

# 基于多尺度字典的红外与微光图像融合

薛模根, 刘存超, 徐国明, 袁宏武

(陆军军官学院 光电技术与系统实验室, 安徽 合肥 230031)

**摘要:** 基于人类视觉系统及信号的过完备稀疏表示理论, 提出了一种基于多尺度字典的红外与微光图像融合方法。首先把输入的红外与微光图像按照高斯金字塔模型分解, 用 DCT 字典作为初始字典按照四叉树的结构进行分解, 对于各尺度的字典按照 K-SVD 算法独立训练更新, 构造出多尺度学习字典。其次在该字典下利用改进的 OMP 算法得到输入源图像各自的稀疏系数, 然后按照最优化融合图像与输入源图像的欧氏距离、融合图像方差的准则, 建立一个融合图像稀疏系数的最优化函数, 最后通过求解该函数的  $l_1$  范数得到融合图像。实验结果表明: 该算法的融合效果优于小波变换法、Laplacian 塔型方法以及 PCA 方法等传统融合方法。

**关键词:** 图像融合; 稀疏表示; 多尺度字典; 四叉树; K-SVD 算法; 最优化函数

中图分类号: TP911.73 文献标识码: A 文章编号: 1001-8891(2013)11-0696-06

## Infrared and Low Light Level Image Fusion Based on Multi-scale Dictionary

XUE Mo-gen, LIU Cun-chao, XU Guo-ming, YUAN Hong-wu

(Laboratory of Electro & Optical technology and system, Army Officer Academy of PLA, Hefei 230031, China)

**Abstract:** A novel infrared and low light level image fusion algorithm based on multi-scale sparse representation is introduced on the basis of the Human Visual System and over-complete sparse representation theory in this paper. Firstly, infrared and low light level images are decomposed according to the Gaussian pyramid model. Then a multi-scale learned dictionary is obtained by using an efficient quadtree decomposition of the DCT dictionary which is considered as the initial dictionary and each scale dictionary independent training update using K-SVD algorithm. We use the improved OMP algorithm with the dictionary to get the input sparse coefficients of source images. And then we get an optimization function of the fusion image sparse coefficients by optimizing the Euclidean distances between fused image and each input, weighted by their own variance. Finally, we obtain the fusion image by solving the  $l_1$  norm of the function. The experimental results show that the proposed method exhibits considerably higher fusion performance than the typical methods such as the wavelet transform method, the Laplacian pyramid method and Principal Component Analysis method.

**Key words:** image fusion, sparse representation, multi-scale dictionary, quadtree, K-SVD algorithm, optimum function

## 0 引言

红外热像仪和微光夜视仪是当前 2 种主要的夜视设备。其中, 红外热像仪是依靠目标(背景)本身红外辐射分布不同而成像, 它可以提供温度梯度较大或与背景有较大热对比的低可视目标的红外图像, 但其对场景的亮度变化不敏感; 微光夜视仪是利用光的反射而成像, 所获取的微光图像具有较高的时空分辨

率, 但在雨、雾、霾等恶劣天气条件下, 存在目标和细节缺失等缺点<sup>[1]</sup>。由于红外与微光图像的信息具有互补性及冗余性, 因此通过对两者的融合, 可以有效提高目标的可识别度和图像的清晰度。目前, 红外与微光图像常用的融合方法有: 加权平均法、PCA (Principal Component Analysis) 变换法<sup>[2]</sup>、图像金字塔法<sup>[3]</sup>、小波融合方法<sup>[4]</sup>等。加权平均法、PCA 变换法消弱了图像的对比度, 尤其是只出现在一幅图像中

收稿日期: 2013-06-05; 修订日期: 2013-07-20.

