

文章编号:1007-2780(2015)06-1045-07

字典学习中字典尺度对 DICOM 图像压缩的影响

西霞^{1*}, 陈菲¹, 贾小林¹, 刘雨娇¹, 杨勇²
(1. 西南科技大学 计算机科学与技术学院, 四川 绵阳 621010;
2. 四川省绵阳市中心医院, 四川 绵阳 621000)

摘要:随着医院数字化医疗进程的加快,医学影像的数据量日益增大,医学影像资料的存储空间和获取速度受到很大的限制。文章在研究主流字典学习算法基础上,提出使用不同尺度的 MOD、K-SVD、ILS-DLA、RLS-DLA 字典算法对 DICOM 图像进行压缩存储,以及恢复再现的方法。与经典的 JPEG 和 JPEG2000 压缩算法相比,字典学习算法压缩和恢复效果较好,特别是采用较小尺度的字典时,压缩效果更为突出;当压缩比为 20 时,采用 4×4 尺度的 RLS-DLA 字典,论文算法的峰值信噪比(PSNR)较 JPEG 算法高出 7.8 dB,比 JPEG2000 算法高出 1 dB。
关键词:字典学习;图像压缩;DICOM 图像;字典尺度
中图分类号:TP391.41 **文献标识码:**A **doi:**10.3788/YJYXS20153006.1045

Effects of dictionary scale on dictionary learning for DICOM image compression

YOU Xia¹, CHEN Fei¹, JIA Xiao-lin¹, LIU Yu-jiao¹, YANG Yong²
(1. *School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang 621010, China;*
2. *Mianyang Central Hospital, Mianyang 621000, China*)

Abstract: With the accelerated developing of hospital digital medical, the amount of medical imaging data grows dramatically, which affects the data storage space and access speed. This paper proposes a new design which uses different scales dictionaries of MOD, K-SVD, ILS-DLA, RLS-DLA for digital imaging and communications in medicine (DICOM) image compression storage and restore methods based on dictionary learning. Compared with the traditional algorithms JPEG and JPEG2000, the pro-

收稿日期:2015-01-22;修订日期:2015-03-24.
基金项目:国家自然科学基金面上项目(No. 61471306);四川省科技厅项目(No. 16ZC1720, No. 2014JY0230);西南科技大学研究生创新基金(No. 14ycxj0058);四川省教育厅重点项目(No. 12ZD1109);绵阳网络融合工程实验室开放基金(No. 12zxwk11)
Supported by National Natural Science Foundation of China(No. 61471306);Sichuan Provincial Science and Technology Support Project(No. 16ZC1720, No. 2014JY0230);Postgraduate Innovation Fund Project by Southwest University of Science and Technology(No. 14ycxj0058);Major program of Education Department of Sichuan Province(No. 12ZD1109);Open funding Program of Network Convergence Laboratory of Mianyang(No. 12zxwk11)
* 通信联系人, E-mail: youzi_2011@yeah.net

posed method has better performance, especially when the dictionary scale is smaller. For example, when the compression ratio is 20, using 4×4 dictionary scale, the peak signal to noise ratio (PSNR) of the proposed method is 7.8 dB higher than that of JPEG, and 1dB than JPEG2000.

Key words: dictionary learning; image compression; DICOM image; dictionary scale

1 引言

随着计算机硬件技术的发展,医院能够采用大量的数字成像设备进行检查,如,计算机 X 线检查、计算机体层摄影(CT)检查、磁共振(MR)成像检查、数字超声成像检查等等,这些检查使医学数据量急剧上升。为了有效管理这些数据,越来越多的医院进行了数字化的改造,数字化医院的建立对影像归档及通信系统(Picture Archiving and Communication Systems, PACS)的需求不断增加。PACS 系统中存储影像的文件都遵循 DICOM 3.0 标准,为 DICOM 文件。通常对医学图像进行无损压缩,其恢复图像较清晰,但压缩比较低,仅为 2.5 倍左右^[1]。所以,研究出针对 DICOM 文件中的医学图像压缩方法对 PACS 系统显得尤为重要。

