**基于联合稀疏表示和剪切波的医学图像融合**

张林发1 张榆锋1 王琨1 韩素雅1

1（云南大学信息学院 昆明 650091）

**摘要：**

**关键词：**

**A novel method for medical image fusion**

Zhang Linfa1, Zhang Yufeng1, Wang Kun1, Han Suyang1

1(Department of Electronic Engineering, Yunnan University, Kunming 650091, China)

**Abstract:**Image fusion combines multiple images to incur a single image with excellent quality, retaining the features of original images.

In parallel, the introduction of the matching pursuit, and the basis pursuit denoising gave rise to the ability to address the image denoising problem as a direct sparse decomposition technique over redundant dictionaries.

**Keywords:**

## 1 引言

随着医学影像技术的发展，衍生出多种成像方式用在临床以提供诊断信息，如核磁共振成像（MRI）、电子计算机断层扫描成像(CT)、单光子发射计算机断层成像(SPECT)、正电子发射型计算机断层成像(PET)等。但由于其各自成像机理的不同，导致各种成像技术各有优缺点[1]（如SPECT能够显示细胞和分子的生物学活动，但是缺乏组织结构信息，MIR则相反，见图1）。图像融合技术能够将两张不同成像模式的图像合成为一张，融合图像能够提供丰富的互补诊断信息，提高疾病诊断精度[2]。同时，图像融合技术的使用，能够使医生对病情的研判不再需要分别参考不同成像模式的病理图像，大大提高诊断效率[3]。

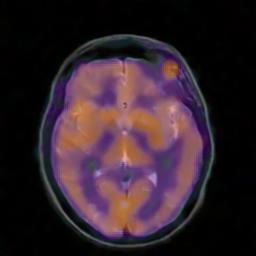
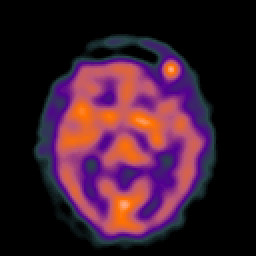
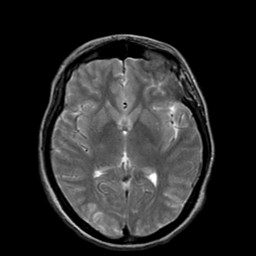
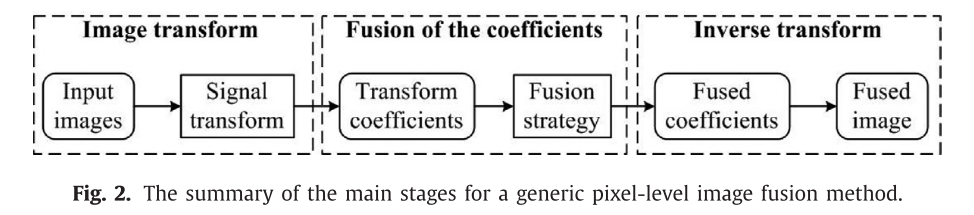


图1

近年来，许多医学图像融合的方法被提出。如李等人提出一种基于引导滤波的图像融合方法[4]。刘云等人将深度卷积神经网络应用于医学图像融合[5]，而且刘云还对深度卷积神经网络在像素级融合应用上做了总结[6]。基于多尺度变换的方法发展迅速，包括基于金字塔的、基于小波的、基于MGA的(如NSCT[10]、NSST[11])[3]。此处给出最新的关于多尺度变换方法的例子，如朱等人提出基于草图-纹理分解（CTD）和应用稀疏表示的图像融合方法[8]。杜娇等人提出一种基于局部金字塔滤波的多尺度分解（LLF）的图像融合方法[9]。一般来说，图像融合方法由图像的分解和重构，图像融合规则、图像质量评估组成[1]。小波变换及相关的多尺度变换方法只使用了有限的字典去提取原图像特征，导致相关方法都有其优点和局限性[18],如小波擅长对图像细节的表示[19]，但是不能很好地表示图像的曲线和边缘信息[16]。脊波更适合描述图像中线的特征[20]等等。曲波被提出用于表示图像[49],以及曲波在遥感图像的融合应用[50]。曲波和轮廓波变换能够捕捉图像几何结构信息，更适合于处理二维信号[51]。在轮廓波变换中，拉普拉斯金字塔第一次用于捕获点的不连续性，方向滤波器组将点的不连续性连成线性结构[16]。由于轮廓波在表示空间结构的有效性，轮廓波变换被成功应用于医学图像融合[52]、遥感图像融合[53][54]。为了克服轮廓波变换不具有移动不变性特性，轮廓波非下子采样（NSCT）被提出，但是需要更多的时间消耗[16]。由于轮廓波方向滤波器组是固定的，导致其不能表示复杂空间结构[55]。剪切波方向滤波器组没有方向限制[57]，作为一种新方法应用于图像融合[56]。在和传统的多尺度方法相比，稀疏表示的方法能够更加准确地表示图像信息[18]。

