中国科学: 信息科学　2013 年　第 43 卷　第 2 期: 226–242
www.scichina.com　info.scichina.com
《中国科学》杂志社 SCIENCE CHINA PRESS

论 文

图像视频信号的压缩采样与稀疏重建

尹宝才, 施云惠*, 丁文鹏, 胡永利, 李敬华

北京工业大学计算机学院, 多媒体与智能软件技术北京市重点实验室, 北京 100124
* 通信作者. E-mail: syhzm@bjut.edu.cn

收稿日期: 2011–12–20; 接受日期: 2012–05–30
国家重点基础研究发展计划 (批准号: 2011CB302703)、国家自然科学基金 (批准号: 60825203, 61033004, 60973056, 61170103, 61003182) 和北京市自然科学基金 (批准号: 4102009, 4112007) 资助项目

摘要　视觉传感器通常不知道它们 "看到" 的现象之下的物理过程, 以远远超出图像视频信号有效维度的 Shannon/Nyquist 采样率获取图像视频数据, 从而导致了对图像视频信号的存储、传输等数字处理的巨大压力. 压缩感知 (compressive sensing, CS) 理论表明: 在某个线性变换域下稀疏的信号, 可以利用少量的观测数据精确地重建, 或在噪声情况下鲁棒地重建. 压缩感知是实现图像视频信号有效维度采样的理论基础, 为图像视频信号的采样、处理和识别等领域带来了前所未有的突破. 本文对图像视频信号领域压缩感知面临的基本问题: 压缩采样、稀疏重建模型及其优化求解算法的研究进展进行了综述. 在采样方面, 分析了图像视频信号随机观测矩阵和有结构观测矩阵的性能; 在稀疏重建模型方面, 从图像视频信号的稀疏先验性出发, 介绍了分析型的重建模型和合成型重建模型的构建方法; 在优化求解方面, 针对重建模型, 介绍了约束优化问题和无约束优化问题两类求解算法. 以此为基础, 分析了在图像视频领域压缩感知的理论与应用的进一步发展所面临的问题和挑战, 展望了未来的发展方向.

关键词　采样　压缩感知　稀疏表示　随机观测　最优化

1　引言

随着信息技术和计算机互联网飞速发展, 图像视频信号已成为人类获取视觉信息的最主要载体, 人眼和照相机等视觉传感器是感知成像物理世界和搜集图像视频信息的工具, 而视觉传感器通常不知道他们 "看到" 的现象之下的物理过程, 以远远超出图像视频信号有效维度的 Shannon/Nyquist 采样率获取图像视频数据. 因此数字化过程产生的大量数据对图像视频信号的存储、传输和识别等处理都产生了巨大的压力. 以有效维度对图像视频信号高效采样已成为以存储、传输和识别等数字化处理为核心的图像视频应用中亟待解决的基础科学问题之一.

2004 年由 Candes 等 [1~6] 提出了一种新的采样理论 —— 压缩感知 (compressive sensing /compressed sampling, CS) , 压缩感知理论表明: 在某个变换域下稀疏的信号, 可以利用优化方法由与变换非一致的少量观测数据精确重建. 这种非自适应的压缩采样将信号中包含的信息凝聚在少量的观测数据上, 大幅度地降低了精确重建原始信号所需要的采样数目. 图像视频信号通常不是稀疏的, 在某变换域是近似稀疏信号, 如图像信号的小波系数成指数衰减. 压缩感知理论的出现改变了图像信息

引用格式: 尹宝才, 施云惠, 丁文鹏, 等. 图像视频信号的压缩采样与稀疏重建. 中国科学: 信息科学, 2013, 43: 226–242, doi: 10.1360/112011-1423

的获取、传输和识别方式, 在压缩成像系统 [7]、图像识别 [8~10]、医学成像处理 [11,12] 和图像视频压缩 [13~16] 等方面应用获得了巨大的成功.

一个稀疏的信号 $\boldsymbol{x} \in \mathbb{R}^N$ 在观测矩阵 $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ 下的采样 $\boldsymbol{y} \in \mathbb{R}^M$ $(M < N)$, 当 $M \geqslant C.K.\log N$ (C 是一个常数) 时, 信号 x 可以通过求解如下优化问题来精确重建

$$\begin{cases} \min\limits_{\boldsymbol{x}} f(\boldsymbol{x}) \\ \text{subject to } \boldsymbol{\Phi x} = \boldsymbol{y}. \end{cases} \tag{1}$$

正则化函数 $f(\boldsymbol{x})$ 的选择依赖于信号的先验假设. 如果图像视频信号 $\boldsymbol{x}$ 在与 $\boldsymbol{\Phi}$ 非一致的变换域 $\boldsymbol{D}$ 下是 (近似) 稀疏信号, $f(\boldsymbol{x})$ 的一个适当的选择是 $\|\boldsymbol{Dx}\|_1$, $\| \ \|_1$ 是 l_1 范数; 如果图像视频信号 $\boldsymbol{x}$ 是分片常数时, 全变分 $\mathrm{TV}(\boldsymbol{x})$ 可以提供高精确的重建结果, $\mathrm{TV}(\boldsymbol{x})$ 的定义在本文的第 3 部分给出.

2 图像视频信号的压缩采样

压缩采样的目的是将图像视频信号编码为与稀疏矩阵非一致的随机观测的观测数据. 在图像视频的压缩感知框架下, 观测矩阵包括: 随机观测矩阵和有结构的观测矩阵 [17~23].

2.1 随机观测矩阵

按照某种概率分布模型生成的随机矩阵被称为随机观测矩阵, 常用的随机观测矩阵包括如下几种 [2].

1) 高斯矩阵 (Gaussian matrix). 矩阵的每一个元素是由独立正态分布的 $N(0,1)$ 生成.

2) 正交高斯矩阵 (orthogonalized Gaussian matrix). 经过正交化处理的高斯矩阵, 其列是正交的.

3) Bernoulli 矩阵 (Bernoulli matrix). 矩阵元素是由 Bernoulli 分布产生的, 矩阵的每一个元素是由 0 或 1 构成的.

随机观测矩阵与任何稀疏矩阵具有通用的非一致关系, 与其他类型的矩阵相比精确重建原始信号所需要采样数几乎是最小的, 然而在实际应用中存在两个本质的缺欠: 1) 由 $O(N^2)$ 独立随机变量生成, 因此需要大量的存储空间; 2) 随机的完全无结构特性导致了较高的运算复杂性.

2.2 有结构的观测矩阵

1) 正交变换矩阵. Fourier 变换和 DCT 变换等正交矩阵也可以作为观测矩阵 [13,17], 在实际应用中, 一般按照某种方式选取正交矩阵的若干列, 由这些列构成观测矩阵, 所获得观测数据就是相应的变换系数, 而对于列的选择等价于变换系数的选择. 随机方式可以是不同频段以不同的概率随机采样, 也可以是采用一致的概率采样.

2) Toeplitz 矩阵. 文献 [18~20] 提出了一种 Toeplitz 结构的观测矩阵, Toeplitz 矩阵具有如下形式:

$$\boldsymbol{T} = \begin{bmatrix} t_n & t_{n-1} & \dots & t_1 \\ t_{n+1} & t_n & \dots & t_2 \\ \ddots & \ddots & \ddots & \\ t_{2n-1} & t_{2n-2} & \dots & t_n \end{bmatrix}. \tag{2}$$

矩阵 $\boldsymbol{T}$ 满足 $\boldsymbol{T}_{i,j}=\boldsymbol{T}_{i+1,j+1}$, $\boldsymbol{T}_{i,j}$ 表示 (i,j) 位置的元素, Toeplitz 的矩阵元素是独立同分布的.

3) Scrambed block Hadamard resemble 观测矩阵 [23] 具有如下形式:

$$\boldsymbol{\Phi}=\boldsymbol{Q}_M\boldsymbol{W}\boldsymbol{P}_N, \tag{3}$$

这里 $\boldsymbol{W}$ 是 $N\times N$ 的块对角阵, 定义如下

$$\boldsymbol{W}=\begin{bmatrix}\boldsymbol{W}_{\mathrm{H}} & & & \\ & \boldsymbol{W}_{\mathrm{H}} & & \\ & & \ddots & \\ & & & \boldsymbol{W}_{\mathrm{H}}\end{bmatrix}, \tag{4}$$

$\boldsymbol{W}_{\mathrm{H}}$ 表示 Hadamard 矩阵, $\boldsymbol{Q}_M$ 是一个算子, 用 $\boldsymbol{Q}_M$ 左乘某个矩阵的操作表示随机选取该矩阵的 M 行. 而算子 $\boldsymbol{P}_N$ 是一个置换阵, 用 $\boldsymbol{P}_N$ 右乘一个矩阵的操作表示随机排序该矩阵的 N 列.

不论是随机的观测矩阵还是有结构的观测矩阵, 在设计和应用时, 需要考虑如下几个方面: (i) 近似最优性. 用于精确重建的观测数目接近理论的边界 $O(K\log N)$; (ii) 普适性. 观测矩阵能够与不同类型的稀疏矩阵相匹配. (iii) 高效的存储与计算. 只需要较少的空间存储观测矩阵和较低的运算复杂度对图像视频信号观测.

