

基于自适应 K -SVD 字典的视频帧稀疏重建算法

钱 阳, 李 雷, 袁安安

(南京邮电大学 视觉认知计算与应用研究中心, 江苏 南京 210023)

摘 要: 压缩感知理论的一个重要前提是找到信号的稀疏域, 其直接影响着算法的重构精度, 研究快速高效的信号稀疏表示方法具有重大的现实意义。为了提高字典训练速度与性能, 基于传统的 K -SVD 算法, 提出了一种自适应 K -SVD 字典学习算法 (Adaptive K -SVD)。该算法交替执行稀疏编码阶段和字典更新阶段。在稀疏编码阶段, 通过引入自适应稀疏约束机制, 以获得更稀疏的表示系数, 从而进一步提高字典的更新效率; 而在字典更新阶段, 则使用经典 K -SVD 的字典更新方式来实现字典原子的逐列更新。将所提算法应用于压缩感知理论的视频帧稀疏表示中, 实现视频帧的稀疏重建。仿真对比实验结果表明, 所提算法比经典的 K -SVD 算法的字典训练速度更快, 稀疏表示性能更优, 且能有效减少压缩感知的重构误差。

关键词: K -SVD 算法; 自适应 K -SVD 算法; 字典学习; 稀疏表示; 压缩感知

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2017)06-0036-05

doi: 10.3969/j.issn.1673-629X.2017.06.008

An Adaptive K -SVD Dictionary Learning Algorithm for Video Frame Sparse Reconstruction

QIAN Yang, LI Lei, YUAN An-an

(Center for Visual Cognitive Computation and Its Application, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Finding sparsifying transforms is an important prerequisite of compressed sensing, which directly affects the reconstruction accuracy. It has practical significance to research the fast and efficient signal sparse representation method. Based on the traditional K -SVD algorithm, an adaptive K -SVD dictionary learning algorithm has been proposed to improve the speed and performance of dictionary training which is an iterative one that alternates between sparse coding and dictionary update steps. In the sparse coding stage, an adaptive sparsity constraint has been utilized to obtain sparser representation coefficient, which has further improved the efficiency of the dictionary update stage. And in the dictionary update stage, the dictionary atoms are updated column by column using the classic K -SVD dictionary update method. With the novel adaptive dictionaries as sparse representation for video frame compressed sensing, comparative experimental results demonstrate that the proposed adaptive K -SVD dictionary learning algorithm achieves better performance than traditional K -SVD algorithm in terms of running time. In addition, the new method has better signal sparse representation performance, and also can reduce the reconstruction error of compressed sensing.

Key words: K -SVD algorithm; adaptive K -SVD algorithm; dictionary learning; sparse representation; compressed sensing

0 引言

近年来, 以信号的稀疏性先验求解图像反问题吸引着学者们的广泛关注^[1-2], 尤其是压缩感知 (Com-

pressed Sensing, CS) 领域^[3]。CS 理论主要包括三个阶段: 信号的稀疏表示、观测矩阵的选取和信号重构。其中寻找信号最佳的稀疏域, 是压缩感知理论应用的

收稿日期: 2016-05-29

修回日期: 2016-09-08

网络出版时间: 2017-04-28

基金项目: 国家自然科学基金资助项目 (61070234, 61071167, 61373137, 61501251); 江苏省 2015 年度普通高校研究生科研创新计划项目 (KYZZ15_0235); 南京邮电大学引进人才科研启动基金资助项目 (NY214191)

作者简介: 钱 阳 (1991-), 女, 硕士生, 研究方向为非线性分析及应用; 李 雷, 博士, 教授, 研究方向为智能信号处理与非线性科学及其在通信中的应用。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20170428.1702.016.html>