目前图像压缩标准有:JPEG 和 JPEG2000 等方式。JPEG 标准的正式名称为:信息技术-数字压缩和连续色调静止图像的编码。其实际上是包括无损编码模式的,但是在大多数产品中并不支持,典型使用的 JPEG 是一种基于离散余弦变换(DCT)的有损压缩方法,其压缩过程中,较高的压缩比会使得图像产生马赛克失真。JPEG2000 是在 2000 年为了取代原来的 JPEG 标准提出的图像压缩标准。其采用离散小波变换(DWT),避免了 JPEG 采用的 DCT 变换造成的马赛克失真;在数据编码上采取与 JPEG 不同的思维,使用一样的数据编码可以使用多种手段解码来得到不同质量的图像,从而增加了可扩展性和可编辑性。但是,因其编码的核心部分的算法,已经被大量注册专利,用 JPEG 2000 存在版权和专利的风险,开发出免授权费的商用编码器是不太可能的。

1999 年 K. Engan 等人提出了最优方向法(MOD)^[2],人们开始逐步探寻获取字典的方法;2006 年, M. Aharon 等人提出 K 奇异值分解(K-SVD)字典学习算法^[3],因 K-SVD 在去噪、模式

识别等的优秀表现,人们开始关注字典学习领域;2008 年 O. Bryt 等人 K-SVD 算法用于人脸图像压缩^[4],是第一次将稀疏表示和字典学习用于图像压缩的一个探索;2010 年 K. Skretting 和 K. Engan 提出了递推最小二乘字典学习算法(RLS-DLA)^[5],2011 年,他们探索了 RLS-DLA 在自然图像上的压缩效果,并与 K-SVD 算法进行对比^[6],在最终效果中,取得了优于 JPEG2000 的效果。在以上探讨的过程中,只是针对自然图像进行研究,字典的原子均为 8×8 的方块。本文针对医学常见的 DICOM 格式的颅脑 CT 图像,使用不同尺度的字典进行压缩对比。

本文第 2 部分先介绍图像信息压缩框架,然后介绍本文采用的字典学习的基本思想,最后给出压缩效果的评价指标;第 3 部分是实验过程及实验结果分析;第 4 部分总结全文并提出进一步研究的方向。

2 不同尺度字典的图像压缩方案

2.1 基于不同尺度字典的图像压缩框架

不同尺度字典的图像压缩方案主要涉及字典学习、利用字典进行压缩和解压缩。其中,字典学习也称为字典训练,主要包括:

(1)对图像的不同尺度(4×4 、 8×8 、 16×16 的图像块)的分解,得到不同尺度的图像块作为训练集;

(2)在训练集的基础上,初始化字典;

(3)使用稀疏编码算法得到初始的稀疏解矩阵,分别利用不同字典学习算法更新字典原子。

(4)重复(3),得到训练的字典。

利用字典进行压缩,主要包括以下 3 步:

(1)将带压缩的图像信息进行不同尺度的分解;

(2)使用稀疏编码算法得到稀疏解矩阵,求解过程中,设置目标峰值信噪比(PSNR);

(3)对稀疏解进行熵编码,得到压缩图像位流。

解压缩的过程是压缩过程的逆过程。压缩和解压缩的简要流程如图 1 所示。

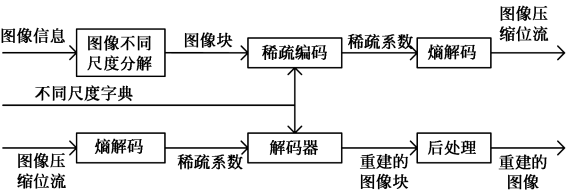


图 1 图像信息的压缩和解压缩过程
Fig. 1 Process of Image compression and decompression

2.2 字典学习算法

字典学习主要需要通过训练集得到一个字典,信号通过字典和相应的系数来表示。一个训练集 $B = \{b_i \in \mathbf{R}^N\}_{i=1}^M$,字典学习中学习得到的字典 $D \in \mathbf{R}^{N \times K}$,需要表示的向量为 $X \in \mathbf{R}^{K \times M}$,重建的矩阵为 $\tilde{X} = DW$,重建的误差为 $R = X - \tilde{X} = X - DW$,使得重建误差 R 最小,使用成本函数 $f(\cdot)$,这可以表述为最小化问题^[5,7]:

$$\arg \min_{D,W} f(D,W) = \arg \min_{D,W} \|X - DW\|_F^2. \tag{1}$$

通常,字典学习可以通过 3 步来实现,分别为:

