



文章编号:1001-9081(2017)06-1722-06

DOI:10.11772/j.issn.1001-9081.2017.06.1722

改进耦合字典学习的脑部 CT/MR 图像融合方法

董 侠, 王丽芳*, 秦品乐, 高 媛

(中北大学 计算机与控制工程学院, 太原 030051)

(*通信作者电子邮箱 wsm2004@nuc.edu.cn)

摘要:针对目前使用单字典表示脑部医学图像难以得到精确的稀疏表示进而导致图像融合效果欠佳,以及字典训练时间过长的问题,提出了一种改进耦合字典学习的脑部计算机断层成像(CT)/磁共振成像(MR)图像融合方法。该方法首先将 CT 和 MR 图像对作为训练集,使用改进的 K 奇异值分解(K-SVD)算法联合训练分别得到耦合的 CT 字典和 MR 字典,再将 CT 和 MR 字典中的原子作为训练图像的特征,并使用信息熵计算字典原子的特征指标;然后,将特征指标相差较小的原子看作公共特征,其余为各自特征,并分别使用“平均”和“选择最大”的规则融合 CT 和 MR 字典的公共特征和各自特征得到融合字典;其次,将配准的源图像编纂成列向量并去除均值,在融合字典的作用下由系数重用正交匹配追踪(CoefROMP)算法计算得到精确的稀疏表示系数,再分别使用“2 范数最大”和“加权平均”的规则融合稀疏表示系数和均值向量;最后通过重建得到融合图像。实验结果表明,相对于 3 种基于多尺度变换的方法和 3 种基于稀疏表示的方法,所提方法融合后图像在亮度、清晰度和对比度上都更优,客观参数互信息、基于梯度、基于相位一致和基于通用图像质量指标在三组实验条件下的均值分别为:4.1133、0.7131、0.4636 和 0.7625,字典学习在 10 次实验条件下所消耗的平均时间为 5.96 min。该方法可以应用于临床诊断和辅助治疗。

关键词: 医学图像融合; K 奇异值分解; 系数重用正交匹配追踪; 稀疏表示; 字典训练

中图分类号: TP391.41 **文献标志码:** A

CT/MR brain image fusion method via improved coupled dictionary learning

DONG Xia, WANG Lifang*, QIN Pinle, GAO Yuan

(School of Computer Science and Control Engineering, North University of China, Taiyuan Shanxi 030051, China)

Abstract: The dictionary training process is time-consuming, and it is difficult to obtain accurate sparse representation by using a single dictionary to express brain medical images currently, which leads to the inefficiency of image fusion. In order to solve the problems, a Computed Tomography (CT) / Magnetic Resonance (MR) brain image fusion method via improved coupled dictionary learning was proposed. Firstly, the CT and MR images were regarded as the training set, and the coupled CT and MR dictionary were obtained through joint dictionary training based on improved K-means-based Singular Value Decomposition (K-SVD) algorithm respectively. The atoms in CT and MR dictionary were regarded as the features of training images, and the feature indicators of the dictionary atoms were calculated by the information entropy. Then, the atoms with the smaller difference feature indicators were regarded as the common features, the rest of the atoms were considered as the innovative features. A fusion dictionary was obtained by using the rule of “mean” and “choose-max” to fuse the common features and innovative features of the CT and MR dictionary separately. Further more, the registered source images were compiled into column vectors and subtracted the mean value. The accurate sparse representation coefficients were computed by the Coefficient Reuse Orthogonal Matching Pursuit (CoefROMP) algorithm under the effect of the fusion dictionary, the sparse representation coefficients and mean vector were fused by the rule of “2-norm max” and “weighted average” separately. Finally, the fusion image was obtained via reconstruction. The experimental results show that, compared with three methods based on multi-scale transform and three methods based on sparse representation, the image visual quality fused by the proposed method outperforms on the brightness, sharpness and contrast, the mean value of the objective parameters such as mutual information, the gradient based, the phase congruency based and the universal image quality indexes under three groups of experimental conditions are 4.1133, 0.7131, 0.4636 and 0.7625 respectively, the average time in the dictionary learning phase under 10 experimental conditions is 5.96 min. The proposed method can be used for clinical diagnosis and assistant treatment.

Key words: medical image fusion; K-means-based Singular Value Decomposition (K-SVD); Coefficient Reuse Orthogonal Matching Pursuit (CoefROMP); sparse representation; dictionary training

收稿日期:2016-11-18;修回日期:2017-02-03。 基金项目:山西省自然科学基金资助项目(2015011045)。

作者简介: 董侠(1992—),女,山西临汾人,硕士研究生,主要研究方向:医学图像融合、机器学习; 王丽芳(1977—),女,山西长治人,副教授,博士,主要研究方向:机器视觉、大数据处理、医学图像处理; 秦品乐(1978—),男,山西长治人,副教授,博士,主要研究方向:机器视觉、大数据处理、三维重建; 高媛(1972—),女,山西太原人,副教授,硕士,主要研究方向:大数据处理、医学图像处理、三维重建。



0 引言

在医学领域,医生需要对同时具有高空间和高光谱信息的单幅图像进行研究和分析,以便对疾病进行准确诊断和治疗^[1]。这种类型的信息仅从单模态图像中无法获取,例如:计算机断层成像(Computed Tomography, CT)能够捕捉人体的骨结构,具有较高的分辨率,而磁共振成像(Magnetic Resonance, MR)能够捕捉人体器官的软组织如肌肉、软骨、脂肪等细节信息。因此,将CT和MR图像的互补信息相融合以获取更全面丰富的图像信息,可为临床诊断和辅助治疗提供有效帮助^[2]。

