



文章编号:1001-9081(2018)04-1134-07

DOI:10.11772/j.issn.1001-9081.2017092291

## 基于自适应联合字典学习的脑部多模态图像融合方法

王丽芳\*, 董侠, 秦品乐, 高媛

(中北大学 大数据学院, 太原 030051)

(\*通信作者电子邮箱 wsm2004@nuc.edu.cn)

**摘要:**针对目前全局训练字典对于脑部医学图像的自适应性不强,以及使用稀疏表示系数的 $L_1$ 范数取极大的融合方式易造成图像的灰度不连续效应进而导致图像融合效果欠佳的问题,提出一种基于自适应联合字典学习的脑部多模态图像融合方法。该方法首先使用改进的K奇异值分解(K-SVD)算法自适应地从已配准的源图像中学习得到子字典并组合成自适应联合字典,在自适应联合字典的作用下由系数重用正交匹配追踪(CoefROMP)算法计算得到稀疏表示系数;然后将稀疏表示系数的“多范数”作为源图像块的活跃度测量,并提出“自适应加权平均”与“选择最大”相结合的无偏规则,根据稀疏表示系数的“多范数”的相似度选择融合规则,当“多范数”的相似度大于阈值时,使用“自适应加权平均”的规则,反之则使用“选择最大”的规则融合稀疏表示系数;最后根据融合系数与自适应联合字典重构融合图像。实验结果表明,与其他三种基于多尺度变换的方法和五种基于稀疏表示的方法相比,所提方法的融合图像能够保留更多的图像细节信息,对比度和清晰度较好,病灶边缘清晰,客观参数标准差、空间频率、互信息、基于梯度指标、基于通用图像质量指标和平均结构相似指标在三组实验条件下的均值分别为:71.0783、21.9708、3.6790、0.6603、0.7352 和 0.7339。该方法可以应用于临床诊断和辅助治疗。

**关键词:**脑部多模态图像融合; K奇异值分解; 自适应联合字典; 系数重用正交匹配追踪; 稀疏表示; 多范数; 无偏规则

中图分类号: TP391.41 文献标志码:A

## Multi-modal brain image fusion method based on adaptive joint dictionary learning

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

(School of Data Science and Technology, North University of China, Taiyuan Shanxi 030051, China)

**Abstract:** Currently, the adaptivity of global training dictionary is not strong for brain medical images, and the “max- $L_1$ ” rule may cause gray inconsistency in the fused image, which cannot get satisfactory image fusion results. A multi-modal brain image fusion method based on adaptive joint dictionary learning was proposed to solve this problem. Firstly, an adaptive joint dictionary was obtained by combining sub-dictionaries which were adaptively learned from registered source images using improved K-means-based Singular Value Decomposition (K-SVD) algorithm. The sparse representation coefficients were computed by the Coefficient Reuse Orthogonal Matching Pursuit (CoefROMP) algorithm by using the adaptive joint dictionary. Furthermore, the activity level measurement of source image patches was regarded as the “multi-norm” of the sparse representation coefficients, and an unbiased rule combining “adaptive weighed average” and “choose-max” was proposed, to chose fusion rule according to the similarity of “multi-norm” of the sparse representation coefficients. Then, the sparse representation coefficients were fused by the rule of “adaptive weighed average” when the similarity of “multi-norm” was greater than the threshold, otherwise the rule of “choose-max” was used. Finally, the fusion image was reconstructed according to the fusion coefficient and the adaptive joint dictionary. The experimental results show that, compared with the other three methods based on multi-scale transform and five methods based on sparse representation, the fusion images of the proposed method have more image detail information, better image contrast and sharpness, and clearer edge of lesion, the mean values of the objective parameters such as standard deviation, spatial frequency, mutual information, the gradient based index, the universal image quality based index and the mean structural similarity index under three groups of experimental conditions are 71.0783, 21.9708, 3.6790, 0.6603, 0.7352 and 0.7339 respectively. The proposed method can be used for clinical diagnosis and assistant treatment.

**Key words:** multi-modal brain image fusion; K-means-based Singular Value Decomposition (K-SVD); adaptive joint dictionary; Coefficient Reuse Orthogonal Matching Pursuit (CoefROMP); sparse representation; multi-norm; unbiased rule

收稿日期:2017-09-22;修回日期:2017-10-18。 基金项目:山西省自然科学基金资助项目(2015011045)。

作者简介:王丽芳(1977—),女,山西长治人,副教授,博士,CCF会员,主要研究方向:机器视觉、大数据处理、医学图像处理;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;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

多模态医学影像在临床医疗诊断中发挥越来越重要的作用。医学图像可分为解剖图像和功能图像,计算机断层成像(Computed Tomography, CT)和磁共振成像(Magnetic Resonance Imaging, MRI)均为解剖图像;单光子发射计算机断层成像(Single Photon Emission Computed Tomography, SPECT)和正电子发射断层成像(Positron Emission Tomography, PET)均为功能图像。解剖图像能够反映器官的解剖结构,分辨率较高;功能图像能够显示人体新陈代谢的功能变化,分辨率较低。单一类型的图像往往无法为病情诊断提供全面充足的信息<sup>[1]</sup>。多模态医学图像融合技术能够将互补的多源信息融合到一幅图像中以改善图像质量并减少随机性和冗余性,有助于提高临床医疗诊断和分析的准确性<sup>[2]</sup>。

