第5节 稀疏表示与字典学习

（Sparse Representation &Dictionary Learning）

内 容

- 稀疏性表示与分类
- 压缩感知
- 字典学习及应用

关键词对照

- 稀疏性表示：Sparse representation
- 压缩感知　：Compressed sensing
- 字典学习　：Dictionary learning

稀疏性表示与分类

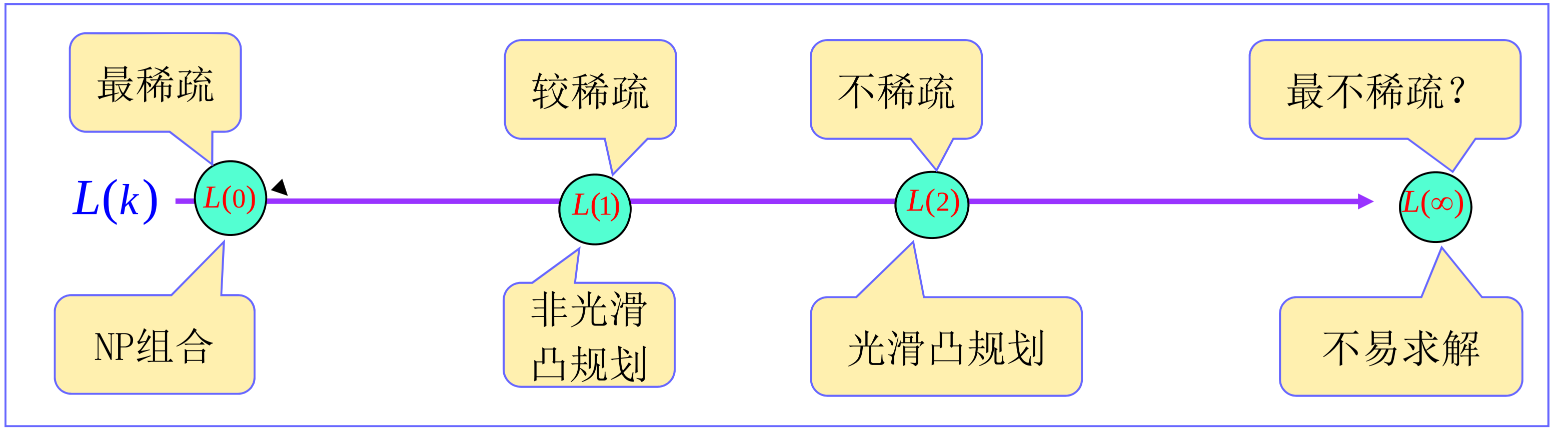
稀疏性问题的数学表述：L_0问题，呈现三种相互等价的形式：

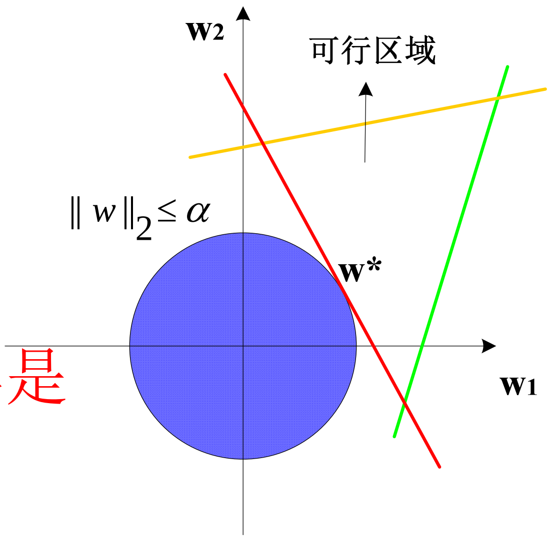
L_0形式： $$\min_{w}\left\{\frac{1}{2}\|Aw-y\|_2^2+\lambda\|w\|_0\right\}$$

- L_0问题为NP组合难问题，对较大规模数据无法直接求解；为使求解稀疏性问题变得可行，当前一般性的作法是将L_0问题放松到L_1问题,从而将一个组合优化问题放松到一个凸优化问题来求解。