压缩感知[12]原始介绍，发展迅速。图像融合方法可以归类为三个步骤：图像变换，变换系数融合，反变换[16]。对于基于稀疏表示的方法来说，图像变换中的字典选择和稀疏系数的融合规则是关键[26]。图像融合算法主要包含四个部分，活动层评价、参数组、参数联合、一致性确认[37]。对于像素级融合算法来说，有两个点非常关键，一是活动层评价和参数联合[18]。字典学习的方法面向JSR模型取得极大进步[14]。杨等人第一次将压缩感知用于图像融合[15]，系数选择采取的是最大值原则。通常来说，稀疏稀疏采用加权平均值方法比最大值原则选取融合效果要更好[38]，因为最大值原则使得融合图像不够平滑,造成信息严重信息丢失[42]。基于SR模型的方法比传统的MST方法在主观和客观评价指标有更出色的表现[13]。SR模型有经典稀疏表示模型、非负稀疏表示模型（NNSR）、联合稀疏表示模型（JSR）、组稀疏表示模型（GSR）、鲁棒稀疏表示模型（RSR）、多任务鲁棒稀疏表示模型（MRSR），其中RSR模型更适合于多聚焦图像融合，NNSR和JSR模型更适合于多模态图像融合，GSR模型适合于两者[13]。图像融合方法可以分为三大类：像素级融合、特征级融合、决策级融合[17]，其中像素级融合方法是使用最广泛的方法，由于其直接对传感器采集的数据进行处理，对源图像信息失真度最小。可以归为四大类：多尺度变换的方法、稀疏表示的方法、直接在像素级上进行融合或者使用变换域方法（如主成分空间或IHS颜色空间）、多种分解方式（多尺度变换、稀疏表示、主成分分析、其他变换）的联合 [16]。

稀疏表示在图像融合中的应用[21]。而字典分为已准备好的字典(如轮廓波字典[22]、曲波、小波)（预设的变换矩阵，但是这种字典的使用效果取决于他们自己本身对信号的系数描述[42]）或基于原图像的自适应字典（如使用K-SVD[23][31]、PCA[45]、MOD[46]）随后系数被一些迭代优化方法处理（见2.2）。大多数基于SR的融合方法都没有考虑原图像的互补信息和冗余信息[26]。为了解决这一问题，使用准备好的字典去表示冗余信息，使用基于K-SVD自适应字典去表示互补信息，之后使用OMP去处理系数[26]。