前提和基础,它决定了图像反问题的求解质量。

传统的稀疏表示思路是基于固定正交基的变换,如傅里叶变换、离散余弦变换、小波变换、Curvelet 变换等。这些正交基虽然构造相对简单,计算复杂度低,但不能与图像本身的复杂结构最佳匹配,并不是最优的稀疏变换基。随着字典学习方法^[4-6]的深入研究,人们开始根据信号本身来学习过完备字典,对稀疏编码研究的一个热点是信号在冗余字典下的稀疏分解。诸多研究成果表明,通过学习获得的字典原子数量更多,形态更丰富,具有更稀疏的表示,能更好地与信号或图像本身的结构匹配,为图像带来更大的压缩空间,在图像分类^[7]、图像去噪^[8-9]、图像超分辨率^[10]等方面性能更优。

K -SVD(K -Singular Value Decomposition)^[11]算法是目前一个比较受欢迎的字典学习算法,该算法交替执行稀疏编码阶段和字典更新阶段,并且在字典更新步骤中利用奇异值分解方式逐个更新字典原子,避免矩阵求逆计算的同时也提高了算法的收敛速度,在图像处理中应用极其广泛^[12-14]。然而, K -SVD 算法存在字典训练时间长、计算量大等不足。

针对这一问题,为提高字典学习的收敛速度与性能,提出了一种新的快速字典学习算法-自适应 K -SVD 算法(adaptive K -SVD)。该算法在稀疏编码阶段,将稀疏上界与迭代更新的字典关联,以获得自适应的稀疏约束;而在字典更新阶段,使用经典 K -SVD 的字典更新方式,通过稀疏编码和字典更新两步迭代学习得到字典。将训练的自适应字典用于视频帧的稀疏表示。实验结果表明,提出算法运行速度快,具有更好的稀疏表示性能。

1 预备知识

1.1 压缩感知理论基础

在压缩感知理论中,若被测信号 $\mathbf{x} \in R^{n \times 1}$ 在某正交基或紧框架 $\Psi = (\psi_1, \psi_2, \dots, \psi_n)^T$ 上是稀疏的或是可压缩的,则可用一个与稀疏变换基不相关的 $m \times n$ 维($m \ll n$)观测矩阵 Φ 对稀疏变换向量 $\Theta = \Psi^T \mathbf{x}$ 进行线性观测,得到观测向量 $\mathbf{y} \in R^{m \times 1}$,而后利用优化算法从低维观测向量 \mathbf{y} 中高概率地重构出原始信号 \mathbf{x} ^[15],其观测模型如式(1)所示:

$$\mathbf{y} = \Phi \mathbf{x} = \Phi \Psi \Theta \quad (1)$$

信号的稀疏性是压缩感知理论最基本的前提,它决定了 CS 非自适应采样过程的有效性。常见的稀疏基包括 DCT 基、小波基、FFT 基等,这些基构造简单,计算复杂度低,方便分析,然而不能处理图像以及更高维数据的奇异性,并非最优的稀疏基。近年来,基于过完备字典的信号稀疏分解方法发展迅猛,如何构造出

更高效的冗余字典已成为学者们研究的重点。

设计出一个平稳的、与稀疏基 Ψ 不相关的测量矩阵 Φ 是 CS 理论应用的关键,Candès 和 Tao 指出,观测矩阵 Φ 只有满足了约束等距性(Restricted Isometry Property, RIP)条件,才能保证准确地重构出原始信号。常用的观测阵有 Gauss 随机矩阵、哈达玛矩阵、伯努利矩阵等。

信号的重构过程即为从观测值 \mathbf{y} 中恢复稀疏信号 $\hat{\Theta}$,然后通过变换 $\hat{\mathbf{x}} = \Psi \hat{\Theta}$ 得到原始信号的逼近,其数学模型可描述如下:

$$\min_{\Theta} \|\mathbf{y} - \Phi \Psi \Theta\|_2^2 + \lambda \|\Theta\|_0 \quad (2)$$

其中, λ 为正则化参数。

目前重构算法主要集中于贪婪追踪算法、凸优化算法和组合算法等。

1.2 字典学习方法

基于过完备字典的稀疏表示是一种全新的信号表示理论。近年来,以设计简单、高效、通用性强的字典为目标的字典学习方法吸引着学者们的广泛关注,其中最受欢迎的是由 Michal Aharon、Michael Elad 提出的 K -SVD 字典学习算法。

