

香农熵加权稀疏表示图像融合方法研究

李奕^{1,2} 吴小俊¹

摘要 针对传统稀疏表示同步超分图像融合模型中对于 LL (Low-low frequency)、LH (Low-high frequency)、H (High frequency) 三部分等比例加权, 不能突出重点信息之不足, 本文提出一种香农熵多视角加权稀疏表示同步超分图像融合方法. 该方法引入香农熵加权技术, 针对 LL、LH、H 三部分根据图像特征进行自适应加权, 突出重点频段的影响, 从而提高了图像融合的效果. 在多组不同类型图像上进行了实验, 实验结果表明所提方法无论从融合视觉效果还是评价指标上均显示出有效性.

关键词 香农熵, 多视角加权, 稀疏表示, 图像融合

引用格式 李奕, 吴小俊. 香农熵加权稀疏表示图像融合方法研究. 自动化学报, 2014, 40(8): 1819–1835

DOI 10.3724/SP.J.1004.2014.01819

Image Fusion Based on Sparse Representation Using Shannon Entropy Weighting

LI Yi^{1,2} WU Xiao-Jun¹

Abstract In order to overcome the shortcoming caused by the equal weighting on LL (Low-low frequency) components, LH (Low-high frequency) components of low resolution images and H (High frequency) components in the sparse representation image fusion model, a novel image fusion method of Shannon entropy multi-view weighting based on sparse representation is proposed. The proposed method can assign different weights to LL, LH and H components, and adaptively enhance the influences of the important components. Thus, the image fusion effect can be improved effectively. Experimental results obtained on different kinds of images indicate that the proposed method can achieve the promising performance in terms of both visual quality and quantitative evaluation metrics.

Key words Shannon entropy, multi-view weighting, sparse representation, image fusion

Citation Li Yi, Wu Xiao-Jun. Image fusion based on sparse representation using Shannon entropy weighting, *Acta Automatica Sinica* 2014, 40(8): 1819–1835

图像融合是通过一种算法将两幅或多幅图像合成一幅新图像. 其主要思想是采用一定的方法, 把具有不同成像机理的多个成像信息融合成一个新图像, 从而使融合图像可信度更高、模糊性更小、可理解性更好.

在众多的融合方法中, 小波理论融合方法依然是目前研究的热点之一. 文献 [1] 较早将小波理论应用于图像融合领域中. 由于离散二进制小波变换在提取图像低频信息的同时, 获得了三个方向的高频细节信息, 在理论上较传统的金字塔融合方法具

有更好的效果. 因而, 文献 [2] 提出采用小波变换的方法进行图像融合处理. 随着小波技术运用于图像融合领域的不断发展, 文献 [3] 对一维小波和二维小波变换的基础知识, 以及小波变换融合方法的框架结构分别进行了总结, 并将此类方法用于多聚焦、遥感、医学图像的融合处理中. 在这些理论的推动下, Li 等^[4] 提出将不可分离小波框架变换用于彩色遥感图像融合中. 进一步研究发现, 小波变换不擅长于表达长边界. 据此, 文献 [5] 提出将小波变换与 Curvelet 变换相结合, 发挥比单一变换更佳的融合性能. 此外, 很多学者都在关注小波变换与其他方法结合的研究问题, 例如比较著名的研究有: 文献 [6] 提出非下采样 Contourlet 变换图像融合方法. 对于低频信息, 采用方向模值的方法选取融合准则, 高频信息根据方向的频带限制对比度和方向向量标准差设置融合策略. 而文献 [7] 提出了一种将小波四叉树和非下采样方向滤波相结合的 CP (Contourlet packet) 方法. 该变换具备了小波包尺度精细的特点以及 NSDFB (Nonsampled directional filter bank) 的可转化特性. 而近年来, 小波变换的研究较

收稿日期 2013-04-10 录用日期 2014-02-17
Manuscript received April 10, 2013; accepted February 17, 2014
国家自然科学基金 (60973094, 61103128, 61373055), 教育部科学技术研究重大项目 (311024)
Supported by National Natural Science Foundation of China (60973094, 61103128, 61373055), and Key Grant Project of Chinese Ministry of Education (311024)
本文责任编辑 潘泉
Recommended by Associate Editor PAN Quan
1. 江南大学物联网工程学院 无锡 214122 2. 青岛大学国际教育合作学院 青岛 266071
1. School of Internet of Things Engineering, Jiangnan University, Wuxi 214122 2. College of International Cooperation in Education, Qingdao University, Qingdao 266071

多地集中在多分辨率和多尺度方面. 文献 [8] 中进行了专门的研究, 发现短滤波比长滤波能提供更好的融合效果. 而针对多尺度问题这方面的研究, 也成为近年来研究的重点问题, 比较经典的多尺度方法有 Laplacian 金字塔^[9]、Morphological 金字塔^[10]、DWT (Discrete wavelet transform)^[11]、Gradient 金字塔^[12]、SWT (Stationary wavelet transform)^[13]、DTCWT (Dual-tree complex wavelet transform)^[14]、NSCT (Non sub-sampled contourlet transform)^[15-16]. 2006 年, Candes^[17] 提出“压缩感知”理论, 抛弃了传统的采样方法必须满足采样率至少是最大频率二倍的限制条件, 提出了一种新型的压缩感知理论^[18] 采样方法, 成为香农理论发展具有里程碑意义的成果. 而在探索信号压缩处理的领域中, 稀疏表示是一个伴生问题. 在早期的对信号进行压缩的操作中, 就蕴含这一思想. 传统的信号压缩方法大多是将原始信号变换到某一种变换域中, 利用变换域系数是稀疏的这一特性, 即信号在变换域中的表示只有有限个非零的系数, 这些有限的系数就包含了原始信号的信息, 在信号的传输过程中, 只将非零的系数值以及位置信息进行传输, 这些非零的系数值就是早期的稀疏系数, 这就达到了信息压缩的效果. 在该领域, 文献 [19] 阐述了有关稀疏表示的相关理论. 伴随着压缩感知技术的提出以及应运而生的稀疏表示技术在图像分类^[20]、文本检测^[21] 等领域的广泛运用, 推动该技术不断发展. 但由于将稀疏理论运用到图像融合领域中需要有效地解决图像超分辨率和字典学习等问题. 因此该领域的研究除在较著名的文献 [22-24] 中提及外, 还很少有涉及. 在图像融合同步超分研究中比较新的成果如文献 [22], 其提出了一种采用稀疏表示方法进行图像融合的思想. 在图像融合过程中增加高频部分, 通过稀疏表示理论实现图像融合和超分辨率同步完成. 可以有效避免将两过程分离带来的显著问题^[22]: 先超分再融合过程中, 超分阶段的人为因素会影响图像融合阶段得到的融合图像质量; 先融合再超分过程中, 融合阶段的人为因素会干扰超分阶段工作, 不利因素甚至会被放大. 除此之外, 在超分辨率研究方面^[24-26], 针对由于仪器缺陷和不完善的光学成像问题带来的难以获取到高空间频率上的高光谱图像问题, 文献 [24] 利用稀疏表示和光谱混合模型实现单张高光谱图像的超分辨率. 不同波长高光谱观测的小补丁可以表示为预训练字典原子的线性加权组合. 这样超分辨率问题就转变为在稀疏限制下的最小平方问题.

针对上述多数方法虽然在某种类型或某些类型的图像融合中表现出一定的有效性, 但大量研究发现其依然面临着一些问题: 1) 在图像融

合过程中, LL (Low-low frequency)、LH (Low-high frequency)、H (High frequency) (以下简称为 LL、LH、H) 三部分通常采用等比例融合, 不能针对不同类型图像的具体特征有针对性地突出重点频段以提高融合效果. 例如由于医学图像往往对细节要求较高, 这就需要在融合中增加相应频段部分的权重. 2) 在稀疏表示字典学习过程中, 通常采用将 LH 和 H 部分字典放在一起进行学习的方法, 并人为认定两者近似相等, 这样有可能会降低学习的准确性.

针对该挑战, 由于熵加权和多视角加权技术能够根据最大熵理论、较准确地确定各个视角信息的权重, 在该方面表现出的独特优势受到了广泛关注. 因而, 本文将引入该技术对上述如何将融合图像的 LL、LH、H 三部分进行有效加权的图像融合领域研究的新问题做相应的探讨.

本文主要完成的工作有: 首先, 对香农熵加权方法进行理论分析, 为本文方法进一步研究提供了理论依据; 其次, 提出香农熵多视角加权稀疏表示同步超分图像融合方法模型; 再次, 在此基础上提出该方法的算法; 最后, 通过实验验证本文方法的有效性. 本文方法具有以下两个特点:

1) 采用香农熵自适应加权方法确定 LL、LH、H 三部分频段的权重. 当参数 λ 的取值较大时, 趋向于权重信息均等化; 相反, 则权重信息可以各有侧重. 这样可根据不同类型图像的特征, 突出重点频段, 以达到更佳的融合效果. 这样一来, 文献 [22] 中的方法成为本文方法的一种特例.

2) LH 和 H 可以分开学习, 各自拥有不同的初始字典集合、目标函数, 以提高学习准确性.

本文结构如下: 第 1 节介绍了稀疏表示技术; 第 2 节介绍了香农熵加权方法、提出方法模型和算法; 第 3 节进行了实验研究并进行了相关结果讨论; 最后, 对本文进行了总结.