- (1)初始化学字典;
- (2)保持已有字典 D 不变,求解稀疏解 W ;
- (3)保持已有稀疏解 W 不变,更新字典 D 。

其中,在(2)中主要用到的求解稀疏解的算法有:匹配追踪算法(MP)^[8]、正交匹配追踪算法(OMP)^[9]、顺序匹配追踪(ORMP)^[10]等。其中,MP 是一种复杂度较低的贪婪算法:在每一次迭代求解过程中,先选择与当前误差最相关的一个原子,其次基于该原子求解稀疏解对应的元素值,最后根据求解得到的稀疏解更新重构误差 R 。OMP 是 MP 的改进算法,在选择原子的过程中,使得当前误差与当前所选的字典原子相互正交,这也正是 OMP 的由来。

在(3)中主要用到字典学习算法包括:最优方向法(MOD)、K 奇异值分解法(K-SVD)、迭代最小二乘法(ILS-DLA)^[11]、递推最小二乘法(RLS-DLA)、在线字典学习法(ODL)^[12]等。其中,MOD 算法通过交替使用以上(2)和(3)步方式学习字典:在第 k 次迭代过程中,基于 $D_{(k-1)}$ 稀疏编码每一个

x_i 求解出对应的 ω_i ,并构成稀疏解矩阵 $W_{(k)}$;然后通过表达式(2)所示更新字典矩阵 D 。

$$D_{(k)} = \arg \min_D \|X - DW_{(k)}\|_F^2 = XW_{(k)}^T (W_{(k)}W_{(k)}^T)^{-1}. \tag{2}$$

与 MOD 算法每次迭代都更新整个字典 D 不同的是,K-SVD 按原子乱序依次更新每个原子来更新整个字典 D 。ILS 算法是 MOD 算法的扩展,其将训练集分为无限制的块、无限制的重叠的块、受限制的重叠的块进行最小二乘迭代更新。RLS 算法是在 ILS 算法的基础上引入遗忘因子 λ ,使得其在递推的过程中逐渐脱离初始化学字典对最后结果的影响。

本文主要 OMP 算法求解稀疏解,选择 MOD、K-SVD、ILS-DLA、RLS-DLA 四种字典学习算法在不同原子尺度下与 JPEG 和 JPEG2000 进行实验对比。

2.3 压缩效果的评价指标

实验评估指标主要从图像的客观保真度和主观保真度进行评价。

图像客观保真度常用一个指标是峰值信噪比^[13](Peak Signal to Noise Ratio, PSNR),其单位为 dB。PSNR 通常通过均方误差(Mean Square Error, MSE)进行定义,定义式(3),两个 $m \times n$ 的单色图像 I 和 K ,其中 I 是原图, K 是 I 经过压缩过后的图,那么 MSE 定义为公式(4):

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right), \tag{3}$$

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i,j) - K(i,j)\|^2, \tag{4}$$

其中: MAX_I 是表示图像点颜色的最大数值,本文中每个采样点用 8 位表示,那么就是 255。由式(3)和式(4)可知,PSNR 值越大,就代表失真越少,图像质量越高。

3 实验和结果分析

本文在真实图像上进行图像压缩。实验采用绵阳市中心医院提供的 DICOM 文件格式的颅脑 CT 图像。先进行预处理,提取图像信息后,对图像信息进行字典学习和压缩效果的比较。