作者简介: 薛模根(1964-), 男, 教授, 主要从事图像融合及图像处理研究。

基金项目: 安徽省自然科学基金, 编号: 1208085QF126。

的有用信号; 图像的金字塔分解是冗余分解, 在高频融合中信息损失大且分解无方向; 小波变换不能最稀疏表示图像结构中的直线和曲线的奇异性。

受人类视觉皮层神经元响应的稀疏性启发, Olshausen 和 Field 提出了一种有效的信号表示方法——稀疏表示<sup>[5]</sup>。该方法是一种自适应的数据表示方法, 它能根据图像的特点自适应地选取与其最相关的原子来进行表示。由于多尺度字典相对于单尺度字典能够最优匹配图像中的各成份结构, 具有更强的稀疏表示能力。基于此, 本文提出了一种基于多尺度字典稀疏表示的红外与微光图像融合方法。

## 1 基于稀疏表示的多尺度字典的构造

### 1.1 图像的稀疏表示

稀疏表示的基本思想是用过完备字典取代传统信号表示中的正交基<sup>[6]</sup>。由于过完备字典的冗余性, 信号能够表示为过完备字典中少数原子的线性组合, 其中利用原子最少(即最稀疏)的表示称为稀疏表示。稀疏表示的数学模型为:

$$\hat{\alpha} = \arg \min \|\alpha\|_0 \quad \text{s.t.} \quad \|y - D\alpha\| \leq \varepsilon \quad (1)$$

在稀疏求解中,  $D$  中的列向量  $\{d_1, d_2, \dots, d_m\}$  称为原子,  $D \in \mathbb{R}^{n \times k}$  ( $n < k$ ) 称为字典, 由于  $D$  的行数小于其列数, 字典  $D$  是超完备字典或者称为冗余字典,  $y \in \mathbb{R}^n$  为原信号(本文中即为图像),  $\alpha$  为信号的表示;  $\varepsilon$  为逼近误差,  $\|\cdot\|_0$  为  $l_0$  范数, 它表示向量中非零元素的个数。

图像稀疏表示本质上可以看成是从字典  $D$  中搜寻到最稀疏的解  $\alpha$ , 即向量  $\alpha$  中所含有的非零元素个数最少的过程, 而该过程本质上是组合优化问题。目前国内外研究求解该问题的算法可以归入以下 3 类: 贪婪追踪算法<sup>[7]</sup>、凸松弛算法<sup>[8]</sup>、组合算法<sup>[9]</sup>。

字典构造是一种寻找稀疏表示下最优基的构造, 字典的优劣是原始信号能否尽可能稀疏表示的关键<sup>[10]</sup>。字典的构造不仅要满足稀疏表示唯一性的约束, 也要获得更完备更精确的信号表示。字典构造的方法主要分为 2 类, 一类是构造一个由  $g'$  组成的字典:

$$g' = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{-\pi \xi t^2} \quad (2)$$

式中:  $g'$  是由基函数  $g(t)$  转换得到;  $s$ 、 $u$  和  $\xi$  分别表示膨胀量、转换参量和调频参量。

这类方法虽可保证以较高频率恢复信号, 但无法选择满足重构目标特点字典; 另一类是从训练样本中直接生成字典。字典是从与待重构目标相似的样本集中产生, 由于具有与待重构目标相似的特征, 目标信号可以更容易通过这类字典稀疏表示。

### 1.2 多尺度字典的构造

相对于传统从训练样本中直接生成的单尺度学习字典, 本文提出了一种从训练样本中生成多尺度学习字典的方法。在提出多尺度字典框架之前, 首先介绍图像的四叉树分解模型, 如图 1 所示。这是一个典型的数据保存、表示、处理的数据结构, 广泛应用于图像表示、边缘检测、图像压缩中, 在多尺度脊波分解<sup>[11]</sup>中已经得到应用。假设一个图像块的大小为  $n$ , 按照四叉树模型其子图像块的大小为  $n_s = n/4^s$ ,  $s=0, \dots, S-1$  表示树的深度。因此不同尺度的字典  $D_s \in \mathbb{R}^{n_s \times k_s}$  由  $k_s$  个大小为  $n_s$  的原子组成。多尺度字典  $D \in \mathbb{R}^{n \times k}$  是 1 个字典集, 由不同尺度的字典分布在四叉树不同分支处的字典组成。一个多尺度字典  $D \in \mathbb{R}^{n \times k}$  在四叉树模型中有  $4^s$  个分支, 该多尺度字典中的原子数  $k = \sum_{s=0}^{S-1} 4^s k_s$ 。

