

基于改进的快速稀疏编码的图像特征提取

尚 丽^{1,2*}, 苏品刚¹, 周 燕^{1,3}

(1. 苏州市职业大学 电子信息工程学院, 江苏 苏州 215104;

2. 中国科学技术大学 自动化系, 合肥 230026; 3. 苏州大学 电子信息学院, 江苏 苏州 215006)

(* 通信作者电子邮箱 shangli0930@126.com)

摘 要:考虑图像特征系数的最大化稀疏分布和特征基的正交性,在快速稀疏编码(FSC)模型的基础上,提出一种改进的FSC模型。该模型利用迭代法解决了基于 L_1 范数的归一化最小二乘法和基于 L_2 范数的约束最小二乘法的凸优化问题,能够实现完备基和过完备基的学习,有效提取出图像的最佳特征,且比标准稀疏编码(BSC)模型的收敛速度快。分别利用自然场景图像和掌纹图像作为训练数据进行特征提取测试,并进一步利用提取的特征基进行图像重构实验,同时与BSC模型的图像重构结果进行对比,实验结果证实了所提出的改进FSC模型能够快速、有效地实现图像的特征提取。

关键词:快速稀疏编码;最小二乘法; L_1 范数; L_2 范数;特征提取;图像重构

中图分类号: TN391.41 **文献标志码:** A

Image feature extraction based on modified fast sparse coding algorithm

SHANG Li^{1,2*}, SU Pin'gang¹, ZHOU Yan^{1,3}

(1. School of Electronic Information Engineering, Suzhou Vocational University, Suzhou Jiangsu 215104, China;

2. Department of Automation, University of Science and Technology of China, Hefei Anhui 230026, China;

3. School of Electronic and Information, Soochow University, Suzhou Jiangsu 215006, China)

Abstract: On the basis of the Fast Sparse Coding (FSC) model, considering the maximum sparse distribution of feature coefficients and the orthogonality of feature bases of an image, a Modified FSC (MFSC) model was proposed in this paper. This FSC algorithm was based on iteratively solving two convex optimization problems: L_1 -norm based regularized least square problem and L_2 -norm based constrained least square problem, and it can realize the learning of complete bases and overcomplete bases, as well as efficiently extract the features of images. Moreover, the convergence speed of FSC is quicker than that of Basic Sparse Coding (BSC). The images of natural scene and palmprint were used to test the property of FSC algorithm proposed by the authors in feature extraction, and then the extracted features were utilized to image reconstruction. Compared with reconstructed images obtained by BSC, the experimental results verify the validity of the modified FSC in quickly extracting image features.

Key words: Fast Sparse Coding (FSC); least square; L_1 -norm; L_2 -norm; feature extraction; image reconstruction

0 引言

稀疏编码(Sparse Coding, SC)模型能够成功模拟哺乳动物初级视觉系统主视皮层V1区神经元的感受野,是高维数据的一种有效的稀疏表示方法^[1-4]。最早的SC模型由Olshausen等^[5]在1996年提出,一般记作标准的SC(Basic Sparse Coding, BSC)模型,该模型提取的特征基具有方向性、空间性和选择性,已广泛应用于图像处理、语音处理及模式识别等领域^[2-4]。但是该模型通常用于观测数据维数等于输出数据维数(即完备基)的情况,而且当输入数据维数较大时,BSC模型的收敛速度较慢。然而,生理实验结果表明^[6-9],在大脑皮层中,初级视觉系统主视皮层V1区的编码具有超完备(特征基的维数大于输出神经元的个数)性质。超完备表示比完备表示更有效率,可以更灵活地提取信号中的内在结构^[10-12]。基于上述认识,同时考虑特征系数最大化稀疏分布和特征基的正交性,本文提出了一种改进的快速稀疏编码

