基于多尺度字典学习的图像融合方法

王 珺，彭进业，何贵青，冯晓毅

(西北工业大学 电子信息学院，陕西 西安 710072)

摘 要: 将小波域多尺度分析的思想和图像域单尺度稀疏表示的思想有效结合，提出基于多尺度字典学习的图像融合方法。首先将训练图像变换到小波域，分别对各个子带系数训练字典；根据训练的字典求解并融合源图像各个子带的稀疏表示系数；经过逆小波变换重构融合图像。提出的方法综合了学习字典的稀疏特性和小波分析的多分辨率特性。实验结果表明较现有基于图像域字典学习的融合方法和基于小波域多尺度分析的融合方法均具有更优的融合效果。

关 键 词: 图像融合，多尺度字典学习，稀疏表示，K-SVD
中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-2758(2013)05-0793-05

图像融合是指把同一对象经多个传感器的不同成像或经单一传感器的多次成像通过特定方法加以综合，以获得更全面、更准确描述的技术，该技术在军事、医学成像、遥感成像、机器视觉、安全监控等领域有广阔的应用前景[1]。

由于基于稀疏表示的图像融合方法表现出了比经典的基于 CVT(Curvelet Transform)、SWT(Stationary Wavelet Transform)、NSCT(Nonsubsampled Contourlet Transform)等融合方法更好的融合效果[2-5]，从而成为当前图像融合领域的一个非常活跃的研究方向。基于稀疏表示的图像融合方法都包含 4 个基本步骤：①构造字典；②求解源图像的稀疏表示；③通过特定规则融合系数；④重构融合图像。其中，步骤①和步骤③是决定融合效果的最关键步骤，是值得重点研究的内容。由于学习型字典比非学习型字典有更好的融合效果[2,3]，本文重点研究基于稀疏表示的图像融合方法中学习型字典的构造。目前基于字典学习的图像融合方法都是在图像域单一尺度下进行字典学习，这种学习型字典虽然可以精细地拟合数据，但是不能多尺度地分析数据。由于图像在不同尺度、不同方向下常常包含不同的特征，这些特征往往是图像融合需要区分和保留的突出信息，并且鉴于以往多尺度分析思想成功应用于图像融合领域，因此，我们认为在小波域多个尺度上训练的字典，将既能保持稀疏表示数据的能力，又具备多尺度分析的性能，进而在图像融合问题中能够在多尺度下用少量的系数提取显著的特征，达到比单尺度下基于字典学习的图像融合方法或者小波多尺度融合方法更优的融合效果。这就是本文提出的基于多尺度字典学习的图像融合方法的基本思路。

1 多尺度字典学习方法

稀疏表示理论基于以下假设：自然信号可以由字典中“少量”原子的线性组合表示或近似表示。对给定的信号 y，可以用如下近似的方法找出其稀疏表示

$$\min_{\alpha} \|\boldsymbol{\alpha}\|_0^0 \quad \text{subject to} \quad \|\boldsymbol{y}-\boldsymbol{D\alpha}\|_2 \leqslant \varepsilon \quad (1)$$

式中，$\boldsymbol{y} \in \mathbf{R}^{n\times 1}$，$\boldsymbol{D} \in \mathbf{R}^{n\times K}$ 表示一个字典，其中每一列称作一个原子，$\boldsymbol{\alpha} \in \mathbf{R}^{K\times 1}$ 表示向量，$\|\boldsymbol{\alpha}\|_0^0$ 表示向量 $\boldsymbol{\alpha}$ 中非零元素的个数，ε 表示允许偏差的精度，解决这个优化问题的过程通常称作“稀疏编码”。

上述模型的一个根本问题是字典的选择。目前主要有两种方法构造字典：分析变换方法和学习方法。分析变换方法学习到的字典没有明确的矩阵形

收稿日期: 2013-02-28 **基金项目:** 国家自然科学基金(61075014)、航天支撑基金(2011XW080001C080001)及西北工业大学博士论文创新基金(CX201318)资助
作者简介: 王 珺(1987—)，女，西北工业大学博士研究生，主要从事图像融合及稀疏表示的研究。

式,但是具有快速方法和高度结构化的构架,并且大多数字典可以多尺度地分析图像。如傅里叶变换、DCT 变换、小波、曲波和轮廓波等。与此相比,学习到的字典通常具有显示矩阵的表达,并且能更精细地拟合数据,从而在许多应用中性能更好[6]。但是这种非结构化的字典计算代价高,原子的大小也限制了这种字典的应用。典型学习方法有 PCA、MOD 和 K-SVD 方法。