该算法解决的是如式(3)所示的优化问题:

$$\min_{D, A} \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_F^2 \quad \text{s.t.} \quad \forall i, \|\mathbf{a}_i\|_0 \leq T_0 \quad (3)$$

其中, $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N \in R^{n \times N}$ 为训练集; $\mathbf{D} \in R^{n \times K}$ ($K > n$) 为稀疏表示的字典; $\mathbf{A} = \{\mathbf{a}_i\}_{i=1}^N \in R^{K \times N}$ 为稀疏系数矩阵; T_0 为稀疏表示系数中非零分量数目的上限。

K -SVD 算法的具体步骤如下:

算法 1: K -SVD 算法。

初始化:随机初始化归一化字典矩阵 $\mathbf{D}^{(0)} \in R^{n \times K}$

While 停止迭代条件不满足;

1) 稀疏编码阶段。

固定字典 \mathbf{D} ,使用任意一种追踪算法来求解稀疏表示系数 \mathbf{a}_i ($i = 1, 2, \dots, N$)。

$$\min_{\mathbf{a}_i} \|\mathbf{x}_i - \mathbf{D}\mathbf{a}_i\|_2^2 \quad \text{s.t.} \quad \forall i, \|\mathbf{a}_i\|_0 \leq T_0 \quad (4)$$

2) 字典更新阶段。

固定稀疏系数矩阵 \mathbf{A} ,对于 $k = 1, 2, \dots, K$

(1) 定义使用到原子 \mathbf{d}_k 的样本的索引为: $\omega_k = \{i \mid 1 \leq i \leq N, \mathbf{a}_i(k) \neq 0\}$;

(2) 计算 $\mathbf{E}_k = \mathbf{X} - \sum_{j \neq k} \mathbf{d}_j \mathbf{a}_{ij} = \mathbf{X} - (\mathbf{D}\mathbf{A} - \mathbf{d}_k \mathbf{a}_k)$, \mathbf{E}_k

为除第 k 个原子以外的其他所有原子产生的表示误差;

(3) 使用 ω_k 对 \mathbf{E}_k 进行约束,得到 \mathbf{E}_k^R ,其中 $\mathbf{E}_k^R = \mathbf{E}_k \Omega_k$, Ω_k 是 $(\omega_k(i), i)$ 为 1,其余都为 0 的 $N \times |\omega_k|$ 矩阵;

(4) 对 E_k^R 进行奇异值分解: $E_k^R = U\Delta V^T$ 。

更新字典: $d_k = u_1, a_k = \Delta(1, 1) \cdot v_1$ 。

2 自适应 K-SVD 算法

尽管 K-SVD 算法不需要矩阵求逆计算,且在字典更新阶段联合更新系数矩阵与字典原子,但存在字典训练时间长、计算量大等缺陷。为此,对经典的 K-SVD 算法进行改进,提出了自适应 K-SVD 算法。该算法交替执行稀疏编码和字典更新这两个阶段。

2.1 稀疏编码阶段

不同于 K-SVD 算法的稀疏编码方式,新算法利用字典的相干性将稀疏约束上限与迭代更新的字典关联,以获得自适应的稀疏约束上限,从而反复减少重构误差。

定义 T_j 为每次迭代过程中的稀疏约束上界:

$$T_j = \frac{1}{2} \left(1 + \frac{1}{\mu(D_j)} \right), j = 1, 2, \dots, P \quad (5)$$

其中, $\mu(D) \in [0, 1]$ 表示字典的相干性,其描述了完备字典中原子间的最大相似程度,公式如下^[16]:

$$\mu(D) = \max | \langle d_i, d_j \rangle |, i \neq j \quad (6)$$

当 $\mu(D)$ 值很大时,字典原子间相似程度很强,反之很弱。

为了充分说明 T_j 的合理性,引入如下定理^[16-17]:

定理 1: 给定字典 $D \in R^{n \times K} (K > n)$, 其相干性为 $\mu(D)$, 假设 $x = Da$ 有稀疏解 a , 其稀疏度 S 若满足:

$$S < \frac{1}{2} \left(1 + \frac{1}{\mu(D)} \right) \quad (7)$$

则可推导出以下结论:

- (1) 解 a 必定是最稀疏的;
- (2) 式(3)中的 l_0 问题也可以等价于 l_1 问题;
- (3) 任意一种追踪算法(如 OMP)都能从字典 D 中找出最佳的 S 项原子的线性组合。

由定理 1 可见,所定义的 T_j 能够保证稀疏信号精确恢复,是合理可行的。

使用 T_j 代替式(4)中的 T_0 ,则稀疏编码阶段即为求解如下的优化问题:

$$\min_{a_i} \|x_i - D_j a_i\|_2^2 \text{ s.t. } \|a_i\|_0 \leq T_j \forall i = 1, 2, \dots, N; j = 1, 2, \dots, P \quad (8)$$

该问题可以通过任意一种追踪算法(如 BP、OMP)进行求解。

2.2 字典更新阶段

在该阶段应用经典的 K-SVD 字典更新方式,根据稀疏表示系数 A ,对字典 D 中的原子进行迭代更新,字典列的更新结合稀疏表示系数的一个更新,使字典和稀疏表示系数同步更新。此阶段求解的是如下的优

化问题:

$$\min_A \|X - DA\|_2^2 \quad (9)$$

经典的 K-SVD 算法采用 SVD 来求解上述优化问题。

给定一组视频序列,假设其由 I 帧 $W \times L$ 图像组成,则这组视频序列可表示为^[18]:

$$\text{squ} = f_r(x, y) \quad (10)$$

其中, $1 \leq x \leq L, 1 \leq y \leq W, 1 \leq r \leq I$ 。

将视频序列的单帧图像分成 $\sqrt{n} \times \sqrt{n}$ 的块,且假设 W, L 均是 \sqrt{n} 的整数倍,则每帧图像可分成 WL/n 块。其中,第 r 帧第 t 块图像 $f_r^t(x, y)$ 所对应的向量 x_r^t 可采用行堆叠或列堆叠的方法构成,即:

$$x_r^t = (f_r^t(1, 1), \dots, f_r^t(1, \sqrt{n}), f_r^t(2, 1), \dots, f_r^t(2, \sqrt{n}), \dots, f_r^t(\sqrt{n}, 1), \dots, f_r^t(\sqrt{n}, \sqrt{n}))^T \quad (11)$$

其中, $x_r^t \in R^{n \times 1}, 1 \leq t \leq WL/n, 1 \leq r \leq I$ 。

对图像块 x_r^t 进行 CS 观测,测量值为 $y_r^t = \Phi x_r^t + \epsilon_r^t$ (ϵ_r^t 为观测过程中的随机噪声)。利用所提自适应 K-SVD 算法训练出冗余字典 $D = D_p$,并将其用于视频帧图像块的稀疏表示,即 $x_r^t = Da_r^t$ ($D \in R^{n \times K}, a_r^t \in R^{K \times 1}$ 表示第 r 帧第 t 块图像的稀疏表示系数),然后分别对每一个图像块进行压缩感知重构操作,最后按照顺序将重构出的图像块合并成整个图像,从而恢复出视频序列的第 r 帧。

综上,基于自适应 K-SVD 字典学习算法的视频帧稀疏重建算法的具体实现步骤如下:

算法 2: Adaptive K-SVD-CS 算法。

输入: 训练样本集 $\{x_i\}_{i=1}^N$, 最大迭代次数 P , 观测矩阵 Φ , 第 r 帧第 t 块图像的线性观测值 $y_r^t (1 \leq t \leq WL/n)$;