(Modified Fast Sparse Coding, MFSC)策略,并采用迭代法解决了基于 L_1 范数的归一化最小二乘法和基于 L_2 范数的约束最小二乘法两个凸优化问题,能够实现完备基和过完备基的学习,并且比BSC模型有更快的收敛速度。本文采用自然场景图像和掌纹图像分别对提出的MFSC算法进行了测试,探讨了该算法在图像特征提取中的有效性和实用性,为图像特征提取提供了一种新途径。

1 BSC 算法

大脑主视皮层中稀疏编码的目的是利用少数未知基向量的加权线性组合近似表示自然界输入^[5-9]。这些基向量可以捕获输入数据的高层模式,使图像重构的误差尽可能小,同时使得响应系数尽可能地稀疏化^[13-15]。令 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T$ 表示 $n \times L$ 维的输入集合, $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m]$ 表示 $n \times m$ 维的特征基矩阵, $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m]^T$ 表示 $m \times L$ 维的输出集合,则图像的BSC模型(即Olshausen模型)优化准则^[5]表

收稿日期: 2012-09-04; **修回日期:** 2012-10-27。 **基金项目:** 国家自然科学基金资助项目(60970058); 江苏省自然科学基金资助项目(BK2009131); 江苏省“青蓝工程”项目; 2010苏州市职业大学创新团队项目(3100125)。

作者简介: 尚丽(1972-),女,安徽砀山人,副教授,博士,主要研究方向:人工智能、模式识别、数字图像处理; 苏品刚(1971-),男,江苏苏州人,副教授,主要研究方向:毫米波焦平面成像、测控; 周燕(1980-),女,江苏苏州人,讲师,博士研究生,主要研究方向:语音信号处理、模式识别。

示为:

$$\min \frac{1}{2} \|X(x, y) - \sum_{j=1}^m a_j(x, y) s_j\|_F^2 + \lambda \sum_{j=1}^m f\left(\frac{s_j}{\sigma}\right) \quad (1)$$

$$\text{s.t.} \quad \|a_j\|^2 \leq c$$

其中: (x, y) 表示像素坐标; 第一项表示图像重构误差; 第二项表示稀疏特性; 参数 $\lambda > 0$; σ 表示一个尺度常数; a_j 表示矩阵 A 中的第 j 列向量; s_j 表示矩阵 S 中的第 j 行向量; $f(\cdot)$ 表示稀疏惩罚函数, 在式(1)中选择为 $|s|$ 形式。采用轮流更新的方法实现特征基和稀疏系数的学习, 更新公式^[3-5]如下:

$$\begin{cases} \Delta s_i = \sum_j a_j(x, y) X(x, y) - \\ \sum_j \left\{ \sum_{\hat{y}} [a_i(x, y) a_j(x, y)] \right\} s_j - \left(\frac{\lambda}{\sigma} \right) f'\left(\frac{s_i}{\sigma}\right) \\ \Delta a_i(x_n, y_m) = \eta < s_i [X(x_n, y_m) - \sum_j s_j a_i(x_n, y_m)] > \end{cases} \quad (2)$$

2 MFSC 算法

2.1 目标函数

为了减少特征基训练时间, 所提出的 MFSC 模型建立如下:

$$\min_{A, S} J(A, S) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^L (X_{ij} - \sum_{k=1}^m a_k s_{kj})^2 + \delta \sum_{k=1}^m \sum_{j=1}^L f(s_{kj}) + \beta \sum_i (A_{ik}^T A_{ik}) \quad (3)$$

$$\text{s.t.} \quad \|a_j\|^2 \leq c$$

稀疏惩罚函数 $f(s)$ 可以选择下面形式中的任意一种:

$$f(s) = \begin{cases} \|s\|_1, & L_1 \text{ 形式函数} \\ (s^2 + \varepsilon)^{\frac{1}{2}}, & \varepsilon\text{-}L_1 \text{ 形式函数} \\ \log(1 + s^2), & \text{对数函数} \end{cases} \quad (4)$$