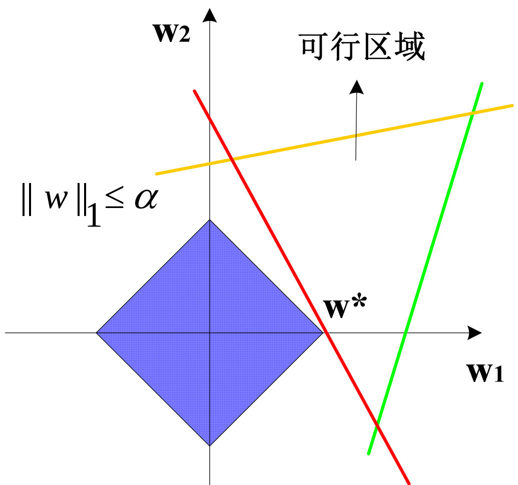
L_1形式： $$\min_{w}\left\{\frac{1}{2}\|Aw-y\|_2^2+\lambda\|w\|_1\right\}$$

稀疏性表示与分类





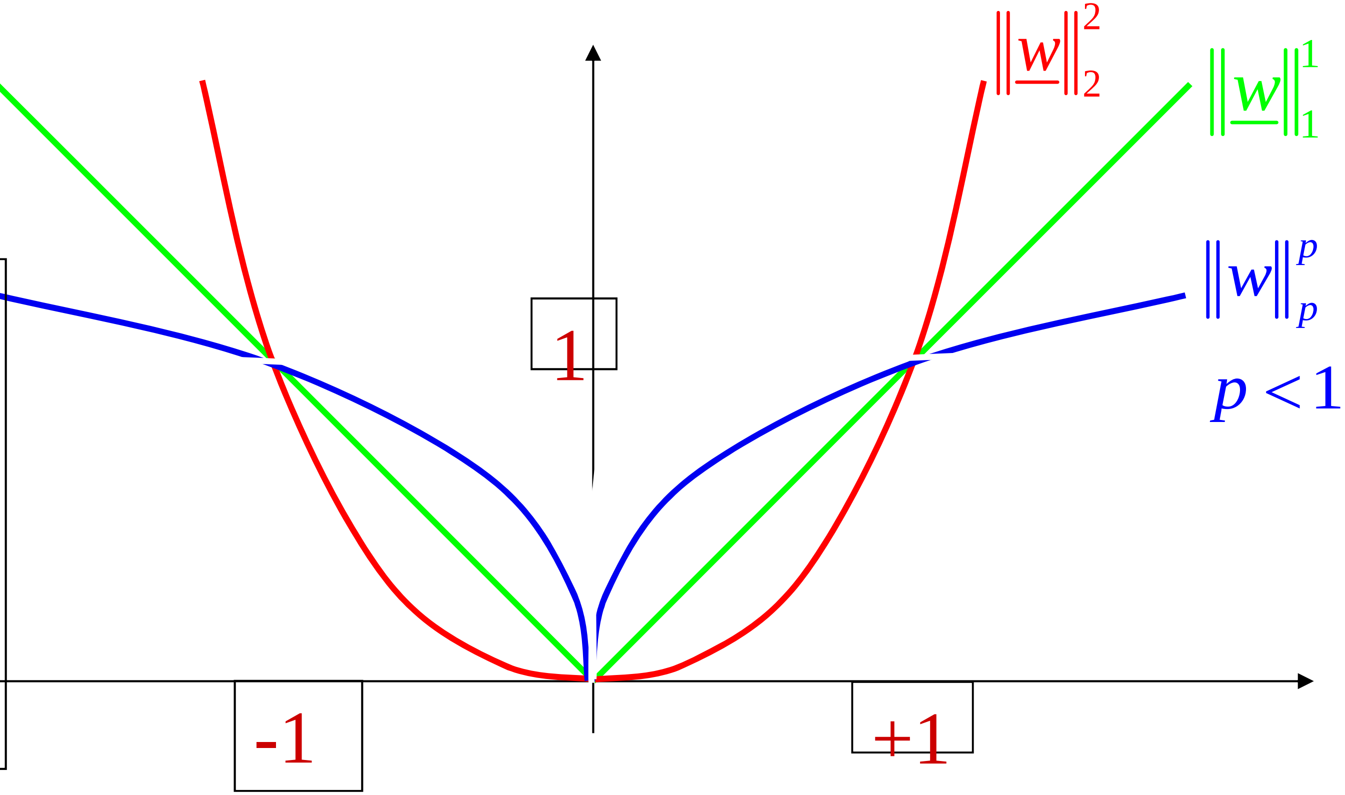
2-范数投影是非稀疏解



1-范数投影是稀疏解

稀疏性表示与分类

在选择**1**范数最小化的情况下，所求得的向量有相当一部分元素为零



Ron Rubinstein：An Introduction to Sparse Representation And the K-SVD Algorithm

