

# 用块稀疏贝叶斯学习算法重构识别体域网步态模式

吴建宁\*, 徐海东

(福建师范大学 数学与计算机科学学院, 福州 350007)

(\* 通信作者电子邮箱 jianingwu@fjnu.edu.cn)

**摘要:**针对低功耗体域网步态远程监测终端非稀疏加速度数据重构和步态模式识别性能优化问题,提出了一种基于块稀疏贝叶斯学习的体域网远程步态模式重构识别新方法,该方法基于体域网远程步态监测系统架构和压缩感知框架,在体域网传感节点利用线性稀疏矩阵压缩原始加速度数据,减少传输数据量,降低其功耗,同时在远程终端基于块稀疏贝叶斯学习算法充分利用加速度数据块结构内在相关性,获取加速度数据内在稀疏性,有效提高非稀疏加速度数据重构性能,为准确识别步态模式提供可靠的数据支撑。采用 USC-HAD 数据库中行走、跑、跳、上楼、下楼五种步态运动的加速度数据验证新方法的有效性,实验结果表明,基于所提算法的加速度数据重构性能明显优于传统压缩感知重构算法性能,使基于支持向量机多步态分类器识别准确率可达 98%,显著提高体域网远程步态模式识别性能。所提新方法不仅有效提高非稀疏加速度数据重构和步态模式识别性能,并且也有助于设计低功耗、低成本的体域网加速度数据采集系统,为体域网远程监测步态模式变化提供一个新方法和新思路。

**关键词:**块稀疏贝叶斯学习算法;压缩感知;体域网;步态模式识别

**中图分类号:**TP393;TP391 **文献标志码:**A

## Block sparse Bayesian learning algorithm for reconstruction and recognition of gait pattern from wireless body area networks

WU Jianning\*, XU Haidong

(College of Mathematics and Computer Science, Fujian Normal University, Fuzhou Fujian 350007, China)

**Abstract:** In order to achieve the optimal performance of gait pattern recognition and reconstruction of non-sparse acceleration data from Wireless Body Area Networks (WBANs)-based telemonitoring, a novel approach to apply the Block Sparse Bayesian Learning (BSBL) algorithm for improving the reconstruction performance of non-sparse accelerometer data was proposed, which contributes to achieve the superior performance of gain pattern recognition. Its basic idea is that, in view of the gait pattern and Compressed Sensing (CS) framework of WBAN-based telemonitoring, the original acceleration-based data acquired at sensor node in WBAN was compressed only by spare measurement matrix (the simple linear projection algorithm), and the compressed data was transmitted to the remote terminal, where BSBL algorithm was used to perfectly recover the non-sparse acceleration data that assumed as block structure by exploiting intra-block correlation for further gait pattern recognition with high accuracy. The acceleration data from the open USC-HAD database including walking, running, jumping, upstairs and downstairs activities were employed for testing the effectiveness of the proposed method. The experiment results show that with acceleration-based data, the reconstruction performance of the proposed BSBL algorithm can significantly outperform some conventional CS algorithms for sparse data, and the best accuracy of 98% can be obtained by BSBL-based Support Vector Machine (SVM) classifier for gait pattern recognition. These results demonstrate that the proposed method not only can significantly improve the reconstruction performance of non-sparse acceleration data for further gait pattern recognition with high accuracy but also is very helpful for the design of low-cost sensor node hardware with lower energy consumption, which will be a potential approach for the energy-efficient WBAN-based telemonitoring of human gait pattern in further application.

**Key words:** Block Sparse Bayesian Learning (BSBL) algorithm; Compressed Sensing (CS); Wireless Body Area Network (WBAN); gait pattern recognition

## 0 引言

近年来,随着无线传感网、可穿戴加速度计、互联网技术的快速发展,基于无线体域网(Wireless Body Area Network, WBAN)远程监测人体步态模式技术在人体步态运动监测、老

年人步态康复评价及预防跌倒等领域得到广泛关注和应  
用<sup>[1]</sup>。其技术实现路线主要是:1)体域网传感节点采集加速  
度数据,并将其传送至路由节点;2)路由节点将采集的加速  
度数据传送至互联网,并经互联网传送至远程数据终端,识别  
步态模式变化评价人体步态功能状况<sup>[2]</sup>。近年来,相关研究

**收稿日期:**2014-12-19;**修回日期:**2015-01-23。 **基金项目:**福建省自然科学基金资助项目(2013J01220);福建省高等学校教学改革研究项目(JAS14674);福建师范大学本科教学改革项目(I201302021)。

**作者简介:**吴建宁(1969-),男,福建福州人,副教授,博士,主要研究方向:生物医学信号处理、无线人体传感网与医学应用;徐海东(1991-),男,福建莆田人,硕士研究生,主要研究方向:无线传感网。

较多集中探索远程步态模式识别方法(诸如隐性马尔可夫模型、支持向量机(Support Vector Machine, SVM)步态分类器)<sup>[3-6]</sup>,而未能全面考虑体域网远程步态模式识别技术在实际应用的状况。研究发现,体域网传感节点低功耗、远程数据终端准确获取采集数据是体域网步态模式远程监测技术实际应用中两个关键环节,也是目前亟须解决的关键问题,对于远程步态模式准确识别和进一步的步态功能评价至关重要<sup>[7-10]</sup>。