对 A 和 S 采用轮流迭代更新的方法对目标函数式(2)进行训练。注意到, 更新 A 时, 优化问题即是一个具有二次约束的最小二乘法问题。采用一般的凸优化求解法即可求解^[8]。但是, 一般的凸优化求解方法收敛的速度很慢, 训练过程较长。为了提高最优特征基的收敛速度, 当固定特征基时, 本文采用 L_1 归一化最小二乘问题实现特征符号搜索以实现稀疏系数的学习; 对任意稀疏惩罚函数, 采用拉格朗日对偶 (Lagrange Dual) 法求解 L_2 约束最小二乘问题以实现特征基的学习。在学习 MFSC 特征基时, 这种优化算法比梯度下降方法更有效。

2.2 特征系数的学习

固定特征基矩阵 A , 采用 L_1 形式的函数作为特征系数 $\{s_j^{(i)}\}$ 的惩罚函数时, 目标函数式(3)的最小化问题则转化为分别对每一个 $s_j^{(i)}$ 求解最小化问题 (即 L_1 归一化最小二乘优化问题):

$$\min_{s_j^{(i)}} \|X^{(i)} - \sum_{j=1}^m \vec{a}_j s_j^{(i)}\|_F^2 + (2\sigma^2\beta) \sum_j |s_j^{(i)}| \quad (5)$$

其中: 如果在最佳值时 $s_j^{(i)}$ 的符号已知, 且仅考虑非零系数时, 式(5)即是一个标准的无约束二次优化 (Quadratic Optimization, QP) 问题。因此, $s_j^{(i)}$ 符号的确定是一个关键。一旦确定 $s_j^{(i)}$ 的符号, 式(5)的优化问题被大幅度简化为求解下面的优化问题:

$$\min_z F(z) = \|y - Az\|_F^2 + \gamma \|z\|_1 \quad (6)$$

其中: y 代表观测向量, z 代表输出向量; γ 是一个常数。针对式

(6), 特征系数的符号搜寻步骤如下所述:

1) 初始化 $z = 0, \theta = 0$, 活跃集合 Q 为空集 $\{\}; \theta_i \in \{-1, 0, 1\}$ 标记 z_i 的符号 ($\text{sign}(z_i)$)。

2) 系数 z 为零时, 标签 i 记为 $i = \arg \max_i \left| \frac{\partial \|y - Az\|_F^2}{\partial z_i} \right|$ 。 z 不为零时, 如果 $\frac{\partial \|y - Az\|_F^2}{\partial z_i} > \gamma$,

则令 $\theta_i = -1$; 如果 $\frac{\partial \|y - Az\|_F^2}{\partial z_i} < -\gamma$, 则令 $\theta_i = +1$ 。这两种情况下的标签 i 被记入活跃集合 Q (记作 $Q := \{i\} \cup Q$)。

3) 搜索特征符号: 令 \hat{A} 表示 A 的一个子矩阵, 该子矩阵的每一列位置对应活跃集合 Q 中的标签 i ; 同时令 \hat{z} 和 $\hat{\theta}$ 分别为对应活跃集合 Q 的 z 和 θ 的子向量。对应无约束优化问题 (即 $\|y - \hat{A}\hat{z}\|^2 + \gamma \hat{\theta}^T \hat{z}$) 更新 z :

$$\hat{z}_{\text{new}} = (\hat{A}^T \hat{A})^{-1} (\hat{A}^T y - \gamma \frac{\hat{\theta}}{2}) \quad (7)$$

去掉 \hat{z} 中的零元素, 更新 $\theta = \text{sign}(z)$ 。

4) 检测最佳条件:

① 利用等式 $\frac{\partial \|y - Az\|_F^2}{\partial z_i} + \gamma \text{sign}(z_i) = 0 (\forall z_i \neq 0)$ 检测非零系数。如果该等式不成立, 则转到 3); 反之, 检测条件 ② 是否成立。

② 利用不等式 $\frac{\partial \|y - Az\|_F^2}{\partial z_i} \leq \gamma (\forall z_i \neq 0)$ 检测零系数。如果条件 ② 不成立, 则转到 2); 否则返回 z 为优化问题的解。