1 稀疏表示

如果将信号 x 表示为 Ds (D 表示字典集, s 表示稀疏系数), 信号的稀疏表示^[19-21] 问题即是字典学习和稀疏系数寻找确定问题. 稀疏系数的寻找和确定的问题通常采用式 (1) 所示的最优化问题来解决. 在此基础上, 稀疏表示的问题实质上是在已知字典的前提下如何通过确定稀疏系数 s 来较完整的表达信号 x 的过程.

$$\begin{aligned} \hat{s} &= \arg \min_s \|s\|_F \\ \text{s. t. } &Ds = x \text{ 或 } \|Ds - x\|_2^2 \leq \varepsilon \end{aligned} \quad (1)$$

式 (1) 中 ε 表示错误容忍度, $\|\cdot\|_F$ 表示 F 范式. F 取 0 时, 表示 l_0 范式, 如式 (2) 所示, 是计算准确

度较高的一种确定稀疏系数的方法,但是存在计算复杂的缺陷以及 NP 难问题,因此在实际应用中较难找到实际解决方法. 当 F 取 1 时,表示 l_1 范式,如式 (3) 所示,相对于 l_0 范式在计算复杂度上有所改善,其中较具代表性的方法有匹配搜索 (Matching pursuit, MP)^[27] 和正交匹配搜索 (Orthogonal matching pursuit, OMP)^[28]. 除此比较著名的是凸优化^[29] 问题,如 BP 搜索方法. 当 F 取 2 时,表示 l_2 范式,由于其限制约束较为松散,在实际中较少应用.

$$\min_{\alpha} \|s\|_0 \quad \text{s. t.} \quad \|Ds - x\|_2^2 \leq \varepsilon \quad (2)$$

$$\min_{\alpha} \|s\|_1 \quad \text{s. t.} \quad \|Ds - x\|_2^2 \leq \varepsilon \quad (3)$$

2 香农熵多视角加权稀疏表示同步超分图像融合方法

2.1 问题提出

文献 [22] 较早将稀疏表示用于图像融合和超分辨率同步完成,有效克服将两个过程人为分开带来的相互干扰的缺陷,成为图像融合领域的贡献性成果. 该方法基于 LH、H 部分拥有相同稀疏度、初始字典集合、目标函数的假设提出,说明 LH、H 部分存在一定的相似性^[22]. 进而,为了验证一幅图像的 LL、LH、H 部分是否存在部分信息相似重复的问题,本文进行了专题实验研究. 实验中选择第 3 节 CT、Goldhill、Tank、Remote 图像,选择直方图^[30] 方法计算相似度. 由于 LH 和 H 部分相似程度较大,我们对这两者进行了重点研究. 结果如图 1 所示. 实验结果可以发现: 1) 同一幅图像的 LH 和 H 频段部分存在一定的信息相似成分,而且不同图像的相似程度存在一定差异. 因而,无论什么图像均人为设定每一项融合系数均取 1,将 LL、LH、H 三频段等比例加权,会出现因忽略 LL、LH 与 H 的相关性,而造成信息重复的情况出现. 2) 若融合过程中不能合理安排三部分的权重,例如在应该适度提高 H 部分权重的情况下,还按照等比例加权的方法,人为设定

H 的权重为 1,本应增加的 H 部分信息未能增加,而造成的融合图像中 H 部分信息损失的问题,反映在实验结果上就是融合图像性能的降低. 为进一步分析,我们在第 3.3 节在多组具体图像上进行了专题实验研究,实验结果显示多数情况下按照本文加权方法得到的融合图像在性能上优于等比例加重的结果,这也印证了上述推断.

为了解决上述问题,有效实现自适应权重学习,我们进行试探性探讨. 从图像分辨率的角度,可以将一幅图像分解为 LL、LH、H 三部分. 多视角学习的主要目的是从不同的视角来观察同一事物^[31],从此处得到启示,图像融合中 LL、LH、H 三频段恰恰正是从三个角度来观察同一幅待融合的图像,这与多视角学习的基本理论是相符的,因此我们可以将三个频段图像视为同一幅图像的三个视角,也是本文方法的动机之一. 基于此,可以把前面所提的不同频段加权问题转换为多视角加权问题. 本文将该技术引入到稀疏表示图像融合方法中,得到能够自适应权重学习的新方法. 为能较好地自适应学习,提高融合图像的质量,本文根据均值和平均梯度两个指标来试探性地构造新的多视角加权学习准则.

2.2 香农熵多视角加权

根据上述分析,为得到合理的权重,将融合得到的 LL、LH 以及 H 部分视为融合图像的三个视角,引入视角划分权重系数 w_k ,为进一步对视角划分权重系数 w_k 加以调控,令 $\sum_{k=1}^K \omega_k = 1$ ($w_k \geq 0$). 若将视角划分权重看作概率分布,则可用香农熵表示为

$$f(w_k) = - \sum_{k=1}^K w_k \ln w_k \quad (4)$$

根据上述分析,结合香农熵理论和超分辨率要求,在图像融合评价指标中均值和平均梯度分别反映融合后图像的平均亮度与清晰度两方面性能. 因而,尝试从均值、平均梯度两个角度来构造香农熵多视角加重的目标函数.

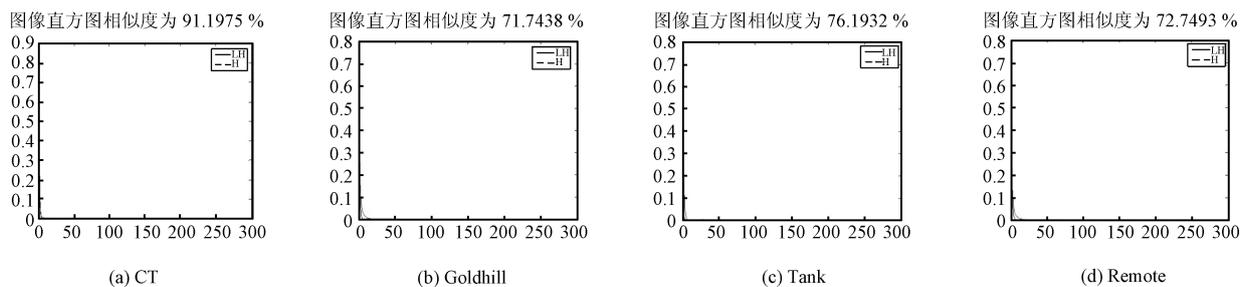


图 1 LH 和 H 部分直方图相似度

Fig. 1 The histogram of LH and H

均值如式 (5), 反映图像的平均亮度.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (5)$$

这里 x_i 表示图像灰度矩阵的灰度值, n 表示像素点的总数.

平均梯度如式 (6), 能敏感地反映图像对细节的表达能力.

$$G = \frac{1}{(M-1)(N-1)} \times \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} \sqrt{\frac{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2}{2}} \quad (6)$$

M, N 表示图像 $F(x, y)$ 的行、列数.

因而, 本文构造的香农熵多视角加权目标函数, 如式 (7). 这里采用哪些评价指标来构造目标函数仍然是一个开放性的问题, 我们也将进一步深入探讨.

$$\begin{aligned} \min J = & -\alpha \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K \omega_k F_{k,i} - \\ & \beta \frac{1}{M \times N} \sum_{k=1}^K \omega_k \left(\sum_{i=1}^N \sum_{j=1}^M ((F_{ij,k} - \hat{F}_{ij,k})^2 + \right. \\ & \left. (F_{ij,k}^T - \check{F}_{ij,k})^2 \right)^{\frac{1}{2}} + \\ & \lambda \sum_{k=1}^K \omega_k \ln \omega_k \quad (7) \\ \text{s. t. } & \sum_{k=1}^K \omega_k = 1, \quad K = 3 \end{aligned}$$

其中

$$\begin{aligned} F_{1,i} &= X_{LL} \\ F_{2,i} &= X_{LH} \\ F_{3,i} &= X_H \\ \hat{F}_k &= [O[F_{ij,k}]], \quad i \neq 2 \\ \check{F}_k &= [O[F_{ij,k}^T]], \quad i \neq 2 \end{aligned}$$

$$\check{F}_k = [O[F_{ij,k}^T]], \quad j \neq 2$$

式 (7) 中, ω_k 表示加权项, $F_{ij,k} - \hat{F}_{ij,k}$ 表示行逼近, $F_{ij,k}^T - \check{F}_{ij,k}$ 表示列逼近, F 表示融合图像, K 表示视角数. 正则化参数有 α, β, λ , 其中 α, β 表示目标函数中两项间的重要程度, 取值越大说明其对应项越突出; λ 取值越大, 视角间趋于平均化, 取值越小越能突出视角的重要程度, 通过大量实验发现: 当 λ 趋于无穷时, 各视角均等化, 此时本文方法退化为文献 [22] 中的方法. 进而, 通过定理 1 给出权重计算方法.

定理 1. 已知 $\sum_{k=1}^K \omega_k = 1, K = 3$, 则目标函数式 (7) 取得极值时, 应满足必要条件 (8).

证明. 根据约束条件 $\sum_{k=1}^K \omega_k = 1$, 引入拉格朗日乘子 η 得到相应的优化目标函数式如下:

$$\begin{aligned} J = & -\alpha \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K \omega_k F_{k,i} - \beta \frac{1}{M \times N} \sum_{k=1}^K \omega_k \\ & \sum_{i=1}^N \sum_{j=1}^M \sqrt{((F_{ij,k} - \hat{F}_{ij,k})^2 + (F_{ij,k}^T - \check{F}_{ij,k})^2)} + \\ & \lambda \sum_{k=1}^K \omega_k \ln \omega_k - \eta \sum_{k=1}^K (\omega_k - 1) \quad (9) \\ \text{s. t. } & \sum_{k=1}^K \omega_k = 1, K = 3 \end{aligned}$$

其中

$$\begin{aligned} F_{1,i} &= X_{LL} \\ F_{2,i} &= X_{LH} \\ F_{3,i} &= X_H \\ \hat{F}_k &= [O[F_{ij,k}]], \quad i \neq 2 \\ \check{F}_k &= [O[F_{ij,k}^T]], \quad j \neq 2 \end{aligned}$$

根据 $\frac{\partial J}{\partial \omega_k} = 0$, 可得式 (10).

$$\omega_k = \frac{\exp \left(\frac{\frac{1}{N} \alpha \sum_{i=1}^N F_{k,i} + \beta \frac{1}{M \times N} \left(\sum_{i=1}^N \sum_{j=1}^M ((F_{ij,k} - \hat{F}_{ij,k})^2 + (F_{ij,k}^T - \check{F}_{ij,k})^2)^{\frac{1}{2}} \right)}{\lambda} \right)}{\sum_{k=1}^K \exp \left(\frac{\frac{1}{N} \alpha \sum_{i=1}^N F_{k,i} + \beta \frac{1}{M \times N} \left(\sum_{i=1}^N \sum_{j=1}^M ((F_{ij,k} - \hat{F}_{ij,k})^2 + (F_{ij,k}^T - \check{F}_{ij,k})^2)^{\frac{1}{2}} \right)}{\lambda} \right)} \quad (8)$$

$$\omega_k = \exp \frac{\frac{1}{N} \alpha \sum_{i=1}^N F_{k,i} + \beta \frac{1}{M \times N} \sum_{i=1}^N \sum_{j=1}^M ((F_{ij,k} - \hat{F}_{ij,k})^2 + (F_{ij,k}^T - \check{F}_{ij,k})^2)^{\frac{1}{2}} + \eta - \lambda}{\lambda} \quad (10)$$

又已知 $\sum_{k=1}^K \omega_k = 1$ 消除 η 即可得到式 (8) 的表达式. \square

根据经典的极大熵原理, 该方法通过引入熵技术使目标函数在达到最优的同时熵尽可能大. 由于熵的极大化会使各概率分量尽可能地趋于均衡, 因而在各个视角上引入了极大熵加权的概念, 可以使最终获取的划分合理, 并且有效地降低那些不具有聚类特性的视角对算法优化迭代时的干扰, 获取最为理想的效果.