近年来,随着压缩感知理论<sup>[3]</sup>的提出,基于稀疏表示的方法<sup>[4-5]</sup>被广泛用于图像融合领域。Mitianoudis 等<sup>[6]</sup>首次使用滑窗技术将源图像划分成图像块并用于图像融合,实验结果表明该方法能够很好地提取源图像的局部显著特征和保持平移不变性。由于在稀疏域中使用活跃度测量具有可靠性和空间滑窗技术的平移不变性,基于稀疏表示的图像融合方法一般优于传统多尺度变换的方法。Yang 等<sup>[7]</sup>首次将稀疏表示的方法用于多聚焦图像融合,该方法使用冗余的离散余弦变换(Discrete Cosine Transform, DCT)字典表示源图像,并使用“ $L_1$  范数最大”的规则融合稀疏表示系数。DCT 字典属于解析字典,具有快速实现的优点,但这种字典仅限于特定类型的信号表示,并且在选择解析基时需要先验知识。Aharon 等<sup>[8]</sup>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。相对于解析字典,学习的字典包含更丰富的图像特征信息,表示能力更强,但是时间效率较低。Liu 等<sup>[9]</sup>从 40 幅自然图像中随机采样 10 万个图像块作为字典训练数据,低频子带对应的稀疏表示系数使用“ $L_1$  范数最大”的规则进行融合。Yin 等<sup>[10]</sup>使用  $K$ -SVD 算法在 USC-SIPI 数据集<sup>[11]</sup>上学习字典。上述方法学习的字典均属于全局训练字典,全局训练字典使用的训练数据来自一些公开的图像数据库,是由外部预先收集的图像组成。在实际应用中,收集合适的图像数据集并不总是可行的。此外,不同类型图像的图像内容千差万别,全局训练字典对于不同类型的图像表现出来的性能差异较大。

医学图像具有对比度低、灰度分布不均匀等特点,全局训练字典对于医学图像的自适应性不强,难以适应多模态医学图像融合的需要。因此,在基于稀疏表示的多模态医学图像融合方法中,如何构建一个自适应于源图像的过完备字典是关键。宗静静等<sup>[12]</sup>将两幅已配准的源图像组成一个整体进行联合稀疏表示,获得一个共同稀疏部分和不同稀疏部分,其中对不同稀疏部分同样使用“ $L_1$  范数最大”的规则进行融合。该方法直接从已配准的源图像本身学习字典可以提高字典对于源图像的自适应性。但是,从图像融合规则来看,传统的使用稀疏表示系数的  $L_1$  范数取极大的方式在融合多模态医学图像时,易造成融合结果的灰度不连续效应。

针对上述问题,本文提出一种基于自适应联合字典学习

的脑部多模态图像融合方法。首先,为实现功能图像和解剖图像的融合,该方法使用亮度-色调-饱和度(Intensity-Hue-Saturation, IHS)变换将已与解剖图像配准的功能图像变换到 IHS 空间,再将其对应的 I 分量(代表灰度图像)和解剖图像作为待融合的源图像;其次,不同于解析字典和全局训练字典,提出的自适应联合字典的学习不需要任何先验知识,也不需要外部预先收集的图像数据,而是将直接从待融合的源图像中学习得到的子字典组合成自适应联合字典来提高字典对于每一对输入图像的自适应性。同时,在融合阶段,提出稀疏表示系数的“多范数”综合地衡量源图像块的活跃度,并使用一种结合“自适应加权平均”与“选择最大”的无偏规则融合稀疏表示系数。对脑部多模态图像进行的大量实验结果表明,本文方法比对比方法可以保留更多的图像细节信息,取得较好的图像融合效果。

## 1 稀疏表示和字典学习

稀疏表示的基本假设是一个信号  $\mathbf{x} \in \mathbb{R}^n$  可以近似地表示为一个超完备字典  $\mathbf{D} \in \mathbb{R}^{n \times m}$  ( $n < m$ ) 中“少量”原子的线性组合, $n$  和  $m$  分别表示信号的维数和字典的大小<sup>[13]</sup>。也就是说,信号  $\mathbf{x}$  可以表示为  $\mathbf{x} \approx \mathbf{D}\boldsymbol{\alpha}, \boldsymbol{\alpha} \in \mathbb{R}^m$  是未知的稀疏表示系数。稀疏系数  $\boldsymbol{\alpha}$  可以由式(1)计算:

$$\begin{aligned} & \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \\ \text{s.t. } & \|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \varepsilon \end{aligned} \quad (1)$$

其中: $\varepsilon \geq 0$  表示允许偏差的精度,问题(1)的求解过程称为“稀疏编码”。

在字典学习中,假设有  $J$  个大小为  $\sqrt{n} \times \sqrt{n}$  的训练图像块,将这些图像块拉成列向量并组成训练数据集  $\{\mathbf{x}_i\}_{i=1}^J, \mathbf{x}_i \in \mathbb{R}^n$ , 则字典学习模型可以表示为:

$$\begin{aligned} & \min_{\mathbf{D}, \{\boldsymbol{\alpha}_i\}_{i=1}^J} \|\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i\|_F^2 \\ \text{s.t. } & \forall i \|\boldsymbol{\alpha}_i\|_0 \leq \tau \end{aligned} \quad (2)$$

其中: $\mathbf{D} \in \mathbb{R}^{n \times m}$ ;  $\|\cdot\|_F$  表示 F 范数, 定义为  $\|\mathbf{X}\|_F = \sqrt{\sum_j X_{ij}^2}$ ;  $\tau$  是稀疏向量  $\{\boldsymbol{\alpha}_i\}_{i=1}^J$  的稀疏度。

## 2 自适应联合字典

为提高字典对于每一对输入医学图像的自适应性,本文提出使用改进的  $K$ -SVD 算法<sup>[14]</sup>分别对两幅已配准的源图像进行字典学习得到两个子字典再组合成自适应联合字典,使每一幅源图像均可以用字典中相同的原子表示。自适应联合字典的学习过程如图 1 所示。