近年相关研究发现,体域网传感节点能量主要消耗于大量采集数据传输和复杂计算,探寻有效减少传感节点数据传输量且计算开销低的数据压缩方法是解决传感节点低功耗的有效途径。为此,一些学者基于生理数据的非线性、非平稳性、随机性等特点,探寻有效可行的数据压缩、重构解决方案,其基本思路主要依据数据压缩降低传感节点数据传输量,然后通过重构被压缩数据获取原始数据,用于进一步数据处理。近年来相继提出了基于神经网络的数据压缩算法<sup>[7]</sup>、基于小波变换的数据压缩算法<sup>[8]</sup>、分布式的数据压缩算法<sup>[9]</sup>以及相应改进的传统数据压缩算法<sup>[10]</sup>。但这些算法所处理数据均需满足奈奎斯特采样率,压缩算法复杂,难以在传感节点低功耗基础上准确重构数据,限制了体域网步态模式远程监测技术的实际应用。近期,一些学者尝试依据生理数据稀疏性,探寻新的压缩感知生理数据解决方案<sup>[11]</sup>。其基本思路是:依据生理数据稀疏性,通过稀疏测量矩阵线性投影压缩数据,以远低于奈奎斯特采样率获取数据,降低传感节点数据传输量和计算开销;然后构建优化学习重构算法准确恢复数据。目前,该方法已在一些具有良好稀疏性生理数据(如心电图数据)的远程监测应用研究中取得较好效果。研究发现,压缩感知重构性能依赖于数据稀疏性,将严重限制压缩感知方法在大多数非稀疏生理数据处理中的推广和应用。目前,一些学者针对非稀疏生理数据,进一步完善传统压缩感知重构算法性能。诸如,Zhang等学者针对非稀疏性生理数据,提出了块稀疏贝叶斯学习(Block Sparse Bayesian Learning, BSBL)压缩感知重构算法,其基本思想是假设非稀疏生理数据由块结构(groups/blocks structure)构成,采用贝叶斯学习算法寻找数据块内在稀疏性,提高非稀疏数据重构性能。该算法在非稀疏生理数据(如脑电数据)远程监测应用研究中得到较好效果<sup>[12]</sup>,为非稀疏生理数据压缩感知处理应用提供了新的思路和方法。当前,应用压缩感知技术压缩、重构非稀疏加速度数据,在体域网步态模式远程监测应用研究中受到广泛关注和重视,相关研究成果仍未见报道。

本文针对体域网加速度数据非稀疏特点,提出了基于块稀疏学习算法的加速度数据压缩感知框架,旨在降低传感节点功耗基础上,准确重构非稀疏加速度数据,为进一步准确识别步态模式提供可靠数据。其基本思路是:依据稀疏二进制测量矩阵,在传感节点线性压缩加速度数据,降低其数据传输量和功耗;在远程数据终端,通过设计加速度块稀疏贝叶斯学习重构算法,准确获取原始加速度数据。采用包含多种步态模式加速度数据的美国 Southern California 大学人体日常行为数据库(University of Southern California Human Activity Dataset, USC-HAD),评估上述所提算法有效性。研究中选用正常行走、跑、跳、上楼梯、下楼梯5种步态运动模式,设计分类性能优越的支持向量机多步态分类器,准确识别加速度数

据重构的5种步态模式。此外,文中将所提算法与传统小波分析方法进行比较,进一步验证所提算法有效性。

## 1 基于BSBL的加速度数据压缩感知重构算法

块稀疏贝叶斯学习算法是近年来新发展的一种压缩感知重构技术解决方案,该技术充分利用数据块内相关性(intra-block correlation),改善非稀疏数据重构性能,可有效解决非稀疏加速度数据压缩感知重构问题。

### 1.1 压缩感知技术

压缩感知技术是近年新发展的数据压缩处理解决方案<sup>[11]</sup>,其核心思想是依据数据稀疏性,在远低于奈奎斯特采样率情况下,有效实现数据的压缩和重构。其技术核心环节为数据稀疏化表示、测量矩阵和数据重构3个部分。技术实现路线是采用与原始数据稀疏表示的稀疏基不相关的测量矩阵,将高维稀疏数据线性投影到一个低维空间上,实现数据压缩;然后通过求解一个优化问题将包含少量投影压缩数据以高概率重构,恢复原始数据。即对于长度为 $N$ 的原始数据 $X \in \mathbf{R}^{N \times 1}$ ,通过测量矩阵 $\Phi \in \mathbf{R}^{M \times N} (M \ll N)$ 线性投影,可得压缩数据

$$Y = \Phi X; Y \in \mathbf{R}^{M \times 1} \quad (1)$$

一般情况下,原始数据 $X$ 可稀疏表示为: $X = \Psi\theta$ ,  $\Psi$ 为稀疏基, $\theta$ 为相应稀疏系数(稀疏度 $K \leq N$ ),式(1)可表示为:

$$Y = \Phi\Psi\theta = \Omega\theta \quad (2)$$

式中传感矩阵 $\Omega = \Phi\Psi \in \mathbf{R}^{M \times N}$ 。若原始数据具有良好时域稀疏性,则 $\Psi = I$ ,  $I$ 为单位矩阵。

由于式(1)中测量值个数远小于未知解个数( $M \ll N$ ),难以从压缩数据 $Y$ 中求解出原始数据 $X$ 。为此,采用数据稀疏分解算法,将压缩感知的重构问题转化为依据压缩数据 $Y$ 和传感矩阵 $\Omega$ ,求解式(2)欠定方程组,得到稀疏系数 $\theta$ ,实现数据重构。为保证算法有效性,传感矩阵 $\Omega$ 满足有限等距条件(Restricted Isometry Property, RIP),即对于任意具有稀疏度 $K$ 的原始数据 $X$ 和常数 $\delta_k \in (0, 1)$ ,传感矩阵 $\Omega$ 若满足:

$$1 - \delta_k \leq \frac{\|\Omega X\|_2^2}{\|X\|_2^2} \leq 1 + \delta_k \quad (3)$$

则可以通过 $l_0$ 范数下求解式(2)最优化问题:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \|\theta\|_0 \quad (4)$$

s. t.  $\Omega\theta = Y$

式中 $\|\theta\|_0$ 表示 $\theta$ 中非零值个数。由于 $l_0$ 范数最优化问题是一个NP-Hard问题,可将最小 $l_0$ 范数问题转化为最小 $l_1$ 范数问题<sup>[11]</sup>,得到同解,故式(4)转化为:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \|\theta\|_1 \quad (5)$$

s. t.  $\Omega\theta = Y$

即可得到稀疏系数 $\theta$ ,然后通过求解 $X = \Psi\theta$ ,可从 $M$ 维压缩数据 $Y$ 准确重构恢复 $N$ 维原始数据 $X$ 。

### 1.2 基于块稀疏贝叶斯学习算法重构加速度数据

如上所述,传统压缩感知重构算法难以准确恢复非稀疏加速度数据,必将影响后续步态模式识别性能。为有效提高非稀疏加速度数据重构性能,假设加速度数据由块结构组成,利用块稀疏贝叶斯学习算法探寻加速度数据块内相关性,捕获加速度数据块内在稀疏性,准确重构非稀疏加速度数据。即将原始加速度数据 $X \in \mathbf{R}^{N \times 1}$ 任意分割成 $g$ 个块结构