稀疏表示的求解

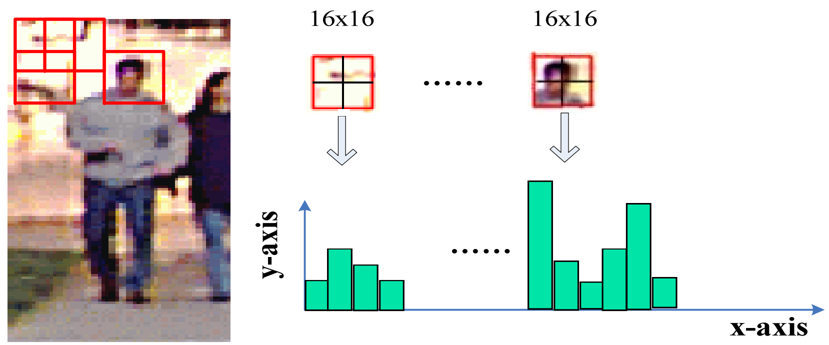
- “Greedy”贪心 方法. 每个循环中选择一个字典元素
  - Step 1: 找到一个能够最好的表示输入信号的元素
  - Step 2 :找到剩余元素中一个能够最好的表示输入信号的元素
  - 表示误差达到一个足够小的阈值或者所选元素足够多的时候停止
- 改进方法: Orthogonal Matching Pursuit (OMP), Optimized OMP (OOMP)