假设两幅已配准的源图像  $\mathbf{I}_{1c}, \mathbf{I}_2 \in \mathbb{R}^{M \times N}$ ,  $\mathbf{I}_{1c}$  表示功能图像,  $\mathbf{I}_2$  表示解剖图像。使用 IHS 变换<sup>[15-16]</sup>将  $\mathbf{I}_{1c}$  由 RGB 空间变换到 IHS 色彩空间,再将  $\mathbf{I}_{1c}$  对应的亮度分量  $\mathbf{I}_1$ (代表灰度图像)与  $\mathbf{I}_2$  用于自适应联合字典学习。需要注意的是对于两幅源图像均为灰度图像(如 CT 和 MRI)的情况,不需要 IHS 变换的处理步骤,见图 1 虚线框外的部分。然后使用步长为 1, 大小为  $\sqrt{w} \times \sqrt{w}$  的滑动窗口对  $\mathbf{I}_1$  和  $\mathbf{I}_2$  进行分块,对于源图像  $\mathbf{I}_1$  和  $\mathbf{I}_2$ , 均有  $L$  ( $L = (M - \sqrt{w} + 1) \times (N - \sqrt{w} + 1)$ ) 个图像块, 分别记为  $\{\mathbf{P}_1^i\}_{i=1}^L$  和  $\{\mathbf{P}_2^i\}_{i=1}^L$ 。将两幅源图像的第  $i$  个图像块  $\mathbf{P}_1^i$



和  $P_2^i$  拉成  $w$  维的列向量记为  $x_1^i$  和  $x_2^i$ , 再将各列向量  $\{x_i^i\}_{i=1}^L$  和  $\{x_2^i\}_{i=1}^L$  分别组成矩阵  $X_1 \triangleq (x_1^1, x_1^2, \dots, x_1^L) \in \mathbb{R}^{w \times L}, X_2 \triangleq (x_2^1, x_2^2, \dots, x_2^L) \in \mathbb{R}^{w \times L}$ 。

在字典学习代价函数(2)中加入支撑完整的先验信息<sup>[17]</sup>, 以源图像  $I_1$  为例, 将  $X_1$  训练的优化问题表示如下:

$$\begin{aligned} & \min_{D_1, A} \|X_1 - D_1 A\|_F^2 \\ \text{s.t. } & \|A\|_0 \leq \tau, A \odot M = 0 \end{aligned} \quad (3)$$

其中: 字典  $D_1 \in \mathbb{R}^{w \times S}$ , 稀疏表示系数  $A \in \mathbb{R}^{S \times L}$ 。 $\odot$  代表点乘, 掩膜矩阵  $M$  由元素 0 和 1 组成, 定义为  $M = \{\mid A \mid = 0\}$ , 等价于若  $A(i, j) = 0$  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。式(3)的求解过程分为稀疏编码和字典更新两个阶段:

首先, 随机矩阵初始化字典  $D_1^{(0)}$ , 通过求解式(4)来实现对系数矩阵  $A$  的更新:

$$\begin{aligned} A &= \arg \min_A \|X_1 - D_1 A\|_F^2 \\ \text{s.t. } & A \odot M = 0 \end{aligned} \quad (4)$$

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(4)可以转换为式(5):

$$\alpha_i = \arg \min_{\alpha_i} \|x_{1i} - D_{1i} \alpha_i\|_2^2 \quad (5)$$

其中:  $D_{1i}$  是  $D_1$  对应  $A$  的非零支集的子矩阵;  $\alpha_i$  是  $A$  第  $i$  列的非零部分。问题(5)由系数重用正交匹配追踪(Coefficient Reuse Orthogonal Matching Pursuit, CoefROMP) 算法<sup>[14]</sup>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(3)的优化问题可以转化为:

$$\begin{aligned} \{D_1, A\} &= \arg \min_{D_1, A} \|X_1 - D_1 A\|_F^2 \\ \text{s.t. } & A \odot M = 0 \end{aligned} \quad (6)$$

则式(6)的补偿项可以写为:

$$\begin{aligned} \|X_1 - D_1 A\|_F^2 &= \|X_1 - \sum_{j=1}^S d_{1j} \alpha_j^T\|_F^2 = \\ &= \left\| \left( X_1 - \sum_{j \neq k} d_{1j} \alpha_j^T \right) \odot (\mathbf{1}_w \cdot m_j^T) - d_{1k} \alpha_k^T \right\|_F^2 = \\ &= \|E_k - d_{1k} \alpha_k^T\|_F^2 \end{aligned} \quad (7)$$

其中:  $d_{1k}$  表示字典  $D_1$  中待更新的第  $k$  列;  $\alpha_k^T$  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  $\alpha_k^T$  中的零元素在正确的位置。掩膜矩阵  $\mathbf{1}_w \cdot m_j^T$  是将行向量  $m_j^T$  复制  $w$  次得到尺寸为  $w \times L$  的秩为 1 的矩阵, 利用掩膜矩阵  $\mathbf{1}_w \cdot m_j^T$  可以有效地去除  $X_1 - \sum_{j \neq k} d_{1j} \alpha_j^T$  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(Singular Value Decomposition, SVD) 得到  $E_k = U \Delta V^T$ , 使用矩阵  $U$  的第一列更新字典  $D_1$  中的原子  $d_{1k}$ , 同时将稀疏系数矩阵  $A$  中的  $\alpha_k^T$  更新为矩阵  $V$  的第一列与  $\Delta(1, 1)$  的乘积。

最后, 循环执行上述两个步骤, 直至达到预设的迭代次数为止, 输出字典  $D_1$ , 字典  $D_2$  的训练过程同字典  $D_1$ , 然后将得到的子字典  $D_1$  和  $D_2$  组合成自适应联合字典  $D = [D_1, D_2]$ ,  $D \in \mathbb{R}^{w \times 2S}$ 。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, 本文使用 CoefROMP 算法根据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<sup>[17]</sup>。

### 3 多模态医学图像融合过程

图像融合方式在很大程度上决定了融合图像的质量。传统基于稀疏表示的医学图像融合方法采用稀疏表示系数的  $L_1$  范数取极大的融合方式易造成融合图像的灰度不连续效应。这是因为对于多模态医学图像的同一区域, 有可能在一幅源图像中灰度较高, 而在另一幅源图像中灰度较低, 但在两幅源图像中都是很“平坦”的, 即区域内部灰度变化很小, 但这两幅源图像在此区域仍会存在细节丰富程度的差异。由于差异很小, 稀疏表示系数  $L_1$  范数取极大的融合方式很难准确地判断具有较多细节的源图像块, 得到错误判定结果的可能性增大, 同时对有噪声的情况非常敏感, 从而导致对此类类型区域的不同图像块的判定结果不一致而最终造成融合图像的灰度不连续效应。