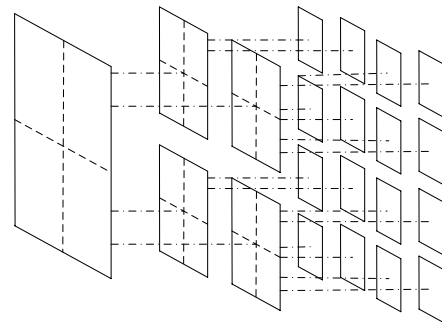


图 1 四叉树分解结构图

Fig.1 The quadtree decomposition framework

多尺度字典的构造是在 K-SVD 算法<sup>[12]</sup>的基础上形成的。为了建立一个多尺度字典, 把输入的源图像进行高斯金字塔图像分解, 该金字塔图像通过下式得到:

$$I_p = (Y * B_p) \downarrow_{s_p} \quad (3)$$

式中:  $Y$  表示输入的源图像;  $B_p$  表示标准方差为  $\sigma^2 \lg(p)/\lg(5)$  的高斯核;  $\downarrow_{s_p}$  表示下采样;  $s_p = (0.8)_p$ ;  $p$  表示分解的层数。

从高斯金字塔图像中抽取大小为  $\sqrt{n} \times \sqrt{n}$  的图像块序列  $\{Z_i\}_{i=1}^N$ 。每个图像块都可以被多尺度字典  $D \in \mathbb{R}^{n \times k}$  稀疏表示, 字典  $D$  包含有  $k$  个原子,  $s$  个不同尺度的字典。源图像中的图像块  $n$  按照四叉树的结构分解, 树的每一层图像块的大小为  $n_s = n/4^s$  ( $s=0, \dots, S-1$ )。本文通过以下方法生成一个多尺度自适应学习字典。

首先把  $D_s \in \mathbb{R}^{n_s \times k_s}$  中的每一个原子按照四叉树的分解结构图进行分解, 如图 2。

其次, 选用 DCT 字典作为多尺度字典的初始字典

典, 选择高斯金字塔中图像块作为样本图像, 按照 K-SVD 算法中字典更新的方法独立地训练出不同尺度的字典, 最后把同一尺度不同位置处更新后的原子取平均组成一个原子, 如图 3。

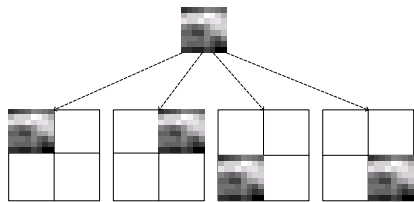


图2 尺度为2的原子分解

Fig.2 Decomposition of an atom with a 2-scale dictionary

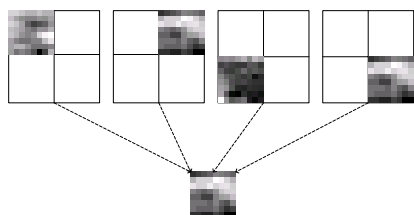


图3 尺度为2的原子的合成

Fig.3 Composition of an atom with a 2-scale dictionary

按照上述方法, 本文构造出了  $S=3$  的多尺度自适应学习字典, 如图 4。

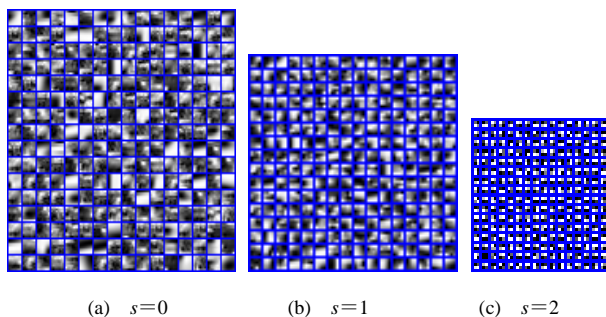


图4 一个3尺度自适应学习字典

Fig.4 An adaptive learned 3-scale dictionary

## 2 融合算法

假设红外与微光图像在多尺度字典  $D$  下, 利用贪婪追踪算法中改进的正交匹配追踪算法——GramSchmidt 算法<sup>[13]</sup>求得的稀疏系数为  $\alpha$  和  $\beta$ , 融合图像  $F=D\gamma$  在该字典下的系数为  $\gamma$ 。因此, 融合图像的优劣等价于系数  $\gamma$  从  $\alpha$  和  $\beta$  中选择的好坏, 若构造的系数  $\gamma$  最优, 即融合图像  $F$  最优。