Gauss 等类型的随机观测矩阵与任何稀疏矩阵具有通用的非一致关系, 与其他类型的观测矩阵相比, 精确重建原始信号所需要采样数几乎是最小的, 在最优性能和普适性方面表现较好, 而该类矩阵是由 $O(N^2)$ 独立随机变量生成, 是随机且完全无结构, 因此, 该类矩阵不适于低存储能力和弱运算能力的成像系统, 该类矩阵多见于自然图像的采样过程. 对于医学成像系统, 完全随机观测在物理上是不实现的, 通常采用部分 Fourier 观测. 目前, 图像视频混合编码框架中的变换主要是 DCT 和小波变换, 为了与图像视频混合编码框架结合, 提高图像视频压缩效率, 通常采用 DCT 变换/小波变换为观测, 解码端以变换系数为观测数据, 稀疏优化重建, 可获得了高于反变换的重建质量 [13], Toeplitz 矩阵是由 $O(N)$ 独立随机变量生成, 该结构的矩阵运算可以通过 FFT 实现, 因此可以获得更快的压缩采样和重构速度, 同时 Toeplitz 矩阵重建性能与完全随机矩阵相当. 由于 Hadamard 矩阵的元素是二进制的, scrambed block Hadamard resemble 的观测易于物理实现, 分块的结构可带来低存储代价、快速计算和并行计算等诸多优势.

3 图像视频信号的稀疏重建模型

对给定的观测矩阵, 信号的重建质量取决于信号稀疏化方式, 如何实现信号的高效稀疏表示是关键. 另一方面, 一个信号有时域 x 和变换域 s 两种表现形式, 因此, 有两种方式刻画该信号的稀疏性: 1) 将信号向某个基下投影而获得稀疏表示 $\boldsymbol{Bx}$; 2) 将信号表示成基向量的线性组合 $\boldsymbol{\Psi s}$, 由此导致两类稀疏重建模型, 式 (5) 和 (6) 分别为标准的分析模型和标准的合成模型.

$$\begin{cases}\min\limits_{\boldsymbol{x}}\ \|\boldsymbol{Bx}\|_1\\ \text{subject to}\ \ \|\boldsymbol{\Phi x}-\boldsymbol{b}\|_2\leqslant\varepsilon,\end{cases} \tag{5}$$

$$\begin{cases}\min\limits_{\boldsymbol{s}}\ \|\boldsymbol{s}\|_1\\ \text{subject to}\ \ \|\boldsymbol{\Phi\Psi s}-\boldsymbol{b}\|_2\leqslant\varepsilon,\end{cases} \tag{6}$$

其中 ε 是常数. 随着压缩感知在图像视频应用的迅猛发展, 超越稀疏性的图像视频信号特性也可以作为先验信息来提高图像的重建质量, 为此出现了许多复杂的模型, 然而, 这些复杂模型可视为合成模型 (5) 和分析模型 (6) 的推广.

3.1 分析模型 (analysis model)

对于给定的观测数据, 信号的重建质量本质上取决于信号的稀疏表示能力, 为此, 我们围绕图像视频信号的稀疏表示问题对分析模型进行讨论.

3.1.1 基于图像梯度信息稀疏重建模型

1) TV 稀疏表示重建模型. 对于 $n\times n$ 的图像信号 $\boldsymbol{x}$, x_{ij} 是 (ij) 位置的像素值, 定义如下算子:

$$D_{h;ij}\boldsymbol{x}=\begin{cases} x_{i+1,j}-x_{i,j}, & i<n,\\ 0, & i=n,\end{cases}\qquad D_{v;ij}\boldsymbol{x}=\begin{cases} x_{i,j+1}-x_{i,j}, & j<n,\\ 0, & j=n,\end{cases}\tag{7}$$

$$D_{ij}\boldsymbol{x}=\begin{pmatrix} D_{h;ij}\boldsymbol{x}\\ D_{v;ij}\boldsymbol{x}\end{pmatrix}.\tag{8}$$

图像相邻像素间具有较强相关性, 导致了图像信号的大部分区域 $D_{ij}\boldsymbol{x}$ 的能量

$$\|D_{ij}\boldsymbol{x}\|_2=\sqrt{(D_{h;ij}\boldsymbol{x})^2+(D_{v;ij}\boldsymbol{x})^2}$$

近似于 0, 因此全变分 $\mathrm{TV}(\boldsymbol{x})=\sum_{ij}\|D_{ij}\boldsymbol{x}\|_2$ 非常小. 图像的 $\mathrm{TV}(\boldsymbol{x})$ 稀疏表示重建模型为求解与观测数据 b 相符合的最小全变分问题 [2]:

$$\min \mathrm{TV}(\boldsymbol{x}) \text{ subject to } \|\boldsymbol{\Phi}\boldsymbol{x}-\boldsymbol{b}\|_2\leqslant\varepsilon.\tag{9}$$

模型 (3) 是图像视频应用中最常用的模型.

2) 基于回归模型指导的稀疏重建模型. Wu 等 [24] 在分析图像视频信号结构特征的基础上, 针对观测数据的全局特性, 建立了图像的分片自回归模型 PAR (piece auto-regressive model)

$$a_{i,j}=\arg\min_{\boldsymbol{a}}\sum_{i,j}(x_{i,j}-\sum_{(u,v)\in\boldsymbol{S}}a_{u,v}x_{i-u,j-v})^2.\tag{10}$$

其基本思想是: 任意 $x_{i,j}$ 可以由其邻域的像素 $\{x_{i-u,j-v}\}_{(u,v)\in S}$ 的线性组合 $\sum_{(u,v)\in\varGamma}a_{u,v}x_{i-u,j-v}$ 逼近, $a_{i,j}$ 是组合系数, $\boldsymbol{a}$ 是由 $a_{i,j}$ 构成的向量, $\varGamma$ 用于表示 (i,j) 位置像素的邻域, 通过求解上述优化问题 (10), 获得图像信号的稀疏表示. 在此基础上建立了基于回归模型指导的稀疏重建模型

$$\begin{cases}\min\limits_{\boldsymbol{x}}\sum\limits_{i,j}(x_{i,j}-\sum\limits_{(u,v)\in\varGamma}a_{u,v}x_{i-u,j-v})^2\\ \text{subject to } \|\boldsymbol{\Phi}\boldsymbol{x}-\boldsymbol{b}\|_2^2\leqslant\varepsilon.\end{cases}\tag{11}$$

在实际计算中, PAR 模型 (10) 与重建模型 (11) 求解是交替进行的. 与 TV 类似, PAR 模型是对信号的稀疏表示的一种方式. 由于 $a_{i,j}$ 是通过学习获得, 可以获得比 $\mathrm{TV}(\boldsymbol{x})$ 更好的稀疏性. 然而, 这种固定结构的 PAR 模型仍然无法适应多变的图像局部特征, 一个自适应纹理结构的 PAR 模型有望进一步改善图像视频信号的重建性能.

3.1.2 基于正交变换的稀疏重建模型

由于图像视频信号在某些正交变换下的系数是稀疏的, 同时正交特性为重建模型的数值求解带来了极大的便利, 出现了许多基于 DCT, DFT 和 DWT 等正交变换的稀疏重建模型. 基于正交变换的稀疏重建模型与标准的分析模型 (5) 具有同样的形式, 此时 $\boldsymbol{B}$ 是正变换. 在这些模型中, 由于 DWT 具有多分辨特性, 更加适于刻画图像视频这类非平稳信号, 因此基于 DWT 的稀疏重建模型得到了广泛的应用.

3.1.3 基于梯度与正交变换的混合稀疏重建模型

正交变换的稀疏表示方法和 TV 稀疏表示方法是从两个完全不同的角度去刻画信号的稀疏性, 将二者结合起来能够捕捉更丰富的图像结构和高频细节信息. Yang [25] 提出了 TV-Wavelet 混合重建模型:

$$\begin{cases} \min\limits_{x} \mathrm{TV}(\boldsymbol{x}) + \lambda \ \|\boldsymbol{B}\boldsymbol{x}\|_1 \\ \text{subject to} \ \ \|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{b}\|_2 \leqslant \varepsilon, \end{cases} \tag{12}$$

其中 $\boldsymbol{B}$ 是小波变换矩阵, λ 是参数. 文献 [26] 基于图像边信息, 利用各向异性特性的加权 TV 算子代替传统的 TV 算子, 从而建立加权 TV-Wavelet 重建模型:

$$\begin{cases} \min\limits_{\boldsymbol{x}} \sum\limits_{\alpha} g_\alpha \|D_\alpha(\boldsymbol{x})\| + \lambda \ \|\boldsymbol{B}\boldsymbol{x}\|_1 \\ \text{subject to} \ \ \|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{b}\|_2 \leqslant \varepsilon, \end{cases} \tag{13}$$

其中 $\sum_\alpha g_\alpha |D_\alpha(\boldsymbol{x})| = \sum g_{(i,j)\to(k,l)} \left|x_{(i,j)} - x_{(k,l)}\right|$, α 表示相邻对 $(i,j) \to (k,l)$, 其中是 D_α 和 g_α 是 α 方向上的梯度算子和加权参数, g_α 的值可根据边的信息确定.