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_{d_1} & x_{d_1+1} & x_{d_1+2} & \cdots & x_{d_2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{d_{g-1}+1} & x_{d_{g-1}+2} & \cdots & x_{d_g} & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}^T \quad (6)$$

每个块内元素之间具有某种相关性,即块内相关性(intra-block correlation)。基于加速度数据块内相关性,块稀疏贝叶斯学习算法可探寻加速度数据内在稀疏性,有效改善数据重构性能。非稀疏加速度数据块稀疏贝叶斯学习重构模型框架如下。

假设加速度数据分割块  $x_i$  均满足多元高斯分布:

$$\rho(x_i; \gamma_i, \mathbf{B}_i) \sim N(0, \gamma_i \mathbf{B}_i); i = 1, 2, \dots, g \quad (7)$$

其中:  $\mathbf{B}_i$  为一个未知正定矩阵,用于刻画块内元素之间相关性;  $\gamma_i$  为用于控制原始数据块稀疏的未知参数,当  $\gamma_i = 0$  时,相应块  $x_i$  为 0。假设数据块相互独立,根据式(7),加速度数据  $\mathbf{X}$  的先验分布定义为  $p(x) \sim N(0, \Sigma_0)$ ,  $\Sigma_0$  为一个主对角线矩阵,即

$$\Sigma_0 = \begin{pmatrix} \gamma_1 \mathbf{B}_1 & 0 & \cdots & 0 & 0 \\ 0 & \gamma_2 \mathbf{B}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma_{g-1} \mathbf{B}_{g-1} & 0 \\ 0 & 0 & \cdots & 0 & \gamma_g \mathbf{B}_g \end{pmatrix} \quad (8)$$

同样,假设噪声服从  $\rho(v; \lambda) = N(0, \lambda \mathbf{I})$ , 其中  $\lambda$  为噪声方差,  $\mathbf{I}$  为单位矩阵。加速度数据  $\mathbf{X}$  的后验分布为

$$p(\mathbf{X} | \mathbf{Y}; \{\gamma_i, \mathbf{B}_i\}_{i=1}^g) = N(\mathbf{u}_x, \Sigma_x) \quad (9)$$

其中:  $\mathbf{u}_x = \Sigma_0 \Phi^T (\lambda \mathbf{I} + \Phi \Sigma_0 \Phi^T)^{-1} \mathbf{Y}$ ,  $\Sigma_x = (\Sigma_0^{-1} + \lambda^{-1} \Phi^T \Phi)^{-1}$ 。

可利用第二类最大似然估计方法 (type-II maximum likelihood estimation)<sup>[13]</sup> 估计参数  $\{\gamma_i, \mathbf{B}_i\}_{i=1}^g, \lambda$ , 求解上述参数可转化为求解下列最小化代价函数<sup>[14]</sup>:

$$L(\lambda, \{\gamma_i, \mathbf{B}_i\}_{i=1}^g) \triangleq \log |\lambda \mathbf{I} + \Phi \Sigma_0 \Phi^T| + \mathbf{Y}^T (\lambda \mathbf{I} + \Phi \Sigma_0 \Phi^T)^{-1} \mathbf{Y} \quad (10)$$

当参数  $\{\gamma_i, \mathbf{B}_i\}_{i=1}^g, \lambda$  确定后,可得到原始加速度数据  $\mathbf{X}$  最大后验估计 (maximum-a-posterior estimate), 即:

$$\hat{\mathbf{X}} = \Sigma_0 \Phi^T (\lambda \mathbf{I} + \Phi \Sigma_0 \Phi^T)^{-1} \mathbf{Y} \quad (11)$$

即可重构原始加速度数据。

目前,基于 BSBL 模型框架的重构迭代算法主要有: BSBL-EM (Expectation Maximization)、BSBL-BO (Bound Optimization) 和 BSBL-L1, 每种算法设计不同获取  $\gamma_i, \mathbf{B}_i, \lambda$  参数的学习规则, BSBL-EM 算法性能最优,但运行速度最慢; BSBL-L1 运行速度最快,但性能差于其他两个算法; BSBL-BO 算法性能和速度介于两者之间。为此,本文采用 BSBL-BO 算法,基于压缩数据  $\mathbf{Y}$ 、传感矩阵  $\Phi$ 、 $g$  个块结构,重构原始加速度数据  $\mathbf{X}$ 。其算法主要步骤如下:

步骤 1 计算学习规则  $\lambda \leftarrow (\|\mathbf{Y} - \Phi \mathbf{u}_x\|_2^2 + \text{Tr}(\Sigma_x \Phi^T \Phi)) / M$ ,  $\mathbf{u}_x^i \in \mathbf{R}^{d_i \times 1}$  是  $\mathbf{u}_x$  中的第  $i$  块 (在不考虑系统存在噪声环境的情况下,可以设置  $\lambda = 10^{-10}$ )。

$$\text{步骤 2 计算学习规则 } \gamma_i \leftarrow \sqrt{\frac{\mathbf{x}_i^T \mathbf{B}_i^{-1} \mathbf{x}_i}{\text{Tr}((\Phi^i)^T (\Sigma_y^*)^{-1} \Phi^i \mathbf{B}_i^i)}},$$

其中,  $\Phi^i \in \mathbf{R}^{M \times d_i}$  为  $\Phi$  的子矩阵,  $\Sigma_y^* = \lambda \mathbf{I} + \Phi \Sigma_0^* \Phi^T$ ,  $\Sigma_0^* \triangleq \Sigma_0 |_{\gamma=\gamma^*}$ ,  $\gamma^*$  为  $\gamma$  中的一个给定值。

步骤 3 当分割块数目相同时,计算学习规则  $\mathbf{B} \leftarrow \frac{1}{g} \sum_{i=1}^g \frac{\Sigma_x^i + \mathbf{u}_x^i (\mathbf{u}_x^i)^T}{\gamma_i}$ , 其中  $\mathbf{B}_i = \mathbf{B}(\forall i)$ ,  $\mathbf{u}_x^i \in \mathbf{R}^{d_i \times 1}$  是  $\mathbf{u}_x$  中的第  $i$  块,  $\Sigma_x^i \in \mathbf{R}^{d_i \times d_i}$  是  $\Sigma_x$  主对角线上的第  $i$  块。若块数目不相同,则直接跳到步骤 4。