对于融合图像的优劣, 其客观评价能一定程度的反应出。客观评价的大小表明了融合图像中信息量的多少, 表明了融合图像与源图像的相似程度。为了生成一个评价价值优的融合图像, 本文采用融合图像的方差值以及融合图像和源图像的差异 2 个准则来建

立目标函数。方差是由图像灰度值相对于均值的分散测定, 其值越大则图像的灰度阶调变化范围越大, 图像的方差越大, 图像的信息量也越大, 分辨率越高; 融合图像与源图像差异越小, 融合图像的质量越高。基于此, 建立一个目标优化函数, 满足融合图像与源图像的差别尽可能小, 而其方差尽可能大, 如下式:

$$y^* = \arg \min_y \|\nabla x_1(y - x_1)\|_{l_2}^2 + \|\nabla x_2(y - x_2)\|_{l_2}^2 \quad (4)$$

式中:  $x_1$  和  $x_2$  分别表示红外和微光图像;  $y$  表示融合图像。残留量  $y - x_1$  和  $y - x_2$  表征了源图像与融合图像的差别,  $\nabla x_1$  和  $\nabla x_2$  是残留量的权重。由于  $x_1 = D\alpha$ ,  $x_2 = D\beta$ ,  $y = D\gamma$ , 因此式(4)等价于:

$$J = \|\nabla x_1 D(\gamma - \alpha)\|_{l_2}^2 + \|\nabla x_2 D(\gamma - \beta)\|_{l_2}^2 \quad (5)$$

令:  $\Phi_1 = \nabla x_1 D$ ,  $\Phi_2 = \nabla x_2 D$ , 则:

$$J = \|\Phi_1 \gamma - \Phi_1 \alpha\|_{l_2}^2 + \|\Phi_2 \gamma - \Phi_2 \beta\|_{l_2}^2 \quad (6)$$

上式, 目标函数可以看成是一个二次方程, 为了进一步的分析, 上式等价于:

$$J = \gamma^T (\Phi_1^T \Phi_1 + \Phi_2^T \Phi_2) \gamma - 2((\Phi_1 \alpha)^T \Phi_1 + (\Phi_2 \beta)^T \Phi_2) \gamma + (\Phi_1 \alpha)^T \Phi_1 \alpha + (\Phi_2 \beta)^T \Phi_2 \beta \quad (7)$$

这里, 我们定义  $P$ ,  $Q$ ,  $C$  分别为:

$$\begin{aligned} P &= \Phi_1^T \Phi_1 + \Phi_2^T \Phi_2 \\ Q &= \Phi_1 \alpha \Phi_1^T + \Phi_2 \beta \Phi_2^T \\ C &= (\Phi_1 \alpha)^T \Phi_1 \alpha + (\Phi_2 \beta)^T \Phi_2 \beta \end{aligned} \quad (8)$$

则式(7)转化为:

$$J = \gamma^T P \gamma - 2Q^T \gamma + C \quad (9)$$

因此, 融合图像的最优化稀疏系数为:

$$\gamma^* = \arg \min_{\gamma} \frac{1}{2} \gamma^T P \gamma - Q^T \gamma + \frac{1}{2} C, \quad \|\gamma\|_{l_p} < L \quad (10)$$

