**	0.381	0.724	0.565	0.505
FOR. : 7	**0.763**	0.621	0.619	0.603	0.602
ETH. : 7	**0.605**	0.371	0.522	0.478	0.428

Classification error rate (%)

Gallery #	PCA	NPE	LPP	ℓ^1-graph-SL	Fisherfaces [7]
USPS : 10	37.21	33.21	30.54	21.91	**15.82**
FOR. : 10	27.29	25.56	27.32	**19.76**	21.17
ETH. : 10	47.45	45.42	44.74	38.48	**13.39**

Classification error rate (%)

Labeled #	ℓ^1-g	LLE-g	LE-g	MFA	PCA
USPS : 10	**25.11**	34.63	30.74	34.63	37.21
FOR. : 10	**17.45**	24.93	22.74	24.93	27.29
ETH. : 10	**30.79**	38.83	34.54	38.83	47.45

[1] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Trans. IP*, 17(1):53–69, January 2008.

Experiments

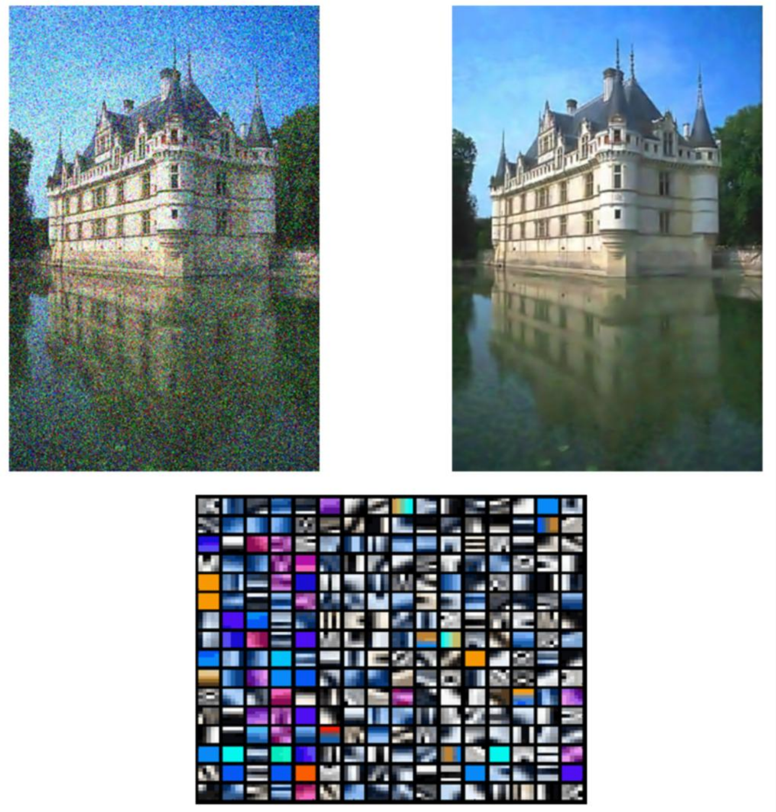
Sparse Modeling for Image Reconstruction

Dictionary

Sparse representatio
n

$$(\boldsymbol{A}^*, \boldsymbol{D}^*) = \arg\min_{\boldsymbol{A},\boldsymbol{D}} \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\|_F^2 + \lambda \|\boldsymbol{A}\|_p$$

L0 or L1
regularization



Dictionary learned from a standard set of color images
[1]

[1] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Trans. IP*, 17(1):53–69, January 2008.

Experiments

Sparse Modeling for Image Classification

Bikes

Backgroun
d

Dictionary learned from a standard set of color images
[1]

[1] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Learning discriminative dictionaries for local image analysis. In *Proc. IEEE CVPR*, 2008

Experiments

Learning to Sense

Dictionary

$$
\begin{aligned}
(\boldsymbol{A}^*, \boldsymbol{D}^*, \Phi^*) \; = \; & \underset{\boldsymbol{A}, \boldsymbol{D}, \Phi}{\arg\min} \; \|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{A}\|_F^2 + \lambda_1 \|Y - \Phi \boldsymbol{D}\boldsymbol{A}\|_F^2 \\
+ \; & \lambda_2 \left\| (\Phi \boldsymbol{D})^T (\Phi \boldsymbol{D}) - I \right\|_F^2 + \lambda_3 \|\boldsymbol{A}\|_p \, .
\end{aligned}
$$

Sensing matrices

Independently < Simultaneously learning the dictionary and sensing matrices

[1] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Learning discriminative dictionaries for local image analysis. In *Proc. IEEE CVPR*, 2008

Conclusions

- Conclusions