此外还有一些类似加权 TV-Wavelet 重建模型, 此类模型充分利用边信息, 在一定程度上改善了信号的重建质量和重建速度, 边信息是通过对信号的初始估计中获得的, 边信息估计的精度直接影响到重建质量, 如果初始估计的精度较差, 会导致错误的边信息, 从而降低重建性能, 这是此类模型的瓶颈.

3.2 合成模型 (synthesis model)

设 $\boldsymbol{\Psi}$ 是由 K 个图像基元 $\{\boldsymbol{\Psi}_j\}_{j=1}^K$ 构成的稀疏矩阵, 一个图像信号可以表示基元的线性组合形式: $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s}$, 或者近似的线性组合形式: $\boldsymbol{x} \approx \boldsymbol{\Psi}\boldsymbol{s}$, 向量 $s \in \mathbb{R}^k$ 是信号 $\boldsymbol{x}$ 的组合系数, 通常称这些基元的集合 $\{\boldsymbol{\Psi}_j\}_{j=1}^K$ 为稀疏字典.

这些基元 $\{\boldsymbol{\Psi}_j\}_{j=1}^K$ 可以通过某种特殊的变换导出, 称之为解析稀疏字典, 解析的稀疏字典因其信号的稀疏表示计算方面表现出简单和快速等优势因此被广泛应用. 该类字典性能取决于字典对信号的稀疏表示能力. 如果 $\{\boldsymbol{\Psi}_j\}_{j=1}^K$ 是通过学习方式获得的, 称之为非解析稀疏字典. 这类字典是从一类信号信息中学习得到的, 会给出信号更好的稀疏表示, 并且可以很方便地用于合成模型, 而这正是分析模型所欠缺的.

3.2.1 解析的稀疏字典表示模型

1) 正交基和冗余基. 在合成模型中, 稀疏矩阵 $\boldsymbol{\Psi}$ 的选择是至关重要的, 与分析模型相似, 可以选择 DCT、Fourier 变换和小波变换等正交变换, 也可以选择 Ridgelet 变换、Curvelet 变换和 Contourlet

变换等. Ridgelet 变换、Curvelet 变换和 Contourlet 变换是一类由基函数生成的冗余函数系统, 除了具有多尺度分析和时频定位等与传统小波变换相同的特性外, 还具有方向变换和各向异性的尺度变换特性, 因此与小波变换相比, 该类冗余函数系统能更容易捕捉到图像结构的特征. 另一方面, 从函数逼近的角度出发, Fourier 变换或小波变换等正交变换和 Ridgelet, Curvelet 和 Contourlet 等非正交变换都可应用于对图像信号的逼近. 不同的是, 前者的任务是针对图像信号特性, 从一个给定的基库中挑选好的基, 后者是在一个适当的冗余函数系统中挑选最好的函数来线性组合逼近原信号, 而这一特性恰好与图像信号的时频多变性相适应. 因此不难发现, 使用 Fourier 变换或小波变换等正交变换表示图像, 不能保证图像总能获得好的稀疏表示, 而上述冗余函数系统却为图像的好的稀疏表示提供了更大的可能, 因此在合成模型中常常采用冗余基作为稀疏基.

2) 多空间冗余基. 由于图像视频信号通常具有显著的非平稳特性, 若采用传统的单一空间变换方法进行信号特性描述, 难以刻画图像视频信号在时空等变换域的特性, 因此, 可以用多个空间变换把图像信号分解, 以达到稀疏表示的目的. 现存的空间变换, 如余弦变换、小波、曲线波等以及所构造的不规则波使得刻画图像几何特征的多个空间表示的选择成为可能. 文献 [27] 提出了多空间冗余基构造方法, 首先对图像的几何特性进行分析, 选择了小波 $\boldsymbol{T}_{\mathrm{W}}$ 和不规则波 $\boldsymbol{T}_{\mathrm{I}}$ 来表示图像 $\boldsymbol{x}$, 图像在这两个变换下的系数为 $\boldsymbol{s}_{\mathrm{W}}$ 和 $\boldsymbol{s}_{\mathrm{I}}$, $\boldsymbol{s}_{\mathrm{W}} = \boldsymbol{T}_{\mathrm{W}}\boldsymbol{x}$ 和 $\boldsymbol{s}_{\mathrm{I}} = \boldsymbol{T}_{\mathrm{I}}\boldsymbol{x}$. 假设 $\boldsymbol{T}_{\mathrm{W}}$ 和 $\boldsymbol{T}_{\mathrm{I}}$ 的反变换为 $\tilde{\boldsymbol{T}}_{\mathrm{W}}$ 和 $\tilde{\mathrm{T}}_{\mathrm{I}}$, 求解问题 (14) 可获得图像信号在小波 $\boldsymbol{T}_{\mathrm{W}}$ 和不规则波 $\boldsymbol{T}_{\mathrm{I}}$ 所构成的冗余基下的稀疏表示

$$\begin{cases} \min\limits_{\boldsymbol{s}_{\mathrm{W}},\boldsymbol{s}_{\mathrm{I}}} ||\boldsymbol{s}_{\mathrm{W}}||_1 + ||\boldsymbol{s}_{\mathrm{I}}||_1 \\ \text{subject to } \boldsymbol{x} = \tilde{T}_{\mathrm{W}}\boldsymbol{s}_{\mathrm{W}} + \tilde{T}_{\mathrm{I}}\boldsymbol{s}_{\mathrm{I}}. \end{cases} \tag{14}$$

3.2.2 非解析的稀疏字典表示模型

学习的目的就是寻找能够稀疏表示训练样本的图像基元 $\{\boldsymbol{\Psi}_j\}_{j=1}^{K}$. 其复杂度高于预先确定的解析字典 [8,28~35].

1) 简单字典. 由于一个测试信号可以表示为训练图像样本的线性组合, 因此由训练图像可以构成一个最简单的字典, 此时, 字典的基元不需要学习获得, 可以直接来自训练样本. Wright 等 [8] 提出了一种稀疏表示的人脸表示方法, 以稀疏表示为特征实现人脸鲁棒高效的识别. 基本思想是: 将一个测试人脸图像表示为一系列人脸训练图像的线性组合, 人脸训练图像是由多人的多幅不同图像组成, 假设第 i 人有 n_i 个人脸图像, 记第 i 人的人脸图像样本为 $\boldsymbol{A}_i = [\boldsymbol{v}_{i,1}, \boldsymbol{v}_{i,2}, \ldots, \boldsymbol{v}_{i,n_i}]$, 第 i 人的一个新测试图像 x 可以表示成该人的人脸图像样本的线性组合

$$\boldsymbol{x} = \alpha_{i,1}\boldsymbol{v}_{i,1} + \alpha_{i,2}\boldsymbol{v}_{i,2} + \cdots + \alpha_{i,n_i}\boldsymbol{v}_{i,n_i}. \tag{15}$$

定义 $\boldsymbol{\Psi} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k]$, $\boldsymbol{x}$ 也可表示成所有人的人脸样本的线性组合形式

$$\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s}, \tag{16}$$

这里 $\boldsymbol{s} = [0, \ldots, 0, \alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n_i}, 0, \ldots, 0]^{\mathrm{T}}$ 是组合系数, 除了与第 i 个人相关系数外, 其他系数均为 0, 由此, $\boldsymbol{x}$ 在简单字典 $\boldsymbol{\Psi}$ 下获得了稀疏表示. 稀疏系数可以通过求解如下优化问题获得:

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{subject to } \|\boldsymbol{\Psi}\boldsymbol{s} - \boldsymbol{x}\|_2 \leqslant \varepsilon. \tag{17}$$

这种模型有许多有效的数值解法. 这种简单字典的构造方法可以直接推广到自然图像的处理中 [28], 所不同的是: 人脸具有相似的结构, 人脸字典是基于整幅图像的字典, 而自然图像具有复杂的时频特

性, 因此基于图像块建立字典. 简单字典无需要学习, 然而当样本数目非常庞大时, 信号的重建速度太慢.

2) $\boldsymbol{K}$-SVD 字典. 从某种意义下 K-means 聚类的向量量化技术是训练字典一种简单的算法, 对于给定的基元集 (码本), 将样本分配到各个分组中, 使得每个样本与某个分组的基元距离最小, 然后按照组内最小化平均距离的原则优化基元. 然而这样每一个样本只能由一个基元表示. Aharon 等 [29] 在传统的 K-means 聚类方法基础上, 结合稀疏表示机制, 提出了一种广义 K-means 方法 ——K-SVD 算法. 其基本思想为对于图像数据样本 $\boldsymbol{x}=\{\boldsymbol{x}_i\}_{i=1}^N$ 寻找一个能够稀疏表示这些样本的基元集合 Ψ. K-SVD 方法是求解如下优化问题:

$$\min_{\boldsymbol{\Psi},\boldsymbol{S}} \|\boldsymbol{X}-\boldsymbol{\Psi}\boldsymbol{S}\|_F \quad \text{subject to} \ \forall i \ \ \|\boldsymbol{s}_i\|_1 \leqslant T_0, \tag{18}$$

这里 $\|\boldsymbol{X}-\boldsymbol{\Psi}\boldsymbol{S}\|_F = \sum_{i=1}^N \|\boldsymbol{x}_i-\boldsymbol{\Psi}\boldsymbol{s}_i\|_2^2$, T_0 是一个参数, $\boldsymbol{s}_i$ 是矩阵 $\boldsymbol{S}$ 中的第 k 列, $\|\boldsymbol{s}_i\|_1 \leqslant T_0$ 是稀疏表示的约束项.