基于稀疏表示的目标检测

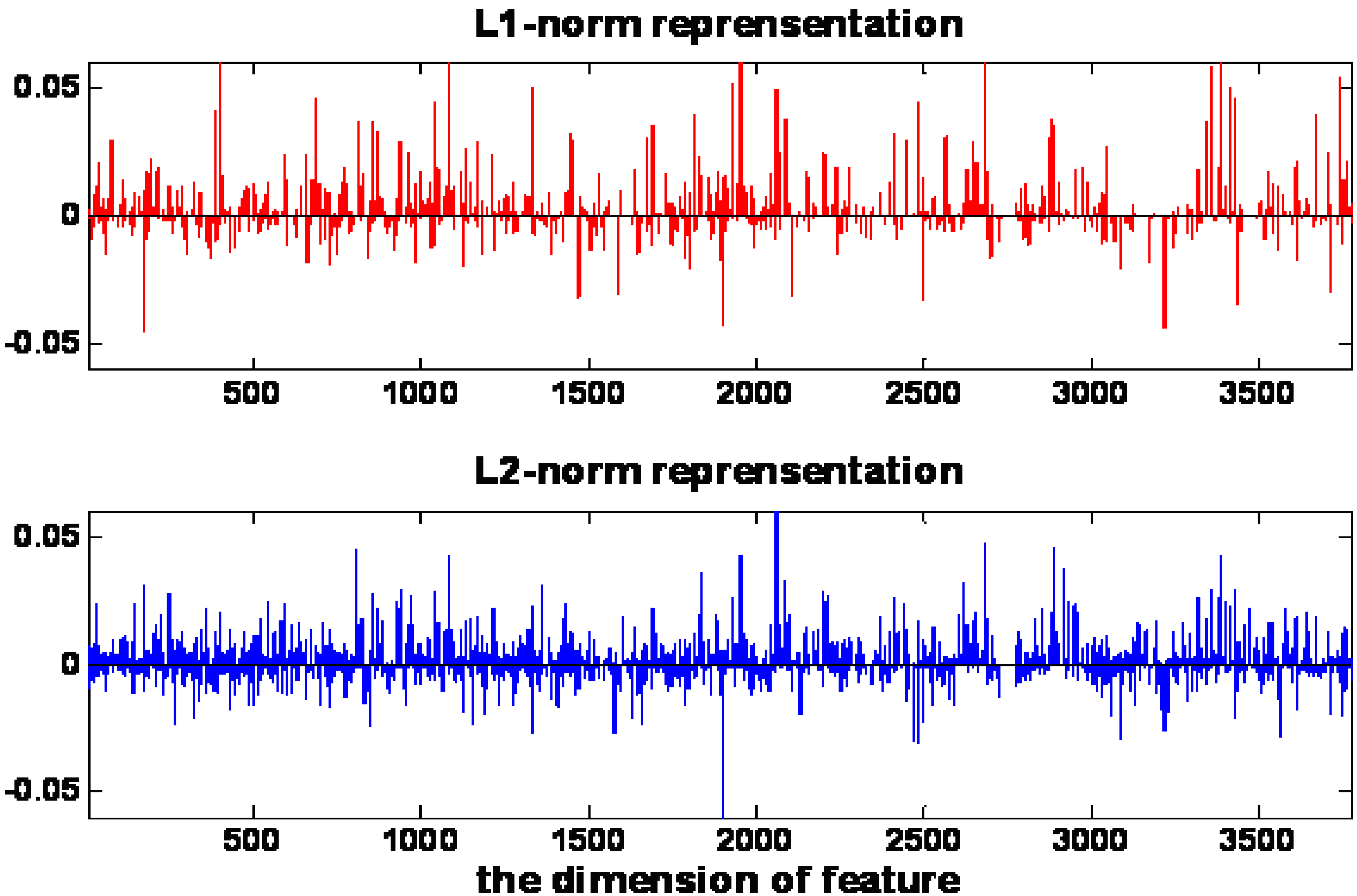
✓基于L1-norm最小化学习的线性分类器(L-LML)

- 针对整张的64x128大小的人体训练图片，提出一种线性分类算法，意在用于特征选择和目标分类。
- 首先提出L1-norm Minimization Learning (LML) 构建线性分类器的法向量。每个训练样本是一个3780维的HOG特征向量。