为解决此问题, 考虑以下两个方面: 一是如何衡量源图像块的活跃度, 另一个是建立什么样的规则融合稀疏表示系数。对于第一个问题, 仅仅使用  $L_1$  范数难以全面地计算源图像块的活跃度, 本文提出将稀疏表示系数的“多范数”作为源图像块的活跃度测量以保留更多的图像细节信息; 对于第二个问题, 寻求能够使融合图像保留源图像所有视觉信息的理想的融合规则是困难的<sup>[18]</sup>。考虑到“加权平均”的融合规则可以有效抑制噪声, “选择最大”的融合规则可以保留源图像最重要的信息, 本文提出“自适应加权平均”与“选择最大”两者相结合的无偏规则。根据稀疏表示系数的“多范数”的相似度来选择融合规则, 对于相似度大于阈值, 即图像块细节丰富程

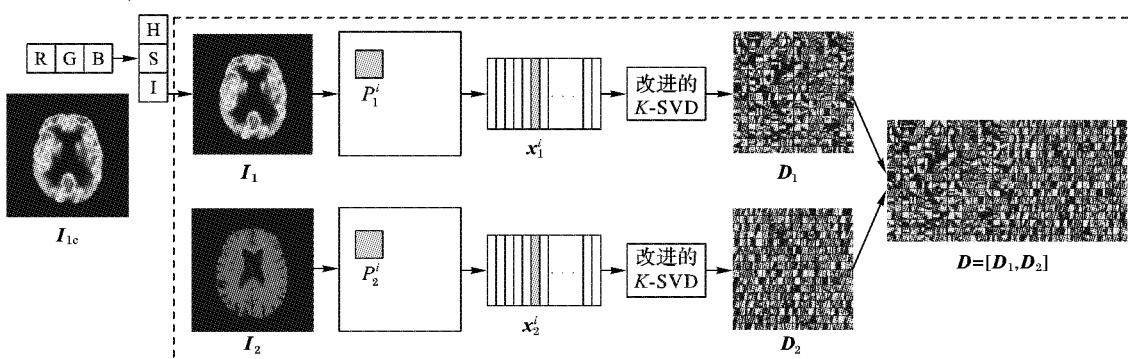


图1 自适应联合字典的学习过程  
Fig. 1 Learning process of adaptive joint dictionary



度差异较小的情况,使用“自适应加权平均”的规则;反之则使用“选择最大”的规则融合稀疏表示系数。

### 3.1 稀疏编码

如图2所示,假设已经配准的脑部多模态医学源图像 $I_{1c}$ , $I_2 \in \mathbb{R}^{M \times N}$ ,与第2章叙述的处理过程相同,将功能图像 $I_{1c}$ 经过IHS变换的处理得到其对应的亮度分量 $I_1$ 。当两幅源图像均为灰度图像时,融合过程不需要IHS变换及其逆变换的处理。将亮度分量 $I_1$ 与解剖图像 $I_2$ 经过滑窗技术的处理分别得到 $L$ 个列向量 $\{x_1^i\}_{i=1}^L$ 和 $\{x_2^i\}_{i=1}^L$ 。求解 $x_1^i$ 和 $x_2^i$ 的稀疏表示系数,公式表示如下:

$$\alpha_1^i = \arg \min_{\alpha} \|\alpha\|_0 \\ \text{s. t. } \|x_1^i - D\alpha\|_2 < \varepsilon \quad (8)$$

$$\alpha_2^i = \arg \min_{\alpha} \|\alpha\|_0 \\ \text{s. t. } \|x_2^i - D\alpha\|_2 < \varepsilon \quad (9)$$

其中: $D$ 是第2章提出的自适应联合字典。

### 3.2 多范数无偏融合规则

稀疏系数向量 $\alpha$ 的 $L_1$ 范数表示向量中各个元素的绝对值之和,可以反映细节信息的重要性,对应的值越大,图像块的显著性越好;稀疏系数向量 $\alpha$ 的 $L_0$ 范数表示向量中非零项的个数,可以反映细节信息的集中度,对应的值越大,图像块包含的细节信息越多。为全面地计算源图像块的活跃度以保留更多的图像细节信息,对于每一个源图像块,使用加权指数分别对 $L_1$ 范数和 $L_0$ 范数进行约束,再将两者的乘积作为两种范数对源图像块的综合度量结果,公式表示如下:

$$R_1^i = (\|\alpha_1^i\|_1)^\beta \cdot (\|\alpha_1^i\|_0)^\gamma \quad (10)$$

$$R_2^i = (\|\alpha_2^i\|_1)^\beta \cdot (\|\alpha_2^i\|_0)^\gamma \quad (11)$$

其中: $\alpha_1^i$ 和 $\alpha_2^i$ 分别为源图像 $I_1$ 和 $I_2$ 的第*i*个图像块的稀疏表示系数; $R_1^i$ 和 $R_2^i$ 分别为两种范数对源图像 $I_1$ 和 $I_2$ 的第*i*个图像块的综合度量结果; $\|\cdot\|_1$ 和 $\|\cdot\|_0$ 分别表示稀疏系数的 $L_1$ 范数和 $L_0$ 范数; $\beta > 0$ 和 $\gamma > 0$ 分别表示两种范数的加权指数,若某一加权指数为0,则表示不考虑其对应的范数。