而对于 K-means 的方法而言, 此时 s_i 中只有一非零向量时, 即存在某个 k 使得 $\boldsymbol{s}_i=\boldsymbol{e}_k$. K-means 方法本质是求解如下优化问题:

$$\min_{\boldsymbol{\Psi},\boldsymbol{S}} \|\boldsymbol{X}-\boldsymbol{\Psi}\boldsymbol{S}\|_F^2 \quad \text{subject to} \ \forall i \ \boldsymbol{s}_i=\boldsymbol{e}_k \ \text{for some} \ k. \tag{19}$$

问题 (7) 的近似解法可通过迭代方式获得. 当只考虑字典 $\boldsymbol{\Psi}$ 的一列 $\boldsymbol{\Psi}_k$ 和与之相适应的稀疏表示 $\boldsymbol{s}_T^k$ 时, $k=1,2,\ldots,K$, $\boldsymbol{s}_T^k$ 是矩阵 $\boldsymbol{S}$ 中的第 k 行, 最小化的目标项可以写成如下形式:

$$\|\boldsymbol{X}-\boldsymbol{\Psi}\boldsymbol{S}\|_F^2 = \left\|\boldsymbol{X}-\sum_{i=1}^K \boldsymbol{\Psi}_j\boldsymbol{s}_T^j\right\|_F^2 = \left\|(\boldsymbol{X}-\sum_{i\neq k}^K \boldsymbol{\Psi}_j\boldsymbol{s}_T^j)-\boldsymbol{\Psi}_k\boldsymbol{s}_T^k\right\|_F^2 = \left\|\boldsymbol{E}_k-\boldsymbol{\Psi}_k\boldsymbol{s}_T^k\right\|_F^2, \tag{20}$$

其中 $\boldsymbol{E}_k=\boldsymbol{X}-\sum_{i\neq k}^K \boldsymbol{\Psi}_j\boldsymbol{s}_T^j-\boldsymbol{\Psi}_k\boldsymbol{s}_T^k$, 为了使得 $\|\boldsymbol{X}-\boldsymbol{\Psi}\boldsymbol{S}\|_F^2$ 最小, 需要 $\boldsymbol{\Psi}_k\boldsymbol{s}_T^k$ 逼近 $\boldsymbol{E}_k$, 而 $\boldsymbol{\Psi}_k\boldsymbol{s}_T^k$ 是秩为 1 的矩阵, 最直接的方法是对 $\boldsymbol{E}_k$ 进行 SVD 分解, 然而无法保证 $\boldsymbol{s}_T^k$ 的稀疏性, 为此定义 $\omega_k=\{i|0\leqslant i\leqslant N, \boldsymbol{s}_T^k(i)\neq 0\}$, 只考虑 s_T^k 非零元素, 在 E_k 中选取与 ω_k 相应的列构成 E_k^R, 即存在一个矩阵 Ω_k, 使得 $\boldsymbol{s}_k^R=\boldsymbol{s}_T^k\boldsymbol{\Omega}_k$, $\boldsymbol{E}_k^R=\boldsymbol{E_k}\boldsymbol{\Omega}_k$. 对 $\boldsymbol{E}_R$ 进行 SVD 分解: $\boldsymbol{E}_k^R=\boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}$, 用 $\boldsymbol{U}$ 的第一列 $\boldsymbol{u}_1$ 替换字典的第 k 列 $\boldsymbol{\Psi}_k$, 用 $\boldsymbol{\Delta}(1,1)$ 和 $\boldsymbol{V}$ 的第一行 $\boldsymbol{v}_T^1$ 乘积替换 $\boldsymbol{s}_k^R$. K-SVD 算法由如下两个阶段交替进行实现:

(i) 稀疏编码阶段. 对于任意样本 x_i, 计算字典 Ψ 下的稀疏表示 s_i, 可通过贪心算法求解下面的优化问题获得

$$\min_{\boldsymbol{s}_i} \|\boldsymbol{x}_i-\boldsymbol{\Psi}s_i\|_F \quad \text{subject to} \ \forall i \ \ \|\boldsymbol{s}_i\|_1 \leqslant \boldsymbol{T}_0.$$

(ii) 字典更新阶段.

For $k=1,2,\ldots,K$

计算: $\omega_k, \boldsymbol{\Omega}_k\boldsymbol{E}_k^R$

奇异值分解: $\boldsymbol{E}_k^R=\boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}$

更新子字典的第 k 列:$\boldsymbol{\Psi}_k=\boldsymbol{u}_1$

计算稀疏表示系数: $\boldsymbol{s}_k^R=\Delta(1,1)\boldsymbol{v}_T^1$, 由 $\boldsymbol{s}_k^R$ 和 $\boldsymbol{\Omega}_k$ 获得 $\boldsymbol{s}_T^k$

End for

K-SVD 是一种简单、灵活的字典训练方法, 在图像去噪方面获得较大的成功, 然而在理论上不能保证该算法是收敛的, 在实际应用中有时收敛到局部极小.

3) Sparse coding 字典. 尽管 K-SVD 中引入稀疏机制, 但没有突破 K-means 框架, 无法高效地稀疏表示图像信息, 神经科学的理论与实验研究表明: 神经系统只使用很少的 active 神经元对感知信息编码, 这种策略被称为 sparse coding. 基于稀疏编码的字典类似于视觉皮质的神经元感知场 (perceptive field of neutrons in the visual cortex), 它可以使得自然图像信号的隐含的结构清晰地展示. 文献 [30~32] 将图像表示成由一个基本函数生成的基函数的线性组合, 要求组合系数满足稀疏性, 这类字典只能近似表达基函数生成的空间的数据, 对时频复杂多变的图像信号无法准确刻画. Lee 等 [33] 提出了一种高效的 sparse coding 方法, 该方法通过对自然图像样本数据学习, 捕捉存在于数据中高层的语义特征, 从而获得图像基元. 考虑输入图像向量的训练集合 $\boldsymbol{X}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_m\}$ 和与之相应的未知的系数向量集合 $\boldsymbol{S}=\{\boldsymbol{s}_1,\boldsymbol{s}_i,\ldots,\boldsymbol{s}_m\}$, 对于字典 $\boldsymbol{\Psi},\boldsymbol{S}$ 的最大后验估计是求解如下优化问题

$$\begin{cases}\displaystyle\min_{\{\boldsymbol{x}_i\},\{\boldsymbol{s}_i\}}\sum_i\frac{1}{2\sigma^2}\left\|\boldsymbol{x}_i-\sum_j\boldsymbol{\Psi}_j\boldsymbol{s}_i^j\right\|_2^2+\beta\sum_i\sum_j\varphi(\boldsymbol{s}_i^j)\\ \text{subject to }\|\boldsymbol{\Psi}_j\|^2\leqslant c,\ \forall j=1,2,\ldots,n.\end{cases}\tag{21}$$

稀疏约束项 $\varphi(\cdot)$ 用来度量系数的稀疏性, 通常采用 l_1 范数描述. c 是一个对字典向量能量约束的常数, 通过求解上述优化问题可以获得稀疏字典, 该问题是在两类变量集合上优化求解, 数值求解的代价是巨大的, Lee 等将问题 (21) 归结为交替求解如下两个凸优化问题.

凸优化问题 A. 对于给定的字典 $\boldsymbol{\Psi}$, 求解稀疏表示系数 $\{\boldsymbol{s}_1,\boldsymbol{s}_2,\ldots,\boldsymbol{s}_m\}$

$$\min_{\{\boldsymbol{s}_i\}}\sum_i\left\|\boldsymbol{x}_i-\sum_j\boldsymbol{\Psi}_j\boldsymbol{s}_i\right\|_2^2+2\sigma^2\beta\sum_j\left|s_i^j\right|.\tag{22}$$

每次单独优化求解 $\boldsymbol{s}_i$, 对于问题 A, 采用 Feature-sign 搜索算法.

设稀疏表示系数为 $\boldsymbol{S}=\{\boldsymbol{s}_1,\boldsymbol{s}_i,\ldots,\boldsymbol{s}_m\}$, 问题 (21) 的等价形式为

$$\begin{cases}\displaystyle\min_{\boldsymbol{\Psi},\boldsymbol{S}}\frac{1}{2\sigma^2}\|\boldsymbol{X}-\boldsymbol{\Psi}\boldsymbol{S}\|_F^2+\beta\sum_{i,j}\varphi(s_{ij})\\ \displaystyle\text{subject to }\sum_i\boldsymbol{\Psi}_{i,j}^2\leqslant c,\qquad\forall j=1,\ldots,n,\end{cases}\tag{23}$$

其中 $\boldsymbol{x},\boldsymbol{\Psi}$ 和 $\boldsymbol{S}=\{s_{ij}\}$ 分别是输入数据矩阵、字典矩阵和稀疏表示矩阵.

凸优化问题 B. 对于任意稀疏约束, 求解字典 Ψ.

$$\begin{cases}\displaystyle\min_{\boldsymbol{\Psi}}\frac{1}{2\sigma^2}\|\boldsymbol{X}-\boldsymbol{\Psi}\boldsymbol{S}\|_F^2\\ \displaystyle\text{subject to }\sum_i\boldsymbol{\Psi}_{i,j}^2\leqslant c,\qquad\forall j=1,\ldots,n,\end{cases}\tag{24}$$

凸优化问题 B 可以利用 Lagrange-dual 方法求解.