接下来本篇论文由如下部分组成：第二章介绍稀疏表示原理和剪切波，及提出的改进点；第三章展示实验结果和结果分析；最后一章对本论文作出总结。

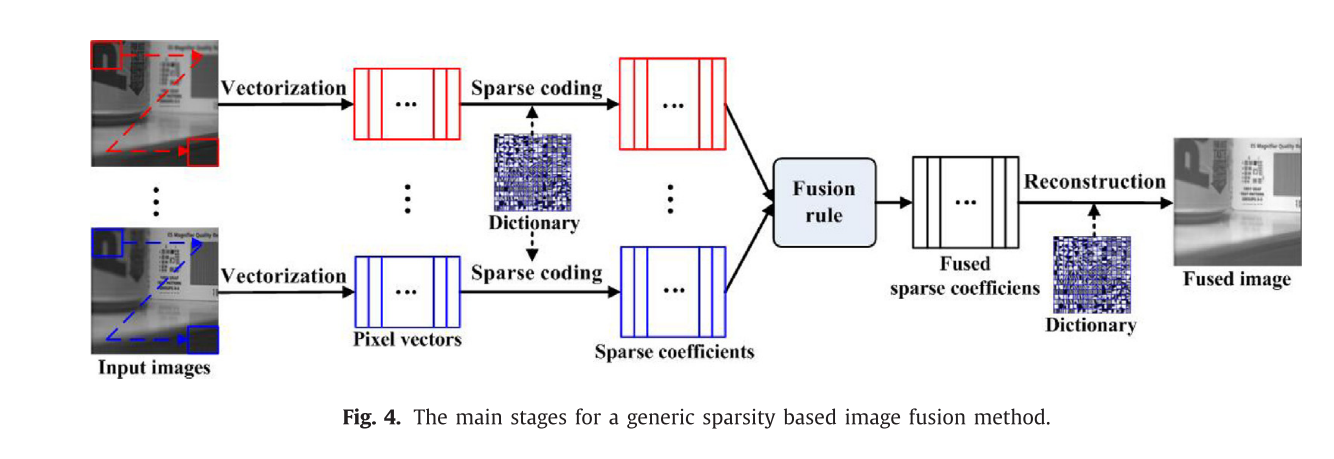
## 2 方法原理

### 2.1 联合稀疏表示

稀疏表示是一种新的图像表示方法[56],杨等人第一次将压缩感知用于图像融合[15]。为了捕获局部显著特征，同时具有平移不变性，在稀疏表示的方法中，首先将输入图像分成许多重叠小区域，之后小块从完备字典中获取稀疏系数[16]。稀疏表示广泛应用于图像处理，如图像降噪[27]、超分辨率重构[28]、图像恢复[29]。由于图像拍摄来自相同区域，那么待融合的图像之间必然存在关联，他们都存在共同的特征和差异性的特征[39][40][41]。JSR模型被广泛应用,都描述原图像有共同成分和个性成分[42][60][61]。在JSR模型中，首先假设所有图像都有共同成分和个性成分[13]。而且证明JSR能够很好解决共同成分和个性成分分离问题，当原图像可以被稀疏或者可压缩的[44]。联合稀疏这个词由作者第一次在其论文中提出[44]，有三种联合稀疏模型：JSM-1（稀疏共同成分+个性成分）、JSM-2（共同稀疏成分支持）、JSM-3（非稀疏共同成分+稀疏个性成分）[14],其中JSM-1更适合用于解决图像融合问题。然后可以用公式表示成如下：