步骤 4 当分块数目不相同,计算学习规则  $\mathbf{B}_i \leftarrow [\Sigma_x^i + \mathbf{u}_x^i (\mathbf{u}_x^i)^T] / \gamma_i$ , 并重新构造  $\mathbf{B}_i$ :

$$\hat{\mathbf{B}}_i \leftarrow \text{Toeplitz}([1, \bar{r}, \dots, \bar{r}^{d_i-1}]); \forall i$$

其中:  $\bar{r} \triangleq \overline{m_1} / \overline{m_0} = \sum_{i=1}^g m_1^i / \sum_{i=1}^g m_0^i$ ,  $m_0^i, m_1^i$  分别为  $\mathbf{B}_i$  主对角线元素平均值和主对角线下的对角线元素平均值。

当上述步骤满足停止收敛准则时,即可得到重构加速度数据  $\hat{\mathbf{X}} \leftarrow \Sigma_0 \Phi^T (\lambda \mathbf{I} + \Phi \Sigma_0 \Phi^T)^{-1} \mathbf{Y}$ 。

## 2 基于重构加速度数据的 SVM 步态模式识别

为有效评价所提算法重构的加速度数据对后续步态模式识别性能的影响,本文构建了性能优越的支持向量机多步态分类器,来准确识别重构加速度数据步态模式。为有效提高支持向量机分类性能,基于加速度数据非线性、非平稳性、随机性特点,需从重构加速度数据中提取包含更多步态模式差异信息的特征参数。

### 2.1 特征提取

本文选取常用的统计参数:标准差、偏度、峰度和 Person 相关系数,作为刻画走、跑、跳、上楼和下楼 5 种步态模式差异信息的特征参数<sup>[15-16]</sup>,其计算过程如下:

1) 标准差。用于刻画不同运动加速度数据的稳定性,其定义如下:

$$s = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2} \quad (12)$$

其中:  $\bar{X}$  表示样本平均值,  $N$  为样本数。

2) 偏度。用于描述不同运动状态加速度数据分布状况,其定义如下:

$$\text{Skewness} = N \sum_{i=1}^N (X_i - \bar{X})^3 / (N-1)(N-2)s^3 \quad (13)$$

其中:  $\bar{X}$  是样本均值,  $N$  是样本数,  $s$  是样本标准差。

3) 峰度。用于描述不同运动加速度数据分布形态陡缓程度,其定义如下:

$$\text{Kurtosis} = \sum_{i=1}^N (X_i - \bar{X})^4 / f_i / N s^4 \quad (14)$$

其中:  $\bar{X}$  为样本均值,  $N$  为样本数,  $s$  为样本标准差,  $f_i$  为样本间隔。当 Kurtosis = 0, 与正态分布陡缓程度相同; 当 Kurtosis > 0, 为比正态分布的高峰更加陡峭的尖顶峰; 当 Kurtosis < 0, 为比正态分布的高峰稍平的平顶峰。

4) Pearson 相关系数。用于刻画不同运动状态的相关程度,其定义如下:

$$\rho_{x,y} = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}} \quad (15)$$

其中:  $\bar{X}, \bar{Y}$  分别为不同样本  $X$  和  $Y$  的均值。  $N$  为样本数, 其值若越接近于 1 或 -1, 不同运动相关性越强; 反之, 若其值接近于

0,不同运动相关性差。

## 2.2 支持向量机步态模式识别

支持向量机(SVM)分类算法是基于统计学习理论的结构风险最小化准则构建,具有良好泛化性能。文中基于支持向量机二分类算法构建支持向量机多步态分类器,支持向量机二分类算法就是将步态模式数据通过核函数映射到高维特征空间,在被映射的高维特征空间寻找最优分类超平面,也就是,对于加速度数据训练样本集:  $(x_i, y_i), x_i \in \mathbf{R}^n, y_i \in \{\pm 1\}, i = 1, 2, \dots, n$ , 分类输入数据  $x_i$  通过函数  $\phi(x_i): \mathbf{R}^n \rightarrow H$  被映射到高维特征空间  $H$ , 在  $H$  找到一个最优线性分类面  $f$ , 将步态模式样本映射到其相应的类标志空间  $y_i \in \{-1, +1\}$ 。在  $H$  中假设分类面为  $w \cdot x + b = 0$ ,  $w$  为分类面法向量,  $b$  为常数项,通常假设分类间隔为  $2/\|w\|^2$ , 若分类间隔最大即为所寻找的最优超平面。也就是,通过寻找最优参数  $w$  和  $b$ , 可得到最优分类超平面。考虑到加速度数据线性不可分,文中采用一阶范数软间隔构造最优分类面。为了有效减少训练误差,引入松弛变量  $\xi_i$ , 通过求解下列优化问题,寻找最优分类面。

$$\min \frac{1}{2} w \cdot w + C \sum_{i=1}^l \xi_i \quad (16)$$

$$\text{s. t. } y_i(w \cdot x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, \forall i$$

其中  $C$  为调节分类面最大间隔与样本训练最小误差的惩罚参数。将求解最优分类面转化为一个二次规划问题,通过构造拉格朗日函数将二次规划问题转化为下列对偶问题,可得到最优解。

$$\max W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \phi(x_i) \cdot \phi(x_j) \quad (17)$$

$$\text{s. t. } \sum_{i=1}^l y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, \forall i$$

其中非负拉格朗日乘子  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)$  就是所求的位于决策边界上步态样本  $x_i$  的支持向量。利用满足 Mercer 定理条件的核函数  $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$  计算  $\phi(x_i) \cdot \phi(x_j)$ , 求解出线性分类面的权向量  $w$  的最优解<sup>[17]</sup>:

$$w = \sum_{i \in SVs} \alpha_i y_i \phi(x_i) \quad (18)$$

为此,步态分类器决策函数定义为:

$$y = \text{sign}\left(\sum_{i \in SVs} \alpha_i y_i K(x_i, x) + b\right) \quad (19)$$