文献 [34,35] 在图像的超分辨重建中引入 sparse coding, 较大幅度地提升了图像的超分辨率重建性能.

4 图像视频信号的稀疏重建模型的优化求解算法

在图像和视频应用中, 压缩感知稀疏重建模型是多样的, 既有简单的基本模型也有较为复杂的模型. 复杂模型的求解可以借鉴基本模型求解方法, 许多学者针对一些基本模型提出了许多高效的算法, 为此我们以全变分和 L_1 范数两类基本模型为例, 从约束优化问题和无约束优化问题两方面介绍一些在图像视频应用中通用的算法, 我们首先针对如下两种有约束优化问题介绍一些优化求解算法:

$$\begin{cases} \mathbf{BP}_{\varepsilon} \quad \min \ \|\boldsymbol{x}\|_1 \\ \text{subject to} \ \ \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2 \leqslant \varepsilon, \end{cases} \tag{25}$$

$$\begin{cases} \mathbf{TV}_{\varepsilon} \ \min \ \|\mathrm{TV}(\boldsymbol{x})\|_1 \\ \text{subject to} \ \ \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2 \leqslant \varepsilon. \end{cases} \tag{26}$$

然后, 讨论如下两种情形的无约束问题的解法:

$$\mathbf{QP}_{\lambda} \quad \min \ \lambda \|\boldsymbol{x}\|_1 + \frac{1}{2} \ \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2, \tag{27}$$

$$\mathbf{TV}_{\lambda} \ \min \ \lambda \mathrm{TV}(\boldsymbol{x}) + \ \frac{1}{2} \ \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2. \tag{28}$$

为了便于后面算法的描述, 在论文的后半部分, 优化变量不论是时域还是变换域, 都统一采用 $\boldsymbol{x}$ 表示, 同理模型中的矩阵用 $\boldsymbol{A}$ 表示.

4.1 有约束优化问题

4.1.1 贪心算法

贪心算法包括 matching pursuit (MP) [36~39], orthogonal matching pursuit (OMP) [40], stagewise OMP (StOMP) [41] 和 CoSaMP[42]. 该类算方法的基本思想是通过迭代的方式寻找信号 x 的支集 Ω 其中, 对每一次迭代, 按照贪婪的机制寻找一个或者若干个支集成员. 以 MP 算法为例, 简单地介绍贪心算法. MP 算法就是通过迭代的方式寻找观测信号 $\boldsymbol{b}$ 在已知的向量组 (或字典) 的稀疏的线性组合 $\sum_{i=1}^{N} s_i \boldsymbol{\alpha}_i$, $s_1, s_2, \ldots, s_N$ 是组合系数, 即选择 $s_1, s_2, \ldots, s_N$ 中 M 个非零的系数 $s_{\gamma_i} \neq 0, i = 0, 1, 2, \ldots, M-1$, $\gamma_i \in \Omega \subset \{1, 2, \ldots, N\}$ 使得 $\boldsymbol{b} \approx \sum_{i=0}^{M-1} s_{\gamma_i} \boldsymbol{\alpha}_{\gamma_i}$. 每一次迭代, 在字典 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_N$ 中, 选择与当前残差向量投影最大的 $\boldsymbol{\alpha}_{\gamma_n}$, 迭代公式如下:

$$\begin{cases} \boldsymbol{R}^0 = \boldsymbol{b}, \\ \boldsymbol{R}^n = \langle \boldsymbol{R}^n, \boldsymbol{\alpha}_{\gamma_n} \rangle \boldsymbol{\alpha}_{\gamma_n} + \boldsymbol{R}^{n+1}, \\ \boldsymbol{\alpha}_{\gamma_n} = \arg\max\limits_{\gamma_i} \left| \langle \boldsymbol{R}^n, \boldsymbol{\alpha}_{\gamma_i} \rangle \right|, \\ s_{\gamma_n} = \langle \boldsymbol{R}^n, \boldsymbol{\alpha}_{\gamma_n} \rangle, \end{cases} \tag{29}$$

其中 $\boldsymbol{R}^n$ 是 n 次迭代的残差向量. 而 OMP 是 MP 方法的直接拓展, 主要差别是所求的稀疏表示系数是稀疏信号在字典上的正交投影, StOMP 和 CoSaMP 是每次考虑多个最优的投影. 该类算法复杂度较低, 在低水平噪声下是鲁棒的; 在某些情况下, 该类算法的收敛性很难保证.

4.1.2 二阶锥规划

当观测是无损时, 原重建问题是等式约束优化问题, 很容易转化为 LP 线性规划问题, 利用 Primal-dual 算法求解. 然而通常观测是有噪的, 如 $\mathrm{BP}_{\boldsymbol{\varepsilon}}$ 和 $\mathrm{TV}_{\varepsilon}$ 问题, 对应于不等式优化约束, 可以转化为二阶锥规划 (second-order cone programs, SOCP) 问题求解 [43,44]. 首先回忆二阶锥规划 (SOCP) 问题的标准形式

$$\begin{cases} \min\limits_{\boldsymbol{z}} \langle \boldsymbol{c}_0, \boldsymbol{z} \rangle \\ \text{subject to } \boldsymbol{A}_0 \boldsymbol{z} = \boldsymbol{b}_0, \\ \qquad\qquad f_i(\boldsymbol{z}) \leqslant 0, \quad i = 1, 2, \ldots, m, \end{cases} \tag{30}$$

这里 $f_i(\boldsymbol{z})$ 是线性约束 $f_i(\boldsymbol{z}) = \langle \boldsymbol{c}_i, \boldsymbol{z} \rangle + \boldsymbol{d}_i$ 或者是二阶锥 $f_i(\boldsymbol{z}) = \frac{1}{2}(\|\boldsymbol{A}_i \boldsymbol{z}\|_2^2 - (\langle \boldsymbol{c}_i, \boldsymbol{z} \rangle - \boldsymbol{d}_i))$. $\boldsymbol{A}_i$ 是矩阵, $\boldsymbol{c}_i$ 是向量, $\boldsymbol{d}_i$ 是标量.

1) $\mathbf{BP}_{\boldsymbol{\varepsilon}}$ 问题求解.

$$\begin{cases} \min\limits_{\boldsymbol{x}, \boldsymbol{u}} \sum\limits_m \boldsymbol{u}_m, \\ \text{suject to } \boldsymbol{x} - \boldsymbol{u} \leqslant 0, \\ \qquad\qquad -\boldsymbol{x} - \boldsymbol{u} \leqslant 0, \\ \qquad\qquad \dfrac{1}{2}(\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})\|_2^2 - \varepsilon) \leqslant 0. \end{cases} \tag{31}$$

通过 $\boldsymbol{f}_{u_1} = \boldsymbol{x} - \boldsymbol{u}$, $\boldsymbol{f}_{u_2} = -\boldsymbol{x} - \boldsymbol{u}$, $\boldsymbol{f}_{\varepsilon} = \frac{1}{2}(\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})\|_2^2 - \varepsilon)$, 这样问题 $\mathbf{BP}_{\varepsilon}$ 就转化为 SOCP 问题求解.

2) $\mathbf{TV}_{\varepsilon}$ 问题.

$$\begin{cases} \min\limits_{\boldsymbol{x}, \boldsymbol{t}} \sum\limits_{ij} t_{i,j} \\ \text{subject to } \|D_{ij}\boldsymbol{x}\|_2 \leqslant t_{i,j}, i, j = 1, \ldots, n, \\ \qquad\qquad \|\boldsymbol{\varPhi}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \leqslant \varepsilon. \end{cases} \tag{32}$$

此时定义 $f_{ij} = \frac{1}{2}(\|D_{ij}\boldsymbol{x}\|_2^2 - t_{i,j}^2)$ 和 $f_{\varepsilon} = \frac{1}{2}(\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 - \varepsilon^2)$, 这里 $\boldsymbol{t}$ 是由 $t_{i,j}$ 构成向量, 这样 $\mathbf{BP}_{\boldsymbol{\varepsilon}}$ 问题可以由 SOCP 问题求解. 对于**SOCP**问题可以采用 Primal-dual 方法和 Log-barrier 方法, 而采用 Log-barrier 算法比 Primal-dual 算法更直接, Log-barrier 算法是通过一系列牛顿法来实现, 精度高, 但收敛速度较慢, 不适于大尺度问题.