HOG特征



基于稀疏表示的目标检测

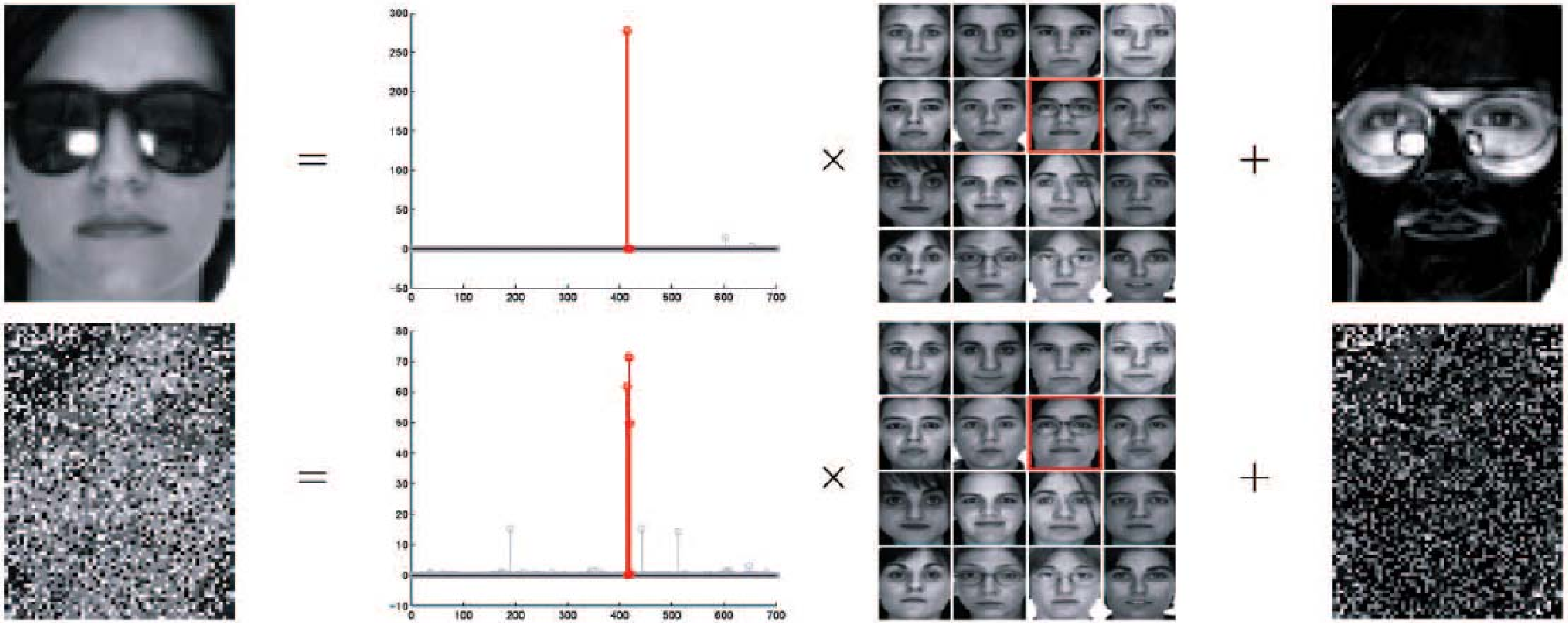


的可视化

的可视化

基于稀疏重构的分类与识别

$\min_{i} \; r_i(y) \doteq \|y - A\,\delta$



$$\min_{w}\left\{\frac{1}{2}\|Aw-y\|_2^2+\lambda\|w\|_1\right\} \qquad \min_{i}\; r_i(y)=\min_{i}\frac{1}{2}\|A\delta_i(w)-y\|_2^2$$

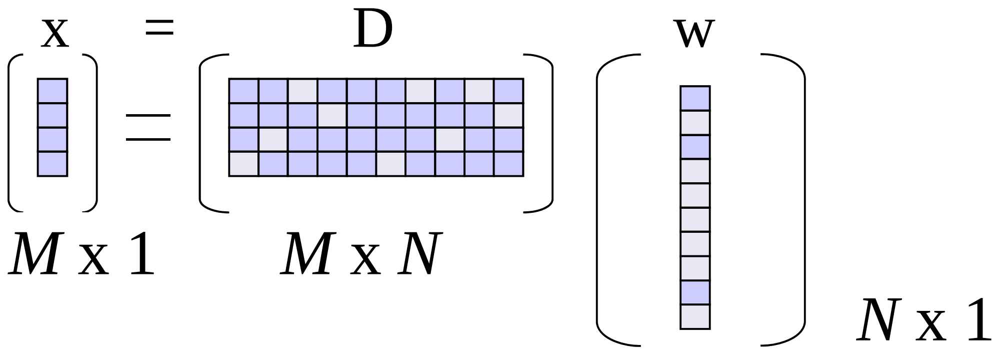
John Wright, Yi Ma, etc., "Robust Face Recognition via Sparse Representation," IEEE Transcations on Pattern Analysis and Machine Intelligence, vol.31, no.2, 2009.