在此基础上,构建 SVM 多步态分类器。目前,构造 SVM 多类分类器的方法主要有两种:一种是直接法,其基本思路就是将二分类算法中的目标函数进行适当修改,将多个分类面参数求解合并到一个最优化问题中,通过求解最优化问题“一次性”实现多类分类。该方法求解复杂,计算复杂度高,实现较困难。另一种是间接法,其基本思路就是将多分类问题分解为多个二分类问题,通过求解多个二分类结果来判别最终分类结果。目前,基于间接法的 SVM 多分类方法主要有:一对一(one-against-one)、一对多(one-against-all),层次支持向量机等多分类方法。本文考虑到不同运动之间差异性,采用一对多方法,其基本思路就是在样本训练时,依次把某个步态模式类别样本归为一类,其他剩余步态模式样本归为另一类,根据  $k$  个类别的样本构造  $k$  个 SVM,仅需训练  $k$  个 SVM 二分类器,依据各分类器决策函数值,将获得最大值分类器的

样本判别为该模式,分类速度较快。

## 3 实验结果

### 3.1 加速度数据获取

本文采用美国 Southern California 大学人体步态运动数据库(USC-HAD)的加速度数据<sup>[18]</sup>,评价所提算法有效性。其数据获取过程如下:挑选 14 名健康受试者,要求每名受试者将加速度传感采集节点佩戴于受试者右前臂,采集节点通过连线实时将采集数据传送至受试者随身携带运动记录节点(motion node),记录采集获取的步态运动数据。采集时,要求每名受试者完成自然行走、跑、跳、上楼、下楼等步态运动,每种步态运动,每名受试者重复采集 5 次,采集频率设置为 100 Hz。图 1 为一名受试者行走期间(1 s)的三维加速度数据( $X$  轴表示人体运动前进方向; $Y$  轴表示人体运动侧向方向; $Z$  轴表示人体运动垂直方向)。

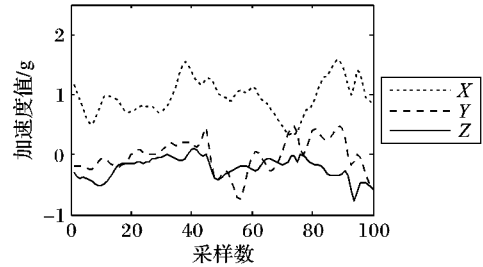


图1 一名受试者行走步态期间的三维加速度数据

### 3.2 基于 BSBL 算法的加速度数据压缩和重构结果

从理论上讲,压缩感知模型要求测量矩阵  $\Phi$  与稀疏矩阵  $\Psi$  乘积必须满足 RIP 性质(即有限等距性质),也就是测量矩阵  $\Phi$  和稀疏矩阵  $\Psi$  不相关。目前研究常采用上述性质的测量矩阵主要有高斯随机矩阵、托普利兹矩阵和稀疏二进制随机矩阵。但就实际应用而言,高斯随机矩阵和托普利兹矩阵计算量大,传感节点功耗大。而稀疏二进制随机矩阵(如式(20)所示)中仅含元素 0 和 1,并且每列含有相同数目的元素 1,其余均为元素 0,有助于降低传感节点计算复杂度和硬件功耗。

$$\Phi = \begin{pmatrix} 1 & 0 & 1 & 0 & \cdots & 1 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & 1 & \cdots & 0 \end{pmatrix} \quad (20)$$

实验中,选用稀疏二进制矩阵作为测量矩阵,每列元素 1 个数选为 10。考虑到稀疏度对压缩重构性能的影响<sup>[16,18]</sup>,压缩率选为 50%,原始加速度数据长度选为 100,稀疏二进制测量矩阵大小为  $50 \times 100$ 。考虑到 BSBL 算法重构性能对数据块划分不敏感特性<sup>[19]</sup>,将加速度数据任意划分为 5 个数据块,每个数据块包含 20 个数据点,设置  $\lambda = 10^{-10}$  (不考虑系统存在噪声),选择最大迭代次数为 25 次作为停止收敛准则。

作为比较,选取 4 个传统压缩感知重构算法:正交匹配追踪(Orthogonal Matching Pursuit, OMP)算法<sup>[20]</sup>、子空间追踪(Subspace Pursuit, SP)算法<sup>[21]</sup>、梯度投影重构(Gradient Projection for Sparse Reconstruction, GPSR)算法<sup>[22]</sup>和 SPGL1-LASSO(Least Absolute Shrinkage and Selection Operator)算法<sup>[23]</sup>。OMP 和 SP 是典型的解决最小化  $l_1$  范数问题的贪婪算法,GPSR 算法是解决最小化  $l_0$  范数问题的凸优化算法,

SPGL1-LASSO 也是解决最小化  $l_1$  范数问题的重构算法,上述重构算法均未利用解向量的块结构信息。实验中,选用  $X$  轴加速度数据,各重构算法均采用 DCT 基作为稀疏基。为了客观评价上述算法重构性能和各种压缩感知重构算法对加速度数据的适用性,选用均方误差和结构相似性<sup>[16,18]</sup>作为客观评价指标,其定义如下:

1) 均方误差:

$$MSE = \|X - \hat{X}\|_2 / \|X\|_2 \quad (21)$$

其中:  $X$  为原始加速度数据,  $\hat{X}$  为重构的加速度数据。

2) 结构相似性

$$SSIM(h, g) = [L(h, g)]^\alpha [C(h, g)]^\beta [S(h, g)]^\gamma; \quad (22)$$

$$\begin{cases} L(g, h) = \frac{2u(g)u(h) + C1}{u(g)^2 + u(h)^2 + C1} \\ C(g, h) = \frac{2d(g)d(h) + C2}{d(g)^2 + d(h)^2 + C2} \\ S(g, h) = \frac{d(g, h) + C3}{d(g)d(h) + C3} \end{cases}$$