2.3 香农熵多视角加权算法

本节给出采用上述香农熵多视角加权的对融合后的 F_{LL}, F_{LH}, F_H 部分进行多视角加权的算法, 能够充分考量融合后的 F_{LL}, F_{LH}, F_H 三部分的相似

程度并以此指导融合过程. 具体算法如下所示:

算法 1. 香农熵多视角加权算法

步骤 1. 根据式 (8) 计算得到 LL、LH、H 视角加权系数 $w(1), w(2), w(3)$;

步骤 2. 按照以下融合策略进行加权融合.

$$F = \lambda_i(w(1)F_{LL} + w(2)F_{LH} + w(3)F_H)$$

其中, λ_i 表示正则化项, 用于调节 F_{LL}, F_{LH}, F_H 三者权重的数量级; F 表示融合图像; F_{LL}, F_{LH}, F_H 分别表示融合图像的 LL、LH、H 频段部分.

2.4 方法框架

本节给出新方法的框架结构, 如图 2 所示. 整体

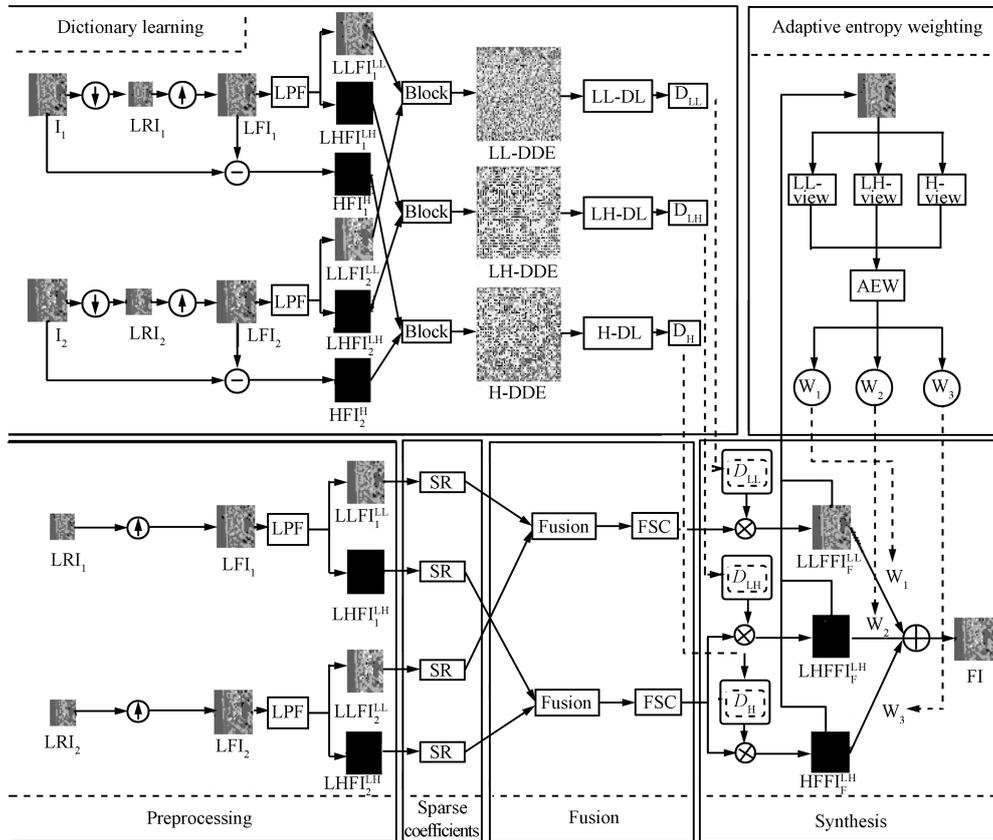


图 2 本文方法框架

Fig. 2 Overview of the proposed method

(LR: low resolution; LF: low frequency; LPF: low-pass filter; LHF: low-high frequency; LLF: low-low frequency component; SR: sparse representation; FSC: fused sparse coefficients; HFF: fused high frequency; LLFF: fused low-low frequency component; LHFF: fused low-high frequency component; D_{LL} : dictionary for LL; D_{LH} : dictionary for LH; D_H : dictionary for H; ODS: original dictionary set; AEW: adaptive entropy weighting; FI: fusion image; DL: dictionary learning)

方法的结构可以分为预处理、稀疏表示、融合、字典学习、香农熵多视角加权、综合 6 部分。其中, 预处理部分主要完成将待融合图像分别分解为 LL、LH 两频段部分, 为下一步工作做好准备。稀疏表示主要完成将待融合图像分解得到的 LL、LH 频段部分表示为稀疏编码形式。融合过程完成 LL、LH 及增加 H 频段融合过程。字典学习描述字典学习过程, 具体分析见第 2.5 节。香农熵多视角加权描述融合图像的 LL、LH、H 三部分如何加权融合的问题, 具体分析见第 2.2 节和第 2.3 节。综合部分完成融合图像的生成及输出。

2.5 字典学习

本节主要介绍传统字典学习方法中存在的问题及本文的改进。在众多的字典学习方法中 K-SVD^[32-35] (Singular value decomposition) 由于其良好的学习性能备受关注, 本文拟采用该方法对字典学习部分进行探讨。传统方法中, 将 D_{LH} 和 D_H 放在一起学习, 该方法虽然体现出学习简便、快捷的特点, 但 D_{LH} 和 D_H 拥有相同的初始字典块集合, 在同一个目标函数求解过程中完成两个字典学习过程, 认定学习后的字典 $D_{LH} = D_H$ 。理论分析及实验研究发现, 此种策略并不是最佳方案, 在本文实验阶段亦对该策略所得之结果进行了分析, 通常这种处理方法会影响最终的融合结果, 即最终获取的融合结果仍不是最佳的结果。为了进一步提升融合效果, 本文尝试将 D_{LH} 和 D_H 进行分离学习。字典分离学习的相关理论研究尚不多见, 在文献 [36] 中有类似的思想被涉及, 该文提出将 D_H 单独进行学习以获取更好的融合效果。据此, 本文进行试探性尝试, 主要抓住分离思想, 具体地, 将使 D_{LH} 和 D_H 各自拥有不同的初始字典块集合、目标函数、进行不断优化, 以分别确定最终的字典 D_{LH} 和 D_H 。如果遇到融合图像是低频图像, 无法从 D_H 中寻找到有关的初始字典块集合的情况, 尝试采用与待融合图像属同一类型的高频图像的字典块集合作为该待融合图像高频部分字典 D_H 学习的初始字典块集合, 通过式 (13) 所示的目标函数寻优求解, 以确定最终的高频部分字典 D_H 。这个问题我们将进一步深入研究。例如: 待融合图像是低频的 CT 图像, 尝试选择满足结构相似度要求的其他高频 CT 图像, 选取这组高频图像中的字典块作为该待融合图像的初始高频字典块集合进行学习。

具体 K-SVD 学习过程采用稀疏编码与字典更新交替进行的策略。针对图像的 LL、LH、H 三部分分别搜集得到的初始字典块集合, 以及稀疏系数 τ_1 、 τ_2 、 τ_3 , 利用式 (11)~(13) 进行计算, 两过程交替进行, 直到求解得到目标函数值取最小时相应的字典

和稀疏系数 a_i , 如图 3。

$$(D_{LL}, A) = \arg \min_{D_{LL}, A} \sum_{i=1}^K \| Z_i^{LL} - D_{LL} \alpha_i \|^2$$

$$\text{s. t. } \forall i \|\alpha_i\|_0 \leq \tau_1 \quad (11)$$

$$(D_{LH}, A) = \arg \min_{D_{LH}, A} \sum_{i=1}^K \| Z_i^{LH} - D_{LH} \alpha_i \|^2$$

$$\text{s. t. } \forall i \|\alpha_i\|_0 \leq \tau_2 \quad (12)$$

$$(D_H, A) = \arg \min_{D_H, A} \sum_{i=1}^K \| Z_i^H - D_H \alpha_i \|^2$$

$$\text{s. t. } \forall i \|\alpha_i\|_0 \leq \tau_3 \quad (13)$$

这里 $Z^{LL} = [Z_1^{LL}, Z_2^{LL}, \dots, Z_k^{LL}]$, $Z^{LH} = [Z_1^{LH}, Z_2^{LH}, \dots, Z_k^{LH}]$, $Z^H = [Z_1^H, Z_2^H, \dots, Z_k^H]$, α_i 表示稀疏系数, τ_i 表示稀疏度。就字典分开学习在融合准确性方面所表现出的优势, 本文在第 3.4 节进行专题实验, 实验中参数设置为: 字典块大小选取 2~8, 字典原子大小为 4~64, 稀疏度均取 6, 字典块也可滑动选取, 训练样本采用满足结构相似度大于 0.8 的同类型图像中随机生成的策略, 训练样本数目与图像大小、分块大小相关。

2.6 方法实现

根据上述分析, 本节提出具体算法 (简称为 SSR-FUSION), 如下所示:

算法 2. SSR-FUSION 算法

预处理: 该部分主要任务是针对不同类型的待融合图像进行预处理。

步骤 1. 判断图像是否属于高频图像, 若是高频图像, 则步骤 2; 否则, 转步骤 3;

步骤 2. 对待融合图像进行下采样处理;

步骤 3. 对待融合图像进行上采样处理。

字典学习: 分别构造三个频段 LL、LH 和 H 初始字典块, 根据各自稀疏度和约束条件, 采用 K-SVD 方法进行学习, 并得到相应字典。

步骤 1. 确定 LL、LH 和 H 三频段初始学习字典块集合和稀疏系数 τ_1, τ_2, τ_3 ;

步骤 2. 根据式 (11)~(13) 计算, 满足目标函数值取最小条件, 返回步骤 3; 否则, 返回步骤 2;

步骤 3. 得到字典 D_{LL}, D_{LH}, D_H 。

稀疏系数产生: 确定待融合图像三个频段 LL、LH、H 的稀疏系数。

步骤 1. 待融合图像进行分块, 分块的大小取 $n \times n$;

步骤 2. 采用 OMP 方法确定待融合图像 LL、LH、H 三部分的稀疏系数。

香农熵多视角加权系数产生: 根据香农熵加权

目标函数 (式 (7)) 和计算方法 (式 (8)) 计算融合图像三个频段 LL、LH、H 的权重系数。

步骤 1. 根据式 (8) 计算权重 w 的值;

步骤 2. 如果计算得到的成本函数 cost 大于错误容忍度, 则返回步骤 1, 否则, 返回步骤 3;