对于上式  $l_p$  范数, 这里我们讨论当  $p=0,1,2$  三种形式。

1)  $l_0$  范数: 由于  $l_0$  范数优化函数是非凸的, 该问题是一个 NP 难问题, 通常很难找到信号的稀疏解, 在实际中很难实现。

2)  $l_1$  范数: 虽然  $l_1$  范数的稀疏性没有  $l_0$  范数强, 但  $l_1$  范数具有唯一性, 而且有很强的鲁棒性; 同时对于  $l_1$  范数的优化问题, 已具有许多强有力的数学工具支持相应的优化问题求解。另外, 由于  $\gamma$  解集非空, 同时又要求其非零系数尽可能的少。因此, 式(10)可

以写为:

$$\gamma^* = \arg \min_{\gamma} \frac{1}{2} \gamma^T P \gamma - Q^T \gamma + \frac{1}{2} C + \lambda \|\gamma\|_{l_1} \quad (11)$$

式中:  $\lambda > 0$  该问题转换为 LASSO 问题<sup>[14]</sup>。

3)  $l_2$  范数: 其稀疏性虽没有前两者好, 但是有很好的凸性和可微性。按照  $l_1$  范数的形式, 上式可以改写为:

$$\begin{aligned} \gamma^* = \arg \min_{\gamma} \frac{1}{2} \gamma^T P \gamma - Q^T \gamma + \frac{1}{2} C + \lambda \|\gamma\|_{l_2}^2 = \\ \arg \min_{\gamma} \frac{1}{2} \gamma^T (P + \lambda I) \gamma - Q^T \gamma + \frac{1}{2} C \end{aligned} \quad (12)$$

该方程转换为不适定问题中 Tikhonov 正则化<sup>[15]</sup>方法, 可通过下式求解:

$$\frac{d}{d\gamma} \left( \frac{1}{2} \gamma^T (P + \lambda I) \gamma - Q^T \gamma + \frac{1}{2} C \right) = 0 \quad (13)$$

$$\gamma^* = (P + \lambda I)^{-1} Q \quad (14)$$

通过对上述 3 种正则化问题进行分析, 不难发现  $l_0$  范数稀疏性最好, 但这是个 NP 难问题, 很难找到有效的解决方法;  $l_2$  范数的稀疏性最差, 但有可行性解决方案;  $l_1$  范数不仅有很好的稀疏性, 而且有行之有效的解决方法。因此, 本文采用  $l_1$  范数求解融合图像的稀疏系数。

### 3 实验结果与分析

为了验证所提出方法的有效性, 选择了 3 种图像

融合领域里广泛使用的图像融合技术——PCA 融合方法、Laplacian 金字塔融合方法和小波融合方法与本文的方法进行对比。本文实验采用了两组源图像大小均为  $360 \times 270$  大小的红外与微光图像。

在图 3 和图 4 中, (a) 表示红外图像, (b) 表示微光图像, (c) 表示 PCA 融合结果; (d) 和 (e) 分别表示小波基为 ‘db8’ 的小波融合结果和 Laplacian 金字塔融合结果, 两者采用了四层小波分解, 低频系数的选择采取加权平均法, 高频系数的选择采取绝对值最大的融合准则; (f) 表示本文算法融合结果。

在图 3 和图 4 中, 微光图像的场景信息丰富, 细节信息清晰; 而在红外图像中, 细节信息纹理模糊, 但是人成像特别清晰。利用红外与微光图像融合技术就可以很好的将两幅图像中信息相互补充, 可以更好的表达出场景中信息。由融合结果可以看出, PCA 融合方法中的目标模糊不清晰, 不利于人眼对目标的观察和识别。与 Laplacian 塔型方法、小波变换方法相比, 本文方法得到的融合结果较前两种方法要好, 既继承了红外图像中的目标信息又较好的保留了微光图像中的场景信息, 图像的信息量更多。

为比较不同融合方法的性能, 本文采用典型的图像融合结果客观评价方法来分析即图像熵、标准方差、融合图像与源图像的互信息<sup>[16]</sup>、从源图像到融合图像的边缘信息传递程度  $Q^{AB/F}$ <sup>[17]</sup>。