4.2 无约束优化问题

4.2.1 固定点连续方法

固定点连续方法 (fixed point continuation, FPC) 是求解 $\mathbf{QP}_{\boldsymbol{\lambda}}$ 问题的线性收敛方法 [45], 主要的思想是采用算子分裂技术将问题转化为求解固定点方程 $\boldsymbol{x} = F(\boldsymbol{x})$, F 是算子, 通过迭代 $\boldsymbol{x}^k = F(\boldsymbol{x}^{k-1})$, 得到该固定点方程的近似解. 如何针对具体问题构造固定点方程是关键. 由凸分析可知: 极小化一个凸泛函 $\varphi(\boldsymbol{x})$, 即寻找一个 $\boldsymbol{x}$, 使得 $0 \in \partial\varphi(\boldsymbol{x})$, $\partial\varphi(\boldsymbol{x})$ 是 $\varphi(\boldsymbol{x})$ 的次梯度, 定义 $T(\boldsymbol{x}) = \partial\varphi(\boldsymbol{x})$, 采用前后向算子分裂法将算子 T 分裂为

$$T = T_1 + T_2. \tag{33}$$

当算子 T_1 和 T_2 是单调, T_1 是单值, $I+\lambda T_2$ 是可逆的, 可导出了一个固定点方程 $\boldsymbol{x}=(I+t_kT_1)^{-1}(I-t_kT_2)\boldsymbol{x}$, 那么, 该不动点方程的近似解可由如下迭代获得:

$$\boldsymbol{x}^k=(I+t_kT_1)^{-1}(I+t_kT_2)\boldsymbol{x}^{k-1}. \tag{34}$$

一定存在 $\mu>0$, 使得 $\mathbf{QP}_{\boldsymbol{\lambda}}$ 问题与

$$\min_{\boldsymbol{x}}\|\boldsymbol{x}\|_1+\frac{\mu}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2 \tag{35}$$

等价. 对于极小化问题 (35), $T_1(\boldsymbol{x})=\partial(\|\boldsymbol{x}\|_1)$ 和 $T_2(\boldsymbol{x})=\nabla(\frac{\mu}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2)$, $(I+t_kT_1)^{-1}$ 是软阈值算子. 由于

$$\partial\Big(\frac{1}{\mu}\|\boldsymbol{x}\|_1\Big)=\frac{1}{\mu}\mathrm{SGN}(\boldsymbol{x}) \tag{36}$$

$$\mathrm{SGN}(\boldsymbol{x})_i:=\begin{cases}1, & x_i>0,\\ \in[-1,1], & x_i=0\\ -1, & x_i<0.\end{cases} \tag{37}$$

于是获得了 $\mathbf{QP}_\lambda$ 问题的固定点迭代算法

$$\boldsymbol{x}^k=\mathrm{sgn}\{\boldsymbol{x}^{k-1}-t_kT_2(\boldsymbol{x}^{k-1})\}\odot\max\Big\{\big|\boldsymbol{x}^{k-1}-t_kT_2(\boldsymbol{x}^{k-1})\big|-\frac{t_k}{\mu},0\Big\}, \tag{38}$$

其中 $\odot$ 表示向量间运算, 其含义为 $(\boldsymbol{x}\odot\boldsymbol{y})_i=\boldsymbol{x}_i\boldsymbol{y}_i$, $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^N$.

在实际中参数 μ 选择对收敛速度都有很大影响, 为此对 μ 采用连续策略 $0<\mu_1<\mu_k<\cdots<\mu_L$, 这样式 (10) 就改写成如下形式:

$$\boldsymbol{x}^k=\mathrm{sgn}\{\boldsymbol{x}^{k-1}-t_kT_2(\boldsymbol{x}^{k-1})\}\odot\max\Big\{\big|\boldsymbol{x}^{k-1}-t_kT_2(\boldsymbol{x}^{k-1})\big|-\frac{t_k}{\mu_k},0\Big\}. \tag{39}$$

FPC 类方法是一阶收敛方法该类算法, 计算代价主要集中在 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的运算上, 算法复杂度较低, 适合求解大尺度问题.

4.2.2 FPC active set 算法

事实上, 在 FPC 迭代求解中, 一定存在一个 $\bar{k}$, 当 $k\geqslant\bar{k}$ 时, 使得 x^k 与真解的符号相同: $\mathrm{sgn}(\boldsymbol{x}^k)=\mathrm{sgn}(\boldsymbol{x}^*)$, 另一方面, $|x_i|=\mathrm{sgn}(x_i)x_i$, $\|\boldsymbol{x}\|_1$ 可以一个较好光滑性的线性函数 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 逼近, $\boldsymbol{c}$ 就是 $\boldsymbol{x}_i$ 符号构成, 这一良好特性有助于在求解过程中充分利用了解的结构信息, 提高解的精度和收敛速度. 基于此, 文献 [46] 提出了 FPC active set (FPC-AS) 算法, 该算法是由两个阶段交替实现的, 同时参数 μ 采用了与 FPC 同样的连续策略.

第 1 阶段采用线性搜索技术生成了一个近似解:

$$\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+\alpha_k\boldsymbol{d}^k,\quad \boldsymbol{d}^k=S(\boldsymbol{x}^k-\lambda\boldsymbol{g}^k,\mu\lambda)-\boldsymbol{x}^k,$$

$\boldsymbol{d}^k$ 是线性搜索方向, $\boldsymbol{g}=\nabla\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|$, $S(y,\nu)=\mathrm{sgn}(y)\odot\max\{|y|-\nu,0\}$.

第 2 阶段定义一个 active set $T^{(k)}$

$$T^{(k)}=\{i\in\{1,2,\ldots,n\},\big|\boldsymbol{x}_i^k\big|=0\}.$$

原问题就转化为求解光滑目标函数的优化问题:

$$\min_{\boldsymbol{x}} \mu \|\boldsymbol{x}\|_1 + \|\mathbf{A}\mathbf{x} - \boldsymbol{b}\|_2^2 \quad \text{subject to } \boldsymbol{x} \in \boldsymbol{\Omega}(\boldsymbol{x}^k), \tag{40}$$

$\Omega(\boldsymbol{x}^k) := \{\boldsymbol{x} \in \mathbb{R}^n : \text{sgn}(\boldsymbol{x}_i)\boldsymbol{x}_i \geqslant 0, i \in \bar{T}^k; \boldsymbol{x}_i = 0, i \in T^k\}$, $\bar{T}^{(k)} = \{i \in \{1, 2, \ldots, n\}, |\boldsymbol{x}_i^k| > 0\}$. 利用成熟优化方法如共轭梯度法和拟牛顿法可以有效求解问题 (40).

在严格的稀疏定义下, 图像信号是不稀疏的, 而是动态范围较大的近似稀疏信号. FPC-AS 算法应用 active set, 可以有效地重建图像信号中幅值较小但是对于表达 b 贡献较大的非零成分. 与 FPC-AS 算法的前驱 FPC 相比, 不仅在收敛速度方面表现较大的优势, 而且在重建稀疏信号的能力方面更为突出.

4.2.3　Bergman 算法

Bergman 迭代方法处理的原问题为

$$\min_{\boldsymbol{x}} \lambda J(\boldsymbol{x}) + \frac{1}{2} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|, \tag{41}$$

$J(\boldsymbol{x})$ 是正则化项, Bergman 迭代公式为

$$\begin{cases} \boldsymbol{x}^{k+1} = \min \lambda J(\boldsymbol{x}) + \dfrac{1}{2} \left\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^k\right\|_2^2, \\ \boldsymbol{b}^{k+1} = \boldsymbol{b}^k + (\boldsymbol{b}^k - \boldsymbol{A}\boldsymbol{x}^{k+1}). \end{cases} \tag{42}$$

此类问题适于求解无约束优化问题, 当 $J(\boldsymbol{x}) = \|\boldsymbol{x}\|_1$, Bergman 迭代法可以求解 $\mathbf{QP}_\lambda$ 问题, 当 $J(\boldsymbol{x}) = \text{TV}(\boldsymbol{x})$, Bergman 迭代可以求解 $\mathbf{TV}_\lambda$ 问题. Bergman 迭代方法所需要外部迭代次数很少, 每次内部迭代需要求解一个子问题, 并且子问题 (42) 与原问题 (41) 具有相同的形式, 对应同样的参数 λ, 该子问题是一个 $\mathbf{QP}_\lambda$ 问题, 通常采用 Barzilai-Borwein 版本的 FPC 求解, 该子问题的求解代价较大 [47~50].

4.2.4　NESTA 算法

Nesterov 将光滑化技术与改进的梯度方法结合推导出了一个一阶的优化求解算法 ——Nesterov 算法 [51], 该算法是近 20 年中优化算法中的收敛性能最优的算法. 最近在信号和图像处理领域出现一些改进的算法 [52]. 文献 [53] 将 Nesterov 算法用于 $\mathbf{BP}_\varepsilon$ 问题的求解, 并构建了一个新的算法 ——NESTA (shorthand for nesterov's algorithm), 该算法是对 Nesterov 算法的推广. 首先回忆 Nesterov 算法, Nesterov 算法适用于求解凸集 Q_p 上光滑的凸泛函极小化问题

$$\min_{\boldsymbol{x} \in Q_p} f(\boldsymbol{x}). \tag{43}$$

Nesterov 算法通过如下迭代方式实现:

$$\boldsymbol{y}_k = \arg\min_{\boldsymbol{x} \in Qp} \frac{L}{2} \|\boldsymbol{x} - \boldsymbol{x}_k\|_2^2 + \langle \nabla f(\boldsymbol{x}_k), \boldsymbol{x} - \boldsymbol{x}_k \rangle, \tag{44}$$

$$\boldsymbol{z}_k = \arg\min_{\boldsymbol{x} \in Qp} \frac{L}{\sigma_p} p_p(\boldsymbol{x}) + \sum_{i=0}^{K} \alpha_i \langle \nabla f(\boldsymbol{x}_i), \boldsymbol{x} - \boldsymbol{x}_i \rangle, \tag{45}$$

$$\boldsymbol{x}_k = \boldsymbol{\theta}_k \mathbf{z}_k + (1 - \boldsymbol{\theta}_k) \boldsymbol{y}_k, \tag{46}$$