经过有限次迭代后, 该算法收敛到一个全局最佳值。

2.3 特征基的学习

固定稀疏系数矩阵 S , 训练特征基 A 时的优化问题转化为如下目标函数:

$$\min_A \|X - AS\|_F^2 + \beta \sum_i (A_{ik}^T A_{ik}) \quad (8)$$

$$\text{s.t.} \quad \sum_{i=1}^n A_{ij}^2 \leq c; \quad \forall j = 1, 2, \dots, m$$

这是一个具有二次约束的最小二乘法优化问题, 通常采用梯度下降优化算法求解即可, 但是采用 Lagrange 对偶法求解的效果会更好。首先考虑式(8)的拉格朗日形式:

$$L(A, \vec{\lambda}) = \text{trace}((X - AS)^T (X - AS) + \beta (A^T A)) + \sum_{j=1}^n \lambda_j \left(\sum_{i=1}^k A_{i,j}^2 - c \right) \quad (9)$$

其中 $\lambda_j \geq 0$ 为对偶变量。式(9)对应的拉格朗日对偶形式推导如下:

$$D(\vec{\lambda}) = \min_A L(A, \vec{\lambda}) = \text{trace}(X^T X - X S^T (S S^T + \text{diag}(\vec{\lambda}))^{-1} (X S^T)^T + \beta A^T A - c \cdot \text{diag}(\vec{\lambda})) \quad (10)$$

利用牛顿方法或者共轭梯度法可以优化式(10), 则推导出特征基的更新形式如下所示:

$$\tilde{A} = [(S S^T + \text{diag}(\vec{\lambda}))^{-1} (X S^T)^T + \beta A]^T \quad (11)$$

利用上述特征系数 S 和特征基 A 的学习步骤, 轮流对 S 和 A 进行更新即可实现图像最佳特征的提取。

3 实验分析

3.1 图像特征提取

选用自然风景图像和掌纹图像分别进行 MFSC 算法测试, 部分原始图像如图 1 所示。其中自然风景图像训练集合同文献[5]中用的 10 幅图像, 每幅图像为 512×512 像素; 掌

纹图像训练集合选用 PolyU 数据库^[4,13]中 100 个人的每人前 3 幅图像组成,每一幅图像尺寸为 384×284 像素。为了减少计算量,每一幅图像被减少到 128×128 像素,然后采用 8×8 大小的子窗口对每一幅自然风景图像采样 5000 次,则得到的数据集合为 64×50000 像素;采用 8×8 大小的子窗口对每一幅掌纹图像采样 200 次,得到的数据集合为 64×60000 像素。用这两类图像数据对 MFSC 算法分别进行训练,考虑特征基维数完备基和超完备基的情况,得到的两类图像特征基分别如图 2~3 所示,分别对应特征基矩阵大小为 64×64 和 64×144 。

