

# 基于目标提取与引导滤波增强的 红外与可见光图像融合

吴一全<sup>1,2,3,4,5,6</sup>, 王志来<sup>1</sup>

<sup>1</sup>南京航空航天大学电子信息工程学院, 江苏 南京 211106;

<sup>2</sup>南京信息工程大学江苏省大数据分析技术重点实验室, 江苏 南京 210044;

<sup>3</sup>浙江工业大学浙江省信号处理重点实验室, 浙江 杭州 310023;

<sup>4</sup>广西师范大学广西多源信息挖掘与安全重点实验室, 广西 桂林 541004;

<sup>5</sup>成都理工大学国土资源部地学空间信息技术重点实验室, 四川 成都 610059;

<sup>6</sup>中国地质科学院矿产资源研究所国土资源部成矿作用与资源评价重点实验室, 北京 100037

**摘要** 为了使融合结果突出目标并发掘更多细节, 提出了一种基于目标提取与引导滤波增强的红外与可见光图像融合方法。首先对红外图像依据二维 Tsallis 熵和基于图的视觉显著性模型提取目标区域。然后对可见光与红外图像分别进行非下采样 Shearlet 变换(NSST), 并对所得低频分量进行引导滤波增强。由增强后的红外图像和可见光图像低频分量基于目标提取的融合规则得到融合图像的低频分量, 高频分量则根据方向子带信息和取大来确定。最后经 NSST 逆变换得到融合图像。大量实验结果表明, 本文方法在增强融合图像空间细节的同时, 有效突出了目标, 并且在信息熵、平均梯度等指标上优于基于拉普拉斯金字塔变换、基于小波变换、基于平稳小波变换、基于非下采样 Contourlet 变换(NSCT)、基于目标提取与 NSCT 变换等。

**关键词** 图像处理; 图像融合; 红外图像; 可见光图像; 非下采样 Shearlet 变换; 目标提取

**中图分类号** TP751 **文献标识码** A

**doi:** 10.3788/AOS201737.0810001

## Infrared and Visible Image Fusion Based on Target Extraction and Guided Filtering Enhancement

Wu Yiquan<sup>1,2,3,4,5,6</sup>, Wang Zhilai<sup>1</sup>

<sup>1</sup>College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 211106, China;

<sup>2</sup>Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, China;

<sup>3</sup>Zhejiang Province Key Laboratory for Signal Processing, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China;

<sup>4</sup>Guangxi Key Lab of Multi-Source Information Mining and Security, Institute of Mineral Resources, Guangxi Normal University, Guilin, Guangxi 541004, China;

<sup>5</sup>Key Laboratory of Geo-Spatial Information Technology, Ministry of Land and Resources, Chengdu University of Technology, Chengdu, Sichuan 610059, China;

<sup>6</sup>Key Laboratory of Metallogeny and Mineral Assessment, Institute of Mineral Resources, Ministry of Land and Resources, Chinese Academy of Geological Sciences, Beijing 100037, China

**Abstract** In order to highlight the fusion result and dig out more details, a fusion method of infrared image and

**收稿日期:** 2017-01-20; **收到修改稿日期:** 2017-03-25

**基金项目:** 国家自然科学基金(61573183)、南京信息工程大学江苏省大数据分析技术重点实验室开放基金资助(KXX1403)、浙江省信号处理重点实验室开放基金(ZJKL\_6\_SP-OP2014-02)、广西多源信息发掘与安全重点实验室开放基金(MIMS16-01)、成都理工大学国土资源部地学空间信息技术重点实验室开放基金(KLGSIT2015-05)、国土资源部成矿作用与资源评价重点实验室开放基金(ZS1406)、江苏高校优势学科建设工程

**作者简介:** 吴一全(1963—), 男, 博士, 教授, 博士生导师, 主要从事遥感图像处理与理解、红外目标检测与识别、视觉检测与图像测量、数字全息等方面的研究。E-mail: nuaimage@163.com

visible image based on the target extraction and guidance filtering enhancement is proposed. Firstly, the two-dimensional Tsallis entropy and graph-based visual saliency model are used to extract the target region of infrared image. Then the visible image and the infrared image are decomposed by non-subsampled shearlet transform (NSST), respectively. The low-frequency components of the visible image and the infrared image are enhanced with guided filtering, respectively. The low-frequency component of the fused image is obtained from the enhanced low-frequency component of the infrared image and the visible image based on the fusion rule of target extraction, and the high-frequency components of the fused image are gained according to the maximization criterion of the directional sub-band information sum. Finally, the fused image is obtained by inverse NSST transform. A large number of experimental results demonstrate that the proposed method can improve the spatial resolution of the fused image, effectively highlight the target, and is superior to the method based on the Laplacian pyramid transform, the method based on wavelet transform, the method based on stationary wavelet transform, the method based on non-subsampled contourlet transform (NSCT), the method based on target extraction and NSCT in the quantitative evaluation indexes such as information entropy and average gradient.