然后计算稀疏表示系数的“多范数” $R_1^i$ 和 $R_2^i$ 的相似度 $S_i$ ,公式表示如下:

$$S_i = \frac{2 \sqrt{R_1^i R_2^i}}{R_1^i + R_2^i} \quad (12)$$

当稀疏表示系数的“多范数”的相似度大于等于阈值

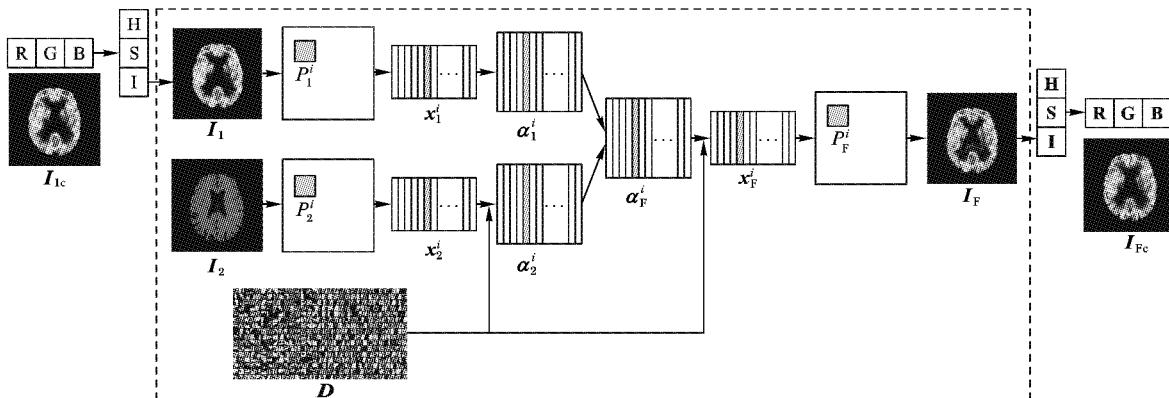


图2 多模态医学图像融合过程  
Fig. 2 Fusion process of multi-modal medical image

( $S_i \geq T$ )时,使用“自适应加权平均”的融合规则:

$$\alpha_F^i = \omega_1 \alpha_1^i + \omega_2 \alpha_2^i \quad (13)$$

其中: $T$ 为阈值; $\alpha_F^i$ 表示融合系数; $\omega_1$ 和 $\omega_2$ 分别表示稀疏表示系数 $\alpha_1^i$ 和 $\alpha_2^i$ 的权重。 $\omega_1$ 和 $\omega_2$ 的计算公式为:

$$\omega_1 = \frac{R_1^i}{R_1^i + R_2^i} \quad (14)$$

$$\omega_2 = 1 - \omega_1 \quad (15)$$

当稀疏表示系数的“多范数”的相似度小于阈值( $S_i < T$ )时,使用“选择最大”的融合规则

$$\alpha_F^i = \begin{cases} \alpha_1^i, & R_1^i \geq R_2^i \\ \alpha_2^i, & \text{其他} \end{cases} \quad (16)$$

则 $x_1^i$ 和 $x_2^i$ 的融合结果为:

$$x_F^i = D\alpha_F^i \quad (17)$$

对所有的图像块作相同的处理以得到所有图像块的融合结果。对于每个块向量 $x_F^i$ ,通过反滑窗的过程重塑成 $\sqrt{w} \times \sqrt{w}$ 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 $I_F$ ,最后经过IHS逆变换得到融合图像 $I_{Fc}$ 。

## 4 实验结果与分析

为验证本文方法的性能,本文进行了大量的仿真实验,现选取三组已配准的脑部多模态医学图像(大小均为 $256 \times 256$ 像素)作为待融合的源图像,脑部图像数据均来源于哈佛大学医学院<sup>[19]</sup>。图3为脑部肿瘤转移、脑部Ⅱ级星形细胞瘤以及脑部阿尔茨海默病的图像。实验的硬件环境为Intel Core i3-2350M @ 2.30 GHz处理器、4 GB运行内存,软件环境为32位Windows 7操作系统、Matlab R2013a。

实验从主观客观两个方面将本文方法与其他八种方法进行比较。选取的对比方法有:1)基于多尺度变换的方法,包括:离散小波变换(Discrete Wavelet Transform, DWT)<sup>[20]</sup>、拉普拉斯金字塔(Laplacian Pyramid, LP)<sup>[21]</sup>、非下采样轮廓波变换(Non-Subsampled Contourlet Transform, NSCT)<sup>[22]</sup>;2)基于稀疏表示的方法,包括:基于DCT字典与多范数无偏规则的方法(method based on DCT dictionary and Multi-Norm-unbiased rule, DCT-MN)、基于K-SVD算法的全局训练字典与多范数无偏规则的方法(method based on the global trained dictionary using K-SVD algorithm and Multi-Norm-unbiased rule, KSVD-MN)、基于多尺度变换和稀疏表示的方法(method based on



Multi-Scale Transform and Sparse Representation, MST-SR) [9]、基于联合稀疏表示的方法 (method based on Joint Sparse Representation, JSR) [12]、基于自适应联合字典与  $L_1$  范数最大规则的方法 (method based on Adaptive Joint Dictionary and max- $L_1$ -norm rule, AJD-L1)。本文提出的基于自适应联合字典与多范数无偏规则的方法记为 AJD-MN (method based on Adaptive Joint Dictionary and Multi-Norm-unbiased rule)。

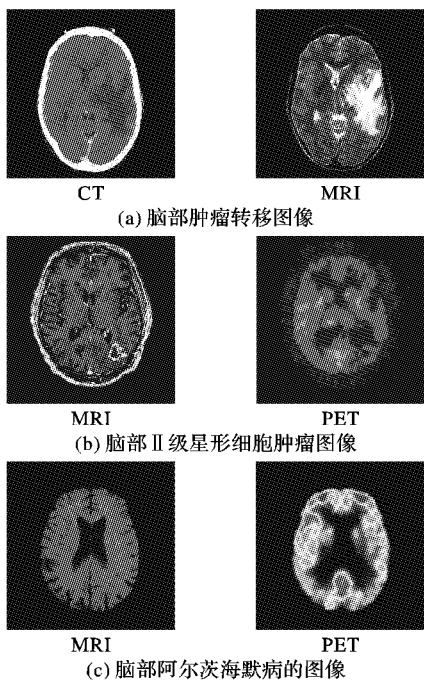


图 3 三对脑部多模态医学图像

Fig. 3 Three pairs of brain multi-modal medical images

对于 DWT 和 LP 方法, 分解水平均设为 3, DWT 方法的小波基设为“db6”。NSCT 方法使用“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均为 1, 图像块大小均为  $8 \times 8$ , 字典大小均为  $64 \times 512$ , 误差  $\varepsilon = 0.01$ , 稀疏度  $\tau = 8$ 。KSVD-MN 和 MST-SR 方法的字典训练样本是从 40 幅自然图像 [23] 中随机采样 10 万