目前应用于脑部医学图像融合领域比较经典的方法是基于多尺度变换的方法:离散小波变换(Discrete Wavelet Transform, DWT)^[3]、平稳小波变换(Stationary Wavelet Transform, SWT)^[4]、双树复小波变换(Dual-Tree Complex Wavelet Transform, DTCWT)^[5]、拉普拉斯金字塔(Laplacian Pyramid, LP)^[6]、非下采样Contourlet变换(Non-Subsampled Contourlet Transform, NSCT)^[7]。基于多尺度变换的方法能够很好地提取图像的显著特征,并且计算效率较高,但是多尺度分解水平不易确定,分解层数太少则不能提取足够多的细节信息,分解层数太多则融合过程对图像误配准更加敏感,分解层数需要在细节提取能力和对误配准鲁棒性之间做出折中。另外,传统的融合策略也会导致融合图像对比度丢失。随着压缩感知^[8]的兴起,基于稀疏表示的方法^[9-10]被广泛用于图像融合领域,并取得了极佳的融合效果。Yang等^[11]使用冗余的离散余弦变换(Discrete Cosine Transform, DCT)字典稀疏表示源图像,并使用“选择最大”的规则融合稀疏系数。DCT字典是一种隐式字典,易于快速实现,但是其表示能力有限。Aharon等^[12]提出K奇异值分解(K-means-based Singular Value Decomposition, K-SVD)算法用于从训练图像中学习字典。与DCT字典相比,学习的字典是一种自适应于源图像的显式字典,具有较强的表示能力。学习的字典中将仅从自然图像中采样训练得到的字典称为单字典,单字典可以表示任意一幅与训练样本类别相同的自然图像,但对于结构复杂的脑部医学图像,使用单字典既要表示CT图像又要表示MR图像,难以得到精确的稀疏表示系数。Ophir等^[13]提出小波域上的多尺度字典学习方法,即在小波域上对所有子带分别使用K-SVD算法训练得到所有子带对应的子字典。多尺度字典有效地将解析字典和学习的字典的优势相结合,能够捕捉图像在不同尺度和不同方向上包含的不同特征。但是所有子带的子字典也是单字典,使用子字典对所有子带进行稀疏表示仍然难以得到精确的稀疏表示系数,并且分离的字典学习时间效率较低。Yu等^[14]提出基于联合稀疏表示的图像融合方法兼具去噪功能。这种方法是对待融合的已配准源图像本身进行字典学习,根据联合稀疏模型的第一种模型(Joint Sparsity Model-1, JSM-1)提取待融合图像的公共特征和各自特征,再分别组合并重构得到融合图像。这种方法由于是对待融合源图像本身训练字典所以适用于脑部医学图像,可以得到精确的稀疏表示系数;但是对于每对待融合的源图像都需要训练字典,时间效率低,缺乏灵活性。

针对上述问题,本文提出一种适用于脑部CT/MR图像融合的方法。不同于单字典的样本采集方式,该方法首先将

高质量的脑部CT/MR图像对作为训练集,使用改进的K-SVD算法联合训练分别得到耦合的CT字典和MR字典,再结合空间域的方法融合CT和MR字典得到融合字典。融合字典很好地保留了CT和MR训练图像的特征,使用这样一个融合字典就可以同时准确地表示CT和MR源图像并得到精确的稀疏表示系数。其次,在耦合的特征空间下使用改进的K-SVD算法联合训练字典较其他几种经典字典学习方法提高了字典学习的时间效率。最后,在融合阶段将源图像列向量的均值去掉,这样只保留了图像纹理信息而不是图像的强度值,再使用不同的融合规则组合重构各部分得到融合图像。实验结果表明本文提出的方法在主观视觉和客观参数上性能较优。

1 稀疏表示和字典学习

稀疏表示理论的基本假设是:任何给定的自然信号都可以表示为字典中原子的稀疏线性组合^[15]。对给定的信号 $\mathbf{x} \in \mathbb{R}^n$,信号 \mathbf{x} 的稀疏表示可以描述为 $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$,其中矩阵 $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m) \in \mathbb{R}^{n \times m}$ 是一个字典,每个列向量 $\mathbf{d}_i \in \mathbb{R}^n$ ($i = 1, 2, \dots, m$)是字典 \mathbf{D} 的一个原子,向量 $\boldsymbol{\alpha} \in \mathbb{R}^m$ 是含有一些非零元素的稀疏系数。稀疏系数 $\boldsymbol{\alpha}$ 的计算式如下:

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \quad (1)$$

$$\text{s. t. } \|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \varepsilon$$

式中:向量 $\boldsymbol{\alpha}$ 的稀疏水平是由 l_0 范数衡量的; $\|\boldsymbol{\alpha}\|_0$ 表示向量 $\boldsymbol{\alpha}$ 中非零元素的个数; ε 表示允许偏差的精度。问题(1)的求解过程称为“稀疏编码”。

在稀疏表示中,字典 \mathbf{D} 的选择是非常关键的。构造字典的两种主要方法是:解析方法和基于学习的方法。与解析方法相比,基于学习的方法可以提高字典的性能和灵活性,应用范围更广。

具体地,令 $\{\mathbf{x}_i\}$ ($i = 1, 2, \dots, Num$)是一个包含 Num 个样本的训练集,把这些样本排列成列组成一个矩阵 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{Num}) \in \mathbb{R}^{n \times Num}$ 。字典学习就是从 \mathbf{X} 中学习一个字典 $\mathbf{D} \in \mathbb{R}^{n \times m}$,则 $\mathbf{X} = \mathbf{DA}$, $\mathbf{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_{Num}) \in \mathbb{R}^{m \times Num}$ 是稀疏系数矩阵。最有代表性的字典学习算法是最优方向法(Method of Optimal Directions, MOD)^[16]和K-SVD算法^[12]。这两种学习算法都是通过最小化稀疏逼近误差 $\|\mathbf{X} - \mathbf{DA}\|_F^2$ 学习字典,对应的优化问题如下:

$$\min_{\mathbf{D}, \mathbf{A}} \|\mathbf{X} - \mathbf{DA}\|_F^2 \quad (2)$$

$$\text{s. t. } \forall i, \|\boldsymbol{\alpha}_i\|_0 \leq \tau$$

式中:符号 $\|\cdot\|_F$ 表示F范数; τ 是稀疏矩阵 \mathbf{A} 的稀疏度。MOD和K-SVD都是迭代算法,通过交替地执行稀疏编码和字典更新这两个步骤得到问题(2)的解。