由表 1 和表 2 可以得出, 本文方法的图像熵、标准方差、互信息以及  $Q^{AB/F}$  都取得了较大值。这是由

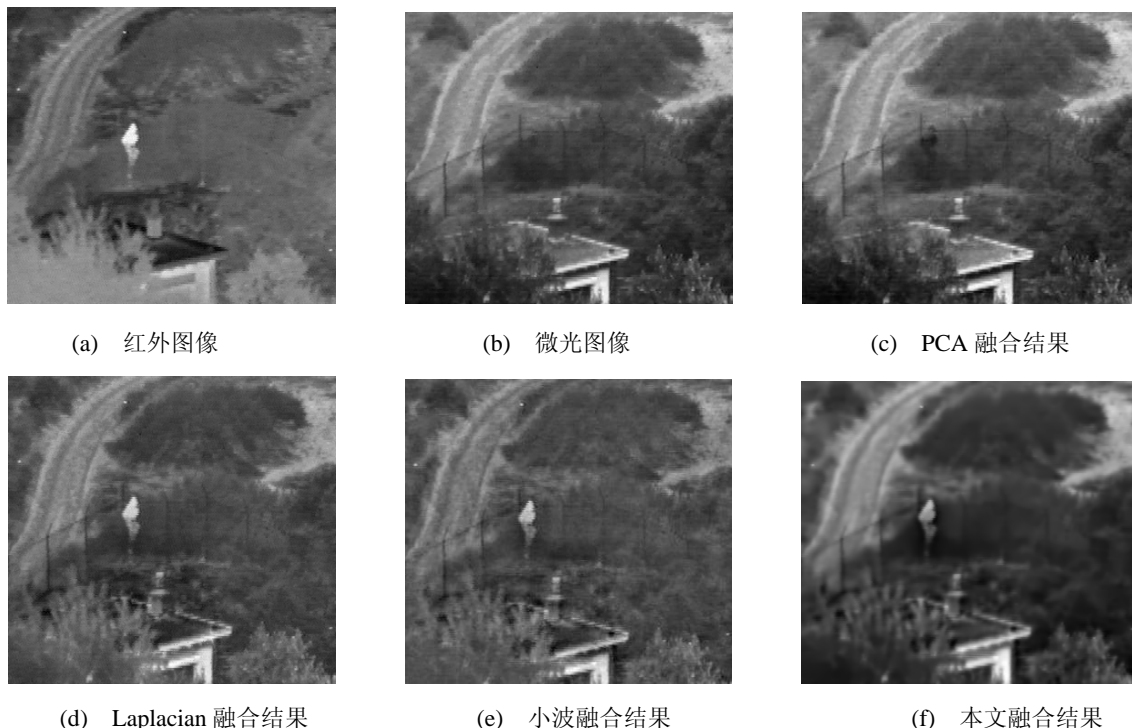


图3 实验 1 融合结果

Fig.3 Results of experiment 1 fusion

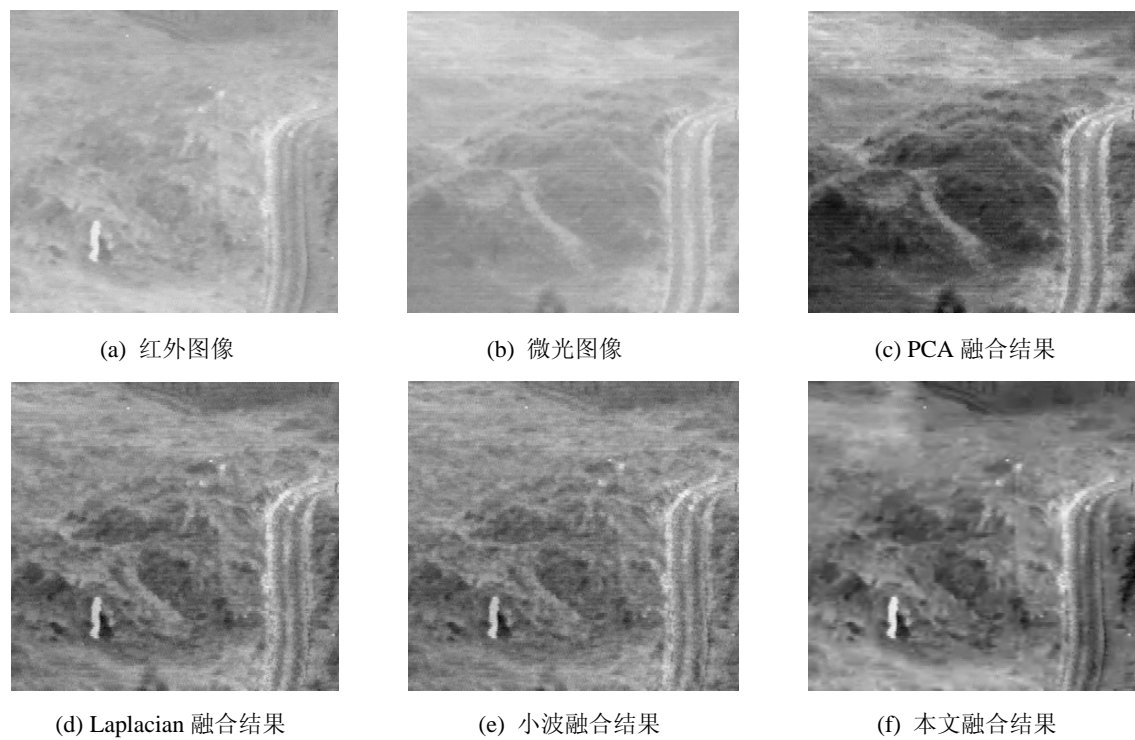


图 4  实验 2 融合结果      Fig.4  Results of experiment 2 fusion

表 1  不同融合算法的定量分析结果

Table 1  The quantitative analysis results of different fusion algorithms

	图像熵	$Q^{AB/F}$	互信息	标准方差
Laplacian 融合	6.7017	0.4213	1.4735	31.2996
小波融合	6.6291	0.4161	1.4718	28.9523
本文方法	7.0966	0.4325	2.0748	41.3733

表 2  不同融合算法的定量分析结果

Table 2  The quantitative analysis results of different fusion algorithms

	图像熵	$Q^{AB/F}$	互信息	标准方差
Laplacian 融合	6.8558	0.4010	1.4027	27.3575
小波融合	6.7799	0.3698	1.3728	28.1614