个图像块, MST-SR 方法的多尺度变换选择 LP, 分解水平设为 3。AJD-MN 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; 提出的多范数使用相等的加权指数  $\beta = \gamma = 1$ ; 经多次实验验证, 无偏融合规则的阈值  $T = 0.9$  时融合效果较好。

图 4 为不同的融合方法在三组实验条件下的运行结果, 从主观视觉效果分析, DWT 方法的融合图像的颜色、亮度和对比度明显失真, 存在块效应和伪影; LP 和 NSCT 方法的融合质量相对较好, 但是融合图像的对比度较低; DCT-MN、KSVD-MN、MST-SR、JSR 和 AJD-L1 方法与上述三种基于多尺度变换的方法 (DWT、LP 和 NSCT) 相比, 融合图像的对比度有很大的提升, 但在组织和病灶区域存在一些伪影; AJD-MN 方法的融合图像的伪影较少, 并且能够保留更多的图像细节信息。综上所述, 本文提出的 AJD-MN 方法的融合图像在颜色、亮度、对比度、清晰度以及细节的保留度上均优于其他八种方法, 病灶清晰, 有助于临床诊断。

为定量地对融合图像的质量进行客观评价, 本文使用六种指标: 标准差 (Standard Deviation, SD) [24]、空间频率 (Spatial Frequency, SF) [25]、互信息 (Mutual Information, MI) [26]、基于梯度的指标  $Q^{AB/F}$  [27]、基于通用图像质量指标  $Q_w$  [28] 和平均结构相似指标 (Mean Structural SIMilarity, MSSIM) [29]。其中: SD 衡量图像信息的丰富程度, 反映融合图像的对比度; SF 反映图像的清晰度; MI 反映融合图像从源图像提取的信息量的多少;  $Q^{AB/F}$  反映融合图像对源图像边缘特性的保留情况;  $Q_w$  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; MSSIM 指标是融合图像分别与两幅源图像的结构相似指标 SSIM 的均值, SSIM 反映融合图像与源图像在亮度、对比度和结构方面的相似度。上述六种指标均是对应的值越大融合图像质量越高。

表 1 为图 4~6 的三组实验中不同融合方法的客观评价指标。从表 1 可看出, DWT 方法在 SD、SF、MI、 $Q^{AB/F}$ 、 $Q_w$  和 MSSIM 六项指标上均不理想, 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; 由于 LP 方法具有较

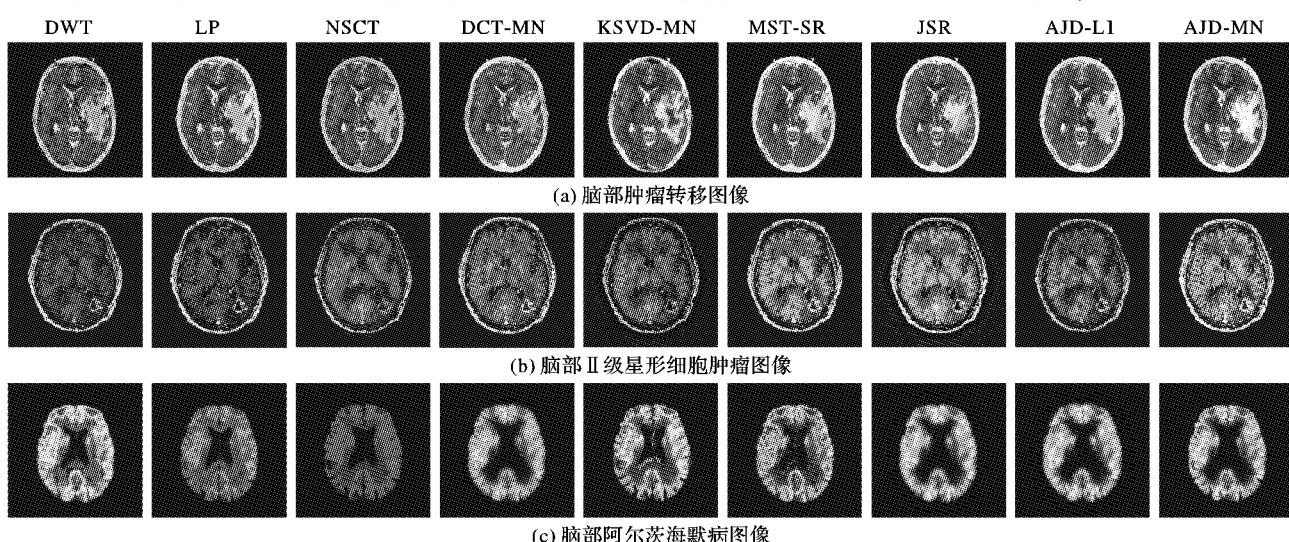


图 4 图像融合结果

Fig. 4 Image fusion results



好的多尺度和多分辨率特性,NSCT 具有多尺度、多方向和平移不变性,因此,LP 和 NSCT 方法在上述六项指标上均明显优于 DWT 方法;上述三种多尺度变换的方法在 SD 和 SF 指标上均低于其他方法,表明融合图像的对比度可能受到了损失;DCT-MN、KSVD-MN、MST-SR 方法在 MI 指标上优于多尺度变换的方法,而在  $Q^{AB/F}$ 、 $Q_w$  和 MSSIM 指标上存在一些低于多尺度变换方法的情况,可能是由于字典对于医学图像的自适应性不强,其表示能力不足以提取足够多的细节信息或者不能精确地表示医学图像的特征;JSR 和 AJD-L1 方法在  $Q^{AB/F}$  和 MSSIM 指标上存在一些低于多尺度变换方法的情况,可能是由于  $L_1$  范数取极大的融合规则导致图像的灰度不连续效应;本文提出的 AJD-MN 方法存在一些次优值,但是均接近于最优值。经计算, AJD-MN 方法的 SD、SF、MI、 $Q^{AB/F}$ 、 $Q_w$  和 MSSIM 六项指标在上述三组实验条件下的均值分别为 71.0783、21.9708、3.6790、0.6603、0.7352 和 0.7339。与其他八种方法相比,本文提出的 AJD-MN 方法在客观评价方面也具有优势。