内 容

- 稀疏性表示与分类
- 压缩感知
- 字典学习及应用

压缩感知（Compressed Sensing）

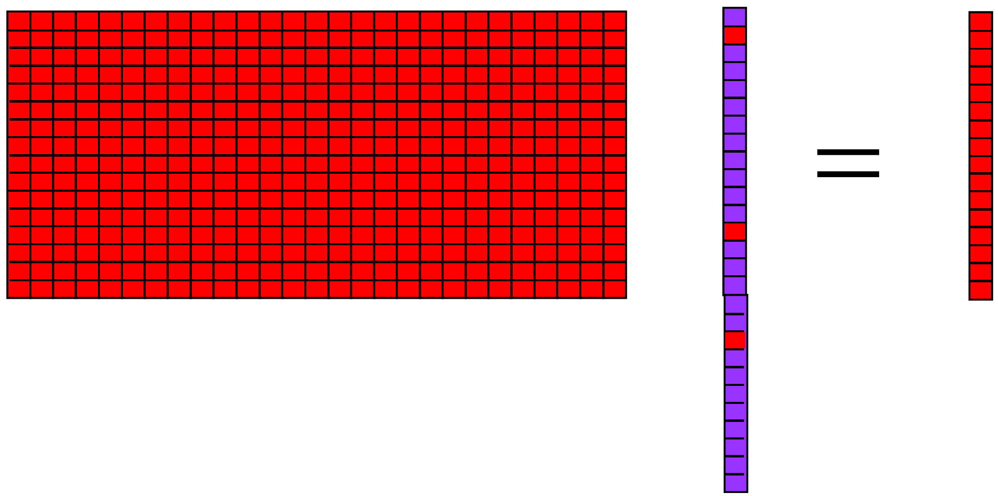
欠定线性方程:



- x：稀疏或者可压缩信号
- D：测量（采样）矩阵，确定或者随机
- w：测量向量

压缩感知（Compressed Sensing）

$$\mathbf{D}\,w = \underline{x}$$



$$\underset{\underline{w}}{Min}\left\|\underline{w}\right\|_0 \quad s.t. \quad \underline{x} = \mathbf{D}\underline{w}$$

内 容

- 稀疏性表示与分类
- 压缩感知
- 字典学习及应用

背景知识

- 线性问题的一般形式：　$x = D \cdot w$　　(1)

  其中　$x \in R^{M}, w \in R^{N}, D \in R^{M \times N}$

- 情况I：如果M >N, 方程(1)是过定的（Over-determined），没有可行解；
- 情况II：如果M <N, 方程(1)是欠定的（Under-determined），有无数可行解；