输出: 字典 $D = D_p$, 重构出的视频序列的第 352×288 帧 \hat{X}_r 。

初始化: 随机初始归一化字典矩阵 $D^{(0)} \in R^{n \times K}$, 通过式(5)获得初始化稀疏约束上限 T_0 。

for $j = 1, 2, \dots, P$

1) 稀疏编码阶段。

固定当前字典 D_{j-1} 和稀疏约束上限 T_{j-1} , 使用 OMP 算法求解式(8), 获得对应于训练样本 X 的稀疏系数矩阵 A_j 。

2) 字典更新阶段。

固定稀疏系数矩阵 A_j , 对于 $k = 1, 2, \dots, K$

(1) 定义使用到原子 d_k 的样本的索引为: $\omega_k = \{i | 1 \leq i \leq N, a_i(k) \neq 0\}$;

(2) 计算 $E_k = X - \sum_{j \neq k} d_j a_j$, E_k 为除第 k 个原子以外的其他所有原子产生的表示误差;

(3) 使用 ω_k 对 E_k 进行约束, 得到 E_k^R , 其中 $E_k^R = E_k \Omega_k$, Ω_k 是 $(\omega_k(i), i)$ 为 1, 其余都为 0 的 $N \times |\omega_k|$ 矩阵;

(4) 对 E_k^R 进行奇异值分析: $E_k^R = U \Delta V^T$ 。

更新字典原子: $d_k = u_1, a_k = \Delta(1, 1) \cdot v_1$ 。

获得更新后的字典 D_j 。

利用式(5)计算稀疏约束上限 T_j 。

end

for $t = 1, 2, \dots, WL/n$

给定观测矩阵 Φ , 第 r 帧第 t 块图像的线性观测值 y_r^t , 稀疏基选取所训练的自适应字典 $D = D_p$, 利用式(1)获得近似稀疏表示系数 \hat{a}_r^t ;

通过 $x_r^t = D a_r^t$ 获得重构的第 t 块图像。

end

将重构出的图像块 \hat{x}_r^t 合并成整个图像 \hat{X}_r , 恢复出视频的第 r 帧。

3 仿真实验及分析

采用格式为 CIF 的标准视频序列 Foreman 进行实验仿真, 随机选取 Foreman 的第 1、6、10、15、23 帧作为测试样本集, 如图 1 所示。



图 1 测试图像集

实验中将大小为 352×288 的视频帧分成不重叠的 8×8 图像块, 利用高斯随机矩阵对每个图像块进行观测, 以获得 CS 观测值。设置所训练的字典原子数为 256, 最大迭代次数为 30。实验平台为 Windows 7, Intel (R) Core (TM) i7-5600U CPU, 2.6 GHz, 8 GB, 所有实验设计基于 Matlab R2011a 编程实现。

3.1 稀疏基选择的视觉效果

为了验证改进的字典学习算法具有更好的稀疏表示性能, 分别采用 K -SVD 算法和所提 Adaptive K -SVD 算法对视频序列的第 23 帧进行稀疏表示, 从图像本身学习过完备字典。实验中设置 K -SVD 算法的稀疏度为 5, 所提稀疏度为稀疏编码阶段稀疏表示系数

中非零分量数目的最大值。通过两种字典学习算法训练出的冗余字典如图 2 所示。

从图 2 可以看出, 采用 K -SVD 算法训练出的字典已经能较好地与图像本身的结构相匹配, 而所提 Adaptive K -SVD 算法训练出的字典形态更丰富, 能与视频帧本身的复杂结构最佳匹配, 稀疏性能更优, 具有更好的应用前景。