步骤 3. 输出 w 计算结果。

融合过程: 根据稀疏系数选取最大的融合准则和三个频段 LL、LH、H 的权重系数进行加权融合, 得到融合图像向量 \mathbf{X}_F 。

按照如下 LL 和 LH 频段部分融合准则^[22] (式 (14) 和 (15)) 确定融合图像:

步骤 1.

$$s(F_{LL}) = \begin{cases} s(A_{LL}), & s(A_{LL}) > s(B_{LL}) \\ s(B_{LL}), & \text{其他} \end{cases} \quad (14)$$

步骤 2.

$$s(F_{LH}) = \begin{cases} s(A_{LH}), & s(A_{LH}) > s(B_{LH}) \\ s(B_{LH}), & \text{其他} \end{cases} \quad (15)$$

步骤 3. 根据融合规则 (式 (16)) 得到融合图像向量 \mathbf{X}_F 。

$\mathbf{X}_F =$

$$\lambda_i(w(1)F_{LL} + w(2)F_{LH} + w(3)F_H) \quad (16)$$

得到融合图像: 根据融合图像向量 \mathbf{X}_F 得到融合图像。

步骤 1. 从 \mathbf{X}_F 得到融合图像的灰度矩阵;

步骤 2. 输出融合图像。

本文提出的方法可以分为待融合图像预处理、字典学习、稀疏系数的生成、香农熵加权系数产生、融合过程、得到融合图像几个阶段。其中, 待融合图像预处理阶段主要实现两幅或多幅待融合图像前期处理工作。根据待融合图像属于高频图像和低频图像分成两种情况: 若属高频的待融合图像, 进行下采样, 将图像分解为 LL、LH、H 频段。这种情况由于超分图像已存在, 既失去了图像超分的意义, 因此只在模拟实验中出现。若待融合图像是低频图像, 直接对其采用低通滤波器进行采样, 得到 LL、LH 两部分频段。字典学习过程分别从预处理后的图像中提取分块大小为 $n \times n$ (n 通常取值 2, 4, 8 等) 的初始字典块, 构成初始字典块集合, 根据式 (11) 和 (12) 迭代运算, 得到字典集合 D_{LL} 和 D_{LH} 。高频字典学习 D_H 过程分为两种情况: 1) 若待融合图像属高频图像, 将图像分解得到 H 部分作为初始高频字典块集合, 根据式 (13) 迭代寻找满足目标函数最小的字典块集合的取值, 得到高频字典 D_H ; 2) 若待融合图像是低频图像, 尝试采用与待融

合图像结构相似度最大的同类型高频图像中, 相同分块大小的字典块作为初始字典块集合, 采用相同的方法进行学习得到高频部分的字典集 D_H 。稀疏系数产生阶段根据目前字典学习情况产生的字典, 采用 OMP 方法进行最优化求解得到 LL、LH、H 三部分频段的稀疏系数。香农熵加权系数产生部分是根据最大熵理论, 通过迭代过程确定目标函数取得最小值时, 相应的权重 $w(1)$, $w(2)$, $w(3)$ 。融合过程中, 根据 LL、LH、H 三部分的融合准则进行图像融合。最后, 将 LL、LH、H 三部分按照式 (16) 加权以得到最终融合结果。根据上述分析, 本文提出的新方法具有以下特点: 1) 引入香农熵多视角加权思想对融合图像的 LL、LH、H 三部分频段进行智能自适应加权, 可以有针对性地突出重点信息, 得到较理想的融合效果; 2) 通过 D_{LH} 、 D_H 独立学习, 提出了一种字典学习新思路, 以提高 D_H 学习的准确性; 3) 图像融合过程和超分过程同步完成。

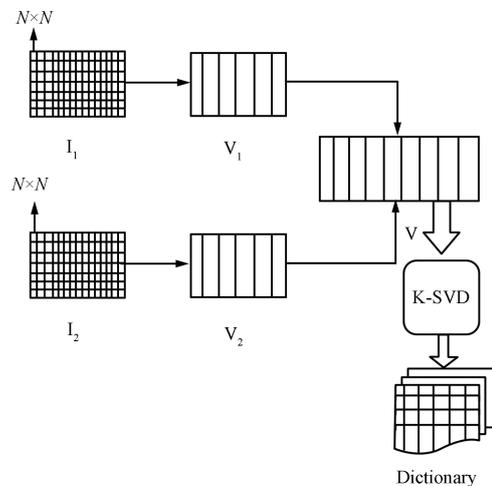


图 3 K-SVD 字典学习

Fig. 3 Dictionary learning based on K-SVD

3 仿真实验及分析

实验部分选择在图像融合领域广泛存在的医学图像、多聚焦图像、遥感图像和小目标图像上进行, 通过与 SRSR (Super-resolution using sparse representation)、NSCT-1、NSCT-2、DWT-1、DWT-2、SR-1、SR-2 方法的比较 (见表 1), 并对实验结果和性能比较进行了深入分析。实验图像来源于图像融合标准图像库, 实验前均进行配准处理。另外, 由于若融合前图像已属于超分辨率图像, 便失去了进一步超分的实际意义, 因而, 本节实验均选择低分辨率图像进行。实验中 NSCT 参数设置为: $\text{pflt} = \text{'Pyrexc'}$, $\text{dfilt} = \text{'Sinc'}$, 分

解层次递阶取值为 [0, 1, 3, 4]; DWT 方法中参数设置为: 最大分解层次为 4, 低频部分融合规则设置为取最大, 高频部分融合规则设置为选取第二幅图像的信息. 客观评价指标采用连续运行 50 次取平均的方法确定. 使用的评价指标见表 2.

实验部分安排如下: 第 3.1 节介绍采用的图像融合评价指标和用于比较的算法; 第 3.2 节介绍本文方法和表 1 所示方法的融合效果和性能指标比较、分析; 第 3.3 节针对平均加权融合问题进行比较、分析; 第 3.4 节针对字典独立学习和合并学习的性能差异进行比较、分析.

3.1 评价指标和比较方法简称

为了更加客观、有效地对各种方法进行性能评估, 本文选用不需要标准图像作为参考图像的评价指标^[37-40] (例如: $Q_0, Q_E, Q_W, Q^{AB/F}$, 均值) 和需要标准图像作为参考的评价指标 (均方根误差 (Root mean square error, RMSE)) 共同进行性能评价, 上述指标的具体含义如表 2 所示.

3.2 不同类型图像仿真实验

3.2.1 医学图像实验

由于针对医学图像处理的问题广泛存在, 而且具有较好的现实寓意. 因而, 本部分实验选取医学 CT 图像进行, 结果见图 4 和 5 及表 3.

表 1 用于比较的不同方法及其名称缩写

Table 1 The abbreviation of different methods used for comparison

	方法简称	融合方法	超分辨率方法 ^[41]
融合和超分同步完成	SSR-FUSION	SSR-FUSION (本文核心方法)	SSR-FUSION
	SRSR ^[22]	SRSR	SRSR
	NSCT-1	NSCT ^[6]	SRSR (本文实现算法)
	NSCT-2	NSCT ^[6]	Bicubic (本文实现算法)
先融合, 再超分	DWT-1	DWT ^[42]	SRSR (本文实现算法)
	DWT-2	DWT ^[42]	Bicubic (本文实现算法)
	SR-1	SR ^[43]	SRSR
	SR-2	SR ^[43]	Bicubic

表 2 本文采用的评价指标

Table 2 The evaluation indices adopted in this paper

指标名称	表达式
Q_0	$Q_0(A, F) = \frac{2\sigma_{af}}{\sigma_a^2 + \sigma_f^2} \cdot \frac{2\bar{\sigma}\bar{f}}{\bar{\sigma}^2 + \bar{f}^2}$ $Q_0(A, B, F) = \frac{Q_0(A, F) + Q_0(B, F)}{2}$
Q_E	$Q_E(A, B, F) = Q_W(A, B, F) \cdot Q_W(A', B', F')^\alpha$
Q_W	$Q_W(A, B, F) = \sum_{w \in W} c(w)(\lambda(w)Q_0(A, F w)) + (1 - \lambda(w))Q_0(B, F w)$
$Q^{AB/F}$	$Q^{AB/F} = \frac{\sum_{n=1}^N \sum_{m=1}^M (Q^{AF}(n, m)w^A(n, m) + Q^{BF}(n, m)w^B(n, m))}{\sum_{n=1}^N \sum_{m=1}^M (w^A(n, m) + w^B(n, m))}$
MEAN	$\mu = \frac{1}{n} \sum_{i=1}^n x_i$
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (R(i, j) - F(i, j))^2}{M \times N}}$