3.1 DICOM 图像预处理

DICOM 文件是按照 DICOM 3.0 标准存储

的医学文件,一个单独 DICOM 文件包括一个文件头(存储有关病人的名字,扫描类型等信息)和图像数据信息。DICOM 图像是指 DICOM 文件中的图像数据,本文研究的对象也正是其图像数据信息。所以对于 DICOM 文件,石晓磊等^[14]的处理方法,通过分解文件,得到需要进行压缩的图像信息存为 BMP 格式。

3.2 字典学习过程

实验选择其中 8 幅 BMP 格式的图像,如图 2。随机选取其中的图像块,当图像块大小为 4×4 、 8×8 时,每幅图像选取 1 500 个图像块,一共 12 000 个图像块用于训练字典,图像块大小为 16×16 时,每幅图像选取 500 个图像块,一共 4 000 个图像块用于训练字典。字典大小分别采用原子大小为 4×4 、 8×8 和 16×16 ,字典原子数为 600 的字典,各个算法设置迭代次数为 200 次,使用 OMP 算法求解稀疏系数。

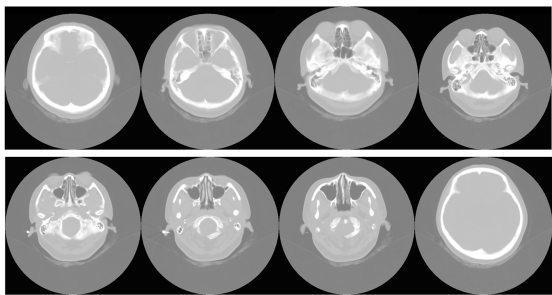


图 2 用作字典学习的图像
Fig. 2 Image for dictionary learning

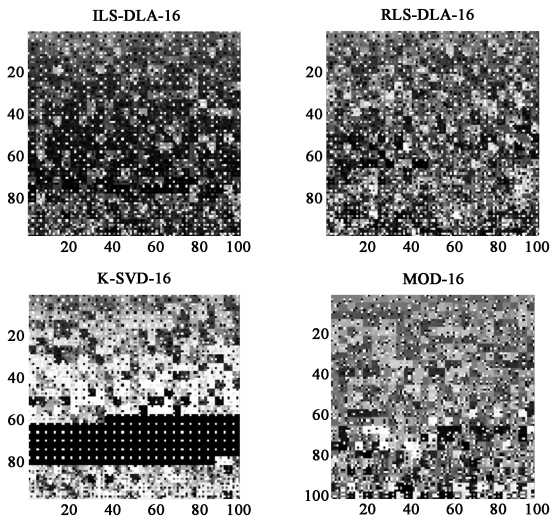


图 3 4 种 4×4 字典经过 200 次迭代的结果
Fig. 3 Four 4×4 dictionary results after 200 iterations

经过 4 种字典学习,字典原子大小为 4×4 、 8×8 和 16×16 得到的字典的如图 3、图 4 和图 5。

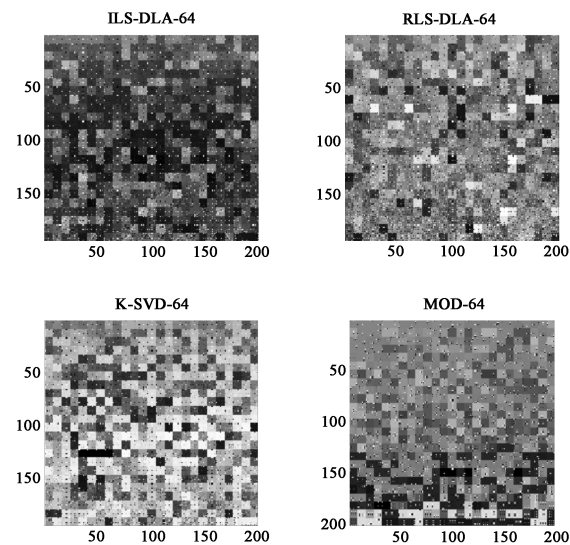


图 4 4 种 8×8 字典经过 200 次迭代的结果
Fig. 4 Four 8×8 dictionary results after 200 iterations