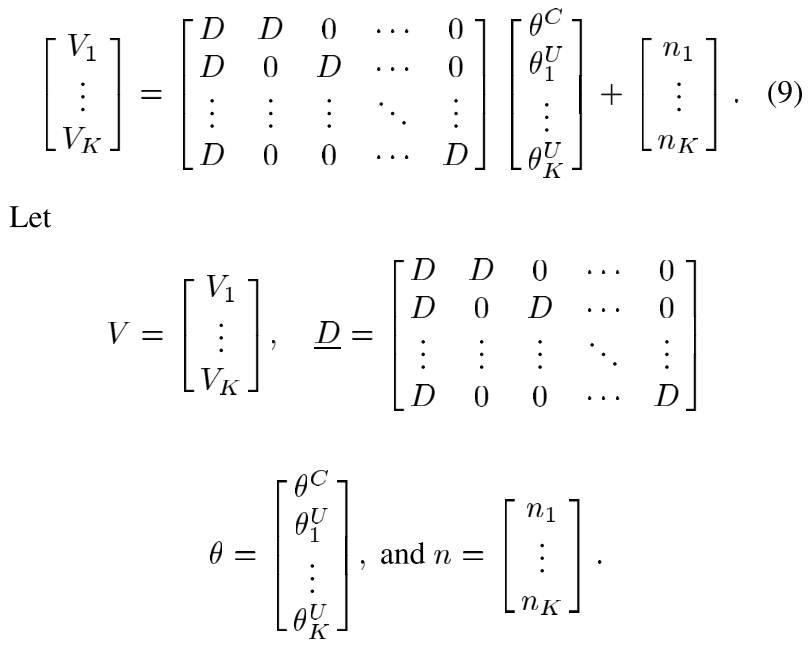


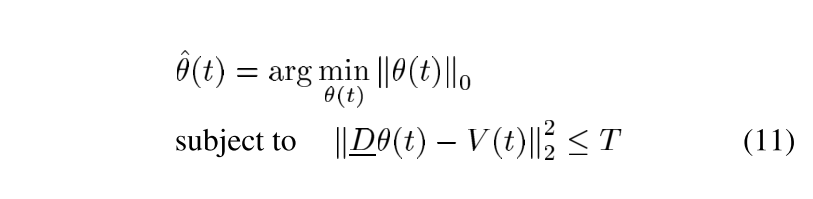
Ir、Ic表示冗余分量和互补分量



JSR模型可以转换为传统模型[12][30]，图像可以表示为下式子：

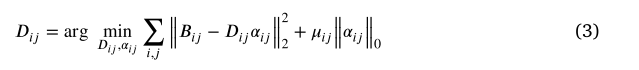






I1、I2是两张图像，Dc1和Dc2表示的是自适应字典。

自适应字典的产生，K-SVD被广泛应用及验证[23][47][48]，K-SVD算法具有噪声抑制能力，能够从噪声图像中获取清晰字典[42]。而且本文中使用K-SVD算法。



### 2.2 融合算法选择

常见的算法有如正交匹配追踪算法（OMP）[24]或者基追踪(BP)[25]或者MP[33]、ILS-DLA[32]、SOMP[34]它们都属于贪婪算法，通过连续选择字典原子实现。对于MP和OMP，不同图像的分解稀疏系数可以对应于字典中不同的原子子集。这类似于使用不同的小波族对每个输入图像块进行分解。因此，融合规则很难设计,我们希望能够不同原图像能够分解成相同的字典原子子集[18]。稀疏表示理论将原图像转化为向量，比如256x256图像会被转化为65536长度的向量，而相关的字典可能会更大，直接应用MP算法会效率不高，而其他方法(比如迭代收缩/阈值（IST）[35]、压缩采样匹配算法（CoSaMP）[36]),即便不存在维度问题，但是这些方法不能保证不用原图像分解成一样的字典原子子集。考虑到图像融合使用到原图像的局部信息，所以论文中使用重叠区域去替代整张图，重叠区域定义为滑动窗，将窗所涵盖的区域作为向量中的元素，减去局部均值相当于小波变换中低频成分[18]。

### 2.3 字典训练方法

字典的建立通常有两种方法，一是基于数学模型（小波变换、曲波变换等等）；二是基于样本学习（如K-SVD），同一类数学函数去构建字典[15]；由于每种数学模型都是面向一个特定结构，对于自然图像的表示能力较差，论文中使用了多种数学模型去建立一个混合字典[18]，混合字典能够很好地表示几种特定结构，但是不能够去表示不同类型的图像。在[14][62][63]中，通过对输入图像相似的样本进行训练获取完备字典，获取对图像自适应表示能力。然而，只有一种字典不能准确反应图像复杂结构，在[64]提出了簇训练样本变成许多结构组，然后训练一个特定的子字典，以这样的方式，一个子字典能够适应特定的机构，使得整个字典拥有更强的表示能力。同样的，分别建立频谱字典和空间细节字典用于多频谱和全色图像的融合[65]。新的系数编码和字典训练方法被提出，大多数方法都是基于传统融合方法，比如基于窗活动层评价[15][18],基于参数联合的选择最大值和加权平均[15][18][62]，系数替换（广泛用于遥感图像融合在不同空间分辨率下）[63][66]。和基于多尺度分解的方法一样，融合方法在融合效果上起着重要作用，在这篇论文中[67],设计了一种有效的基于稀疏表示的多焦点图像融合方法，该方法不仅考虑了每个图像块及其空间邻域的详细信息，而且还考虑了图像的局部特征。

## 3 实验和结果

### 3.1 实验设置

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | SSIM | RMSE | MI | PSNR | SF |
| 0.9/0.1 | 0.0725 | 0.2309 | 0.9658 | 18.1860 | 3.9030 |
| 0.8/0.2 | 0.0729 | 0.2238 | 0.9529 | 17.9757 | 4.6672 |
| 0.6/0.4 | 0.0676 | 0.2136 | 0.8821 | 17.1384 | 7.0816 |
| 0.57/0.43 | 0.0793 | 0.2127 | 0.8720 | 16.9708 | 7.5675 |
| 0.5/0.5 | 0.0774 | 0.2107 | 0.8625 | 16.5706 | 8.5449 |
| 04/0.6 | 0.094 | 0.2093 | 0.8572 | 15.9643 | 10.0262 |
| 0.3/0.7 | 0.0650 | 0.2093 | 0.8391 | 15.3182 | 11.5129 |