其中:  $u(g), u(h)$  为信号  $g$  和  $h$  的均值;  $d(g), d(h)$  为信号  $g$  和  $h$  的方差;  $d(g, h)$  为信号  $g, h$  的协方差;  $L(g, h)$  常用于比较信号  $g$  和  $h$  的均值;  $C(g, h)$  用于比较信号  $g$  和  $h$  的对比度;  $S(g, h)$  比较数据信号  $g$  和  $h$  的结构差异;  $\alpha > 0, \beta > 0, \gamma > 0$  为调整  $L(g, h), C(g, h)$  和  $S(g, h)$  的相对重要性的参数。结构相似性指标的值越大,表明两个数据相似性程度就越高。实验中,为了计算方便,设置  $\alpha, \beta, \gamma$  均为 1,  $C1, C2, C3$  也均为 1。

实验中,首先根据均方误差(即重构误差)来评估各种重构算法性能与测量维数(也就是压缩数据)的变化关系,其结果如图 2 所示。

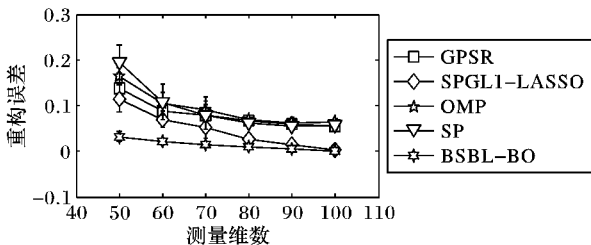


图2 测量维数与重构误差的关系

从图 2 中发现,各种重构算法的重构误差均随着测量维数的增大而减少,相比较而言,BSBL-BO 的重构误差明显低于其他传统压缩感知重构算法的重构误差,并且在测量维数 50~100 变化时保持较稳定。而其他传统压缩感知重构算法的误差随测量维数增加急剧减少,变化极为不稳定。同样,再依据结构相似性,评估各种重构算法性能与测量维数的变化关系,其结果如图 3 所示。

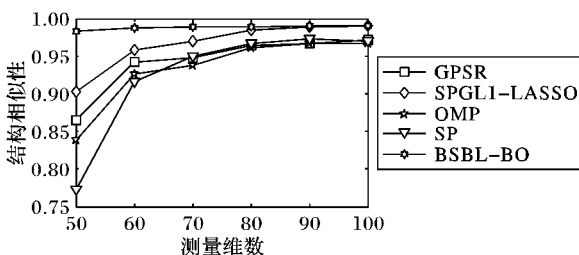


图3 测量维数与结构相似性的关系

从图 3 中能够看到,各种重构算法的结构相似性均随着

测量维数增大而增大,相比较而言,块稀疏学习重构算法(BSBL-BO)结构相似性明显优于其他传统压缩感知重构算法的结构相似性,并且在测量维数 50~100 变化时保持较稳定(约 0.98)。而其他传统压缩感知结构相似性随测量维数增加迅速增加,变化极为不稳定。

这些结果表明,与传统的压缩感知重构算法相比,块稀疏贝叶斯学习重构算法能够有效地从加速度数据块结构中获取非稀疏加速度数据内在稀疏性,有效提高非稀疏加速度数据重构性能。此外,从图 2 和图 3 的结果中可以发现,测量维数越大,各种重构算法性能越优异。但从数据压缩角度来讲,测量维数越大,压缩数据越小,数据计算量大,不利于传感节点低功耗。因根据实际应用情况,在重构性能和测量维数之间折中,来选择测量维数。本文选取测量矩阵的维数  $M$  为 50,即压缩率为 50%。

### 3.3 支持向量机多步态模式识别结果

为评价重构加速度数据在步态模式识别中的有效性,实验中,选取 10 名受试者行走、跑、跳、上楼梯、下楼梯 5 种步态运动模式,分别从  $X, Y, Z$  轴方向上各截取 10 段加速度数据,每段加速度数据包含 512 个数据点(即窗长为 512,相邻两窗长重叠 50%,加窗后的每段加速度数据样本间隔是 5.12 s),步态样本总数为 500,其中每种步态运动样本数均为 100,基于标准差、偏差、峰度和 Person 相关系数 4 个征参数,将每个步态模式定义为 12 维特征向量,即构成一个  $500 \times 12$  的特征矩阵样本数据。由于样本数据有限,采用留一交叉验证法测试支持向量机多分类器识别性能,即依次从每种步态选取 80 个样本数据组成  $400 \times 12$  特征矩阵作为训练集,然后用剩余的  $100 \times 12$  的特征矩阵作为测试集,上述训练和测试过程循环 20 次,取 20 次分类结果的平均值作为最终分类结果。实验选用高斯径向基核函数(Radial Basis Function, RBF):

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2 \times r^2}\right) \quad (23)$$

其中,  $r$  为函数的径向宽度参数。采用网格寻优法确定高斯径向基最优参数。采用台湾大学林智仁教授开发的 LIBSVM 程序代码<sup>[24]</sup>,运行环境为 Intel Core i5-3470 3.20 GHz CPU 和 4.00 GB RAM,在 Matlab7.0 上进行仿真实验。高斯径向基核参数变化过程如图 4 所示,从图 4 可以发现,支持向量机多分类器性能与其算法的相关参数变化密切相关,当  $r = 0.2$ ,惩罚参数  $C = 6$  时,准确率约为 98%,分类性能最优。

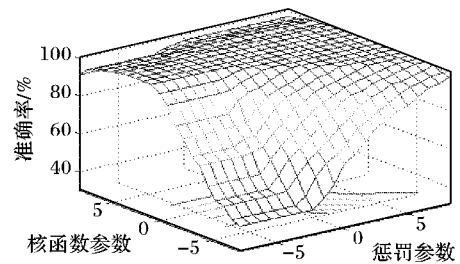


图4 支持向量机分类算法中搜寻最优参数过程

基于最优重构加速度数据的步态模式识别结果如表 1 所示,从表 1 的分类器混淆矩阵中发现,除行走步态模式的分类性能稍差外,其余四种步态运动模式分类准确率均为 100%,

总体平均分类准确率达到 98%, 结果表明, 基于块稀疏贝叶斯学习算法重构的加速度数据能够保持原始加速度数据内在刻画步态运动的细节特征信息, 为支持向量机分类算法提供了更多步态运动模式差异信息, 有效提高支持向量机多步态模式识别性能。