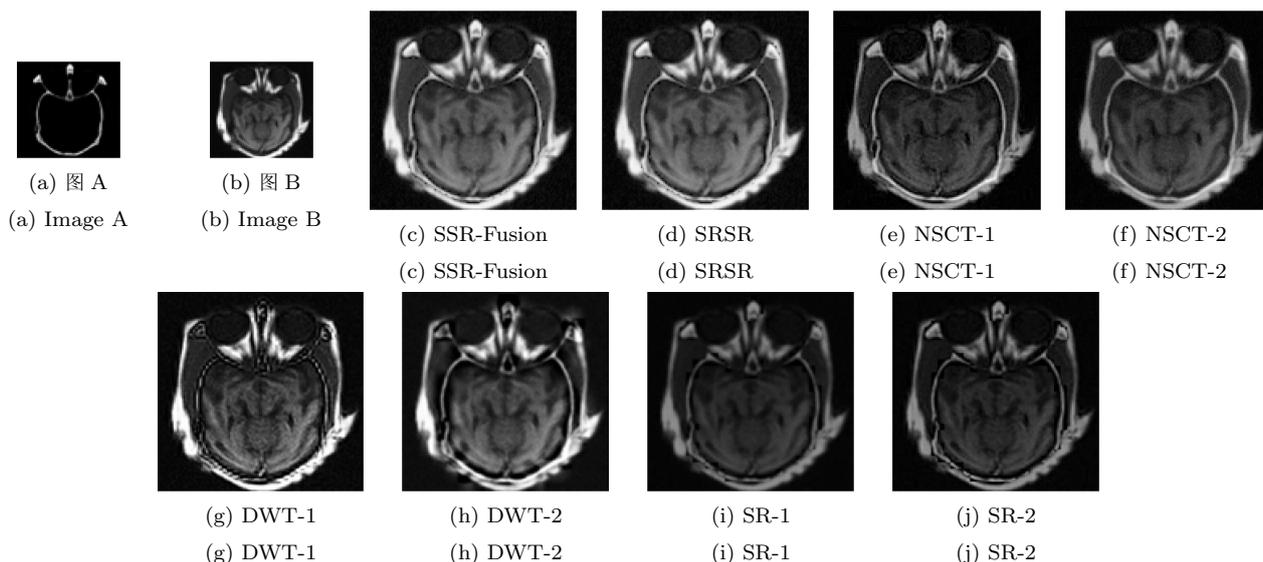


图 4 CT1 图像实验结果

Fig. 4 Experimental results of different methods on CT1 images

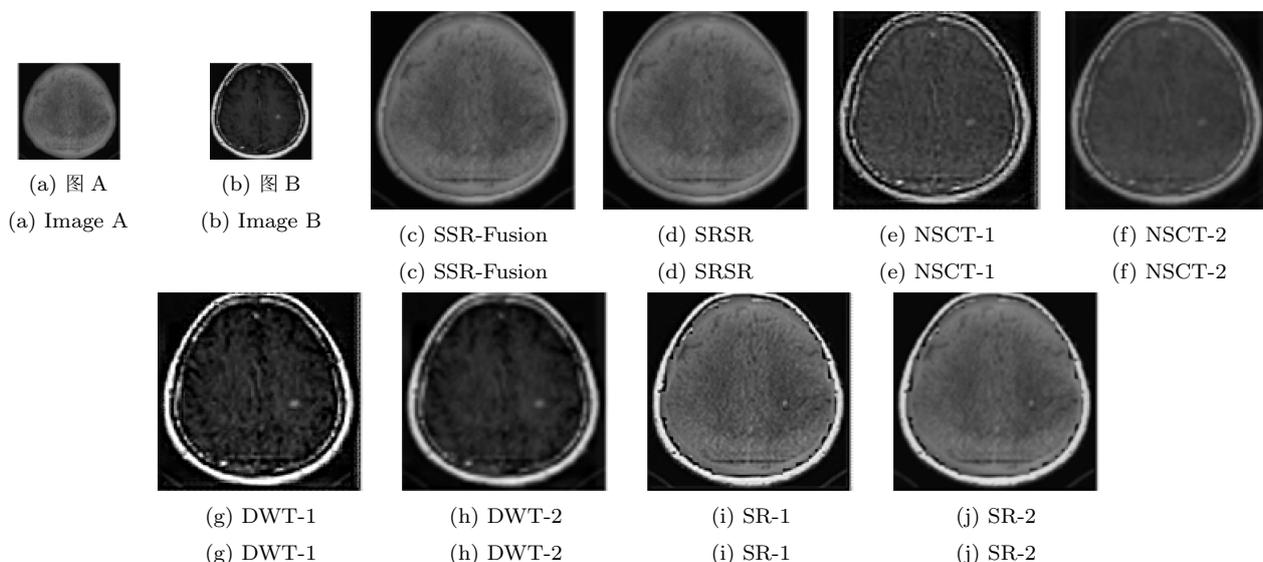


图 5 CT2 图像实验结果

Fig. 5 Experimental results of different methods on CT2 images

3.2.2 多聚焦图像实验

现实中存在对某一场景进行成像时, 由于场景中目标与传感器的距离不同, 无法同时聚焦到这些目标都清晰. 通常处理是分别聚焦到各个目标多次拍摄而得到多幅聚焦点不同的图像. 聚焦点的不同会导致各幅图像中清晰区域和模糊区域不同, 通过图像融合方法可以得到一幅更加清晰的图像. 多聚焦图像是图像融合领域中具有代表性的几大型图像之一, 在该类型图像上的实验效果在一定程度上可以说明方法的有效性. 实验中选取分别为左、右聚焦的两幅图像进行实验, 结果见图 6、图 7 和表 3.

3.2.3 其他类型图像实验

1) 小目标图像. 小目标图像是指图像中主要目标较小且背景宽泛、较模糊的图像, 图像中由于目标小而难于捕捉到清晰、完整的信息, 成为融合领域难题之一. 本节实验中选取两幅小目标图像, 结果见图 8 和表 3. 2) 遥感图像. 凡是记录各种地物电磁波大小的胶片, 都称为遥感影像, 主要指航空像片和卫星像片. 遥感图像是图像融合领域研究较为广泛的一类图像, 在该类型图像上的融合效果在一定程度上能够说明方法的有效性. 因此, 本节实验中选取两幅遥感图像, 结果见图 9 和表 3.

表 3 不同方法实验结果比较
Table 3 Comparison of results obtained by different methods

图像	指标方法	本文方法	SRSR	NSCT-1	NSCT-2	DWT-1	DWT-2	SR-1	SR-2
CT1	Q_0	0.5168	0.5151	0.4948	0.5155	0.3325	0.3083	0.4909	0.514
	Q_E	0.4499	0.4469	0.3039	0.3588	0.1788	0.1308	0.3143	0.3671
	Q_W	0.827	0.85	0.7589	0.7629	0.4659	0.3956	0.8162	0.8404
	$Q^{AB/F}$	0.71	0.6988	0.5373	0.6119	0.2279	0.2174	0.5618	0.6166
	MEAN	31.445	31.4446	17.4447	17.4284	12.3208	12.0833	31.4134	31.1915
	RMSE	22.229	21.3141	63.6353	64.6631	33.5386	33.5386	45.9429	44.5735
CT2	Q_0	0.6635	0.6569	0.6632	0.6813	0.5578	0.6006	0.6323	0.6625
	Q_E	0.2352	0.2353	0.2051	0.1923	0.2831	0.2807	0.2193	0.1696
	Q_W	0.7839	0.7829	0.7475	0.6642	0.8124	0.8295	0.7601	0.7808
	$Q^{AB/F}$	0.4545	0.4731	0.4768	0.3183	0.5218	0.5454	0.4359	0.4452
	MEAN	84.0319	80.4565	57.8145	57.4347	49.7296	47.8192	80.2276	80.1038
	RMSE	24.0871	32.5212	67.0632	62.0935	82.0442	74.2368	37.4871	31.5742
Goldhill	Q_0	0.9878	0.9878	0.9517	0.9824	0.9554	0.982	0.9692	0.9867
	Q_E	0.4719	0.4553	0.4379	0.3929	0.3991	0.4549	0.3897	0.4143
	Q_W	0.873	0.8716	0.8121	0.7984	0.8162	0.8803	0.8423	0.8503
	$Q^{AB/F}$	0.4877	0.4937	0.4331	0.4024	0.4274	0.4981	0.4385	0.4703
	MEAN	113.2378	113.1789	111.6891	112.1849	111.723	112.1872	112.2606	112.7073
	RMSE	14.9878	12.228	12.5313	11.8024	14.997	7.9438	12.1465	9.196
Qinghua	Q_0	0.9951	0.995	0.9772	0.9716	0.9846	0.9957	0.9888	0.9918
	Q_E	0.6595	0.6392	0.5471	0.4514	0.6028	0.669	0.5795	0.4868
	Q_W	0.9513	0.9505	0.8724	0.7731	0.9154	0.9563	0.9339	0.9139
	$Q^{AB/F}$	0.68	0.6797	0.5277	0.3222	0.6331	0.7078	0.624	0.5366
	MEAN	151.8831	148.9453	146.584	147.7187	147.1086	147.4595	148.1698	148.8045
	RMSE	11.9737	11.3116	19.7076	16.4849	17.7055	10.9	15.2923	14.0357
Tank	Q_0	0.9772	0.9768	0.897	0.9447	0.9801	0.9682	0.9324	0.9637
	Q_E	0.4272	0.429	0.4011	0.3191	0.6563	0.3352	0.3478	0.3234
	Q_W	0.9504	0.9497	0.8524	0.8561	0.9597	0.931	0.8895	0.9132
	$Q^{AB/F}$	0.5955	0.5945	0.4534	0.4276	0.7375	0.5281	0.4837	0.5049
	MEAN	128.4998	110.1473	107.6575	108.394	108.2712	108.2511	108.3408	109.1573
	RMSE	14.7141	15.5021	13.067	17.9363	5.0978	12.5656	10.3325	15.4805
Remote	Q_0	0.9023	0.899	0.8592	0.9022	0.8367	0.8704	0.8643	0.8949
	Q_E	0.3346	0.3451	0.352	0.2956	0.3072	0.3124	0.298	0.2614
	Q_W	0.7898	0.7888	0.7606	0.7225	0.7064	0.7645	0.7455	0.7492
	$Q^{AB/F}$	0.4574	0.4811	0.4198	0.3164	0.3964	0.4861	0.4346	0.3843
	MEAN	81.3129	81.1785	74.731	75.0652	75.0513	75.3561	79.7966	80.1594
	RMSE	17.564	23.541	18.8129	24.3336	24.2395	23.1833	23.0761	25.0632