多尺度字典学习有效地将这两种字典的优势结合,一方面使字典更有效、更全面地表示信号,另一方面使字典具有多尺度分析的能力。具体来讲,根据双重稀疏的思想修改(1)式为

$$\underset{\boldsymbol{D},\boldsymbol{\alpha}}{\operatorname{argmin}} \| \boldsymbol{Y} - \boldsymbol{W}_S \boldsymbol{D}\boldsymbol{\alpha} \|_F^0, \text{subject to} \| \boldsymbol{\alpha}_i \|_0^0 \leqslant T, \forall i. \tag{2}$$

式中 $\boldsymbol{D}$ 表示学习的字典,$\boldsymbol{\alpha}$ 表示稀疏的向量,$\boldsymbol{Y}$ 是训练信号,矩阵 $\boldsymbol{W}_S$ 表示小波变换算子的集合或等同于小波变换的字典。上式说明,表示信号的原子可以由固定的多尺度字典构成,因为该字典本身有一定的稀疏性,学习的字典在其基础上又进一步稀疏,所以称为“双重稀疏”。

假设 $\boldsymbol{W}_S$ 是酉矩阵,那么(2)式可以等价的写为

$$\underset{\boldsymbol{D},\boldsymbol{\alpha}}{\operatorname{argmin}} \| \boldsymbol{W}_A \boldsymbol{Y} - \boldsymbol{D}\boldsymbol{\alpha} \|_F^2, \text{subject to}, \| \boldsymbol{\alpha}_i \|_0^0 \leqslant T, \forall i \tag{3}$$

式中 $\boldsymbol{W}_A$ 表示小波分析算子,具体证明过程见文献[7]。上式说明,我们不仅可以在图像域中训练字典,也可以在小波域中训练字典。

对每一子带训练各自的子字典 D_b

$$\forall b, \quad \underset{\boldsymbol{D}_b,\boldsymbol{\alpha}_b}{\operatorname{argmin}} \| (\boldsymbol{W}_A \boldsymbol{Y})_b - \boldsymbol{D}_b \boldsymbol{\alpha}_b \|_F^2$$
$$\text{subject to}, \| \boldsymbol{\alpha}_{i,b} \|_0^0 \leqslant T, \forall i. \tag{4}$$

其中下标 b 表示不同的子带。最终学习的多尺度字典是子带字典的集合。

显然,多尺度字典学习方法是针对单一尺度学习型字典在图像处理中缺乏多尺度、多方向性提出的。该方法一方面通过在小波域中学习字典,从而具有小波分析的多尺度、多方向特点;另一方面,在小波系数的基础上进一步稀疏,从而达到双重稀疏的效果。因此,基于多尺度字典学习方法既能够较小波分析方法具有更好的稀疏表示源图像的能力,又能够较单尺度学习型字典具有多尺度分析的能力,从而使得该方法能够更稀疏更有效地表示图像信息。多尺度字典学习步骤将在第 3 小节给出。

2 基于多尺度字典学习的图像融合方法

2.1 图像的表示与稀疏编码

用 $\boldsymbol{cI}_b^k \in \mathbf{R}^{M\times N}$ 表示第 k 幅原始图像在 b 子带的小波系数,采用滑动窗口策略,把小波系数 $\boldsymbol{cI}_b^k$ 分成大小为 $\sqrt{n}\times\sqrt{n}$ 的块,即 $\boldsymbol{cI}_{b_j}^k$,表示 $\boldsymbol{cI}_b^k$ 中的第 j 块,$j \in [1, L]$,L 表示分块个数。再把每一块拉直作为一个向量 $\boldsymbol{cV}_{b_j}^k$。将 $\boldsymbol{cI}_b^k$ 中所有块构成的向量组成一个矩阵 $\boldsymbol{cV}_b^k$,即源图像的信号,$\boldsymbol{cV}_b^k \in \mathbf{R}^{n\times L}$。有

$$\underset{\boldsymbol{D}_b,\boldsymbol{ca}_b^k}{\operatorname{argmin}} \| \boldsymbol{cV}_s^k - \boldsymbol{D}_b \boldsymbol{ca}_b^k \|_F^2,$$
$$\text{subject to}, \| \boldsymbol{ca}_{b_j}^k \|_0^0 \leqslant T, \forall j \tag{5}$$