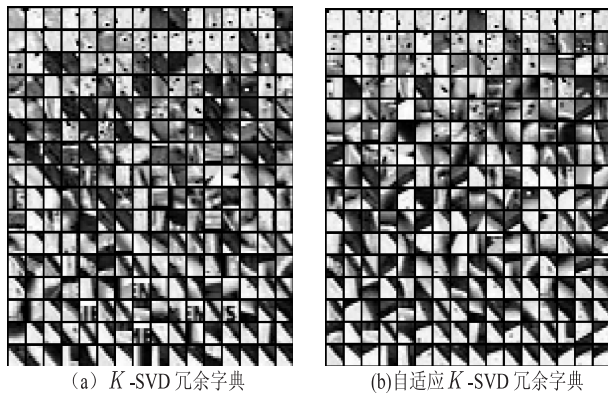


图 2 两种字典学习算法训练出的字典

3.2 在视频帧重构精度上的改进

本节将会展示所提字典学习算法对压缩感知重构性能的影响。分别选用 K -SVD 冗余字典和自适应 K -SVD 冗余字典作为 CS 的稀疏基, 利用 OMP 算法对测试图像集中的每一幅图像进行重构。为了方便描述, 分别记这两种算法为 KSVD-CS 和 Adaptive KSVD-CS。

为了评估各算法的重构性能, 除了选用 PSNR 作为评价标准外, 近年来相关学者提出的 FSIM^[19] 可以用来衡量重构图像的视觉效果。FSIM 值越高, 则重构图像的视觉效果越好。

图 3 从直观上给出了不同采样率下两种算法的运行时间, PSNR 和 FSIM 的对比情况。为了消除随机性, 运行时间、PSNR、FSIM 的数值均取五幅测试帧的平均值, 且每一测试帧的 TIME、PSNR 与 FSIM 均取 10 次执行结果的平均值。

从图 3(a) 可以看出, 所提 Adaptive KSVD-CS 算法的运行时间远远低于 KSVD-CS 算法, 这是因为所提的自适应 K -SVD 字典学习方法在稀疏编码阶段自适应选择稀疏度, 加快了字典的训练速度。

从图 3(b)、(c) 可以看出, 当采样率较低时, 两种算法的重构效果都不算很好, 随着采样率的增加, 两种算法的重构性能都在提升, 且所提算法的重构效果略胜一筹。