背景知识

- 正则化（Regularization）；

$$\min \frac{1}{2} \| x - D \cdot w \|_2^2 + \lambda \| w \|_2^2$$

- 可以求可行解：

$$w^* = \left(D^T D + \lambda \cdot I \right)^{-1} \cdot D^T \cdot x$$

因为 $D^T D + \lambda \cdot I$ 是非奇异的，即使 $D^T D$ 是奇异的

背景知识

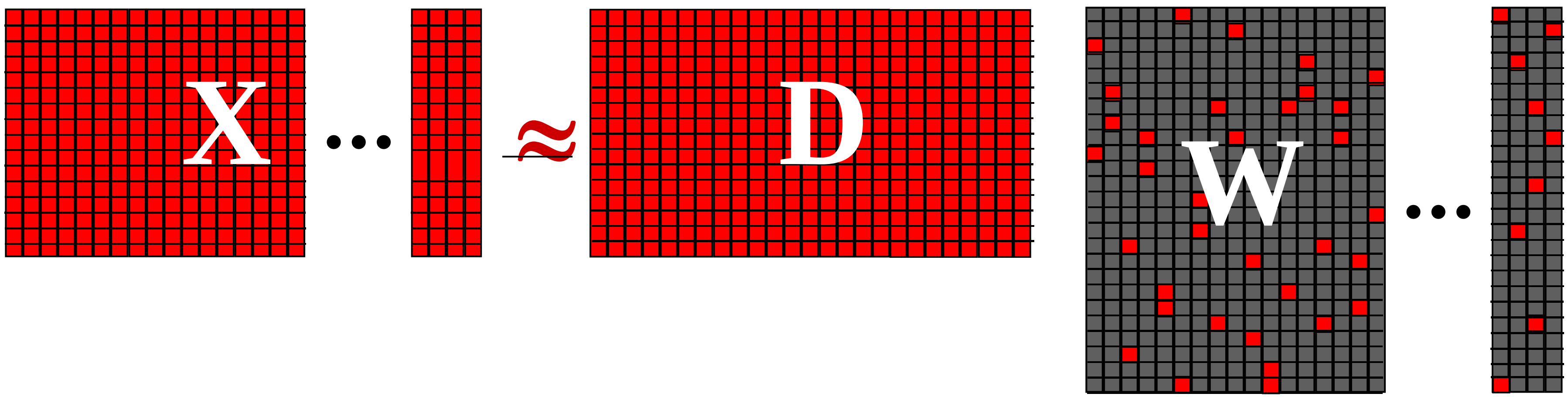
- 正则化（Regularization）；

$$\min \frac{1}{2} \| x - D \cdot w \|_2^2 + \lambda \| w \|_0$$

$$\min \frac{1}{2} \| x - D \cdot w \|_2^2 + \lambda \| w \|_1$$

$$\min \frac{1}{2} \| x - D \cdot w \|_2^2 + \lambda \| w \|_2^2$$

字典学习（Dictionary learning）



$$\operatorname*{Min}_{\mathbf{D},\mathbf{A}} \quad \left\| \mathbf{D}W - \mathbf{X} \right\|_2^2 \quad s.t. \quad \forall j, \left\| \underline{w}_j \right\|_1 \le L$$

字典学习（Dictionary learning）

$$\underline{x} = \mathbf{D}w$$

Dictionary Learning
字典的学习（构造字典）

Sparse representation
稀疏“编码”问题

背景知识

- 线性问题的一般形式： $x = D \cdot w$

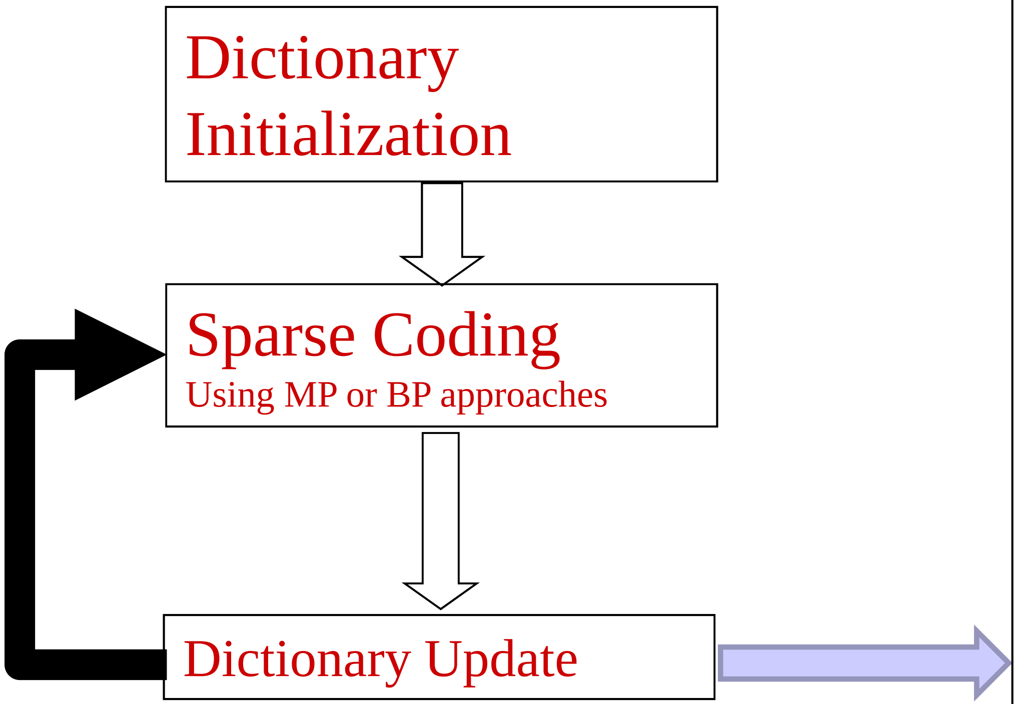
  其中 $x \in R^{M}, w \in R^{N}, D \in R^{M \times N}$