2 改进的耦合字典学习和融合字典

精确的稀疏表示是图像融合成功的关键。经典的模型是从自然图像当中训练一个超完备字典来描述源图像。与自然图像相比,医学图像具有灰度不均匀、对比度低等特点,使用这样的超完备字典并不能精确地表示CT和MR图像。基于此,本文提出使用改进的K-SVD算法联合训练分别得到耦合的CT字典和MR字典,耦合的字典学习能够更高效完整地提取CT和MR图像的特征,然后再结合空间域的方法融合CT字典和MR字典得到融合字典。由于融合字典同时包含了



CT 和 MR 图像的特征,所以使用融合字典表示这两种源图像能够得到更精确的稀疏表示系数,从而提高脑部医学图像融合的质量。本文将提出的基于改进耦合字典学习的脑部 CT/MR 图像融合方法简记为 ICDL(Improved Coupled Dictionary Learning)。图 1 为 ICDL 方法的流程。

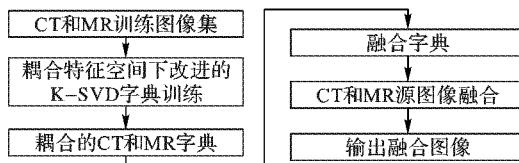


图 1 ICDL 方法的流程

Fig. 1 Flow chart of ICDL method

2.1 改进的耦合字典学习

本文使用的训练集是 8 个已经配准的 CT 和 MR 图像对,这些图像对来自哈佛大学医学院^[17],如图 2 所示。从训练集中采样得到向量对 $\{X^C, X^R\}$,定义 $X^C \triangleq (x_1^C, x_2^C, \dots, x_n^C) \in \mathbf{R}^{d \times n}$ 为 n 个采样的 CT 图像向量组成的矩阵, $X^R \triangleq (x_1^R, x_2^R, \dots, x_n^R) \in \mathbf{R}^{d \times n}$ 为对应的 n 个采样的 MR 图像向量组成的矩阵。根据文献[18] 和 [19],本文的目标是在耦合的特征空间下训练两个字典 $D^C, D^R \in \mathbf{R}^{d \times N}$,使得 CT 和 MR 训练图像在相应的字典作用下具有相同的稀疏表示系数 $A \in \mathbf{R}^{N \times n}$ 。图 3 为 ICDL 方法的字典学习过程。

传统的耦合字典学习问题可表示为:

$$\min_{D^C, D^R, A} \|X^C - D^C A\|_2^2 + \|X^R - D^R A\|_2^2 \quad (3)$$

$$\text{s.t. } \|A\|_0 \leq \tau, \|d_i^C\|_2^2 \leq 1, \|d_i^R\|_2^2 \leq 1, \forall 1 \leq i \leq N$$

其中: A 是 X^C 和 X^R 的联合稀疏编码; d_i^C 和 d_i^R 分别是 D^C 和 D^R 字典原子的第 i 列; τ 是稀疏矩阵 A 的稀疏度。

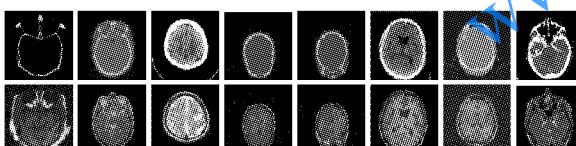
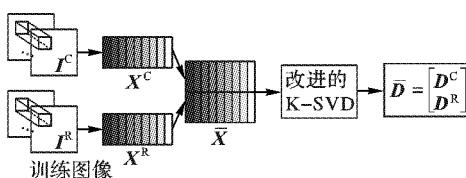
图 2 训练集
Fig. 2 Training set

图 3 ICDL 方法的字典学习过程

Fig. 3 Dictionary learning process of ICDL method

本文的耦合字典训练使用改进的 K-SVD 算法^[20],该算法是在字典学习代价函数式(3)中加入支撑完整的先验信息^[21],交替地更新 D^C, D^R 和 A ,对应的训练优化问题如下:

$$\min_{D^C, D^R, A} \|X^C - D^C A\|_2^2 + \|X^R - D^R A\|_2^2 \quad (4)$$

$$\text{s.t. } \|A\|_0 \leq \tau, A \odot M = 0$$

式中: \odot 代表点乘;掩膜矩阵 M 由元素 0 和 1 组成,定义为 $M = \{A \neq 0\}$,等价于如果 $A(i, j) = 0$ 则 $M(i, j) = 1$,否则为 0。因此 $A \odot M = 0$ 能使 A 中所有零项保持完备。引入辅助变量:

$$\bar{X} = \begin{pmatrix} X^C \\ X^R \end{pmatrix} \quad (5)$$

$$\bar{D} = \begin{pmatrix} D^C \\ D^R \end{pmatrix} \quad (6)$$

则问题式(4)可以等价地转化为:

$$\min_{\bar{D}, A} \|\bar{X} - \bar{D} A\|_2^2$$

$$\text{s.t. } \|A\|_0 \leq \tau, A \odot M = 0 \quad (7)$$

式(7)的求解过程分为稀疏编码和字典更新两个步骤。

首先,随机矩阵初始化字典 D_0^C 和 D_0^R 。通过求解式(8)来实现对系数矩阵 A 的更新:

$$A = \arg \min_A \|\bar{X} - \bar{D} A\|_2^2$$

$$\text{s.t. } A \odot M = 0 \quad (8)$$

分别对系数矩阵 A 中每一列的非零元素进行处理,而保持零元素完备,则式(8)可以转换为式(9):

$$\alpha_i = \arg \min_{\alpha_i} \|\bar{x}_i - \bar{D}_i \alpha_i\|_2^2 \quad (9)$$

式中: \bar{D}_i 是 \bar{D} 对应 A 的非零支集的子矩阵; α_i 是 A 第 i 列的非零部分。问题(9)由系数重用正交匹配追踪(Coefficient Reuse Orthogonal Matching Pursuit, CoefROMP) 算法^[19]求解,由此可得到更新的稀疏系数矩阵 A 。