表 1 基于 SVM 的 5 种步态模式识别的混淆矩阵

步态类别	走	跑	跳	上楼梯	下楼梯	总计
走	18	2	0	0	0	20
跑	0	20	0	0	0	20
跳	0	0	20	0	0	20
上楼梯	0	0	0	20	0	20
下楼梯	0	0	0	0	20	20

实验中, 也将所提算法与基于传统小波分析压缩重构算法进行比较, 进一步验证所提算法处理加速度数据的有效性。基于上述相同加速度数据, 采用 db3 小波对原始加速度数据 3 层分解, 然后采用软阈值方法选取分解后每层高频系数重构数据, 小波压缩重构所得能量剩余为 98.6476, 零系数成分为 78.2051。基于小波分析重构加速度数据的步态模式识别结果如表 2 所示, 从表 2 可以看到, 基于小波分析的压缩重构的多步态分类器平均识别率仅为 92%, 远低于本文所提算法的步态模式识别率 98%, 表明非稀疏加速度小波分析重构算法, 难以准确恢复原始加速度数据, 易丢失步态运动的细节特征信息, 难以为支持向量机分类算法提供更多的步态模式差异信息。

表 2 基于小波分析和 SVM 的多步态模式识别的混淆矩阵

步态类别	走	跑	跳	上楼梯	下楼梯	总计
走	18	2	0	0	0	20
跑	5	15	0	0	0	20
跳	0	1	19	0	0	20
上楼梯	0	0	0	20	0	20
下楼梯	0	0	0	0	20	20

此外, 根据算法复杂度和步态识别时间, 进一步评价本文所提加速度数据压缩感知框架在体域网步态模式远程监测系统架构的实用性。基于上述算法步态识别时间结果如表 3 所示。

表 3 基于不同压缩重构算法的步态识别算法消耗时间比较 s

步态识别算法	压缩数据	重构数据	步态识别
本文算法	4.60	9.15	23.19
基于小波分析步态识别算法	5.55	5.73	28.64

从表 3 的结果来看, 基于稀疏二进制测量矩阵的数据压缩消耗时间明显低于传统小波分析数据压缩消耗时间, 这主要是由于稀疏二进制测量矩阵线性投影压缩加速度数据, 所需计算次数仅为  $N \times d$  次简单加法运算, 所耗费时间复杂度仅为  $O(N)$ 。而传统小波分析压缩数据时, 则需处理大量乘法和加法运算; 基于块稀疏贝叶斯学习的加速度数据重构算法消耗时间稍高于传统小波分析加速度数据重构消耗时间, 其原因主要是所提重构算法主要求解加速度数据块内未知稀疏解, 也就是在求解迭代过程中, 寻找最优块稀疏控制参数  $\gamma$ 、正定矩阵  $B_i$ , 尺度参数  $\lambda$ , 所需时间复杂度为  $O(N^3 + 2N^2 \times$

$M)$ , 重构时间相对较长。但在实际应用中, 可将本文所提重构算法放置体域网远程终端高性能计算机上运行。由于本文采用小样本步态数据, SVM 多步态分类器时间复杂度和空间复杂度均为  $O(N^3)$ , 识别时间主要消耗于寻找最优参数惩罚因子  $C$  和核函数参数  $g$ , 空间复杂度为  $O(N^3)$ 。结果表明, 本文所提加速度数据压缩感知框架较适用于体域网步态模式远程监测系统, 有助于传感节点低功耗和远程数据终端准确获取数据。

## 4 讨论与分析

实验结果表明, 基于块稀疏贝叶斯学习的压缩感知重构算法能够有效处理非稀疏加速度数据。本文算法选用稀疏二进制测量矩阵, 有效降低传感节点传输数据量和计算复杂性; 在远程数据终端采用块稀疏贝叶斯学习算法准确重构非稀疏加速度数据, 为后续步态模式识别提供可靠数据, 有助于体域网步态模式远程监测技术推广和应用。

目前, 传感节点低功耗和远程数据终端准确获取加速度数据仍是体域网步态模式远程监测技术应用中亟须解决的关键问题。当前相关领域学者仍致力于探寻有效解决上述问题的技术手段。本文从体域网步态模式远程监测系统架构和加速度数据非稀疏特点两方面综合考虑, 提出了基于块稀疏学习的加速度数据压缩感知框架, 旨在探讨压缩感知框架对非稀疏加速度数据的适用性, 寻找有效解决体域网传感节点低功耗和远程准确获取加速度数据等关键技术的手段。

从图 2 和图 3 的实验结果来看, 当选用稀疏二进制测量矩阵, 压缩率为 50% 时, 所提算法的重构性能明显优于传统压缩重构算法 (诸如 OMP、SP、GPSR、SPGL1-LASSO) 性能, 表明非稀疏加速度数据块稀疏贝叶斯学习重构算法在稀疏二进制测量矩阵有效压缩数据基础上, 能够充分利用加速度数据块结构内在相关性, 捕获加速度数据块内在稀疏性, 明显改善非稀疏加速度压缩数据重构性能, 而传统压缩重构算法更多依赖于重构数据的先验信息, 未能有效利用加速度数据空间信息和时序结构, 难以有效改善非稀疏加速度数据重构性能。更重要的是, 本文选用的稀疏二进制测量矩阵中每列仅含有少量元素 1, 有助于低功耗传感节点硬件体系架构简化设计<sup>[12]</sup>。

为评价所提算法重构的加速度数据对步态模式识别的影响, 构建了泛化性能优越的支持向量机多步态分类器。从表 1 的实验结果来看, 基于本文所提算法重构加速度数据的步态模式识别性能明显优于基于传统压缩感知重构算法的步态模式识别性能, 表明本文所提取的特征参数 (标准差、偏度、峰度和 Person 相关系数) 能够有效从重构的加速度步态数据中获取步态运动内在变化差异信息, 为支持向量机多步态模式识别算法提供更多不同步态模式间的差异信息。此外, 考虑到所提算法处理的数据假设于独立随机分布特性, 本文选用高斯径向基核函数构建支持向量机多步态模式识别模型, 采用网格搜寻方法寻找识别性能最优的核参数, 从图 4 的实验结果来看, 网格搜寻方法能够充分综合考虑不同参数对步态模式识别性能的影响, 较直观体现多步态模式识别性能变化状况, 可避免以往采用经验方法反复寻找最优参数时出现