**Key words** image processing; image fusion; infrared image; visible image; non-subsampled shearlet transform; target extraction; guided filtering enhancement

**OCIS codes** 100.2960; 110.3080; 300.6550; 350.2660; 110.7410; 330.6110

## 1 引 言

图像融合是将取自同一区域的不同光谱与空间细节信息的两幅或多幅图像通过一定的算法融合到单幅图像中<sup>[1]</sup>。红外图像依据物体的热辐射特性成像,虽然它可以突出目标信息,但其主要呈现的是场景的概貌。可见光图像根据物体的光谱反射特性成像,可准确地展现边缘、纹理细节,并能捕获到图像的背景信息,但其易受照明、气候等外部因素影响。红外与可见光图像融合可综合地利用红外图像与可见光图像背景细节清晰的互补特性,在多源信息挖掘、大数据分析、地质资源勘探等领域有着重要应用<sup>[2]</sup>。

图像融合方法主要可划分为基于空间域和基于变换域两类方法。基于变换域的图像融合方法当前受到的关注较多,这类方法大多是基于拉普拉斯金字塔(LP)变换、主成分分析(PCA)变换、小波(WT)变换、Contourlet变换等多尺度分解的融合方法<sup>[3-5]</sup>。其中,基于小波变换的方法是较为典型的多尺度分解融合方法。而小波变换受方向数的限制无法更稀疏地表示图像的边缘、纹理细节信息。为此,人们相继提出了Contourlet变换、非下采样Contourlet变换(NSCT)<sup>[6]</sup>、Shearlet变换<sup>[7]</sup>等更稀疏的信号表示方法。Shearlet变换作为一种高维信号稀疏表示方法,能更有效地捕获图像的细节信息。非下采样Shearlet变换(NSST)<sup>[8]</sup>由Shearlet变换改进而来,且具备移不变性。吴一全等<sup>[9-10]</sup>证实了其融合结果比其他多尺度分解融合方法包含更多的细节信息。图像融合中,NSST变换虽可较好地实现源图像的多尺度分解与重建,但如何设计融合规则对融合效果也甚为关键。郭明等<sup>[11]</sup>先分割红外图像以提取目标区域,再采用替代法融合低频分量,所得融合图像目标较为突出,但目标区域不够准确,含有一些非目标区域和噪声点。高颖等<sup>[12]</sup>结合区域生长法与Canny算子,提取的目标较为完整、准确,但未充分发掘可见光图像隐藏的背景细节。邢素霞等<sup>[13]</sup>依据Renyi熵分割出目标区域,并在NSCT域对可见光图像进行自适应增强,但其存在对图像细节过增强和噪声放大的问题,导致融合图像某些区域过亮,掩盖了一些细节。Kong等<sup>[14]</sup>提出了基于NSCT变换的图像融合方法,避免了伪Gibbs现象,在捕捉背景信息方面表现较好,但目标不够突出,融合图像的对比较低。

本文针对以往融合方法在光线昏暗、背景和目標灰度相近时,融合图像易出现背景细节信息没有充分发掘和目標不够准确、突出的问题,提出了一种基于目標提取与引导滤波增强的红外与可见光图像融合方法。首先依据二维Tsallis熵和基于图的视觉显著性(GBVS)模型提取红外图像中的目標区域;然后分别对红外和可见光图像进行NSST变换,并对两幅源图像的低频分量进行引导滤波增强,融合图像的低频分量目標区域系数选自增强后的红外图像目標区域的低频分量,而非目標区域系数选自增强后的可见光图像对应区域的低频分量;融合图像的高频分量采用方向子带信息和取大的方法选取融合系数。最后,将本文方法与基于LP变换、基于WT、基于平稳小波变换(SWT)、基于NSCT变换<sup>[15]</sup>、基于目標提取与NSCT变换<sup>[13]</sup>等5种融合方法的实验结果进行了比较,验证了本文方法的优越性。