其次,在字典更新阶段,式(7)的优化问题可以转化为式(10):

$$\{\bar{D}, A\} = \arg \min_{\bar{D}, A} \|\bar{X} - \bar{D} A\|_2^2$$

$$\text{s.t. } A \odot M = 0 \quad (10)$$

则式(10)的补偿项可以写为:

$$\begin{aligned} \|\bar{X} - \bar{D} A\|_2^2 &= \|\bar{X} - \sum_{j=1}^N \bar{d}_j \alpha_j^T\|_2^2 = \\ &\|\left(\bar{X} - \sum_{j \neq k} \bar{d}_j \alpha_j^T\right) \odot (1_d \cdot m_j^T) - \bar{d}_k \alpha_k^T\|_2^2 = \\ &\|E_k - \bar{d}_k \alpha_k^T\|_2^2 \end{aligned} \quad (11)$$

式中: \bar{d}_k 表示字典 \bar{D} 中待更新的第 k 列; α_k^T 表示稀疏系数矩阵 A 的第 k 行; m_j^T 表示掩膜矩阵 M 的第 j 行用来保证 α_k^T 中的零元素在正确的位置。掩膜矩阵 $1_d \cdot m_j^T$ 是将行向量 m_j^T 复制 d 次得到尺寸为 $d \times n$ 的秩为 1 的矩阵,利用掩膜矩阵 $1_d \cdot m_j^T$ 可以有效地去除 $\bar{X} - \sum_{j \neq k} \bar{d}_j \alpha_j^T$ 中那些未用到第 k 个原子所对应样本的列。对误差矩阵 E_k 进行奇异值分解(SVD)得到 $E_k = U \Delta V^T$,使用矩阵 U 的第一列更新字典 \bar{D} 中的原子 \bar{d}_k ,同时将稀疏系数矩阵 A 中的 α_k^T 更新为矩阵 V 的第一列与 $\Delta(1, 1)$ 的乘积。

最后,循环执行稀疏编码和字典更新这两个阶段,直至达到预设的迭代次数为止,输出一对耦合的 D^C 和 D^R 字典。由于字典更新阶段同时更新字典和稀疏表示系数的非零元素,使得字典的表示误差更小且字典的收敛速度更快。在稀疏编码阶段,考虑到每次迭代时都忽略前一次迭代的表示,CoefROMP 算法提出利用上次迭代的稀疏表示残差信息进行系数更新,从而更快地得到所要求问题的解^[21]。

2.2 融合字典

为构造一个同时包含 CT 和 MR 两种图像特征的字典,使得这个字典能够同时精确地表示 CT 和 MR 源图像。考虑到 2.1 节中 CT 和 MR 训练图像在一对耦合字典的作用下具有相同的稀疏表示系数,因此将 CT 和 MR 字典中的原子作为训练图像的特征,再结合空间域的方法融合 CT 和 MR 字典得到融合字典。

具体地,令 $D^C, D^R \in \mathbf{R}^{d \times N}, L^C(n)$ 和 $L^R(n)$ ($n = 1, 2, \dots$,



N) 分别代表 CT 字典和 MR 字典的第 n 个原子的特征指标。由于脑部 CT 和 MR 图像是对应于人体同一部位由不同成像设备获取的图像,因此两者之间一定存在着公共特征和各自特征。本文提出:将特征指标相差较大的原子看作各自特征,使用“选择最大”规则融合;特征指标相差较小的原子看作公共特征,使用“平均”的规则融合。融合字典的计算式表示如下:

$$\mathbf{D}_F(n) = \begin{cases} \mathbf{D}^C(n), & L^C(n) > L^R(n) \text{ 且} \\ |L^C(n) - L^R(n)| / |L^C(n) + L^R(n)| \geq \lambda \\ \mathbf{D}^R(n), & L^C(n) \leq L^R(n) \text{ 且} \\ |L^C(n) - L^R(n)| / |L^C(n) + L^R(n)| \geq \lambda \\ (\mathbf{D}^C(n) + \mathbf{D}^R(n)) / 2, & \text{其他} \end{cases} \quad (12)$$

式中:设 $\lambda = 0.25$, $\mathbf{D}_F \in \mathbb{R}^{d \times N}$ 表示融合字典。根据医学图像的物理特性,使用信息熵^[22]作为特征指标。这种方法将稀疏域和空间域的方法结合起来,考虑了医学图像的物理特性计算字典原子的特征指标,与稀疏域的方法相比,具有更加明确的物理意义。

3 脑部 CT/MR 图像融合过程

不失一般性,设已经配准的脑部 CT/MR 源图像 $I_C, I_R \in \mathbb{R}_{MN}$ 。图 4 为脑部 CT/MR 图像融合流程,融合过程如下:

1) 预处理阶段。使用步长为 1 的滑动窗把源图像 I_C, I_R 分为 $\sqrt{m} \times \sqrt{m}$ 大小的图像块,然后将这些提取的图像块编纂成 m 维列向量。对于每幅源图像 I_C, I_R ,都有 $(M - \sqrt{m} + 1)(N - \sqrt{m} + 1)$ 个图像块。将两幅源图像中的第 j 个图像块拉成的列向量记为和 x_C^j 和 x_R^j ,并减去各自的均值:

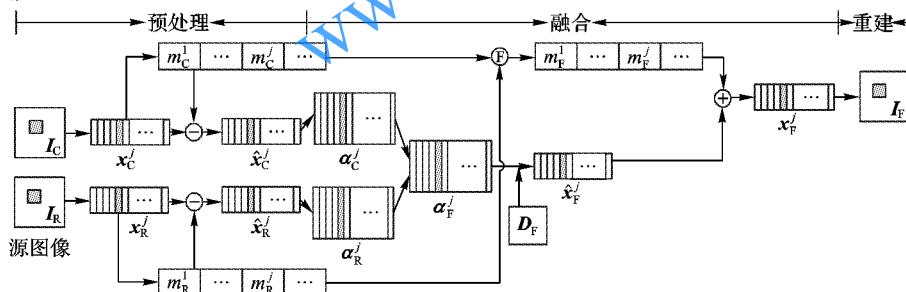


图 4 脑部 CT/MR 图像融合
Fig. 4 Fusion of CT/MR brain images

4 实验结果与分析

本文的实验环境为:32 位 Windows7 操作系统、Matlab R2013a、Intel i3-2350 2.3 GHz 处理器、4 GB 运行内存。为验证本文方法的有效性,选取三组已经配准的脑部 CT/MR 图像进行融合,分别为正常脑部 CT/MR(图 5(a)、(b))、脑萎缩 CT/MR(图 6(a)、(b))、脑肿瘤 CT/MR(图 7(a)、(b)),图片大小均为 256×256 。选取的对比算法有:离散小波变换(DWT)^[3]、平稳小波变换(SWT)^[4]、非下采样 Contourlet 变换(NSCT)^[7]、传统基于选择最大的稀疏表示方法(Sparse Representation “choose-Max”-based, SRM)^[11]、基于 K-SVD 的稀疏表示方法(Sparse Representation based on K-SVD, SRK)^[12]、基于多尺度字典学习的方法(Multi-scale Dictionary Learning, MDL)^[13]。

基于多尺度变换的方法中,对于 DWT 和 SWT 方法,分解水平都设为 3,小波基分别设为“db6”和“bior1.1”。NSCT 方

$$\hat{x}_C^j = x_C^j - m_C^j \cdot \mathbf{1} \quad (13)$$

$$\hat{x}_R^j = x_R^j - m_R^j \cdot \mathbf{1} \quad (14)$$

其中: m_C^j 和 m_R^j 分别是 x_C^j 和 x_R^j 中所有元素的均值; $\mathbf{1}$ 表示一个全 1 的 m 维列向量。

2) 融合阶段。使用 CoefROMP 算法^[17]求解 $\{\hat{x}_C^j, \hat{x}_R^j\}$ 的稀疏表示系数,计算式如下:

$$\alpha_C^j = \arg \min_{\alpha} \|\alpha\|_0 \quad (15)$$

s. t. $\|\hat{x}_C^j - D_F \alpha\|_2 < \varepsilon$

$$\alpha_R^j = \arg \min_{\alpha} \|\alpha\|_0 \quad (16)$$

s. t. $\|\hat{x}_R^j - D_F \alpha\|_2 < \varepsilon$

其中: D_F 是由 2.2 节提出的融合字典。

将稀疏系数的 l_2 范数作为源图像的活跃度测量,则稀疏系数 α_C^j 和 α_R^j 通过以下的融合规则融合:

$$\alpha_F^j = \begin{cases} \alpha_C^j, & \|\alpha_C^j\|_2 \geq \|\alpha_R^j\|_2 \\ \alpha_R^j, & \text{其他} \end{cases} \quad (17)$$

均值 m_C^j 和 m_R^j 使用“加权平均”规则融合:

$$m_F^j = w \cdot m_C^j + (1 - w) \cdot m_R^j \quad (18)$$

其中, $w = 1 / (1 + \exp(-\beta(m_C^j - m_R^j)))$, $\beta > 0$ 。则 x_C^j 和 x_R^j 的融合结果为:

$$x_F^j = D_F \alpha_F^j + m_F^j \cdot \mathbf{1} \quad (19)$$

3) 重建阶段。对所有的图像块都执行上述两个步骤以得到所有图像块的融合结果。对于每个块向量 x_F^j ,通过反滑窗的过程重塑成 $\sqrt{m} \times \sqrt{m}$ 的图像块并放回到对应的像素位置,再对重复像素取平均得到最终的融合图像 I_F 。

法使用“9-7”金字塔滤波器和“c-d”方向滤波器,分解水平设为 $\{2^2, 2^2, 2^3, 2^4\}$ 。基于稀疏表示的方法中滑动步长为 1,图像块大小均为 8×8 ,字典大小均为 64×256 ,误差 $\varepsilon = 0.01$,稀疏度 $\tau = 6$ 。ICDL 方法使用改进的 K-SVD 算法,执行 6 个字典更新周期(Dictionary Update Cycles, DUC)和 30 次迭代。

由图 5~7 可以看出,DWT 方法的融合图像边缘纹理模糊,图像信息失真且存在块效应;与 DWT 方法相比,SWT 和 NSCT 方法的融合质量相对较好,图像的亮度、对比度、清晰度有了很大的提升,但仍存在边缘亮度失真,软组织和病灶区域存在伪影的问题;SRM 和 SRK 方法较基于多尺度变换的方法,图像的骨组织和软组织更加清晰,伪影也有所减少,能很好地识别病灶区域;MDL 方法与 SRM 和 SRK 方法相比,能保留更多的细节信息,图像质量取得进一步的改善,但仍有部分伪影存在;本文提出的 ICDL 方法在图像的亮度、对比度、清晰度和细节的保持度上都优于其他方法,融合图像没有伪影,骨组织、软组织和病灶区域显示清晰,有助于医生诊断。