的过拟合等问题。

作为比较,采用传统小波分析方法压缩、重构研究中的加速度数据,从表2的实验结果来看,基于小波分析重构加速度数据的步态模式识别性能明显低于本文所提算法重构加速度数据的识别性能,表明基于小波分析阈值重构算法难以完整重构非稀疏加速度数据,易丢失与步态运动内在模式变化密切相关的特征信息,难以为支持向量机多步态模式识别模型提供更多的不同步态模式差异信息。

## 5 结语

本文提出了一种基于块稀疏贝叶斯学习的加速度数据压缩感知重构算法,该算法能够有效捕获加速度数据内在块稀疏性,显著提高非稀疏加速度数据压缩感知重构算法性能,准确恢复原始加速度数据,为有效优化远程步态模式识别性能提供可靠数据,有助于远程步态功能评价。

### 参考文献:

- [1] GONG J, WANG R, CUI L. Research advances and challenges of body sensor network [J]. *Journal of Computer Research and Development*, 2010, 47(5): 737–753. (宫继兵, 王睿, 崔莉. 体域网BSN的研究进展及面临的挑战[J]. *计算机研究与发展*, 2010, 47(5): 737–753.)
- [2] CHEN C, ALOIS K, H-ERICH W, *et al.* A review of three-layer wireless body sensor network systems in healthcare for continuous monitoring [J]. *Journal of Modern Internet of Things*, 2013, 2(3): 24–34.
- [3] WANG Z, JIANG M, HU Y, *et al.* An incremental learning method based on probabilistic neural networks and adjustable fuzzy clustering for human activity recognition by using wearable sensors [J]. *IEEE Transactions on Information Technology in Biomedicine*, 2012, 16(4): 691–699.
- [4] LIN B, LIU Y, YU C, *et al.* Gait recognition and walking exercise intensity estimation [J]. *International Journal of Environment Research and Public Health*, 2014, 11(4): 3822–3844.
- [5] MI Z, ALEXANDER A S. A feature selection-base framework for human activity recognition using wearable multimodal sensor [C]// *Proceedings of the 6th International Conference on Body Area Networks*. Beijing: ICST, 2011: 201–208.
- [6] XING X, LIU H, HUANG W. Gait pattern classification with wavelet energy and sample entropy based on acceleration signals [J]. *Chinese Journal of Sensor and Actuators*, 2013, 26(4): 545–549. (邢秀玉, 刘鸿宇, 黄武. 基于加速度的小波能量特征及样本熵组合的步态分类算法[J]. *传感器学报*, 2013, 26(4): 545–549.)
- [7] WANG J, LIN X, WU K. ECG data compression research based on wavelet neural network [C]// *Proceedings of the 2010 International Conference on Computer, Mechatronics, Control and Engineering*. Piscataway: IEEE, 2010, 1: 361–363.
- [8] AHMED Z. ECG signal compression technique based on discrete wavelet transform and QRS-complex estimation [J]. *Signal Processing: an International Journal*, 2010, 4(3): 138–160.
- [9] WU H, TSAI W, CHANG W. Distributed source coding for real-time ECG signal monitoring [J]. *IEICE Transactions on Information and Systems*, 2014, (8): 2157–2165.
- [10] ABBADI N K E, AL-BAKRY A M. New efficient technique for compression of ECG signal [J]. *International Journal of Computer Science Issues*, 2013, 10(4): 139–146.
- [11] DONOHO D. Compressed sensing [J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 289–306.
- [12] ZHANG Z, JUNG T, MAKEIG S, *et al.* Compressed sensing for energy-efficient wireless telemonitoring of non-invasive fetal ECG via block sparse Bayesian learning [J]. *IEEE Transactions on Biomedical Engineering*, 2013, 60(2): 300–309.
- [13] SAAD Q, RANA M B, WAFA I, *et al.* Compressive sensing: from theory to applications [J]. *Journal of Communications and Networks*, 2013, 15(5): 443–456.
- [14] ZHANG Z, RAO B. Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation [J]. *IEEE Transactions on Signal Processing*, 2013, 61(8): 2009–2015.
- [15] YANG J, WANG S, CHEN N, *et al.* Wearable accelerometer based extendable activity recognition system [C]// *ICRA 2010: Proceedings of the 2010 IEEE International Conference on Robotics and Automation*. Piscataway: IEEE, 2010: 3641–3647.
- [16] XU C, GU Q, YAO M. Activity recognition method based on three-dimensional accelerometer [J]. *Computer System Applications*, 2013, 22(6): 132–135. (徐川龙, 顾勤龙, 姚明海. 一种基于三维加速度器的人体行为识别方法[J]. *计算机系统应用*, 2013, 22(6): 132–135.)
- [17] CORTES C, VPMIK V. Support-vector networks [J]. *Machine Learning*, 1995, 20: 273–297.
- [18] MI Z, ALEXANDER A S. USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors [C]// *ACM 2012: ACM International Conference on Ubiquitous Computing Workshop on Situation, Activity and Goal Awareness*. New York: ACM, 2012: 1036–1043.
- [19] ZHANG Z, JUNG T P, MAKEIG S, *et al.* Compressed sensing of EEG for wireless telemonitoring system with low energy consumption and inexpensive hardware [J]. *IEEE Transactions on Biomedical Engineering*, 2013, 60(1): 221–224.
- [20] TROPP J, GILBERT A. Signal recovery from random measurement via orthogonal matching pursuit [J]. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655–4666.
- [21] DAI W, MILENKOVIC O. Subspace pursuit for compressive sensing: closing the gap between performance and complexity [J]. *IEEE Transactions on Information Theory*, 2009, 55(5): 2230–2249.
- [22] FIGUEIREDO M, NOWAK R, WRIGHT S. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems [J]. *IEEE Journal on Selected Topics in Signal Processing*, 2007, 1(4): 586–597.
- [23] SEUNG-JEAN K, KOH K, LUSTIG M, *et al.* A method for large-scale  $\ell_1$ -regularized least squares [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(4): 606–607.
- [24] CHANG C-C, LIN C-J. LIBSVM: A library for support vector machines (Version 3.18) [EB/OL]. [2014–10–12]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html>.