这里 $p_p(\boldsymbol{u})$ 是一个 dual prox-function 函数, L_μ 是 $f_\mu(\boldsymbol{x}_k)$ 的 Lipschitz 常数, $\{\alpha_i\}_{i=1}^K$ 是常数, σ_p 是与 $p_p(\boldsymbol{u})$ 有关的参数, $\boldsymbol{x}_k$ 是第 k 次迭代的结果.

l_1 范数的光滑性较差, $\|\boldsymbol{x}\|_1$ 可以写成 $\max_{\boldsymbol{u}\in Q_p}\langle \boldsymbol{u},\boldsymbol{x}\rangle$ 形式, $Q_p=\{\boldsymbol{u}:\|\boldsymbol{u}\|_\infty\leqslant 1\}$, 对于 $\|\boldsymbol{x}\|_1$ 的光滑逼近为

$$f_\mu(\boldsymbol{x})=\max_{\boldsymbol{u}\in Q_p}\langle \boldsymbol{u},\boldsymbol{x}\rangle-\mu p_p(\boldsymbol{u}), \tag{47}$$

$p_p(\boldsymbol{u})$ 是对于 Q_p 的 dual prox-function 函数. 这样我们将 $\mathbf{BP}_\varepsilon$ 问题转化为求解一个光滑约束问题

$$\min_{\mu\in Q_p} f_\mu(\boldsymbol{x}), \tag{48}$$

这里 $Q_p=\{\boldsymbol{x}:\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2\leqslant\varepsilon\}$, 可以采用 Nesterov 算法求解. 对于目标函数 $f_\mu(\boldsymbol{x})$, 此时 (44) 与 (45) 式对应于如下的 (49) 和 (50) 式:

$$\boldsymbol{y}_k=\arg\min_{x\in Qp}\frac{L_\mu}{2}\|\boldsymbol{x}-\boldsymbol{x}_k\|_2^2+\langle\nabla f_\mu(\boldsymbol{x}_k),\boldsymbol{x}-\boldsymbol{x}_k\rangle, \tag{49}$$

$$\boldsymbol{z}_k=\arg\min_{x\in Qp}\frac{L_\mu}{\sigma_p}p_p(\boldsymbol{x})+\sum_{i=0}^K\alpha_i\langle\nabla f_\mu(\boldsymbol{x}_i),\boldsymbol{x}-\boldsymbol{x}_i\rangle, \tag{50}$$

其中 L_μ 是 $f_\mu(\boldsymbol{x}_k)$ 的 Lipschitz 常数, 利用 Lagrange 法求解式 (49) 和式 (50), 可以获得 $\boldsymbol{y}_k$ 和 $\boldsymbol{z}_k$, 从而获得了 NESTA 算法的迭代公式:

$$\boldsymbol{y}_k=\Big(\boldsymbol{I}-\frac{\lambda_1}{L_\mu}\boldsymbol{A}^*\boldsymbol{A}\Big)\Big(\frac{\lambda_1}{L_\mu}\boldsymbol{A}^*\boldsymbol{b}+\boldsymbol{x}_k-\frac{1}{L_\mu}\nabla f_\mu(\boldsymbol{x}_k)\Big), \tag{51}$$

$$\boldsymbol{z}_k=\Big(\boldsymbol{I}-\frac{\lambda_2}{\lambda_2+L_\mu}\boldsymbol{A}^*\boldsymbol{A}\Big)\Big(\frac{\lambda}{L_\mu}\boldsymbol{A}^*\boldsymbol{b}+\boldsymbol{x}_0-\frac{1}{L_\mu}\sum_{i=0}^K\alpha_i\nabla f_\mu(\boldsymbol{x}_i)\Big), \tag{52}$$

$$\boldsymbol{x}_k=\theta_k\boldsymbol{z}_k+(1-\theta_k)\boldsymbol{y}_k, \tag{53}$$

其中

$$\lambda_1=\max\{0,\varepsilon^{-1}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{q}_1\|_2-L_\mu\},\quad \boldsymbol{q}_1=\boldsymbol{x}_k-L_\mu^{-1}\nabla f_\mu(\boldsymbol{x}_k), \tag{54}$$

$$\lambda_2=\max\{0,\varepsilon^{-1}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{q}_2\|_2-L_\mu\},\quad \boldsymbol{q}_2=\boldsymbol{x}_0-L_\mu^{-1}\sum_{i=0}^K\alpha_i\nabla f_\mu(\boldsymbol{x}_i). \tag{55}$$

事实上, NESTA 算法是一类求解不完整观测问题的有效算法, 有广阔应用前景; 并且在性能不依赖于优化中控制参数的调整的前提下, 这类方法具有快速、高精度和高鲁棒性的特点.

5 存在的问题及挑战

图像压缩采样与稀疏表示及数值求解是图像视频信号的压缩感知理论的重要问题, 国内外学者已经针对上述问题, 开展了大量、系统和深入的研究, 并取得了一些丰富的理论研究成果, 相关的理论在图像视频应用领域取得了初步的应用. 然而在图像视频信号的观测设计, 稀疏表示和数值算法等方面还存在一些问题和挑战.

1) 最优观测设计. 随机分布的观测矩阵因其具有与任意稀疏化矩阵非一致的特性, 因此被认为是一个万能的观测矩阵, 随机观测与采样图像视频信号无关, 因此通常不是最优观测矩阵, 如何利用图像视频的先验信息, 设计最优观测矩阵以最大化采样信息量是图像视频压缩采样的难点问题.

2) 图像视频信号的分组稀疏表示. 图像视频信号在某些变换下不仅近似稀疏, 而且变换系数之间还存在相互协同关系, 稀疏系数特别是非零系数之间还存在较强的相关性, 最新出现的 Group 稀疏 (group sparse) 是通过对稀疏系数分组来刻画这种系数之间的依赖关系, 与传统的稀疏 (flat sparse) 相比, 分组稀疏在一定程度上改善了图像视频信号的重建质量 [54]. 如何利用图像视频信号的分组稀疏特性, 捕捉信号中隐含的丰富结构和细节信息, 是提高图像视频信号重建性能的关键问题.

3) 大尺度高动态范围的稀疏解的数值求解. 图像视频信号是近似稀疏的, 其数据是高动态范围分布的, 在某个变换域的系数包含了许多对图像重建质量起到不可忽略作用的低幅度成分, 大多数算法只能找到幅值较大的成分, 高动态范围的稀疏解的数值解法还处在起步阶段. 另一方面, 随着信息科学的发展, 人们对图像视频的分辨率提出了更高的要求, 大尺度的稀疏求解方法的研究受到了广泛关注. 因此, 研究大尺度高动态范围的稀疏解的数值求解是压缩感知应用面临的一个重大挑战.

4) 图像视频信号的稀疏 – 低秩联合表示. 2008 年 Candes 等 [55,56] 在压缩感知框架下提出了低秩矩阵理论, 该理论是对现有压缩感知理论的发展. 低秩矩阵理论表明: 一个具有低秩特性的矩阵可以由其不完整或有损的信息通过求解最小秩优化问题精确重构. 图像视频信号该理论一经提出, 在图像视频信号等方面获得了应用 [57,58]. 稀疏表示与低秩表示是从不同的角度利用数据冗余性对图像视频建模的方法. 稀疏表示与低秩表示的内在关系和图像视频的稀疏 – 低秩联合表示方法是目前图像视频应用方面面临的一个挑战.

参考文献

1 Candes E, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans Inf Theory, 2006, 52: 489–509

2 Donoho D. Compressed sensing. IEEE Trans Inf Theory, 2006, 52: 1289–1306

3 Candes E, Tao T. Near optimal signal recovery from random projections: Universal encoding strategies? IEEE Trans Inf Theory, 2006, 52: 5406–5425

4 Candes E, Romberg J. Sparsity and incoherence in compressive sampling. Inverse Probl, 2007, 23: 969–985

5 Baraniuk R. Compressive sensing. IEEE Signal Process Mag, 2007, 24: 118–121

6 Baraniuk R, Davenport M, DeVore R, et al. A simple proof of the restricted isometry property for random matrices. Constr Approx, 2008, 28: 253–263

7 Boufounos P, Baraniuk R. 1-Bit compressive sensing. In: Proc. of CISS, Princeton, 2008. 16–21

8 Wright J, Yang A, Ganesh A, et al. Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell, 2009, 31: 210–227

9 Wright J, Ma Y. Dense error correction via l1-Minimization. IEEE Trans Inf Theory, 2009, 57: 3540–3560

10 Wan T, Canagarajah N, Achim A. Compressive image fusion. In: Proc. of ICIP, San Diego, 2008. 1308–1311

11 Lustig M, Donoho D, Pauly J. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn Resonance Med, 2007, 58: 1180–1195

12 Block K, Uecker M, Frahm J. Undersampled radial MRI with multiple coils Iterative image reconstruction using a total variation constraint. Magnc Resonance Med, 2007, 57: 1086–1098