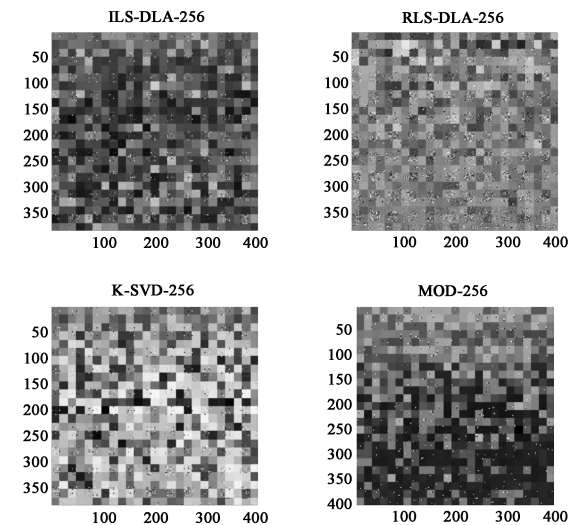


图 5 四种 16×16 字典经过 200 次迭代的结果
Fig. 5 Four 16×16 dictionary results after 200 iterations

从生成的字典可以看出,各种字典学习算法得到的字典有较大区别:

- (1)ILS-DLA 算法在不同尺度下,对比其他算法,其获取到更多的深色图像块作为字典原子的组成;
- (2)RLS-DLA 算法在不同尺度下获取到的

字典分布比较均匀,原子间差距不大;

(3)MOD 算法获取到的字典原子间的差距比较大,每种尺度下都有明显的深色区域和浅色区域,过度区域的原子较少;

(4)K-SVD 算法在 4×4 的字尺度下,获取到的字典原子间的差距最大,从图中可以明显看出有一部分基本都是由黑色组成,而在 8×8 和 16×16 的尺度下,基本全是黑色组成的原子较少,并且后两种尺度的字典较为相似。

3.3 压缩图像的客观保真度

对比实验使用 JPEG、JPEG2000 压缩方法和字典学习的方法进行。JPEG 和 JPEG2000 压缩利用 MATLAB 2011b 自带的 imwrite 函数进行实验;字典学习参考 Skretting K^[6] 的压缩过程,通过设置不同的量化参数进行实验。本文展示 32 号图(见图 8(a))进行压缩实验,在实验的结果中,尺度为 4×4 和 8×8 的字典压缩及 JPEG、JPEG2000 压缩实验结果如图 6。在比特率为 0.4 bpp 时(压缩比为 20)。尺度为 16×16 的各种字典压缩图像的 PSNR 在均 24 dB 以下,主观评价图像不可用,不在本文中列出。

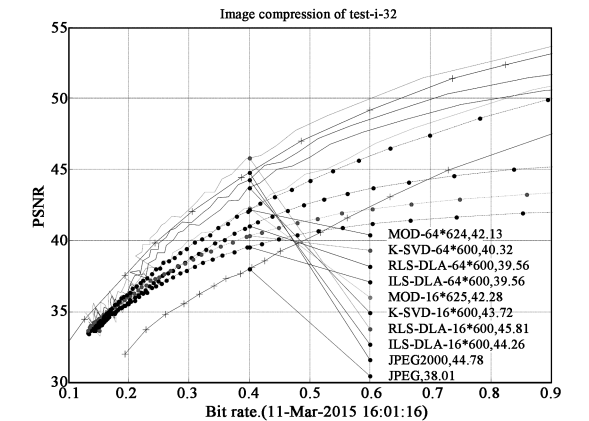


图 6 不同方法压缩结果
Fig. 6 Results of different compression methods

从图 6 实验结果中可以看出:

(1)总体上使用 4×4 字典的压缩效果优于 8×8 字典的压缩效果。在低比特率(低于 0.25 bpp)情况下,字典压缩效果和字典尺度关系不大。

(2)使用 4×4 字典的压缩效果明显优于 JPEG,图像的 PSNR 平均比 JPEG 高 5 dB,并且接近 JPEG2000,在相同比特率下 RLS-DLA 压缩后图像的 PSNR 比 JPEG2000 高 1 dB。