### 3.2 评价指标

### 3.3 和其他图像融合方法对比

### 3.4 进一步讨论

## 4 结论

因此，设计更先进的基于稀疏表示的融合规则，有望成为该领域另一个有趣的方向。

## 参考文献

[1] J. Du, W. Li, K. Lu, and B. Xiao, ‘‘An overview of multi-modal medical image fusion,’’ Neurocomputing, vol. 215, pp. 3–20, Nov. 2016.

[2] A.P. James, B.V. Dasarathy, Medical image fusion: a survey of the state of the art, Inf. Fusion 19 (2014) 4–19.

[3] M. Yin, X. Liu, Y. Liu, and X. Chen, ‘‘Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shear- let transform domain,’’ IEEE Trans. Instrum. Meas., vol. 68, no. 1, pp. 49–64, Jan. 2019.

[4] S. Li, X. Kang, J. Hu, Image fusion based on guided filtering, IEEE Trans. Image Process. 22 (7) (2013) 2864–2875.

[5] Y. Liu, X. Chen, H. Peng, and Z. F. Wang, “Multi-focus image fusion with a deep convolutional neural network,” Inf. Fusion,vol.36 pp. 191–207, Jul. 2017.

[6] L. Yu, C. Xun, W. Zengfu, W. Jane, R.K. Ward, W. Xuesong, Deep learning for pixel-level image fusion: Recent advances and future prospects, Inf. Fusion 42 (2018) 158–173.

[7] D.L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306.

[8] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi,A novel multi-modality image fusion method based on image decomposition and sparse representation, Inf. Sci., vol. 432, pp. 516–529, Mar. 2018.

[9] J. Du, W. Li, and B. Xiao, “Anatomical-functional image fusion by information of interest in local laplacian filtering domain, IEEE Trans. Image Process., vol. 26, no. 12, pp. 5855–5866, Dec. 2017.

[10] Q. Zhang and B.-L. Guo, Multifocus image fusion using the nonsubsampled contourlet transform, Signal Process., vol. 89, no. 7, pp. 1334–1346, 2009.

[11] G. Guorong, X. Luping, and F. Dongzhu, Multi-focus image fusion based on non-subsampled shearlet transform, IET Image Process., vol. 7, no. 6, pp. 633–639, 2013.

[12] D.L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306.

[13] Q. Zhang, Y. Liu, R. S.Blum, J. Han, D. Tao, Sparse representation based multi- sensor image fusion for multi-focus and multi-modality images: A review, Inf. Fusion 40 (2018) 57–75.

[14] Q. Zhang, Y. Fu, H. Li, J. Zou, Dictionary learning method for joint sparse representation-based image fusion, Opt. Eng. 52 (5) (2013) 057006.

[15] B. Yang , S. Li , Multifocus image fusion and restoration with sparse representation, IEEE Trans. Instrum. Meas. 59 (4) (2010) 884–892.

[16] S. Li , X. Kang , L. Fang , J. Hu , H. Yin , Pixel-level image fusion: a survey of the state of the art, Inf. Fusion 33 (2017) 100–112 .

[17] D. Yang, S. Hu, S. Liu, X. Ma, Y. Sun, Multi-focus image fusion based on block matching in 3d transform domain, J. Syst. Eng. Electron. 29 (2) (2018) 415–428 .

[18] B. Yang , S. Li , Pixel-level image fusion with simultaneous orthogonal matching pursuit, Inf. Fus. 13 (1) (2012) 10–19 .

[19] J.L. Starck, D.L. Donoho, E.J. Candès, Very high quality image restoration by combining wavelets and curvelets, in: Proceedings of SPIE, vol. 4478, 2001, pp. 9–19.