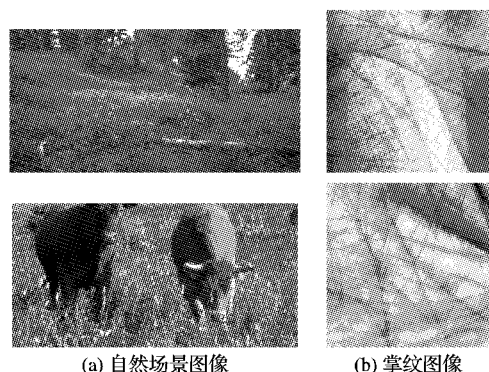


图1 原始图像样本

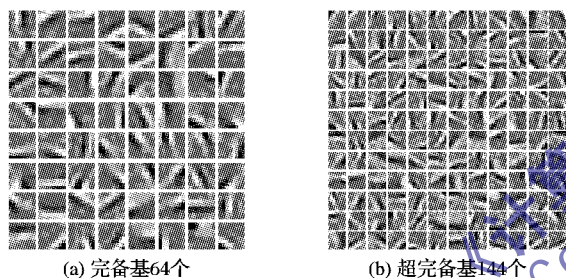


图2 自然图像特征基

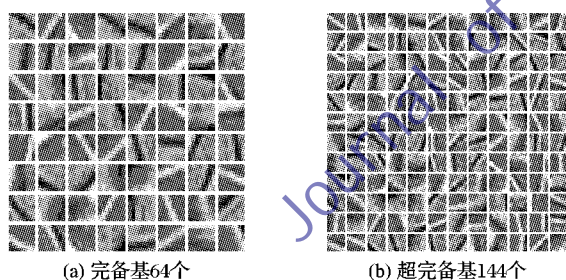


图3 掌纹图像特征基

由图 2~3 可看出:自然风景图像的特征基与文献[4]中 BSC 自然图像特征基的特性相同,具有明显的方向性和空间性;而掌纹图像特征基类似文献[5]中 BSC 训练稀疏像素集合得到的特征基,具有明显的稀疏性。测试中也发现:在相同的实验条件下,MFSC 算法比 BSC 的收敛速度要快;而且特征

基数目越大,二者收敛速度的对比越明显。这表明了与 BSC 算法相比,MFSC 算法能够用来快速、有效地模拟大脑主视皮层 VI 区简单细胞的感受野特性,可有效提取出图像的特征。

3.2 特征基的验证

进一步利用提取到的特征基进行图像重构实验从而验证 MFSC 算法在图像特征提取应用中的有效性。任意选择一幅自然风景图像和一幅掌纹图像(可为不包含在训练集合中的图像),采用随机分块、主分量分析(Principal Component Analysis, PCA)降维的方法进行预处理。考虑两类图像特征基完备和超完备情况以及不同的子图像块数,利用 MFSC 特征得到的部分重构结果如图 4~5 所示。限于篇幅,仅给出子图像块数为 1000,10000 和 50000 时的重构结果。利用信噪比(Signal-to-Noise Ratio, SNR)和均方差(Mean Square Error, MSE)评价重构图像的质量时,不同特征基、不同子图像块对应的重构图像的 SNR 值如表 1~2 所示。可以看出,完备基和超完备基情况下,子图像块数目越大,重构的图像越清晰,SNR 值越大;而且当图像块数目大于 20000 时,重构图像的目视效果和原始图像几乎没有区别。同时也可以看到,子图像块数目相同时,重构掌纹图像的 SNR 值大于自然风景图像的 SNR 值,所得实验数据证实了 MFSC 模型在图像特征提取中的有效性和实用性。