不难看出,(5)式中的优化问题是一个 NP 难问题,通常考虑用近似解代替。本文采用 SOMP (Simultaneous Orthogonal Matching Pursuit) 算法[8] 求解各个源图像的稀疏表示系数。SOMP 算法是 OMP 算法的变形,其中,同时输入需要表示的信号,在每一次迭代时,选择对所有残差最大的原子,最终获得的信号稀疏表示由系数不同的同一组原子构成。

2.2 系数的融合

本文采用较为常用的选取绝对值最大的策略融合图像,即融合图像的系数是由活动水平最大的系数构成。具体步骤如下:

1) 求解 $\boldsymbol{c\alpha}_{b_j}^k$ 的活动水平,它反映了第 k 幅源图像在 b 子带第 j 列(块)的特征

$$\boldsymbol{A}_{b_j}^k = \| \boldsymbol{c\alpha}_{b_j}^k \|_1 \tag{6}$$

2) 选取活动水平最大的系数作为融合图像在 b 子带的系数 $\boldsymbol{c\alpha}_b^F$,$\boldsymbol{c\alpha}_b^F$ 的第 j 列由下式获得

$$\boldsymbol{ca}_{b_j}^F = \boldsymbol{ca}_{b_j^*}^k, b_j^* = \operatorname{argmax}(\boldsymbol{A}_{b_j}^k) \tag{7}$$

3) 可以得到融合图像的子带系数矩阵 $\boldsymbol{c\alpha}_1^F$,$\boldsymbol{c\alpha}_2^F, \cdots, \boldsymbol{c\alpha}_S^F$,$S$ 表示子带总数。

4) 分别计算融合图像每一子带的融合向量 $\boldsymbol{cV}_b^F$

$$\boldsymbol{cV}_b^F = \boldsymbol{D}_b \boldsymbol{c\alpha}_b^F \tag{8}$$

2.3 基于多尺度字典学习的融合方法框架

假设待融合图像已经经过配准,有 K 幅大小为 $M\times N$ 的 $I_1, \cdots, I_K$ 源图像,基于多尺度字典学习的融合方法框架如图 1 所示,其中图 1a) 是多尺度字典学习的示意图,图 1b) 是融合步骤的示意图。以下分述多尺度字典学习步骤和融合过程。

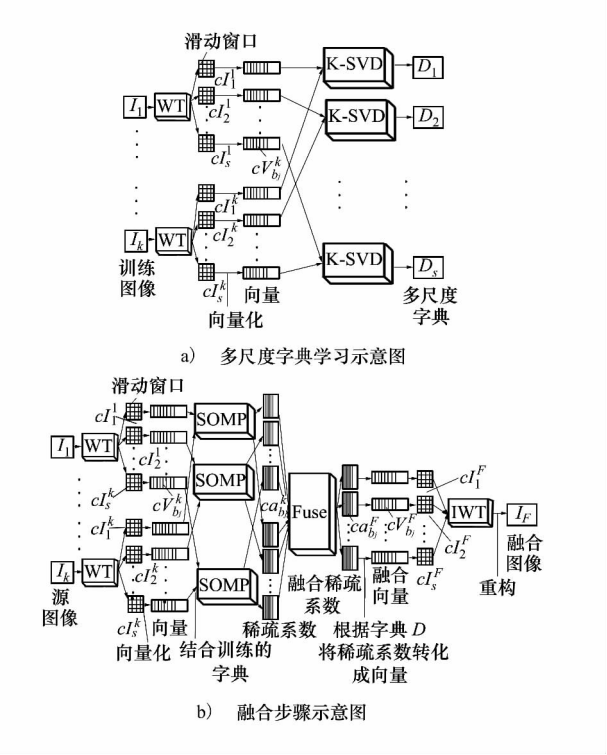


图 1　本文融合方法框架

多尺度字典学习的步骤如下:

1) 选择小波基函数、子字典个数 S(与分解级数和小波类型有关)、每个子字典的原子数 m、字典原子的大小 n 以及训练图像,其中,训练图像可以是待融合图像本身,也可以是与其同类型的图像;

2) 初始化所有子字典 $\boldsymbol{D}_b \in \mathbf{R}^{n\times m}$, $b = 1,2,\cdots,S$。

3) 用 $2D$ 小波变换分解每一幅训练图像,每一幅训练图像分解出 S 个子带系数;