13 Zhang Y, Mei S, Chen Q, et al. A novel image/video coding method based on compressed sensing theory. In: Proc. of ICASSP, Las Vegas, 2008. 1361–1364

14 Kang L W, Lu C S. Distributed compressive video sensing. In: Proc of ICASSP, Taiwan, 2009. 1169–1172

15 Zhang Y, Mei S, Chen Q, et al. A multiple description image/video coding method by compressed sensing theory. In: Proc of ISCAS, Seattle, 2008. 1830–1833

16 Stankovic V, Stankovic L, Cheng S. Compressive image sampling with side information. In: Proc. of ICIP, Cairo, 2009. 3037–3040

17 Ma S, Yin W, Zhang Y, et al. An efficient algorithm for compressed MR imaging using total variation and wavelets. In: Proc. of CVPR, Minneapolis, 2008. 1–8
18 Bajwa W, Haupt J , Raz G , et al. Toeplitz-structured compressed sensing matrices. In: IEEE Workshop on Statistical Signal Processing, Madison, 2007. 294–298
19 Yin W, Morgan S, Yang J F, et al. Practical Compressive Sensing with Toeplitz and Circulant Matrices. Technical Report CAAM. Rice University, 2010
20 Sebert F, Zou Y, Ying L. Toeplitz block matrices in compressed sensing and their applications in imaging. In: Proc. of International Conference on Information Technology and Application in Biomedicine, Cairns, 2008. 47–50
21 Liang D, Xu G, Wang H, et al. Toeplitz random encoding MR imaging using compressed sensing. In: Proc. of ISBI, Boston, 2009. 270–273
22 Do T T, Tran T D, Gan L. Fast compressive sampling with structurally random matrices. In: Proc. of ICASSP, Las Vegas, 2008. 3369–3372
23 Gan L, Do T T, Tran T D. Fast compressive imaging using scrambled hadamard ensemble. In: Proc. of EUSIPCO, Lausanne, 2008
24 Wu X, Dong W, Zhang X, et al. Model-guided adaptive recovery of compressive sensing. In: Proc. of DCC, Snowbird, 2009. 123–132
25 Yang J, Zhang Y, Yin W. A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data. IEEE J Sel Top Signal Process, 2010, 4: 288–297
26 Guo W, Yin W. EdgeCS: An edge guided compressive sensing reconstruction. In: Proc. of VCIP, Huangshan, 2010
27 Liang L, Xie X, Shi G. Nonuniform directional filter banks with arbitrary frequency partitioning. IEEE Trans Image Process, 2011, 20: 283–288
28 Wang J, Shi Y, Kong D, et al. Sparse representation based down-sampling image compression. J Comput Appl Math, 2011, 236: 675–683
29 Aharon M, Elad M, Bruckstein A. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation. IEEE Trans Signal Process, 2006, 54: 4311–4322
30 Olshausen B, Field D. Emergence of simple-cell receptive field properties by a sparse coding for natural images. Nature, 1996, 381: 607–609
31 Olshausen B, Field D. Sparse coding with an overcomplete basis set : A strategy employed by V1. Vision Res, 1997, 37: 3311–3325
32 Olshausen B, Field D. Sparse coding of sensor inputs. Current Opin Neurobiol, 2004, 14: 481–487
33 Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms. Adv Neural Inf Process Syst, 2007, 19: 801–802
34 Yang J, Wright J, Huang T, et al. Image super-resolution via sparse representation. IEEE Trans Image Process, 2010, 19: 2861–2873
35 Turek J, Yavneh I, Elad M. On MMSE and MAP denoising under sparse representation modeling over a unitary dictionary. IEEE Trans Signal Process, 2011, 59: 3526–3535
36 Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Process, 1993, 41: 3397–3415
37 Bergeaud F, Mallat S. Matching pursuit of images. In: Proc ICIP, Arlington, 1995. 53–56
38 Neff R, Zakhor A. Very low bit-rate video coding based on matching pursuits. IEEE Trans Circuits SystVideo Technol, 1997, 7: 158–171
39 Mendels F, Vandergheynst P, Thiran J. Matching pursuit-based shape representation and recognition using scale-space. Int J Imaging Syst Technol, 2006, 16: 162–180
40 Tropp J, Gilbert A. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory, 2007, 53: 4655–4666
41 Donoho D, Tsaig Y, Drori I, et al. Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit. IEEE Trans Inf Theory, 2012 58: 1094–1121
42 Needell D, Tropp J. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. Appl Comput Harmonic Anal, 2009, 26: 301–321
43 Alizadeh F, Goldfarb D. Second-order cone programming. Math Program, 2003, 95: 3–51

44 Goldfarb D, Yin W. Second-order cone programming methods for total variation based image restoration. SIAM J Sci Comput, 2005, 27: 622–645

45 Hale E, Yin W, Zhang Y. Fixed-point continuation for l1-minimization: Methodology and convergence. SIAM J Optim, 2008, 19: 1107–1130

46 Wen Z, Yin W, Goldfarb D, et al. A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization and continuation. SIAM J Sci Comput, 2010, 32: 1832–1857

47 Osher S, Burger M, Goldfarb D, et al. An iterated regularization method for total variation-based image restoration. Multiscale Model Simul, 2005, 4: 460–489

48 Xu J, Osher S. Iterative regularization and nonlinear inverse scale space applied to wavelet-based denoising. IEEE Trans Image Process, 2007, 16: 534–544

49 Rudin L, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. Phys D: Nonlinear Phenom, 1992, 60: 259–268

50 Yin W, Osher S, Goldfarb D, et al. Bregman iterative algorithms for l1 minimization with applications to compressed sensing. SIAM J Imaging Sci, 2008, 1: 143–168

51 Nesterov Y. Smooth minimization of non-smooth functions. Math Program Ser A, 2005, 103: 127–152

52 Beck A, Teboulle M. Fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci, 2008, 2: 183–202

53 Becker S, Bobin J, Candes E. NESTA: A fast and accurate first-order method for sparse recovery. SIAM J Imaging Sci, 2011, 4: 1–39

54 Rao N, Nowak R, Wright S. Convex approaches to model wavelet sparsity patterns. In: Proc. of ICIP, Brussels, 2011. 1917–1920

55 Candes E, Recht B. Exact matrix completion via convex optimization. Found Comput Math, 2009, 9: 717–772

56 Cai J, Candes E, Shen Z. A singular value thresholding algorithm for matrix completion. SIAM J Optim, 2008, 20: 1956–1982

57 Candes E, Li X, Ma Y, et al. Robust principal component analysis? 2009. http://arxiv.org/abs/0912.3599

58 Ji H, Liu C, Shen Z, et al. Robust video denoising using low rank matrix completion. In: Proc. of CVPR, San Francisco, 2010. 1791–1798

Compressive sampling and sparse reconstruction of images/videos

YIN BaoCai, SHI YunHui*, DING WenPeng, HU YongLi & LI JingHua

Key Laboratory of Multimedia and Intelligent Software Technology, College of Computer Science and Technology, Beijing University of Technology, Beijing 100124*, China*
*E-mail: syhzm@bjut.edu.cn

Abstract Vision sensors usually do not account for the physical process of imaging and they acquire image/video samples at the Nyquist rate. The Nyquist rate is significantly higher than the effective dimensions of an image/video, and consequently compression is essential for the image/video prior to storage or transmission. The emerging Compressive Sensing (CS) theory states that a signal can be perfectly reconstructed, or can be robustly approximated in the presence of noise, using a few random measurements, provided that it is sparse in some linear transform domain. CS is the theoretical foundation for capturing a signal with effective information dimensions, and thus represents an unprecedented breakthrough in many fields such as sampling, processing, and recognition of image/video. We review the fundamental problems of CS for image/video including compressive sampling, sparse reconstruction models, and algorithms for the models. For compressive sampling, the construction of random and structural measurement matrices are considered separately and the performance of these two kinds of matrices is evaluated. For sparse reconstruction, models are classified as analysis-based or synthesis-based reconstruction models by the sparse representation prior, features of which are presented. The optimization models can be considered as constrained and unconstrained optimization problems. Some feasible algorithms for these two kinds of optimization problems are explained in detail and the performance of the algorithms is given. In addition, several challenges of compressive sensing technology are presented and future work is discussed.

Keywords sample, compressing sensing, sparse representation, random measurement, optimization

YIN BaoCai received his Ph.D. degree from Dalian University of Technology in 1993. He is a Professor and doctoral supervisor at the College of Computer Science and Technology, Beijing University of Technology. He is a member of the China Computer Federation. His research interests include multimedia, multifunctional perception, virtual reality, and computer graphics.

SHI YunHui received an M.S. degree in computational mathematics and a Ph.D. degree in navigational guidance and control from the Harbin Institute of Technology, in 1994 and 2002, respectively. She is a Professor at the College of Computer Science and Technology, Beijing University of Technology. Her research interests include compressive sensing, multimedia compression, and image/video processing.

DING WenPeng received his Ph.D. degree from the University of Science and Technology in 2009. He is a Lecturer at the College of Computer Science and Technology, Beijing University of Technology. His research interests include multimedia and image/video coding.

HU YongLi received his Ph.D. degree from Beijing University of Technology in 2005. He is an Associate Professor in the College of Computer Science and Technology at Beijing University of Technology. His research interests include computer graphics, pattern recognition, and multimedia technology.