4 结束语

为提高字典学习的收敛速度与性能, 提出了一种自适应 K -SVD 字典学习算法。该算法在稀疏编码阶

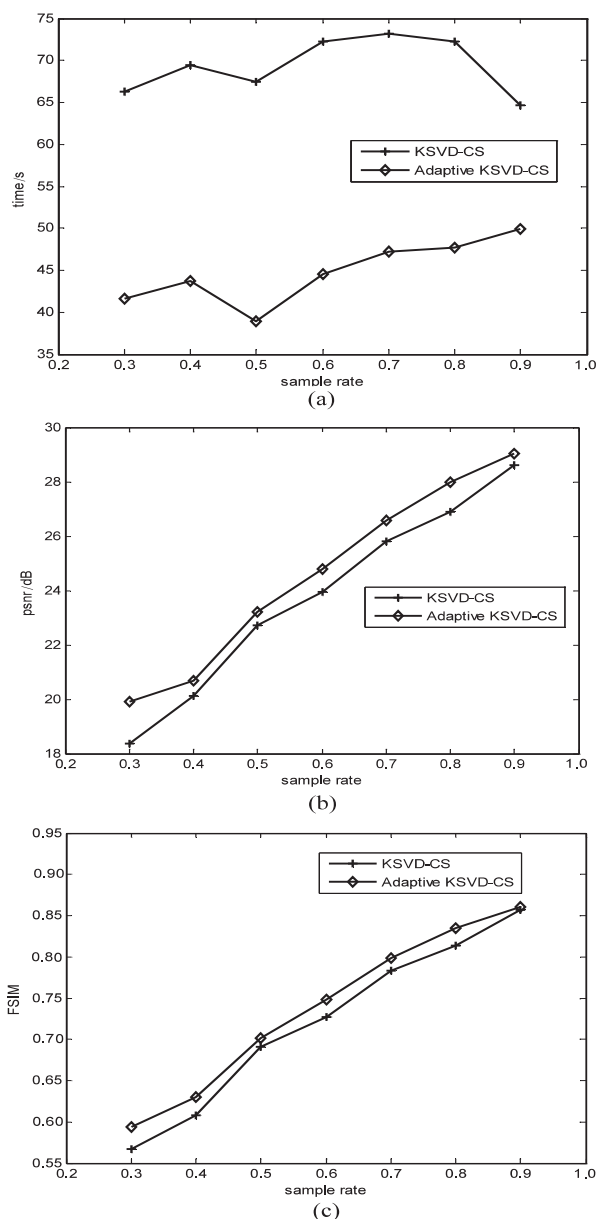


图 3 两种算法重构性能随采样率变化图

段根据当前字典自适应地调整稀疏度,更新字典则使用经典 K -SVD 的字典更新方式,稀疏编码与字典更新两步迭代学习得到字典,并将所学习到的字典用于视频帧的稀疏表示。仿真对比实验表明,所提算法提升了字典训练速度,提高了重构性能。

参考文献:

- [1] Donoho D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(5): 1289–1306.
- [2] Candès E, Tao T. Near optimal signal recovery from random projections: universal encoding strategies[J]. IEEE Transactions on Information Theory, 2006, 52(12): 5406–5425.
- [3] Donoho D L, Tsai Y. Extensions of compressed sensing[J].

Signal Processing, 2006, 86(3): 533–548.

- [4] 练秋生, 石保顺, 陈书贞. 字典学习模型、算法及其应用研究进展[J]. 自动化学报, 2015, 41(2): 240–260.
- [5] Kreutzdelgado K, Murray J F, Rao B D, et al. Dictionary learning algorithms for sparse representation[J]. Neural Computation, 2003, 15(2): 349–396.
- [6] Rubinstein R, Bruckstein A, Elad M. Dictionaries for sparse representation modeling[J]. Proceedings of the IEEE, 2010, 98(6): 1045–1057.
- [7] Bahrampour S, Nasrabadi N, Ray A, et al. Multimodal task-driven dictionary learning for image classification[J]. IEEE Transactions on Image Processing, 2016, 25(1): 24–38.
- [8] Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries[J]. IEEE Transactions on Image Processing, 2006, 15(12): 3736–3745.
- [9] Protter M, Elad M. Image sequence denoising via sparse and redundant representations[J]. IEEE Transactions on Image Processing, 2009, 18(1): 27–35.
- [10] Liu X, Zhai D, Zhao D, et al. Image super-resolution via hierarchical and collaborative sparse representation[C]//Data compression conference. [s.l.]: IEEE, 2013: 93–102.
- [11] Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311–4322.
- [12] Ravishanker S, Bresler Y. MR image reconstruction from highly undersampled k -space data by dictionary learning[J]. IEEE Transactions on Medical Imaging, 2011, 30(5): 1028–1041.
- [13] Bilgin A, Kim Y, Liu F, et al. Dictionary design for compressed sensing MRI[C]//ISMRM. [s.l.]: [s.n.], 2010.
- [14] Xu T T, Yang Z, Shao X. Adaptive compressed sensing of speech signal based on data-driven dictionary[C]//Conference on communications. [s.l.]: [s.n.], 2009.
- [15] 钱 阳, 李 雷. 一种基于新型 KPCA 算法的视频压缩感知算法[J]. 计算机技术与发展, 2015, 25(10): 101–106.
- [16] Donoho D L, Huo X M. Uncertainty principles and ideal atomic decomposition[J]. IEEE Transactions on Information Theory, 2001, 47(7): 2845–2862.
- [17] Donoho D L, Tsai Y. Fast solution of l_1 -norm minimization problems when the solution may be sparse[J]. IEEE Transactions on Information Theory, 2008, 54(11): 4789–4812.
- [18] Liu Sheng, Gu Mingming. K-L transform in video compressed sensing[C]//Proceeding of the 32nd Chinese control conference. Xi'an, China: IEEE, 2013: 4528–4532.
- [19] Zhang L, Zhang L, Mou X, et al. FSIM: a feature similarity index for image quality assessment[J]. IEEE Transactions on Image Processing, 2012, 20(8): 2378–2386.