[20] T. Chen, J.P. Zhang, Y. Zhang, Remote sensing image fusion based on ridgelet transform, in: Proceedings of International Conference on Geoscience and Remote Sensing Symposium, 2005, pp. 1150–1153.

[21] Z. Gao, C. Zhang, Texture clear multi-modal image fusion with joint sparsity model, Optik 130 (2017) 255–265 .

[22] M. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106.

[23] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Trans. Image Process. 15 (12) (2007) 3736–3745.

[24] S. Chen, S.A. Billings, W. Luo, Orthogonal least squares methods and their application to non-linear system identification, Internat. J. Control 50 (1989) 1873–1896.

[25] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, Soc. Ind. Appl. Math. 43 (1) (2001) 129–159.

[26] X.Ma, S. Hu,S. Liu,J. Fang,S. Xu,Multi-focus image fusion based on joint sparse representation and optimum theory, Signal Processing: Image Communication 78 (2019) 125–134 .

[27] W. Dong, X. Li, L. Zhang, G. Shi, Sparsity-based image denoising via dictionary learning and structural clustering, in: CVPR, 2011.

[28] J. Yang, J. Wright, T. S.Huang, Y. Ma, Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (11) (2010) 2864–2873 .

[29] X. Gao, N. Wang, D. Tao, X. Li, Face sketch-photo synthesis and retrieval using sparse representation, IEEE Trans. Circuits Syst. Video Technol. 22 (8) (2012) 1213–1226 .

[30] Y. Liu, S. Liu, Z. Wang, A general framework for image fusion based on multi-scale transform and sparse representation, Inf. Fusion 24 (2015) 147–164.

[31] M. Aharon, M. Elad, A. Bruckstein, The K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, IEEE Transactions on Signal Processing 54 (11) (2006) 4311–4322.

[32] K. Engan, K. Skretting, J.H. Husoy, Family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation, Digital Signal Processing 17 (1) (2007) 32–49.

[33] S.G. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Transactions on Signal Processing 41 (12) (1993) 3397–3415.

[34] J.A. Tropp, A.C. Gilbert, M.J. Strauss, Algorithms for simultaneous sparse approximation. Part I: greedy pursuit, Signal Processing 86 (3) (2006) 572– 588.

[35] M. Figueiredo, R. Nowak, An EM algorithm for wavelet-based image restoration, IEEE Transactions on Image Processing 12 (8) (2003) 906–916.

[36] D. Needell, J. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, Applied and Computational Harmonic Analysis 26 (3) (2009) 301–321.

[37] R. Shen, I. Cheng, A. Basu, Cross-scale coefficient selection for volumetric medical image fusion, IEEE Trans. Biomed. Eng. 60 (4) (2013) 1069–1079.

[38] N. Mitianoudis and T. Stathaki, “Pixel-based and region-based image fusion schemes using ICA bases,” Inf. Fusion, vol. 8, no. 2, pp. 131–142, 2007.

[39] S. M. Mahbubur Rahman, M. Omair Ahmad, and M. N. S. Swamy, Contrast-based fusion of noisy images using discrete wavelet trans- form, IET Image Process., vol. 4, no. 5, pp. 374–384, 2010.

[40] E. Cand’es, X. Li, Y. Ma, and J. Wright, Robust Principal Compo- nent Analysis? preprint, 2009.

[41] D. Baron, M. Duarte, M. Wakin, S. Sarvotham, and R. Baraniuk, “Distributed compressive sensing,” in Proc. Sens., Signal, Inf. Process. Workshop, 2008.

[42] N. Yu, T. Qiu, F. Bi, et al., Image features extraction and fusion based on joint sparse representation, IEEE J. Sel. Top. Signal Process. 5 (5) (2011) 1074–1082.

[43] Z. Xu, Medical image fusion using multi-level local extrema, Inf. Fusion 19 (2014) 38–48.

[44] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, and R. Baraniuk, Distributed compressed sensing of jointly sparse signals, in Proc. Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, Nov. 2005, pp. 1537–1541.

[45] I. T. Jolliffe, Principal Component Analysis. New York: Springer, 2002.

[46] K. Engan, S. O. Aase, and J. H. Husøy, “Multi-frame compression: Theory and design,” in Proc. EURASIP Signal Process., 80, 2000, no. 10, pp. 2121–2140.