(3)使用 4×4 字典的压缩效果在低比特率(低于 0.5 bpp)情况下,均优于 JPEG,图像的 PSNR 平均比 JPEG 高 2 dB,但是都不及 JPEG2000。这是由于 JPEG 是采用的 DCT 变换,其在高压缩比(即低比特率)情况下,会出现明显的马赛克失真,影响其峰值信噪比。而 JPEG2000 采用的是 DWT 变换,避免马赛克失真,并且其采用复杂的编码策略,使其取得优异的压缩效果。

就不同压缩效果,进一步对字典进行分析。如图 7,纵坐标表示每个原子像素的平均值,横坐标表示原子序号。从图中可以看出当尺度为 4×4 比尺度 8×8 的字典原子之间的差距大。结合图 6,可以看出字典原子差距较大的时候,压缩效果越好。特别是 MOD 算法,两个尺度情况下,字典原子差距没有其他 3 种算法的大,因此两个尺度下压缩效果接近。

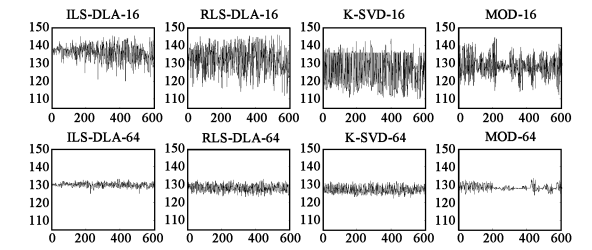
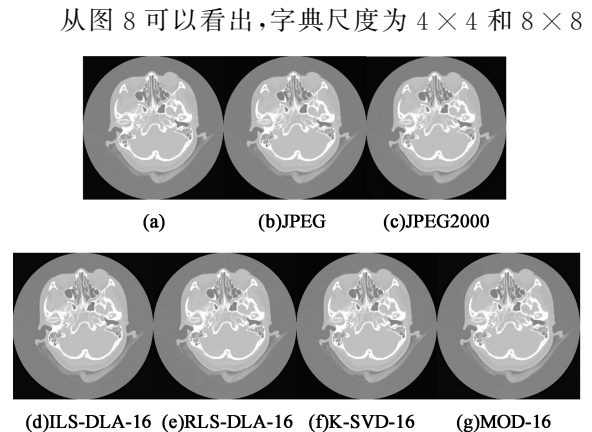
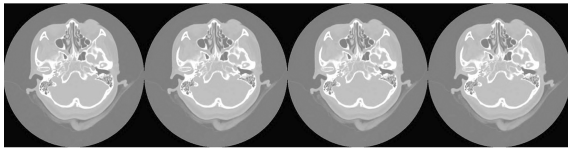


图 7 各字典原子均值对比
Fig. 7 Contrast of each dictionary atomic mean

3.4 压缩图像的主观保真度

下面从主观保真度的角度进行分析。图 8 是在压缩比为 20 时(即比特率为 0.4 bpp)时,32 号原图(图 8(a))和 JPEG、JPEG2000 以及各个字典压缩图(图片命名方式为字典名称-字典尺度)对比。





(h)ILS-DLA-64 (i)RLS-DLA-64 (j)K-SVD-64 (k)MOD-64

图 8 各算法效果对比

Fig. 8 Subjective fidelity contrast of each algorithm

时,字典学习都能够达到 JPEG 压缩效果,可以很好保存原图像的细节,人眼几乎不能分辨原图和压缩图的区别。ILS-DLA-16 和 RLS-DLA-16 解压后的图像亮度比 JPEG2000 高,更接近原始图像。

4 结 论

本文针对 DICOM 图像,使用不同尺度的字

典学习算法进行图像压缩性能的分析 and 对比。实验结果表明:字典原子间差距较小,有较多过度类型原子的时候,字典学习的压缩效果较好;利用字典学习对 DICOM 图像进行压缩,能优于现在使用的 JPEG 算法的压缩效果。并且明字典尺度较小时,压缩效果较好;当字典尺度在 4×4 时,图像压缩和恢复效果明显优于 JPEG,其中 RLS-DLA 的压缩效果优于 JPEG2000。