本文方法	6.9929	0.4128	2.3237	31.3035

于多尺度字典相对于单尺度字典能够最优匹配图像中的各成份结构，具有更强的稀疏表示能力，图像 的表示效果更好；又由于本文的融合算法不仅使得融合 图像的方差尽可能的大（图像中的信息量大）又使得 融合图像与源图像的差异尽可能的小。因此本文算法 得到的融合图像携带的信息量较大，融合效果较好， 与主观评价结果一致。

4  结  论

本文提出了一种基于多尺度字典稀疏表示的红 外与微光图像融合方法。利用高斯金字塔模型、四叉 树结构及 K-SVD 算法建立一个多尺度字典，在该字 典的基础上，建立一个融合图像的方差最大，与源图 像的差距最小的目标函数，通过最优化求解该模型， 得到融合图像。实验结果表明，利用本文算法得到的 融合图像比传统方法得到的融合图像场景清晰、信息 量大，更利于人眼的观察。

参考文献：

[1] 徐铭蔚,李郁峰,陈念年,等. 多尺度融合与非线性颜色传递的微光 与红外图像染色[J]. 红外技术, 2012, 34(12): 722-728.

[2] 时海亮,魏涛,辛向军,等. 基于 PCA 和 NSCT 的多光谱图像和全色 图像的融合[J]. 计算机工程与应用, 2012, 48(10): 212-216.

[3] 张建勋,牛文宾,张凯文.一种改进的基于小波变换的图像融合算法[J]. 重庆理工大学学报:自然科学版, 2012, 26(1): 61-65.

[4] 章坚武,颜欢,包建荣.改进的基于拉普拉斯先验的贝叶斯压缩感知 算法[J]. 电路与系统学报, 2012, 17(1): 34-40.