图 6 Goldhill 图像实验结果

Fig. 6 Experimental results of different methods on Goldhill images

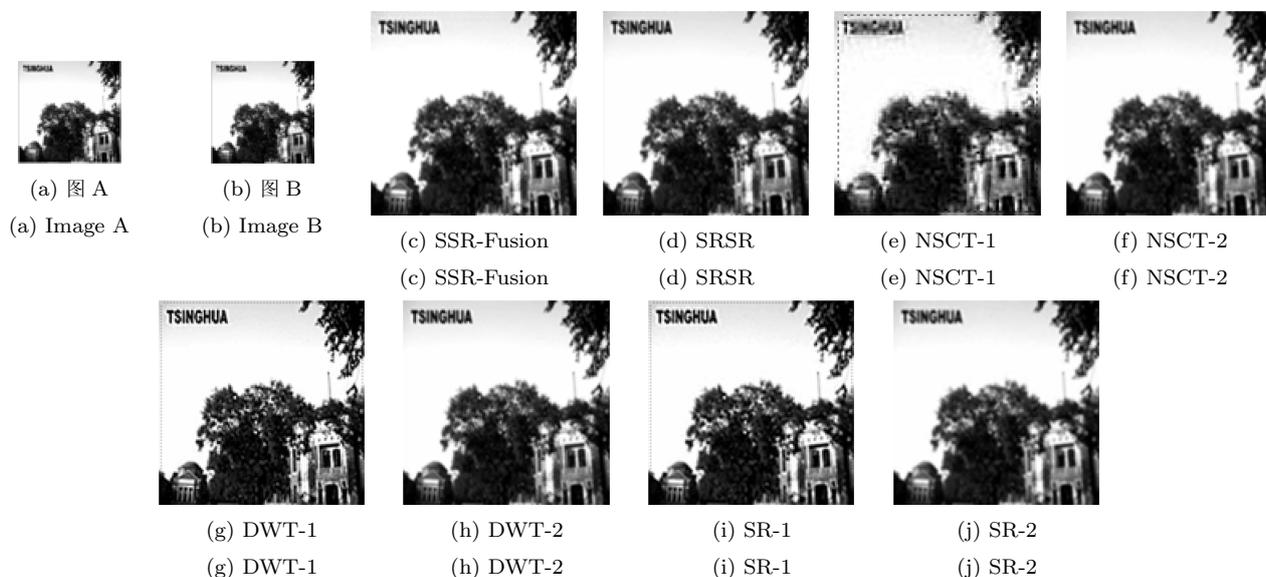
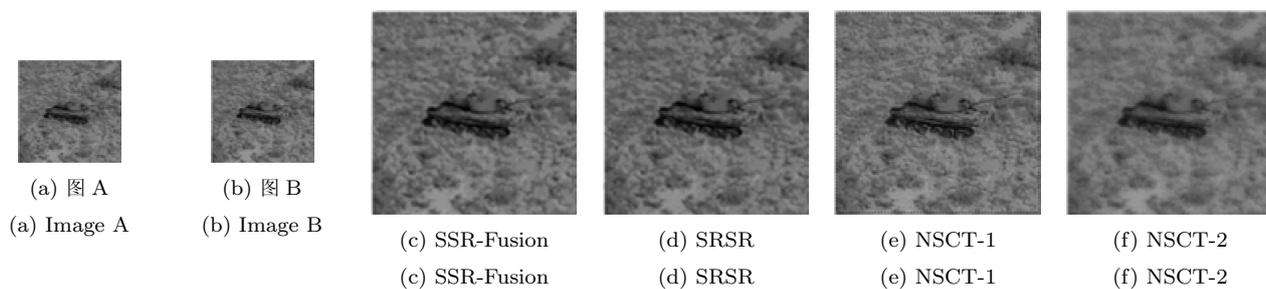


图 7 Qinghua 图像实验结果

Fig. 7 Experimental results of different methods on Qinghua images



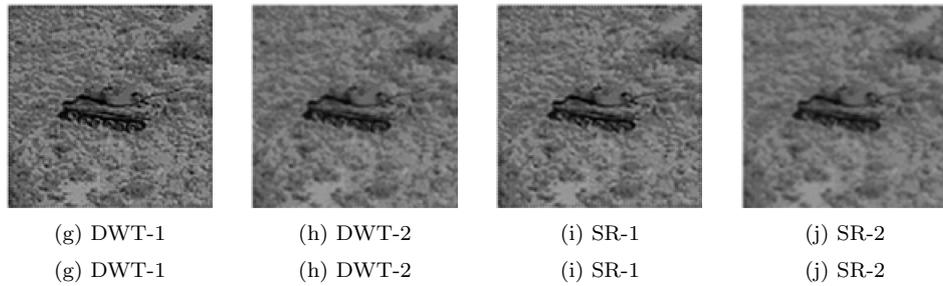


图 8 Tank 图像实验结果

Fig. 8 Experimental results of different methods on Tank images

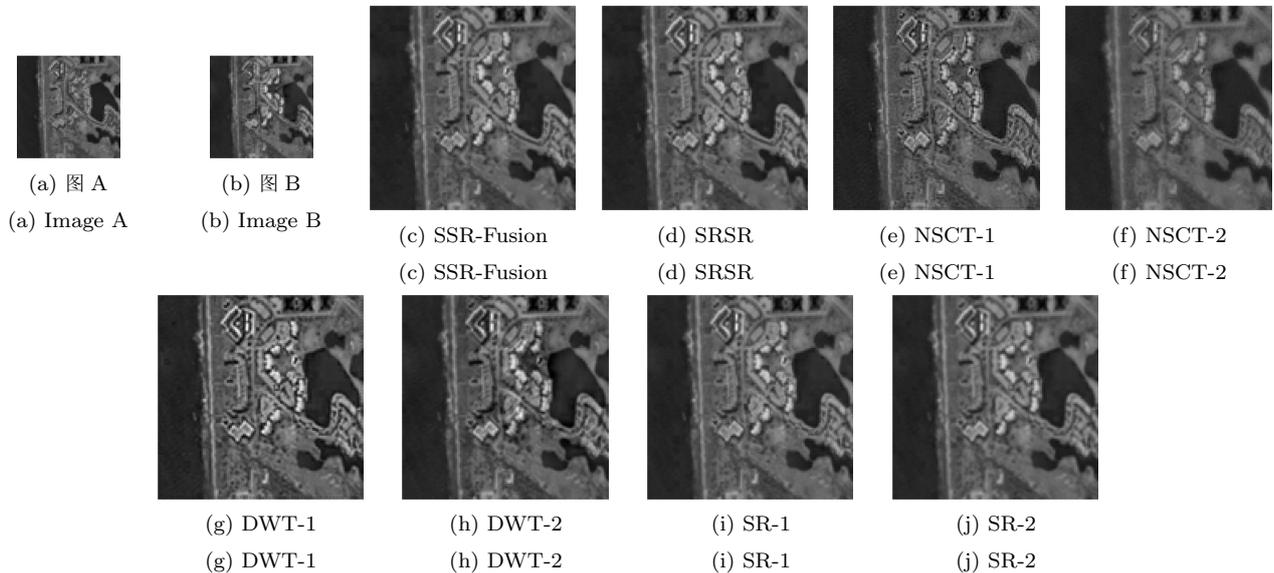


图 9 Remote 图像实验结果

Fig. 9 Experimental results of different methods on Remote images

根据上述实验结果, 进行如下分析.

1) 从视觉效果和客观评价数据进行分析. 首先, 从视觉效果来看, 与其他三类 SR 方法相比较, 本文方法融合得到的图像边缘清晰、清晰度较高、分辨率较高, 而且在一定程度上, 能够较好地突出图像中的细节信息和重要特征. 其次, 从客观评价数据来看: 与 SR 同类方法相比, 第 1 组实验结果中, 本文方法融合得到的结果在 Q_0 、 Q_E 、 $Q^{AB/F}$ 、均值方面均表现出一定的优势, 但在 Q_W 、RMSE 方面略差于 SRSR 方法的结果; 第 2 组实验结果中, 在 Q_0 、 Q_W 、均值、RMSE 上有所提高, Q_E 、 $Q^{AB/F}$ 方面相对 SRSR 还有提高的空间; 第 3 组实验结果中, 本文方法融合得到的结果在 Q_E 、 Q_W 、均值方面均表现出一定的优势, 在 $Q^{AB/F}$ 方面 SRSR 更佳, RMSE 方面 SR-2 更优; 第 4 组实验结果中, 在 Q_0 、 Q_E 、 Q_W 、 $Q^{AB/F}$ 、均值等方面指标较优, RMSE 方面 SRSR 更优; 第 5 组实验结果中本文方法融合得到的结果在 Q_0 、 Q_W 、 $Q^{AB/F}$ 、均值方面均表现出一定的优势, 但在 Q_E 方面略差于 SRSR 方法的结

果, RMSE 方面劣于 SR-1; 第 6 组实验结果中本文方法融合得到的结果在 Q_0 、 Q_W 、均值、RMSE 方面均表现出一定的优势, 但在 Q_E 、 $Q^{AB/F}$ 方面略差于 SRSR 方法的结果.

综上所述, 由于稀疏表示与小波变换存在一定联系, 为了全面、客观地衡量方法的优劣, 文中针对经典小波变换方法也进行了对比实验. 实验结果可以发现在某些实验图像上 NSCT、DWT 融合再进行超分的结果较为突出, 这主要是因为这两种融合方法均是基于像素级进行融合的, 相对于稀疏表示基于字典块的融合方法效果更好. 如何通过字典块间平均化、边缘处理等技术来提高融合性能, 是我们将进一步深入探讨的问题. 但是, 本文方法主要侧重在基于文献 [22] 提出方法的基础上适度增加香农熵自适应加权功能, 使得 LL、LH、H 三部分能有效智能加权, 和三类经典稀疏表示方法 (同步超分方法 SRSR、分别超分方法 SR-1 和 SR-2) 相比在 Q_0 、均值等方面均存在一定程度的提高, 说明本文方法是有效的, 即根据不同图像的实际要求, 本文方法采用

香农熵多视角加权的思想在融合图像的 LL、LH、H 部分进行智能加权, 可以有针对性地突出重点信息, 较好地提高同步超分辨率融合效果. 另外, 如何通过评价准则构造更为合理、有效的香农熵加权目标函数以提高融合性能的问题仍然是个开放性的问题, 我们也将继续深入探讨. 此外, 为了进一步研究具体加权情况对融合性能的影响, 本节给出不同图像三部分频段的加权比例差异部分的举例, 为了方便研究, 均进行取整处理, 结果如表 4 所示.

表 4 不同图像权重比

Table 4 Weights of different images

图像	加权比例
CT1	10.5 : 10.5 : 9.9
CT2	10.5 : 10.5 : 10.3
Goldhill	10.5 : 10.6 : 11.7
Qinghua	10.7 : 10.3 : 9.9
Tank	10.4 : 10.5 : 10.5
Remote	10.3 : 10.4 : 10.3

2) 加权比例分析. 本部分针对表 4 所示的举例结果进行详细分析, 可以得出如下结论: a) 不同图像 LL、LH、H 三频段加权比例均不同, 说明本文方法在根据图像具体频段信息要求进行自适应加权方面具有一定的有效性; b) 由于本文进行了某种构造香农熵加权目标函数的尝试, 是否有更好的构造方法还有待进一步深入探讨; c) 针对相应实验结果进行比较、分析, 发现本文加权方法在某些图像上改善性能明显, 例如 CT1 图实验; 相反在某些图像上并不明显, 如 Tank 图实验, 其加权比例接近于等比例, 这与图像的具体特征和目标函数构造均存在一定关联.

3) 分块性能与融合效果关系分析. 本节根据字典学习分块大小不同是否会对融合效果带来一定的

影响的问题进行了详细实验讨论, 实验中依然采用第 3.2 节所采用的医学、多聚焦、小目标、遥感 4 种类型图像, 并做了相应分析, 结果见图 10. 根据上述结果发现, 就本文实验中采用的医学、多聚焦、小目标、遥感 4 组共 8 张图像来看, 随着分块大小从小到大的变化, $Q_0, Q_E, Q_W, Q^{AB/F}$ 这 4 个融合指标均发生变化, 当分块数目在 [2-8] 范围内时, 融合图像在这 4 个指标上能够取得整体相对较高的水平. 但伴随着分块数目继续增加, 出现下降变化. 本节在第 3.4 节中多组图像上也进行了相似实验, 也可得到相似结论. 因而, 就上述图像而言, 分块数目设置为 2~8 较为合理. 这为稀疏表示字典学习分块大小的进一步研究提供了试探性的结论.