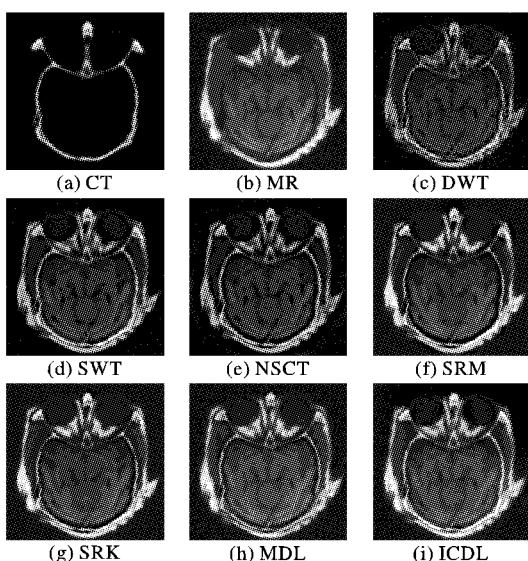


图5 正常脑部的CT/MR融合结果
Fig. 5 CT/MR fusion results of normal brain

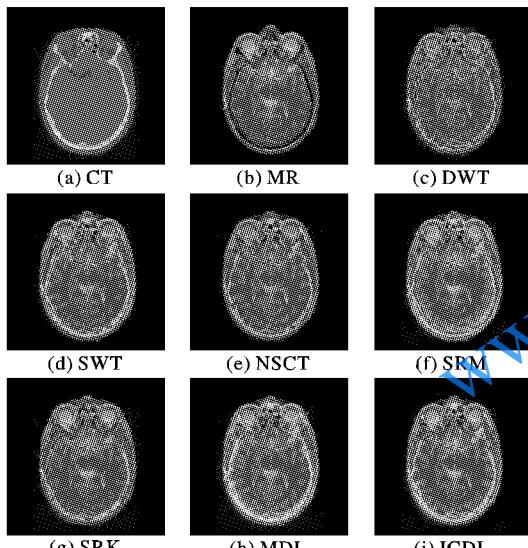


图6 脑萎缩的CT/MR融合结果
Fig. 6 CT/MR fusion results of brain atrophy

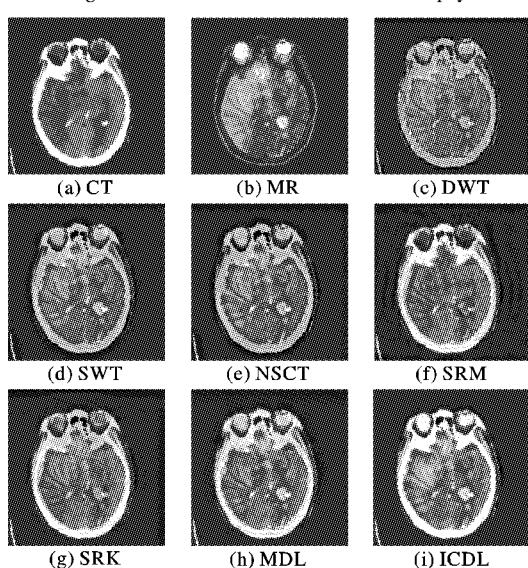


图7 脑肿瘤的CT/MR融合结果
Fig. 7 CT/MR fusion results of brain tumors

为进一步验证本文方法的有效性,使用互信息(Mutual Information, MI)^[23]、基于梯度 $Q^{AB/F}$ ^[24]、基于相位一致 Q_p ^[25]和基于通用图像质量指标 Q_w ^[26]来对融合图像进行客观评价。MI反映融合图像从源图像提取的信息量的多少,提取的信息量越多则融合图像的质量越好; $Q^{AB/F}$ 反映融合图像对源图像边缘特性的保留情况,取值范围在[0,1],越接近1表示融合图像的边缘越清晰; Q_p 反映融合图像多种类型的显著特征,如边缘和角点等; Q_w 反映融合图像在系数相关性、光照和对比度方面与源图像的关联,符合人类视觉系统的特点。

表1~3为三组实验条件下不同融合方法的客观评价指标。

表1 正常脑部CT/MR融合结果性能比较

Tab. 1 Performance comparison of CT/MR fusion results of normal brain

图像	融合方法	评价指标			
		MI	$Q^{AB/F}$	Q_p	Q_w
图5	DWT	2.6120	0.5445	0.2836	0.7098
	SWT	2.6880	0.7307	0.4519	0.7427
	NSCT	2.6277	0.7062	0.4542	0.7459
	SRM	3.0122	0.6986	0.4613	0.7504
	SRK	3.4351	0.7129	0.4687	0.7557
	MDL	3.8501	0.7452	0.4601	0.7615
	ICDL	4.0645	0.7682	0.4738	0.7796

表2 脑萎缩CT/MR融合结果性能比较

Tab. 2 Performance comparison of CT/MR fusion results of brain atrophy

图像	融合方法	评价指标			
		MI	$Q^{AB/F}$	Q_p	Q_w
图6	DWT	3.6143	0.4809	0.2756	0.6845
	SWT	3.6332	0.6311	0.4283	0.7055
	NSCT	3.4966	0.6247	0.4266	0.7140
	SRM	3.7492	0.6132	0.4141	0.7152
	SRK	3.8772	0.6259	0.4240	0.7233
	MDL	4.1913	0.6404	0.4134	0.7103
	ICDL	4.4398	0.6751	0.4397	0.7385

表3 脑肿瘤CT/MR融合结果性能比较

Tab. 3 Performance comparison of CT/MR fusion results of brain tumors

图像	融合方法	评价指标			
		MI	$Q^{AB/F}$	Q_p	Q_w
图7	DWT	2.2467	0.4255	0.2952	0.6957
	SWT	2.9572	0.5957	0.4258	0.7137
	NSCT	3.1088	0.5715	0.4265	0.7209
	SRM	3.3281	0.6132	0.4314	0.7105
	SRK	3.4485	0.6391	0.4471	0.7359
	MDL	3.6381	0.6680	0.4650	0.7529
	ICDL	3.8355	0.6960	0.4774	0.7694

从表1~3可以看出,DWT方法在MI、 $Q^{AB/F}$ 、 Q_p 和 Q_w 四项指标上均不理想,主要是因为离散小波变换具有平移变化性、走样和缺乏方向性等缺陷; SWT和NSCT方法在上述四项指标上均明显优于DWT方法是由于它们具有平移不变性; SRM、SRK和MDL方法在MI指标上优于三层分解水平的多尺度变换的方法,而在其他三项指标上存在一些低于多尺度变换方法的情况,这可能是由于字典的表示能力不足以提取足够的细节信息和 l_1 范数取极大的融合规则导致的图像的灰度不连续效应; 经计算,本文提出的ICDL方法的MI、 $Q^{AB/F}$ 、



Q_p 和 Q_w 四项指标在上述三组实验条件下的均值分别为 4.1133、0.7131、0.4636 和 0.7625。与其他几种方法相比:本文提出的 ICDL 方法在 MI 指标上的值最大,表明该方法的融合图像从源图像提取的信息最丰富;同时在 $Q^{AB/F}$ 、 Q_p 和 Q_w 三项指标上对应的值也最大,表明 ICDL 方法对图像误配准处理得较好,图像融合质量较高,其主客观评价一致。