[5] Olshausen B A, Field D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature*, 1996, 381 (6583): 607-609.

[6] Hu J, Li S, Yang B. Remote sensing image fusion based on HIS transform and sparse representation[C]//Pattern Recognition (CCPR), 2010 Chinese

- Conference on IEEE*, 2010: 1-4.
- [7] Fang Hong, Zhang Quan-Bing, Wei Sui. Image reconstruction based on improved backward optimized orthogonal matching pursuit algorithm[J]. *Journal of South China University of Technology (Natural Science)*, 2008, **36**(8): 23-27.
- [8] 肖永亮, 苏显渝, 薛俊鹏, 等. 基于凸松弛全局优化算法的视觉测量位姿估计[J]. *光电子·激光*, 2011, **22**(9): 1384-1389.
- [9] 韩生廉, 武晓今, 倪萌. 组合优化问题中遗传算法的局限性及其改进模式[J]. *控制与决策*, 2002, **17**(2): 220-223.
- [10] 李修志, 吴键, 崔志明, 等. 复杂交通场景中采用稀疏表示的车辆识别方法[J]. *中国图像图形学报*, 2012, **17**(3): 387-392.
- [11] 邓承志. 基于多尺度脊波字典的图像去噪算法[J]. *计算机工程*, 2010, **36**(23): 207-211.
- [12] Aharon M, Elad M, Bruckstein A. The K-SVD: an algorithm for designing of overcomplete dictionaries for sparse representation[J]. *IEEE Transactions on Signal Processing*, 2006, **54**(11): 4311-4322.
- [13] Candes E J, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. *IEEE Transactions on Information Theory*, 2011, **52**(2): 489-509.
- [14] Tibshirani R. Regression shrinkage and selection via the lasso[J]. *Journal of the Royal Statistical Society: Series B*, 1994, **58**: 267-288.
- [15] HAN Yu, WU Fan, WAN Yong-jian, et al. Application of Tikhonov regularization method in Zernike polynomials fitting[J]. *Opto-Electronic Engineering*, 2010, **37**(7): 60-63.
- [16] Qu Guihong, Zhang Dali, Yan Pingfan. Information measure for performance of image fusion. *Electronics Letters*, 2002, **38**(7): 313-15.
- [17] Xydeas C S, Petrovic V. Objective image fusion performance measure[J]. *Electronics Letters*, 2000, **36**(4): 308-309.

(上接第 686 页)

进行拟合, 得到了吸收带尾特征参数 $\alpha_g$ 、 $E_{\text{Urbach}}$ , 对 AlGaIn 材料在截止波长外的光吸收模型进行了补充, 并在模型中考虑了异质结的极化效应, 最终获得了与实测结果的吻合度更高的仿真结果。此外选取了不同组分和厚度的器件结构进行了仿真计算, 其结果与实测值保持了高吻合度, 验证了仿真设计工作的有效性, 使器件设计能力得到了提升。

#### 参考文献:

- [1] 黄鑫, 罗木昌, 周勋. 背照式 AlGaIn/GaN 基光电探测器的结构设计及性能模拟[J]. *红外与激光工程*, 2011, **40**(11): 2071-2077.
- [2] SILVACO Inc. *ATLAS User's Manual*[M]. 2010: 5.3.7.
- [3] Joachim Piprek. *Nitride Semiconductor Device(Principles and Simulation)*[M]. 2007: 30-32.
- [4] 王勇淮. 极化效应对 AlGaIn/GaN 异质结 pin 光探测器的影响[D]. 西安: 西安电子科技大学, 2011.
- [5] Charles J C. Aluminum gallium nitride-based solar-blind ultraviolet photodetectors[D]. Austin: The University of Texas at Austin, 2002.
- [6] 高频, 张益军. 不同结构的反射式 GaAs 光电阴极的光谱特性比较[J]. *红外技术*, 2011, **33**(7): 429-432.
- [7] 周圣军. 大功率 GaN 基 LED 芯片设计与制造技术研究[D]. 上海: 上海交通大学, 2011.