4) 对所有子带系数,以 step 为滑动步长(step $\in [1,\sqrt{n}]$)、大小为 $\sqrt{n}\times\sqrt{n}$ 的滑动窗口提取块,再依次排列成向量;

5) 分别对每个子带用 K-SVD 算法训练子字典 $\boldsymbol{D}_b$;

6) 输出所有子字典,即子带字典 $\boldsymbol{D}_b \in \mathbf{R}^{n\times m}$, $b = 1,2,\cdots,S$。

融合过程如下:

1) 对 K 幅大小为 $M\times N$ 的源图像进行小波变换,得到所有源图像在各个子带的小波系数 $\{cI_b^k, k = 1,2,\cdots,K; b = 1,2,\cdots S\}$。其中,小波变换类型与分解级数与学习部分一致;

2) 从左上到右下将每一波段的小波系数 cI_b^k 以步长为 1,大小为 $\sqrt{n}\times\sqrt{n}$ 的滑动窗口方法,分解成大小为 $\sqrt{n}\times\sqrt{n}$ 的块,即字典中原子的大小。再将所有块按顺序排列成向量 $cV_{b_j}^k$,它表示第 k 幅源图像的小波系数在第 b 个子带中第 j 块,这样 $cV_{b_j}^k$ 的每一列对应 cI_b^k 的一块。所有向量组成一个矩阵,cV_b^k 的大小为 $(\sqrt{n}\times\sqrt{n}) \times [M-\sqrt{n}+1) \times (N-\sqrt{n}+1)]$;

3) 用 SOMP 算法求解子带块 $\{cV_{b_j}^k\}_{k=1}^K$ 的稀疏表示系数,得 $c\boldsymbol{\alpha}_{b_j}^1, c\boldsymbol{\alpha}_{b_j}^2, \cdots, c\boldsymbol{\alpha}_{b_j}^K$;

4) 按照 3.2 小节得到所有子带的融合向量矩阵 $cV_1^F, cV_2^F, \cdots, cV_S^F$;

5) 分别遍历每一子带的融合向量矩阵。先将每一个向量 cV_b^F 重塑成 $\sqrt{n}\times\sqrt{n}$ 大小的块,再将这些放到 cI_b^F 的对应位置,再取平均;

6) 最后进行逆小波变换,小波类型与第 1 步一样,即得融合图像 I_F。

3　实验结果与分析

为了验证本文方法的性能,将其与两类方法进行比较,分别是基于单尺度字典学习的 OMP[2,3] 图像融合方法以及基于多尺度融合方法 DWT(Discrete Wavelet Transform)。实验中参数均取常用数值,滑动窗口和字典大小分别为 8×8 和 64×256,误差 $\varepsilon = 0.01$,小波类型取 db4 小波,分解级数为 3 级。本实验以 Windows XP 操作系统为平台,使用 MATLAB 7.10 仿真。

实验数据采用图像融合数据库[9] 中的遥感图像图像大小为 256×256。由图 2 可以看出,DWT 方法的融合结果显得较为杂乱,没有体现源图像本身

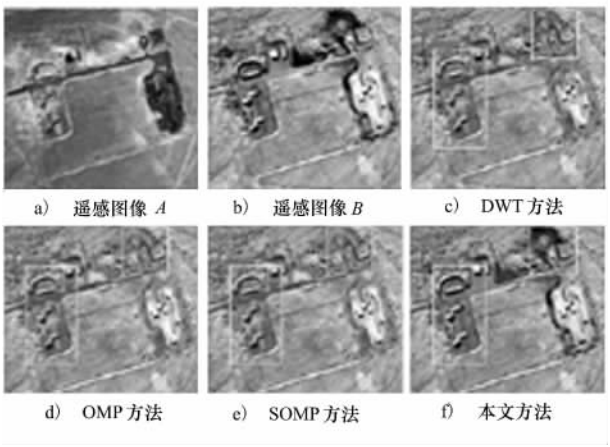


图 2　源遥感图像及几种方法的融合结果

的显著特点; OMP 和 SOMP 方法在平坦区域的边界处融合的较为生硬,出现边界效应,也没有突显遥感图像 B 中右上区的特点; 而本文方法建筑物突出,更有效地提取源图像的主要信息,所以从整体的视觉效果上更优。