本文研究证明了字典学习和稀疏表达在 DICOM 图像压缩领域有良好的表现,对于 DICOM 图像压缩方法的应用具有积极意义。而且论文研究成果为以下两方面的研究奠定了基础:1)对 DICOM 图像进行多尺度字典压缩方法的研究;2)对 DICOM 图像中的感兴趣区域进行无损编码,并与字典学习方法相结合,提高图像压缩的质量。

参 考 文 献:

- [1] 李萍. 适用于 PACS 系统的医学图像压缩算法研究[D]. 郑州:郑州大学, 2013.
Li P. The study of compression method for medical image in PACS [D]. Zhenzhou: Zhenzhou University, 2013. (in Chinese)
- [2] Engan K, Aase S, Husoy J. Method of optimal directions for frame design [C]. IEEE International Conference on Acoustics, Speech and Signal Processing, Phoneix, USA, 1999: 2443-2446.
- [3] 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: 4311-4322.
- [4] Bryt O, Elad M. Compression of facial images using the K-SVD algorithm [J]. *Journal of Visual Communication and Image Representation*, 2008, 19(4): 270-282.
- [5] Skreting K, Engan K. Recursive least squares dictionary learning algorithm[J]. *IEEE Transactions on Signal Processing*, 2010, 58(4): 2121-2130.
- [6] Skreting K, Engan K. Image compression using learned dictionaries by RLS-DLA and compared with K-SVD[C]. IEEE International Conference on Acoustics, Speech and Signal Processing, Prague, Czech Republic, 2011: 1517-1520.
- [7] 霍承富. 超光谱遥感图像压缩技术的研究[D]. 合肥:中国科技大学, 2012.
Huo C F. Research on hyperspectral remote sensing image compression technique [D]. Hefei: University of Science and Technology of China, 2012. (in Chinese)
- [8] Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries[J]. *IEEE Transactions on Signal Processing*, 1993, 41(12): 3397-3415.
- [9] Pati Y, Rezaiifar R, Krislinaprasad P. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition[C]. Proceedings of the 27th Annual Asilomar Conference on Signals, Systems and Computers, Asilomar Grounds, 1993: 40-44.
- [10] Gharavi-Alkhansari M, Huang T S. A fast orthogonal matching pursuit algorithm [C]. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seattle, 1998: 1389-1392.
- [11] Engan K, Skreting K, Husoy J H. A family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation[J]. *Digital Signal Process.*, 2007, 17: 32-49.

[12] Mairal J, Bach F, Ponce J, *et al.* Online dictionary learning for sparse coding[C]. International Conference on Machine Learning, Montreal, Canada, 2009:689-696.

[13] 张健, 李宏升. 基于图论阈值算法的图像分割研究[J]. 液晶与显示, 2014, 29(4):592-597.
Zhang J, Li H S. Image mosaic research based on wavelet and rough set algorithm[J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(4):592-597. (in Chinese)

[14] 石晓磊, 王明泉. DICOM 图像格式与 BMP 图像格式的转换[J]. 微计算机信息, 2010, 26:195-197.
Shi X L, Wang M Q. Transformation of DICOM Digital Medical Image Format into BMP General Image Format [J]. *Microcomputer Information*, 2010, 26:195-197. (in Chinese)

作者简介:西霞(1990—),女,四川简阳人,硕士研究生,主要研究方向:数字图像处理,机器学习。E-mail:youzi_2011@yeah.net

陈菲(1974—),女,四川绵阳人,副教授,硕士研究生导师,主要研究方向:嵌入式系统技术及图像处理。

贾小林(1975—),男,四川绵阳人,副教授,博士,主要研究方向:数据采集与识别技术。

刘雨娇(1991—),女,四川绵阳人,硕士研究生,主要研究方向:计算机视觉。

杨勇(1974—),男,四川绵阳人,主治医师,学士,主要研究方向:影像技术方向。