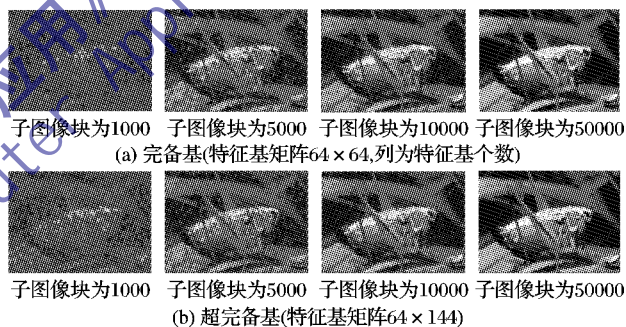


图4 自然图像重构结果

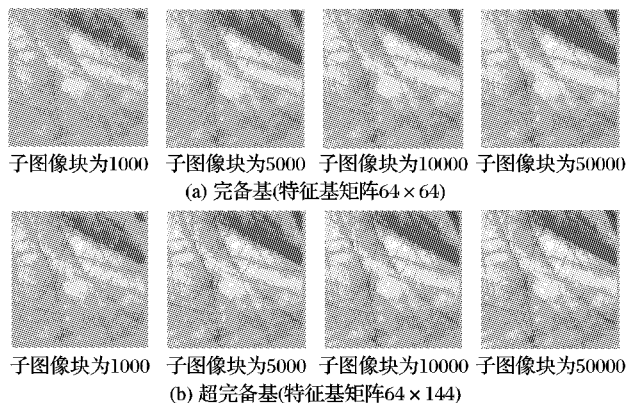


图5 掌纹图像重构结果

表1 不同子图像块数目 N 下重构自然风景图像的 MSE 和 SNR 比较

特征基类别	特征基个数	N=1000		N=5000		N=10000		N=20000		N=30000		N=50000	
		MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB
完备基	64	52.15	6.07	30.26	10.79	19.55	14.59	13.76	17.64	11.86	18.93	10.44	20.04
超完备基	144	51.90	6.11	29.37	11.05	18.30	15.16	12.45	18.51	10.55	19.95	9.27	21.07

表2 不同子图像块数目 N 下重构掌纹图像的 MSE 和 SNR 比较

特征基类别	特征基个数	N=1000		N=5000		N=10000		N=20000		N=30000		N=50000	
		MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB	MSE	SNR/dB
完备基	64	13.28	22.58	8.17	26.80	7.57	27.47	7.29	27.79	7.19	27.92	7.17	28.00
超完备基	144	12.04	23.43	6.77	28.43	6.33	29.02	6.10	29.34	6.03	29.44	5.98	29.51

4 结语

本文主要讨论了改进的快速稀疏编码模型在图像特征提取中的应用。该模型采用 L_1 归一化最小二乘问题实现了特征符号的搜索,从而实现了稀疏系数的更新;采用 L_2 约束最小二乘问题实现了特征基的更新,因而比采用梯度下降算法的标准稀疏编码模型有更快的收敛速度。考虑特征基完备和超完备情况,采用稀疏度不同的自然风景图像和掌纹图像分别进行图像特征提取测试,并利用提取的图像特征基对不同子图像块进行图像重构实验。重构图像的目视效果和相应的信噪比值证明了所提出的改进稀疏编码模型在图像特征提取中的有效性,为后续的图像消噪、模式识别等提供了一种有效的图像预处理途径。

参考文献:

- [1] 肖泉,丁兴号,王守觉,等.基于自适应超完备稀疏表示的图像去噪方法[J].仪器仪表学报,2009,30(9):1886-1890.
- [2] 成昭.信息稀疏表示算法及其在图像恢复中应用的研究[D].湘潭:湘潭大学,2011.
- [3] 麻立波.视觉超完备拓扑表示的稀疏编码计算模型[D].上海:上海交通大学,2008.
- [4] 尚丽,杜吉祥,翟传敏,等.稀疏编码算法概述[J].苏州市职业大学学报,2009,20(1):5-10.
- [5] OLSHAUSEN B A, FIELD D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images [J]. Nature, 1996, 381(6853): 607-609.
- [6] 李志清,施智平,李志欣,等.基于结构相似度的稀疏编码模型[J].软件学报,2010,21(10):2410-2419.
- [7] 路锦正,张启衡,徐智勇,等.超完备稀疏表示的图像超分辨率重构方法[J].系统工程与电子技术,2012,34(2):403-408.
- [8] LEE H, BATTLE A, RAINA R, et al. Efficient sparse coding algorithms [EB/OL]. [2010-10-10]. <http://ai.stanford.edu/~hlee/nips06-sparsecoding.pdf>.
- [9] HEILER M, SCHNORR C. Learning sparse representations by non-negative matrix factorization and sequential come programming [J]. Journal of Machine Learning Research, 2006, 7: 1385-1407.
- [10] 李清勇.视觉感知的稀疏编码理论及其应用研究[D].北京:中国科学院研究生院,2006.
- [11] 晁永国,戴芳,韩舒然,等.改进的非负稀疏编码图像基学习算法[J].计算机工程与科学,2010,32(1):77-79.
- [12] 浦剑,张军平.基于词典学习和稀疏表示的超分辨率方法[J].模式识别与人工智能,2010,23(3):335-340.
- [13] YANG J C, WRIGHT J, HUANG T, et al. Image super-resolution via sparse representation of raw image patches [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2008: 1-8.
- [14] RUBINSTEIN R, BRUCKSTEIN A M, Elad M, et al. Dictionaries for sparse representation modeling [J]. Proceedings of the IEEE, 2010, 98(6): 1045-1057.
- [15] CENSOR Y, ZENIOS S A. Parallel optimization: theory, algorithms and applications [M]. Oxford: Oxford University Press, 1997.
- [16] to image reconstruction using ordered subsets and data subdivisions [J]. IEEE Transactions on Nuclear Science, 2004, 51(1): 130-135.
- [17] MALLAT S. Multifrequency channel decomposition of images and wavelets [J]. IEEE Transactions on ASSP, 1989, 37(12): 2091-2110.
- [18] PENNEC E L, MALLAT S. Image compression with geometrical wavelets [C]// Proceedings of ICIP '2000. Piscataway, NJ: IEEE Press, 2000: 661-664.
- [19] RAMIN E, HAYDER R. Image denoising using translation invariant Contourlet transform [C]// Proceedings of 2005 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2005: 557-560.
- [20] DO M N, VETTERLI M. The Contourlet transform: an efficient directional multi-resolution image representation [J]. IEEE Transactions on Image Processing, 2005, 14(14): 2091-2106.
- [21] 蔡政,陶少华.基于小波系数关系的图像去噪方法[J].计算机应用,2011,31(9):2515-2517.
- [22] 朱呈祥,邹云.基于小波分析的分数阶系统辨识信号降噪的变尺度阈值方法[J].计算机应用,2011,31(2):543-547.
- [23] DONOHO D L, JOHNSTONE J M. Ideal spatial adaptation by wavelet shrinkage [J]. Biometrika, 1994, 81(3): 425-455.
- [24] ZHU H Q, SHU H Z, ZHOU J. Accelerating SAGE algorithm in PET image reconstruction by rescaled block-iterative method [J]. Journal of Southeast University: English Edition, 2005, 21(2): 207-210.
- [25] KADRMAS D J. Statistically regulated and adaptive EM reconstruction for emission computed tomography [J]. IEEE Transactions on Nuclear Science, 2001, 48(3): 790-798.

(上接第650页)

同时稳定求解过程,防止重建图像退化。新算法不仅仍具有加速算法收敛的优点,同时改善了 OSEM 算法在迭代一定次数后重建质量迅速下降的缺陷。

参考文献:

- [1] HERMAN G T. Image Reconstruction from Projections: implementation and application [M]. Berlin: Springer-Verlag, 1979.
- [2] SHEPP L A, VARDI Y. Maximum likelihood reconstruction for emission tomography [J]. IEEE Transactions on Medical Imaging, 1982, 1(2): 113-122.
- [3] HUDSON H M, SLARKIN R. Accelerated image reconstruction using ordered subsets of projection data [J]. IEEE Transactions on Medical Imaging, 1994, 13(4): 601-609.
- [4] FESSLER J A, HERO A O. Space alternating generalized expectation maximization algorithm [J]. IEEE Transactions on Signal Processing, 1994, 42(10): 2664-2677.
- [5] BYRNE C L. Choosing parameters in block-iterative ordered subset reconstruction algorithm [J]. IEEE Transactions on Image Processing, 2005, 14(3): 321-327.
- [6] BYRNE C L. Convergent block-iterative algorithms for image reconstruction from inconsistent data [J]. IEEE Transactions on Image Processing, 1997, 6(9): 1296-1304.
- [7] RAHEJA A, DONIERE T F, DHAWAN A P. Multi-resolution expectation maximization reconstruction algorithm for positron emission tomography using wavelet processing [J]. IEEE Transactions on Nuclear Science, 1999, 46(3): 594-602.
- [8] 何佳伟.正电子发射断层成像的统计迭代方法及加速方法研究[D].太原:中北大学,2011.
- [9] SHENG J H, LIU D R. An improved maximum likelihood approach