除了主观视觉分析,本文还采用国内外较常用的信息熵 E 和 Q_0、Q_W、Q_E 指标进行客观评价。其中,指标 E 用来评价融合图像中所含的信息量,越大越好; Q_0 是源图像与融合图像结构化相似度的一种度量,Q_W 是源图像与融合图像窗口加权的融合质量评价,Q_E 反映了融合图像融合源图像边缘的情况,Q_0、Q_W、Q_E 的值均在 [0,1] 之间,越接近 1 表明融合质量越好[8]。表 1 给出了 4 种融合算法的客观指标,其中粗体表示最优的指标值。观察表中实验的客观评价,本文方法要优于其他几种方法,这与主观评价结果一致。

表 1 几种融合方法的客观评价

融合方法	E	Q_0	Q_W	Q_E
DWT 方法	7.327	0.759	0.776	0.578
OMP 方法	7.231	0.783	0.807	0.656
SOMP 方法	7.279	0.794	0.812	0.661
本文方法	**7.372**	**0.841**	**0.841**	**0.683**

通过对主观视觉评价和客观指标评价的结果分析可以看出,本文方法较单一图像域字典学习的融合方法能更有效表示图像细节信息(如边缘和纹理);较小波域多尺度分析的融合方法能更稀疏地提取源图像显著特征,所以具有更优的融合效果。

4 结 论

本文提出了基于多尺度字典学习的图像融合方法,该方法通过多尺度字典学习,既能有效地稀疏表示源图像的显著特征,又能多尺度地融合图像细节信息。实验结果表明,本文方法较图像域单尺度学习型字典和小波域多尺度分析的融合方法均具有更优的融合效果。由于本文方法需要在多个尺度下分块求解源图像的稀疏表示系数,所以在后续工作中,我们将重点研究图像的分块处理策略,以提高图像稀疏表示的效率,同时,还将从稀疏表示的求解方法、图像融合规则的设计等方面入手,使融合效果更加优越。

参考文献:

[1] Mahyari A G, Yazdi M. Panchromatic and Multispectral Image Fusion Based on Maximization of Both Spectral and Spatial Similarities. IEEE Transa on Geoscience and Remote Sensing, 2011, 49(6): 1-10

[2] Yang B, Li S T. Multifocus Image Fusion and Restoration with Sparse Representation. IEEE Trans on Instrumentation and Measurement, 2010, 4(59): 884-892

[3] Yang B, Li S T. Pixel-Level Image Fusion with Simultaneous Orthogonal Matching Pursuit. Information Fusion, 2012, 1(13): 10-19

[4] 严春满, 郭宝龙, 易 盟. 自适应字典学习的多聚焦图像融合. 中国图象图形学报, 2012, 9(17): 1144-1149
Yan Chunman, Guo Baolong, Yi Meng. Multi-Focus Image Fusion Using Adaptive Dictionary Learning Method. Journal of Image and Graphics, 2012, 9(17): 1144-1149 (in Chinese)

[5] Yu Nannan, Qiu T S, Bi F. Image Features Extraction and Fusion Based on Joint Sparse Representation. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 1074-1082

[6] Ophir B, Lustig M, Elad M. Multi-Scale Dictionary Learning Using Wavelets. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 1014-1024

[7] Rubinstein R, Zibulevsky M, Elad M. Double Sparsity: Learning Sparse Dictionaries for Sparse Signal Representation. IEEE Trans on Signal Processing, 2010, 58(3): 1553-1564

[8] Tropp J A, Gilbert A C, Strauss M J. Algorithms for Simultaneous Sparse Approximation. Part I: Greedy Pursuit. Signal Processing, 2006, 86(3): 572-588

An Image Fusion Algorithm Based on Multi-Scale Dictionary Learning

Wang Jun, Peng Jinye, He Guiqing, Feng Xiaoyi

(Department of Electronics Engineering, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: We combine the multi-scale analysis in wavelet domain with the single-scale sparse representation in image domain and propose an image fusion algorithm based on multi-scale dictionary learning. We transform the trained images into wavelet domain and train the dictionary for each sub-band dictionary. We use the trained dictionary to solve and fuse the sparse representation coefficient of each sub-band of a source image. The fused image is reconstructed through the inverse wavelet domain. Our algorithm combines the sparse character of a learned dictionary with the multi-resolution character of wavelet analysis. The experimental results, given in Fig. 2 and Table 1, and their analysis show that our image fusion algorithm outperforms those based on the learned dictionary in image domain and multi-scale analysis in wavelet domain respectively.

Key words: algorithms, image fusion, image processing, wavelet transforms; image domain, muli-scale analysis, muli-scale dictionary learning, sparse representation, wavelet domain