  w 测量向量是未知的

- 数据回归问题： $x = D \cdot w$, w 是未知的;
- 稀疏表示问题： $x = D \cdot w$, w 是未知的，也是稀疏的;
- 字典学习问题: $x = D \cdot w$, w & D 是未知的，w是稀疏的

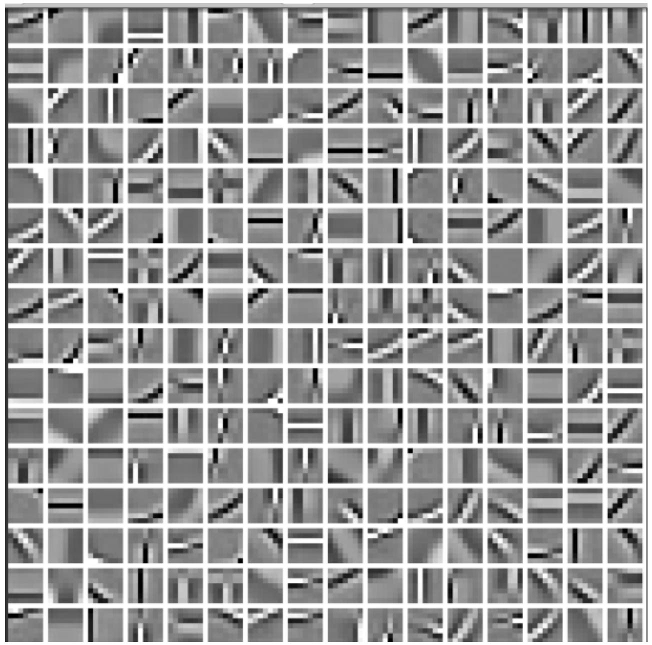
.

字典学习（Dictionary learning）



- Hard Competitive
  - Only the selected dictionary atoms are updated
    - **KSVD** [Aharon, Elad & Bruckstein ('04)]
- Soft Competitive
  - All dictionary atoms are updated based on a ranking
    - Sparse Coding Neural Gas (SCNG) [Labusch, Barth & Martinetz ('09)]

字典学习（重构）



Qixiang Ye,   UCAS

David Doermann, UMD

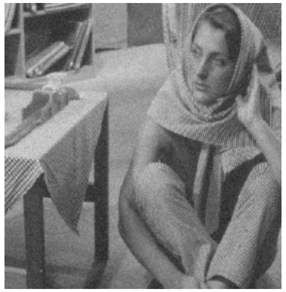
Q. Ye, D. Doermann, "Robust Scene Text Detection Using Integrated Feature Discrimination," IEEE Int'l Conf. Image Processing, 2014.

字典学习（重构）

Q. Ye, D. Doermann, "Robust Scene Text Detection Using Integrated Feature Discrimination," IEEE Int'l Conf. Image Processing, 2014.

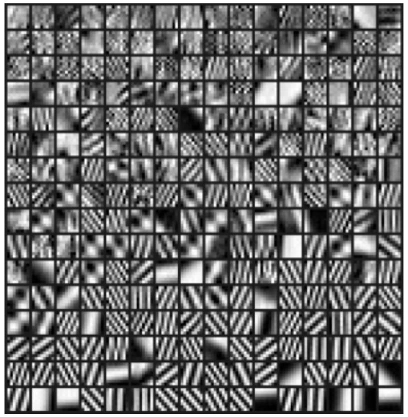
字典学习（Dictionary learning）

Source



Noisy image

Result 30.829dB



Dictionary

Source

Result

Dictionary

[M. Elad, Springer 2010]

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