[47] M. Protter and M. Elad, “Image sequence denoising via sparse and redundant representations,” IEEE Trans. Image Process., vol. 18, no. 1, pp. 27–35, Jan. 2009.

[48] Q. Zhang and B. Li, “Discriminative K-SVD for dictionary learning in face recognition,” in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recog- nition, San Francisco, CA, 2010, pp. 2691–2698.

[49] E.J. Cands , D.L. Donoho , Curvelets and curvilinear integrals, J. Approximation Theor. 113 (1) (2001) 59–90 .

[50] F. Nencini , A. Garzelli , S. Baronti , L. Alparone , Remote sensing image fusion using the curvelet transform, Inf. Fus. 8 (2) (2007) 143–156 . Special Issue on Image Fusion: Advances in the State of the Art.

[51] M.N. Do , M. Vetterli , Contourlets: a directional multiresolution image repre- sentation, in: Proceedings of IEEE International Conference on Image Process- ing, vol. 1, 2002, pp. I–357–I–360 .

[52] L. Yang , B. Guo , W. Ni , Multimodality medical image fusion based on mul-tiscale geometric analysis of contourlet transform, Neurocomputing 72 (1-3) (2008) 203–211 .

[53] J. Saeedi , K. Faez , A new pan-sharpening method using multiobjective parti- cle swarm optimization and the shiftable contourlet transform, ISPRS J. Pho- togramm. Remote Sensing 66 (3) (2011) 365–381 .

[54] K.P. Upla , M.V. Joshi , P.P. Gajjar , An edge preserving multiresolution fusion: use of contourlet transform and MRF prior, IEEE Trans. Geosci. Remote Sens- ing 53 (6) (2015) 3210–3220 .

[55] K.P. Upla , M.V. Joshi , P.P. Gajjar , An edge preserving multiresolution fusion: use of contourlet transform and MRF prior, IEEE Trans. Geosci. Remote Sens- ing 53 (6) (2015) 3210–3220 .

[56] L. Wang , B. Li , L. Tian , Multi-modal medical image fusion using the inter-s- cale and intra-scale dependencies between image shift-invariant shearlet co- efficients, Inf. Fus. 19 (1) (2014) 20–28 .

[57] G. Easley , D. Labate , W.-Q. Lim , Sparse directional image representations us- ing the discrete shearlet transform, Appl. Comput. Harmonic Anal. 25 (1) (2008) 25–46 .

[58] B.A. Olshausen , J.F. David , Emergence of simple-cell receptive field proper- ties by learning a sparse code for natural images, Nature 381 (6583) (1996) 607–609 .

[59] H. Yin , S. Li , Multimodal image fusion with joint sparsity model, Opt. Eng. 50 (6) (2011) 067007.1–067007.10 .

[60] H. Yin , S. Li , Multimodal image fusion with joint sparsity model, Opt. Eng. 50 (6) (2011) 067007.1–067007.10 .

[61] B. Yang , J. Luo , S. Li , Color image fusion with extend joint sparse model, in: Proceedings of International Conference on Pattern Recognition, 2012, pp. 376–379 .

[62] H. Yin , S. Li , L. Fang , Simultaneous image fusion and super-resolution using sparse representation, Inf. Fus. 14 (3) (2013) 229–240 .

[63] S. Li , H. Yin , L. Fang , Remote sensing image fusion via sparse representations over learned dictionaries, IEEE Transactions on Geoscience and Remote Sens- ing 51 (9) (2013) 4779–4789 .

[64] M. Kim , D.K. Han , H. Ko , Joint patch clustering-based dictionary learning for multimodal image fusion, Inf. Fus. 27 (1) (2016) 198–214 .

[65] W. Wang , L. Jiao , S. Yang , Fusion of multispectral and panchromatic images via sparse representation and local autoregressive model, Inf. Fus. 20 (1) (2014) 73–87 .

[66] X.X. Zhu , R. Bamler , A sparse image fusion algorithm with application to pan-sharpening, IEEE Trans. Geosci. Remote Sensing 51 (5) (2013) 2827–2836 .

[67] Q. Zhang , M. Levine , Robust multi-focus image fusion using multi-task sparse representation and spatial context, IEEE Trans. Image Process. (2016) . to be published on.

## 参考文献末尾