表1 九种融合方法的客观指标

Tab. 1 Objective indexes of nine fusion methods

图像	融合方法	评价指标					
		SD	SF	MI	$Q^{AB/F}$	$Q_w$	MSSIM
脑部肿瘤转移图像	DWT	69.1147	24.3978	3.0432	0.6003	0.5827	0.7102
	LP	74.1779	26.6591	3.2611	<b>0.6822</b>	0.6371	0.7491
	NSCT	67.2460	25.2463	3.1476	0.6583	0.6322	0.7176
	DCT-MN	72.0256	25.2630	3.2856	0.6451	0.6489	0.7322
	KSVD-MN	74.7748	26.9669	<b>4.0124</b>	0.6692	0.6429	0.7341
	MST-SR	76.9590	27.4546	3.4237	0.6794	0.6442	0.7483
	JSR	78.1800	26.4162	3.3610	0.6618	0.6494	0.7497
	AJD-L1	77.9314	27.3524	3.3674	0.6762	0.6420	0.7464
脑部Ⅱ级星形细胞肿瘤图像	AJD-MN	<b>80.7010</b>	<b>28.8587</b>	3.5173	0.6750	<b>0.6561</b>	<b>0.7563</b>
	DWT	43.7912	17.9190	2.2539	0.5957	0.6657	0.6127
	LP	48.0477	18.4139	2.3590	0.6633	0.7824	0.6072
	NSCT	42.7933	17.8408	2.3036	0.6520	0.7586	0.6374
	DCT-MN	52.2453	18.0168	3.5994	0.6377	0.7363	0.6727
	KSVD-MN	56.8109	18.8100	3.5360	0.6794	0.7720	0.6718
	MST-SR	56.9328	18.8974	3.5491	0.6899	0.7712	0.6654
	JSR	56.3869	18.7595	3.5636	0.6722	0.7959	0.6756
脑部阿尔茨海默病图像	AJD-L1	56.1256	<b>18.9431</b>	3.8285	0.6883	0.7880	<b>0.6878</b>
	AJD-MN	<b>57.5739</b>	18.7558	<b>4.0032</b>	<b>0.6930</b>	<b>0.8031</b>	0.6761
	DWT	54.2307	17.5914	2.6357	0.5323	0.6254	0.6594
	LP	62.6674	17.8467	3.4750	0.5808	0.6824	0.7215
	NSCT	62.2965	16.6728	3.4782	0.5722	0.6756	0.7111
	DCT-MN	67.9670	14.5476	3.4983	0.5626	0.7249	0.7469
	KSVD-MN	74.2402	<b>19.0024</b>	3.5227	0.5770	0.7417	0.7250
	MST-SR	73.3098	17.7855	<b>3.5571</b>	0.6005	0.7408	0.7201
脑部CT/MR图像	JSR	72.8368	17.9215	3.4919	0.5933	0.7399	0.7558
	AJD-L1	71.7961	18.3454	3.5068	0.5837	0.7357	0.7341
	AJD-MN	<b>74.9599</b>	18.2978	3.5166	<b>0.6130</b>	<b>0.7465</b>	<b>0.7692</b>

注:加粗数据为最优值。

## 5 结语

针对目前全局训练字典对于脑部医学图像的自适应性不强的问题,本文提出使用改进的 K-SVD 算法自适应地从输入源图像中学习子字典并组合成自适应联合字典的方法;同时,

针对稀疏表示系数的  $L_1$  范数取极大的融合方式易造成图像的灰度不连续效应的问题,提出用稀疏表示系数的“多范数”全面地衡量源图像块的活跃度,并使用“自适应加权平均”与“选择最大”相结合的无偏规则融合稀疏表示系数。在实验阶段,分别对脑部 CT/MRI 图像和脑部 MRI/PET 图像进行多次实验,结果表明,与其他八种方法相比,本文方法能够取得最佳的图像融合效果,可以有效辅助临床医疗诊断。

但仍存在需要进一步研究的问题:本文提出的字典学习方法是将源图像中所有图像块用于字典学习过程,这样往往会带来很多冗余信息。如何构造一个非冗余但信息丰富的字典来更好地提高图像融合质量和时间效率,将是下一步的研究重点。

## 参考文献(References)

- [1] LI S T, KANG X D, FANG L Y, et al. Pixel-level image fusion: a survey of the state of the art[J]. Information Fusion, 2017, 33: 100–112.
- [2] JAMES A P, DASARATHY B V. Medical image fusion: a survey of the state of the art[J]. Information Fusion, 2014, 19(3): 4–19.
- [3] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289–1306.
- [4] LIU Y, WANG Z F. Simultaneous image fusion and denoising with adaptive sparse representation[J]. IET Image Processing, 2015, 9(5): 347–357.
- [5] YIN H T. Sparse representation with learned multiscale dictionary for image fusion [J]. Neurocomputing, 2015, 148 (148): 600–610.
- [6] MITIANOUDIS N, STATHAKI T. Pixel-based and region-based image fusion schemes using ICA bases[J]. Information Fusion, 2007, 8(2): 131–142.
- [7] YANG B, LI S T. Multifocus image fusion and restoration with sparse representation[J]. IEEE Transactions on Instrumentation & Measurement, 2010, 59(4): 884–892.
- [8] 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.
- [9] LIU Y, LIU S P, WANG Z F. A general framework for image fusion based on multi-scale transform and sparse representation [J]. Information Fusion, 2015, 24: 147–164.
- [10] YIN H T, LI S T. Multimodal image fusion with joint sparsity model[J]. Optical Engineering, 2011, 50(6): 7007.
- [11] WEBER A. The USC-SIPI image database[EB/OL]. [2017-09-20]. <http://sipi.usc.edu/database/>.
- [12] 宗静静, 邱天爽, 郭冬梅. 联合稀疏表示的医学图像融合及同步去噪[J]. 中国生物医学工程学报, 2016, 35(2): 133–140. (ZONG J J, QIU T S, GUO D M. Simultaneous medical image fusion and de-noising with joint sparse representation[J]. Chinese Journal of Biomedical Engineering, 2016, 35(2): 133–140.)
- [13] 董侠, 王丽芳, 秦品乐, 等. 改进耦合字典学习的脑部 CT/MR 图像融合方法[J]. 计算机应用, 2017, 37(6): 1722–1727. (DONG X, WANG L F, QIN P L, et al. CT/MR brain image fusion method via improved coupled dictionary learning[J]. Journal of Computer Applications, 2017, 37(6): 1722–1727.)