3.3 与 LL、LH、H 等比例加权融合效果对比分析

本节针对 LL、LH、H 三部分采用香农熵多视角加权还是等比例加权对融合结果带来影响的问题进行了实验分析. 实验图像依然选择第 3.2 节所采用的经典医学、多聚焦、小目标、遥感 4 种类型. 实验中在每幅图像的 LL、LH、H 频段部分分别采用香农熵多视角加权、等比例加权融合方法得到相应融合结果, 实验结果如图 11 (d) 和 (e). 从上述实验结果来看, 不仅发现针对不同图像, 均采用 LL、LH、H 频段等比例加权进行融合, 得到的融合图像在清晰度、完整性以及细节信息等方面存在较大差异, 而且针对同一幅图像采用本文方法融合结果 (图 11 (d) 所示) 与等比例方法得到融合结果 (图 11 (e) 所示) 也存在较大差异. 经研究发现主要原因在于: 不同的图像经滤波分解得到的 LL、LH 部分频段, 其中低频部分主要反映图像的轮廓信息, 高频部分主要反映图像的细节信息. 不同的图像侧重点不同, 例如第 3.2 节所采用的第 4 组多聚焦图像 Goldhill 由于背景中很多信息非常接近, 只有通过有效增加高频部

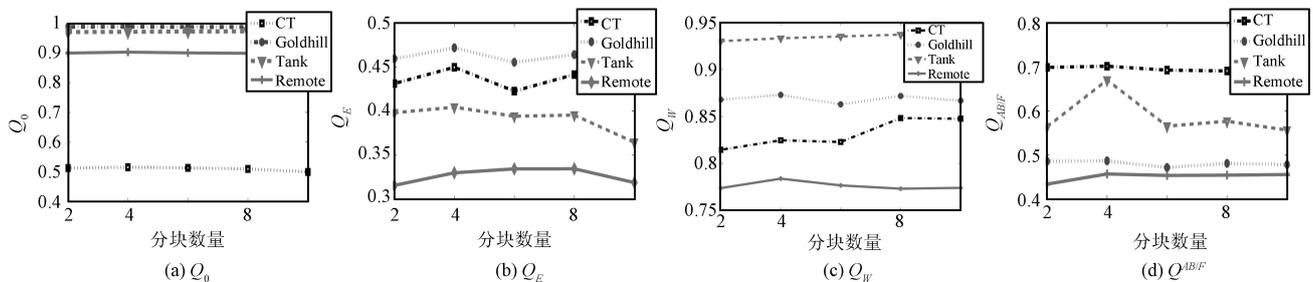


图 10 分块大小与融合效果关系的分析

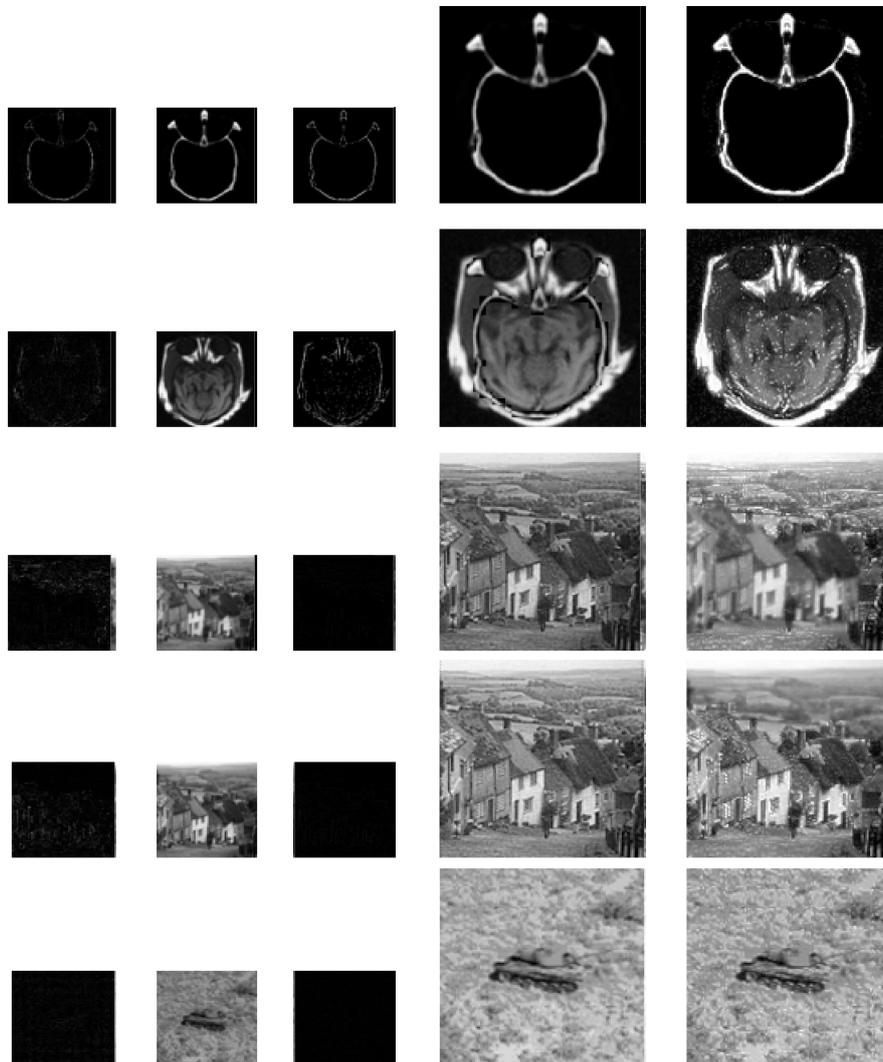
Fig. 10 Analysis of fusion results versus different sizes of image patches

分的信息, 才能将其有效区分, 以提高融合效果. 又例如医学图像多属黑白灰度图像, 高频信息相对非常有限, 单纯通过增加高频信息权重很难达到提高融合效果的目的, 往往需要按照自适应加权比例适度提高 LH、H 比例, 才能提高融合性能. 针对这一点本文进行了上述实验分析, 实验结果表明本文方法在最终的融合阶段, 可以根据实际需要进行有效加权, 得到的融合效果更佳. 这为图像融合领域开拓了一条试探性的思路, 也为不断提高本文方法的实用价值带来了可能.

3.4 LH、H 字典分离学习、合并学习对比分析

本节主要针对在其他情况不变的前提下, LH、H 字典是否在一起学习, 给融合效果带来的影响的问题进行详细比较、分析, 参数设置参见第 2.5 节, 实验结果如图 12 和表 5.

实验中依然采用本文所使用的多聚焦和遥感两组图像, 在其他几组图像和经典的 People、Clock、Barbara、Remote' 四组图像上也进行了相似实验, 并得到类似的结论. 实验数据显示: 在本节所采用的多聚焦和遥感两组图像上, 字典分开学习较合并学习在 Q_0, Q_E, Q_W 三个指标上均显示出一定的提高, 这在其他两组图像上也得到了相似的结论. 针对经典 4 组图像, 除 Barbara 图像四个指标均有所提高外, 其他三组图像字典分开学习方式均在 $Q_E, Q_W, Q^{AB/F}$ 三个指标上显示出一定的优势. 具体可总结为: 字典分开学习可以根据不同图像的 LL、LH、H 三频段部分, 分别设置不同的初始字典集合、目标函数, 独立进行智能学习, 针对性更强, 对提高融合效果可以起到作用. 这一点也反映出字典学习方式对稀疏表示图像融合方法融合效果影响的重要性.



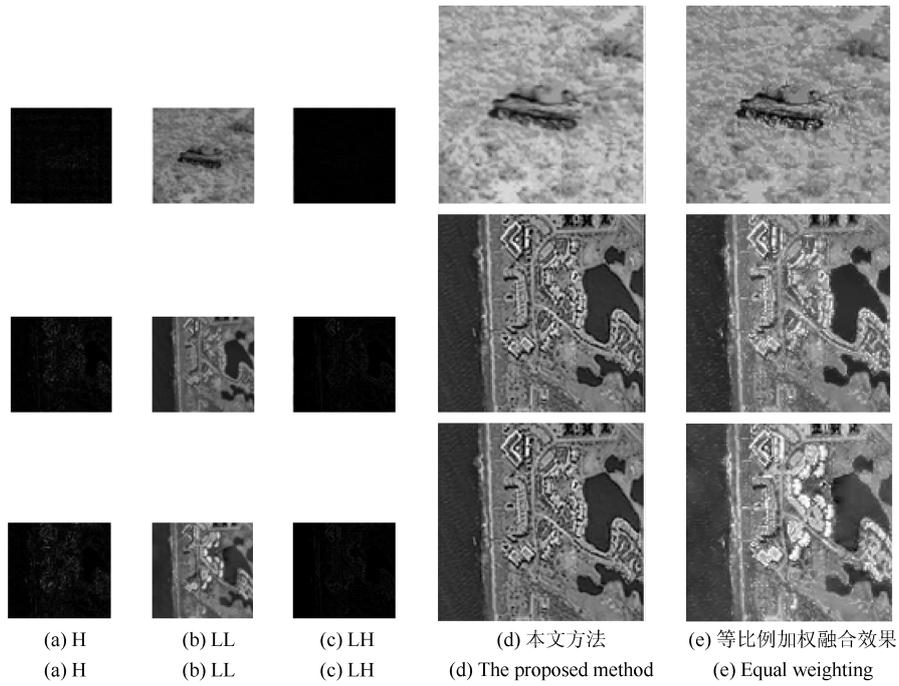


图 11 LL、LH、H 不同多视角加权方法 (香农熵加权、等比例加权) 融合实验比较

Fig. 11 Performance comparison with different weighting ways (Shannon entropy multi-view weighting and equal weighting) on LL, LH, H



(a) 分离字典学习效果

(a) Separating dictionary learning way



(b) 合并字典学习效果

(b) Merging dictionary learning way

图 12 不同字典学习方式 (合并与分离) 性能比较

Fig. 12 Performance comparison with different dictionary learning ways (merging and separating)

表 5 不同字典学习方式 (合并与分离) 性能比较

Table 5 Performance comparison with different dictionary learning ways (merging and separating)

指标 方法	Goldhill		Remote		People		Clock		Barbara		Remote'	
	分离	合并	分离	合并	分离	合并	分离	合并	分离	合并	分离	合并
Q_0	0.9889	0.9877	0.8986	0.8981	0.5418	0.5422	0.9719	0.9722	0.9777	0.9776	0.8800	0.8853
Q_E	0.4631	0.4558	0.3328	0.3279	0.1053	0.0991	0.2705	0.2687	0.4257	0.4257	0.2341	0.2253
Q_W	0.8730	0.8729	0.7758	0.7750	0.4798	0.4773	0.6865	0.6864	0.8232	0.8220	0.6548	0.6422
$Q^{AB/F}$	0.4701	0.4860	0.4537	0.5440	0.3795	0.3697	0.3606	0.3588	0.5412	0.5344	0.3424	0.3329

4 结论

针对传统稀疏表示同步超分图像融合模型中对 LL、LH、H 三部分频段等比例加权, 不能突出重点信息以及 LH、H 使用相同的字典进行学习, 缺乏准确性之问题, 本文提出一种香农熵加权稀疏表示同步超分图像融合方法. 首先, 对待融合图像进行预处理, 分解成为 LL、LH 频段; 其次, 进行独立字典学习, 并生成字典集 D_{LL} 、 D_{LH} 、 D_H ; 再次, 根据构造得到的香农熵加权方法计算 LL、LH、H 三部分频段的权重. 最后, 根据本文提出的算法完成融合过程. 实验结果表明, 本文方法具有以下特点: 1) 通过香农熵加权对融合规则中 LL、LH、H 三部分频段的权重信息进行自适应学习, 有针对性地突出重点; 2) 将 LH 和 H 分离进行学习, 使得到的字典集 D_{LH} 、 D_H 更适合于图像融合要求.