为验证本文提出的字典学习方法的时间效率,分别计算不同的图像融合方法的字典学习阶段在相同条件下进行 10 次实验所消耗时间的平均值。SRK、MDL、CDL 和 ICDL 方法的字典训练时间均值分别为 4.61 min、42.59 min、81.27 min 和 5.96 min。SRK 方法的字典学习中训练数据是从 50 幅自然图像中提取 5 万个 8×8 图像块;MDL 方法的字典学习中训练数据是在小波域中分别对每个子带提取 5 万个 8×8 图像块;传统的耦合字典学习 (Coupled Dictionary Learning, CDL) 方法和本文提出的 ICDL 方法的字典学习中训练数据均是从图 2 的训练集中提取 5 万对 8×8 的图像块。下面对四种字典学习方法的时间复杂度进行分析。假设训练图像的尺寸为 $N \times N$,提取 N^2 个大小为 $\sqrt{n} \times \sqrt{n}$ 的训练图像块。SRK 方法的时间复杂度为 $O(N^2)$;MDL 方法的时间复杂度为 $O((3S + 1)N^2)$,其中 S 为小波变换的分解水平;CDL 方法的时间复杂度为 $O(N^4)$,这是因为两个字典交替地更新影响了字典学习的效率;ICDL 方法的时间复杂度为 $O(N^2)$,这是由于字典更新阶段同时更新字典和稀疏表示系数的非零元素而使零元素保持完备,降低了矩阵的维数,同时使用联合学习的策略,进一步提高了字典学习的时间效率。

5 结语

针对目前基于字典学习的多模态脑部医学图像融合中,使用单字典难以得到精确的稀疏表示进而影响脑部医学图像融合的质量,以及字典训练时间过长的问题,提出了一种改进耦合字典学习的脑部 CT/MR 图像融合方法。该方法分别对正常脑部、脑萎缩和脑肿瘤三组脑部医学图像进行了多次实验,结果表明本文提出的 ICDL 方法与基于多尺度变换的方法,传统基于选择最大的稀疏表示方法、基于 K-SVD 的稀疏表示方法以及多尺度字典学习的方法相比,不仅提高了脑部医学图像融合的质量,而且有效降低了字典训练的时间,能为临床医疗诊断提供有效帮助。但仍存在以下需要进一步研究的问题:

1) 本文提出的融合方法在空间域中使用滑窗技术易使源图像的细节被模糊。因此,进一步研究适用于稀疏表示方法的新的图像融合规则是下一步的工作重点。

2) 本文是在已配准的脑部医学图像的基础上进行的融合研究,对未精确配准的脑部医学图像进行有针对性的图像融合的方法也需要进一步的研究。

3) 本文主要实现了脑部 CT 和 MR 医学图像的高质量融合,但如何对脑部解剖图像和功能图像如 MR 与正电子发射断层成像 (Positron Emission Tomography, PET)、MR 与单光子发射计算机断层成像 (Single Photon Emission Computed Tomography, SPECT) 等进行高质量融合,也将是下一步继续研究的内容。

参考文献 (References)

- [1] BHAVANA V. Multi-modality medical image fusion — a survey [C]// Proceedings of the 2015 International Journal of Engineering Research and Technology. Bangalore, India: Esrsa Publication, 2015: 778–781.
- [2] 李超,李光耀,谭云兰,等.基于非下采样 Contourlet 变换和区域特征的医学图像融合[J].计算机应用,2013,33(6):1727–1731.(LI C, LI G Y, TAN Y L, et al. Medical image fusion of nonsubsampled contourlet transform and regional feature [J]. Journal of Computer Applications, 2013, 33(6): 1727 – 1731.)
- [3] PAJARES G, DE LA CRUZ J M. A wavelet-based image fusion tutorial [J]. Pattern Recognition, 2004, 37(9): 1855 – 1872.
- [4] LI S T, KWOK J T, WANG Y N. Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images [J]. Information Fusion, 2002, 3(1): 17 – 23.
- [5] LEWIS J J, O'CALLAGHAN R J, NIKOLOV S G, et al. Pixel- and region-based image fusion with complex wavelets [J]. Information Fusion, 2007, 8(2): 119 – 130.
- [6] BURT P J, ADELSON E H. The Laplacian pyramid as a compact image code [J]. IEEE Transactions on Communications, 1983, 31(4): 532 – 540.
- [7] ZHANG Q, GUO B L. Multifocus image fusion using the nonsubsampled contourlet transform [J]. Signal Processing, 2009, 89(7): 1334 – 1346.
- [8] DONOHO D L. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289 – 1306.
- [9] YANG B, LI S T. Pixel-level image fusion with simultaneous orthogonal matching pursuit [J]. Information Fusion, 2012, 13(1): 10 – 19.
- [10] LIU Y, WANG Z F. Simultaneous image fusion and denoising with adaptive sparse representation [J]. IET Image Processing, 2015, 9(5): 347 – 357.
- [11] YANG B, LI S T. Multifocus image fusion and restoration with sparse representation [J]. IEEE Transactions on Instrumentation & Measurement, 2010, 59(4): 884 – 892.
- [12] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311 – 4322.
- [13] OPHIR B, LUSTIG M, ELAD M. Multi-scale dictionary learning using wavelets [J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 1014 – 1024.
- [14] YU N N, QIU T S, BI F, et al. Image features extraction and fusion based on joint sparse representation [J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 1074 – 1082.
- [15] 赵井坤,周颖琪,林茂松.基于稀疏表示与非局部相似的图像去噪算法[J].计算机应用,2016,36(2):551 – 555.(ZHAO J K, ZHOU Y Y, LIN M S. Image denoising algorithm based on sparse representation and nonlocal similarity [J]. Journal of Computer Applications, 2016, 36(2): 551 – 555.)
- [16] ENGAN K, AASE S O, HAKON H J. Method of optimal directions for frame design [C]// Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway, NJ: IEEE, 1999: 2443 – 2446.
- [17] JOHNSON K A, BECKER J A. The whole brain atlas [EB/OL]. [2016-10-09]. <http://www.med.harvard.edu/aanlib/home.html>.
- [18] YANG J C, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861 – 2873.