- [14] SMITH L N, ELAD M. Improving dictionary learning: multiple dictionary updates and coefficient reuse[J]. IEEE Signal Processing Letters, 2013, 20(1): 79–82.
- [15] TU T M, SU S C, SHYU H C, et al. A new look at IHS-like image fusion methods[J]. Information Fusion, 2001, 2(3): 177–186.
- [16] DANESHVAR S, GHASSEMIAN H. MRI and PET image fusion by combining IHS and retina-inspired models[J]. Information Fusion, 2010, 11(2): 114–123.
- [17] 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.)
- [18] ZHANG Q, LIU Y, BLUM R S, et al. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: a review[J]. Information Fusion, 2017, 40: 57–75.
- [19] JOHNSON K A, BECKER J A. The whole brain Atlas[DB/OL]. [2017-09-20]. <http://www.med.harvard.edu/aanlib/home.html>.
- [20] PAJARES G, CRUZ J M D L. A wavelet-based image fusion tutorial[J]. Pattern Recognition, 2004, 37(9): 1855–1872.
- [21] BURT P J, ADELSON E H. The Laplacian pyramid as a compact image code[J]. IEEE Transactions on Communications, 1983, 31(4): 532–540.
- [22] ZHANG Q, GUO B L. Multifocus image fusion using the nonsubsampled contourlet transform[J]. Signal Processing, 2009, 89(7): 1334–1346.
- [23] The imagefusion image database [EB/OL]. [2017-09-20]. <http://www.imagefusion.org>.
- [24] SHI W Z, ZHU C Q, TIAN Y, et al. Wavelet-based image fusion and quality assessment[J]. International Journal of Applied Earth Observation and Geoinformation, 2005, 6(3/4): 241–251.
- [25] ESKICIOGLU A M, FISHER P S. Image quality measures and their performance[J]. IEEE Transactions on Communications, 1996, 43(12): 2959–2965.
- [26] QU G H, ZHANG D L, YAN P F. Information measure for performance of image fusion[J]. Electronics Letters, 2002, 38(7): 313–315.
- [27] XYDEAS C S, PETROVIC V. Objective image fusion performance measure[J]. Military Technical Courier, 2000, 36(4): 308–309.
- [28] 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.
- [29] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600–612.

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

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

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

**QIN Pinfei**, 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.

(上接第 1120 页)

- [10] TAN Z, BAI X, WANG B, et al. Fast single-image defogging[J]. FUJITSU Scientific & Technical Journal, 2014, 50(1): 60–65.
- [11] 王园宇, 李元宗, 张天序. 粉尘环境中的图像恢复方法[J]. 华中科技大学学报(自然科学版), 2010, 38(8): 42–45. (WANG Y Y, LI Y Z, ZHANG T X. Method to restore dust degraded images[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2010, 38(8): 42–45.)
- [12] McCARTNEY E J. Optics of the Atmosphere: Scattering by Molecules and Particles[M]. New York: Wiley, 1976: 52.
- [13] NARASIMHAN S G, NAYAR S K. Vision and the atmosphere [J]. Computer Vision, 2002, 48(3): 233–254.
- [14] 武鹏飞, 方帅, 徐青山, 等. 非均匀大气中单幅图像复原[J]. 电子学报, 2013, 41(10): 1895–1902. (WU P F, FANG S, XU Q S, et al. Single image dehazing in inhomogeneous atmosphere [J]. Acta Electronica Sinica, 2013, 41(10): 1895–1902.)
- [15] 刘锦锋, 黄峰. 一种基于输运方程的大气退化图像复原方法 [J]. 光散射学报 2005, 16(4): 364–369. (LIU J F, HUANG F. A restoration method of atmospheric degraded image based radiative transfer equation[J]. The Journal of Light Scattering, 2005, 16(4): 364–369.)
- [16] ANCUTI C O, ANCUTI C, HERMANS C, et al. A fast semi-inverse approach to detect and remove the haze from a single image [C]// ACCV 2010: Proceedings of the 10th Asian Conference on Computer Vision, LNCS 6493. Berlin: Springer, 2010: 501–514.
- [17] OHASHI T, AGHBARI Z, MAKINOUCHI. Fast segmentation of texture image regions based on hill climbing[C]// Proceedings of the 2003 IEEE Pacific Rim Conference on Communications, Computers and signal Processing. Piscataway, NJ: IEEE, 2003: 848–851.
- [18] HE K M, SUN J. Fast guided filter [EB/OL]. [2016-03-05]. <https://arxiv.org/pdf/1505.00996v1.pdf>.
- [19] PELI E. Contrast in complex images [J]. Journal of the Optical Society of America A, 1990, 7(10): 2032–2040.

This work is partially supported by the Natural Science Foundation of Shanxi Province (2012011014–2).

**WANG Yuanyu**, born in 1973, Ph. D., Lecturer. His research interests include robot vision, image processing, pattern recognition.

**ZHANG Yifan**, born in 1993, M. S. candidate. Her research interests include image processing, pattern recognition.

**WANG Yunfei**, born in 1994, M. S. candidate. His research interests include virtual reality.