References

- Chipman L J, Orr T M, Graham L N. Wavelets and image fusion. In: Proceedings of the 4th International Conference on Image Processing. Washington D. C., USA: IEEE, 1995. 248–251
- Li H, Manjunath B S, Mitra S K. Multi-sensor image fusion using the wavelet transform. In: Proceedings of the 1994 IEEE International Conference Image Processing. Austin, USA: IEEE, 1994, 1: 51–55
- Pajares G, Manuel J. A wavelet-based image fusion tutorial. *Parttern Recognition*, 2004, 37(9): 1855–1872
- Li Z H, Jing Z L, Yang X H, Sun S Y. Color transfer based remote sensing image fusion using non-separable wavelet frame transform. *Parttern Recognition Letters*, 2005, 26(13): 2006–2014
- Li S T, Yang B. Multifocus image fusion by combining curvelet and wavelet transform. *Parttern Recognition Letters*, 2008, 29(9): 1295–1301
- Zhang Q, Guo B L. Multifocus image fusion using the non-subsampled contourlet transform. *Signal Processing*, 2009, 89(7): 1334–1346
- Yang S Y, Wang M, Jiao L C, Wu R X, Wang Z X. Image fusion based on a new contourlet packet. *Information Fusion*, 2010, 11(2): 78–84
- Li S T, Yang B, Hu J W. Performance comparison of different multi-resolution transforms for image fusion. *Information Fusion*, 2011, 12(2): 74–84
- Mukane S M, Ghodake Y S, Khandagle P S. Image enhancement using fusion by wavelet transform and Laplacian pyramid. *Computer Science*, 2013, 10(4): 122–126
- Jiang Y, Wang M H. Image fusion with morphological component analysis. *Information Fusion*, 2014, 18: 107–118
- Balakrishnan S, Cacciola M, Udpa L, Rao B P, Jayakumar T, Raj B. Development of image fusion methodology using discrete wavelet transform for eddy current images. *NDT and E International*, 2012, 51: 51–57
- Petrovic V S, Xydeas C S. Gradient-based multiresolution image fusion. *IEEE Transactions on Image Processing*, 2004, 13(2): 228–237
- Borwonwatanadelok P, Rattanapitak W, Udomhunsakul S. Multi-focus image fusion based on stationary wavelet transform and extended spatial frequency measurement. In: Proceedings of the 2009 International Conference on Electronic Computer Technology. Macau, China: IEEE, 2009. 77–81
- Lewis J J, O'callaghan R J, Nikolov S G, Bull D R, Canagarajah N. Pixel-and-region-based image fusion with complex wavelets. *Information Fusion*, 2007, 8(2): 119–130
- Miao Q G, Lou J J, Xu P F. Image fusion based on NSCT and bandelet transform. In: Proceedings of the 2012 Computational Intelligence and Security. Guangzhou, China: IEEE, 2012. 314–317
- Li H F, Chai Y, Li Z F. Multi-focus image fusion based on nonsubsampling contourlet transform and focused regions detection. *Optik-International Journal for Light and Electron Optics*, 2013, 124(1): 40–51
- Candes E S. Compressive Sampling. In: Proceedings of the 2006 International Congress on Mathematicians. Madrid, Spain: European Mathematical Society Publishing House, 2006, 3: 1433–1452
- Liu Fang, Wu Jiao, Yang Shu-Yuan, Jiao Li-Cheng. Research advances on structured compressive sensing. *Acta Automatica Sinica*, 2013, 39(12): 1980–1995 (刘芳, 武娇, 杨淑媛, 焦李成. 结构化压缩感知研究进展. 自动化学报, 2013, 39(12): 1980–1995)
- Ma Xiao-Hu, Tan Yan-Qi. Face recognition based on discriminant sparsity preserving embedding. *Acta Automatica Sinica*, 2014, 40(1): 73–82 (马小虎, 谭延琪. 基于鉴别稀疏保持嵌入的人脸识别算法. 自动化学报, 2014, 40(1): 73–82)
- Zuo Y Y, Zhang B. Robust hierarchical framework for image classification via sparse representation. *Tsinghua Science and Technology*, 2011, 16(1): 13–21
- Zhao M, Li S T, Kwok J. Text detection in images using sparse representation with discriminative dictionaries. *Image and Vision Computing*, 2010, 28(12): 1590–1599
- Yin H T, Li S T, Fang L Y. Simultaneous image fusion and super-resolution using sparse representation. *Information Fusion*, 2013, 14(3): 229–240
- Yang B, Li S T. Pixel-level image fusion with simultaneous orthogonal matching pursuit. *Information Fusion*, 2012, 13(1): 10–19
- Zhao Y Q, Yang J X, Zhang Q Y, Song L, Cheng Y M, Pan Q. Hyperspectral imagery super-resolution by sparse representation and spectral regularization. *Journal on Advances in Signal Processing*, 2011, (1): 87
- Zhao Y, Yang J, Chan J C W. Hyperspectral imagery super-resolution by spatial-spectral joint nonlocal similarity. *Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, PP(99): 1939–1404
- Lian Qiu-Sheng, Zhang Jun-Qin, Chen Shu-Zhen. Single image super-resolution algorithm based on two-stage and multi-frequency-band dictionaries. *Acta Automatica Sinica*, 2013, 39(8): 1310–1320 (练秋生, 张钧芹, 陈书贞. 基于两级字典与分频带字典的图像超分辨率算法. 自动化学报, 2013, 39(8): 1310–1320)
- Moussallam M, Daudet L, Richard G. Matching pursuits with random sequential subdictionaries. *Signal Processing*, 2012, 92(10): 2532–2544
- Hsieh S H, Lu C S, Pei S C. Fast OMP: reformulating OMP via iteratively refining l_2 -norm solutions. In: Proceedings of the 2012 IEEE Statistical Signal Processing Workshop. Ann Arbor, MI: IEEE, 2012. 189–192

- 29 Peng Yi-Gang, Suo Jin-Li, Dai Qiong-Hai, Xu Wen-Li. From compressed sensing to low-rank matrix recovery: theory and applications. *Acta Automatica Sinica*, 2013, **39**(7): 981–994
(彭义刚, 索津莉, 戴琼海, 徐文立. 从压缩传感到低秩矩阵恢复: 理论与应用. *自动化学报*, 2013, **39**(7): 981–994)
- 30 Chen L X, Dobra A. Histograms as statistical estimators for aggregate queries. *Information Systems*, 2013, **38**(2): 213–230
- 31 Deng Z H, Choi K S, Chung F L, Wang S T. EEW-SC: Enhanced entropy-weighting subspace clustering for high dimensional gene expression data clustering analysis. *Applied Soft Computing*, 2011, **11**(8): 4798–4806
- 32 Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, **54**(11): 4311–4322
- 33 Bryt O, Elad M. Compression of facial images using the K-SVD algorithm. *Visual Communication and Image Representation*, 2008, **19**(4): 270–282
- 34 Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A. Discriminative learned dictionaries for local image analysis. In: *Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, USA: IEEE, 2008. 1–8
- 35 Engan K, Skretting K, Husoy J H. Family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation. *Digital Signal Processing*, 2007, **17**(1): 32–49
- 36 Divekar A, Ersoy O. Image fusion by compressive sensing. In: *Proceedings of the 2009 International Congress on Geoinformatics*. Fairfax, VA: IEEE, 2009. 1–6
- 37 Wang Z, Bovik A C. A universal image quality index. *IEEE Signal Processing Letters*, 2002, **9**(3): 81–84
- 38 Piella G, Heijmans H. A new quality metric for image fusion. In: *Proceedings of the 2003 IEEE International Conference on Image Processing*. Barcelona, Spain: IEEE, 2003, **2**: 173–176
- 39 Xydeas C S, Petrovic V. Objective image fusion performance measure. *Electronics Letters*, 2000, **36**(4): 308–309
- 40 Petrovic V. Subjective tests for image fusion evaluation and objective metric validation. *Information Fusion*, 2007, **8**(2): 208–216
- 41 Yang J C, Wright J, Huang T S, Ma Y. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 2010, **19**(11): 2861–2873
- 42 Deng A, Wu J, Yang S. An image fusion algorithm based on discrete wavelet transform and canny operator. In: *Proceedings of the 2011 International Congress on Computer Education, Simulation and Modeling*. Berlin, Heidelberg: Springer, 2011. 32–38
- 43 Yang B, Li S T. Multifocus image fusion and restoration with sparse representation. *IEEE Transactions on Instrumentation and Measurement*, 2010, **59**(4): 884–892



李奕 江南大学物联网工程学院博士研究生。主要研究方向为模式识别, 智能计算及应用。

E-mail: lyqgx@126.com

(**LI Yi** Ph.D. candidate at the School of Internet of Things Engineering, Jiangnan University. Her research interest covers pattern recognition, intelligent computation and their application.)



吴小俊 江南大学物联网工程学院教授。主要研究方向为人工智能, 模式识别和计算机视觉。本文通信作者。

E-mail: wu.xiaojun@aliyun.com

(**WU Xiao-Jun** Professor at the School of Internet of Things Engineering, Jiangnan University. His research interest covers artificial intelligence, pattern recognition, and computer vision. Corresponding author of this paper.)