(下转第 1746 页)



- shop on GEometrical Models of Natural Language Semantics. Stroudsburg, PA: ACL, 2010: 33–37.
- [10] MITCHELL J, LAPATA M. Composition in distributional models of semantics [J]. *Cognitive Science*, 2010, 34(8): 1388–1429.
- [11] SOCHER R, HUANG E, PENNINGTON J, et al. Dynamic pooling and unfolding recursive autoencoders for paraphrase detection [C]// Proceedings of the 2011 International Conference on Neural Information Processing Systems. Cambridge, MA: MIT Press, 2011: 801–809.
- [12] ZANZOTTO F M, KORKONTELOS I, FALLUCCHI F, et al. Estimating linear models for compositional distributional semantics [C]// Proceedings of the 2010 23rd International Conference on Computational Linguistics. Stroudsburg, PA: ACL, 2010: 1263–1271.
- [13] SOCHER R, HUVAL B, MANNING C D, et al. Semantic compositionality through recursive matrix-vector spaces [C]// Proceedings of the 2012 Joint Conference on Empirical Methods in Natural language Processing and Computational Natural Language Learning. Stroudsburg, PA: ACL, 2012: 1201–1211.
- [14] GUEVARA E. A regression model of adjective-noun compositionality in distributional semantics [C]// Proceedings of the 2010 Workshop on GEometrical Models of Natural Language Semantics. Stroudsburg, PA: ACL, 2010: 33–37.
- [15] BARONI M, ZAMPARELLI R. Nouns are vectors, adjectives are matrices: representing adjective-noun constructions in semantic space [C]// Proceedings the 2010 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2010: 1183–1193.
- [16] PAPERNO D, PHAM N, BARONI M. A practical and linguistically-motivated approach to compositional distributional semantics [C]// Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2014: 90–99.
- [17] TAI K S, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long short-term memory networks [C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Stroudsburg, PA: ACL, 2015: 1556–1566.
- [18] ZAREMBA W, SUTSKEVER I. Learning to execute [EB/OL]. [2016-10-09]. <http://www.cs.nyu.edu/~zaremba/docs/Learning%20to%20Execute.pdf>.
- [19] 王智强,李茹,梁吉业,等.基于汉语篇章框架语义分析的阅读理解问答研究[J].*计算机学报*,2016,39(4):795–807.(WANG Z Q, LI R, LIANG J Y, et al. Research on question answering for reading comprehension based on Chinese discourse frame semantic parsing [J]. *Chinese Journal of Computers*, 2016, 39(4): 795–807.)
- [20] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [EB/OL]. [2016-10-09]. <https://core.ac.uk/download/pdf/24794691.pdf>.
- [21] 张志昌,张宇,刘挺,等.基于浅层语义树核的阅读理解答案句抽取[J].*中文信息学报*,2008,22(1):80–86.(ZHANG Z C, ZHANG Y, LIU T, et al. Answer sentence extraction of reading comprehension based on shallow semantic tree kernel [J]. *Journal of Chinese Information Processing*, 2008, 22(1): 80–86.)
- [22] 朱征宇,孙俊华.改进的基于《知网》的词汇语义相似度计算[J].*计算机应用*,2013,33(8):2276–2279.(ZHU Z Y, SUN J H. Improved vocabulary semantic similarity calculation based on HowNet [J]. *Journal of Computer Applications*, 2013, 33(8): 2276–2279.)
- This work is partially supported by the National High Technology Research and Development Program (863 Program) of China (2015AA015407), the Natural Science Foundation of Shanxi Province (201601D102030).
- WANG Yuanlong**, born in 1983, Ph. D., lecturer. His research interests include virtual reality, natural language processing, high performance computing.

(上接第1727页)

- [19] YANG J C, WANG Z, LIN Z, et al. Coupled dictionary training for image super-resolution [J]. *IEEE Transactions on Image Processing*, 2012, 21(8): 3467–3478.
- [20] SMITH L N, ELAD M. Improving dictionary learning: multiple dictionary updates and coefficient reuse [J]. *IEEE Signal Processing Letters*, 2013, 20(1): 79–82.
- [21] 练秋生,石保顺,陈书贞.字典学习模型、算法及其应用研究进展[J].*自动化学报*,2015,41(2):240–260.(LIAN Q S, SHI B S, CHEN S Z. Research advances on dictionary learning models, algorithms and applications [J]. *Acta Automatica Sinica*, 2015, 41(2): 240–260.)
- [22] CVEJIC N, CANAGARAJAH C N, BULL D R. Image fusion metric based on mutual information and Tsallis entropy [J]. *Electronics Letters*, 2006, 42(11): 626–627.
- [23] QU G H, ZHANG D L, YAN P F. Information measure for performance of image fusion [J]. *Electronics Letters*, 2002, 38(7): 313–315.
- [24] XYDEAS C S, PETROVIC V. Objective image fusion performance measure [J]. *Military Technical Courier*, 2000, 36(4): 308–309.
- [25] ZHAO J Y, LAGANIÈRE R, LIU Z. Performance assessment of

combinative pixel-level image fusion based on an absolute feature measurement [J]. *International Journal of Innovative Computing Information & Control* Ijicic, 2006, 3(6): 1433–1447.

- [26] PIELLA G, HEIJMANS H. A new quality metric for image fusion [C]// ICIP 2003: Proceedings of the 2003 International Conference on Image Processing. Washington, DC: IEEE Computer Society, 2003: 173–176.

This work is partially supported by the Natural Science Foundation of Shanxi Province (2015011045).

DONG Xia, born in 1992, M. S. candidate. Her research interests include medical image fusion, machine learning.

WANG Lifang, born in 1977, Ph. D., associate professor. Her research interests include machine vision, big data processing, medical image processing.

QIN Pinle, born in 1978, Ph. D., associate professor. His research interests include machine vision, big data processing, 3D reconstruction.

GAO Yuan, born in 1972, M. S., associate professor. Her research interests include big data processing, medical image processing, 3D reconstruction.