## 2 非下采样 Shearlet 变换

Shearlet 变换是一种新的多尺度分解方法,可通过对基本函数的一系列仿射变换生成基函数。当维数等于 2 时,根据能量有限函数生成的合成膨胀仿射系统形式为

$$\mathbf{M}_{AB}(\psi) = \{\psi_{j,l,k}(x) = |\det \mathbf{A}|^{j/2} \psi(\mathbf{B}^l \mathbf{A}^j x - k) \quad j, l \in \mathbb{Z}, k \in \mathbb{Z}^2\}, \quad (1)$$

式中  $\psi \in L^2(\mathbb{R})$ ;  $\mathbf{A}$  和  $\mathbf{B}$  为  $2 \times 2$  的可逆矩阵。 $j$  为分解尺度,  $l$  表示方向数,  $k$  为平移。若  $\forall f \in L^2(\mathbb{R}^2)$ , 满足  $\sum_{j,l,k} |\langle f, \psi_{j,l,k} \rangle| = \|f\|^2$ , 则  $\mathbf{M}_{AB}(\psi)$  的元素称为合成小波。Shearlet 是合成小波的特例, 矩阵  $\mathbf{A}$  为膨胀矩阵,  $\mathbf{A} = \mathbf{A}_0 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$ ; 矩阵  $\mathbf{B}$  为剪切矩阵,  $\mathbf{B} = \mathbf{B}_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ 。

相较于 Shearlet 变换, NSST 具有移不变性, 能够更准确地捕获图像中的纹理、边缘、细节信息。NSST 变换包括多尺度分解和方向局部化两部分, 具体步骤如下: 1) 采用非下采样金字塔分解源图像, 获得低频、高频分量; 2) 高频分量的方向局部化则通过改进的剪切滤波器实现。它能够将 Shearlet 中的标准剪切滤波器从 Pseudo-Polar 坐标系映射到 Cartesian 坐标系, 免去了降采样操作, 使得 NSST 分解后图像尺寸不变; 最后, 通过傅里叶逆变换进行二维卷积, 得到 NSST 系数。

## 3 基于 Tsallis 灰度熵和 GBVS 模型的红外图像目标提取

### 3.1 基于二维 Tsallis 灰度熵的红外图像分割

红外图像依据热辐射进行成像, 具有良好的目标指向性。首先采用二维 Tsallis 灰度熵<sup>[16]</sup>对红外图像进行阈值分割实现目标的粗提取。然而使用二维 Tsallis 熵对红外图像进行分割提取出来的区域不仅包括目标区域, 还包括一些背景区域和孤立的噪声点等非目标区域, 如图 1 所示。故进一步结合基于视觉注意机制的模型剔除非目标区域, 更精确地提取目标区域。



图 1 基于二维 Tsallis 灰度熵的分割结果

Fig. 1 Segmentation result based on two-dimensional Tsallis gray entropy

### 3.2 基于 GBVS 模型的红外图像目标提取

给定一幅图像, 视觉注意模型根据特定显著度衡量方法提取图像中的显著性区域。GBVS 模型<sup>[17]</sup>是对经典视觉注意模型 Itti 的改进。该模型首先使用高斯金字塔低通滤波器组和 Gabor 金字塔滤波器组对源图像进行滤波, 分别得到多尺度下的亮度特征图 and 方向特征图。GBVS 显著度衡量方法如下: 特征图  $\mathbf{M}_F$  中每个像素点视作图论中图的一个节点, 将该节点与其他节点相连。定义特征图  $\mathbf{M}_F$  中节点  $(i, j)$  和节点  $(p, q)$  之间的差异度为

$$d[(i, j), (p, q)] = \left| -\lg \frac{\mathbf{M}_F(i, j)}{\mathbf{M}_F(p, q)} \right|, \quad (2)$$

式中  $\mathbf{M}_F(i, j)$ 、 $\mathbf{M}_F(p, q)$  分别为特征图中  $(i, j)$ 、 $(p, q)$  处的值。从点  $(i, j)$  到点  $(p, q)$  的有向边被赋予权重:

$$\omega_1[(i, j), (p, q)] = d[(i, j), (p, q)] \cdot F(i - p, j - q), \quad (3)$$

式中  $F(a, b) = \exp\left(-\frac{a^2 + b^2}{2\sigma^2}\right)$ ,  $\sigma$  为可调参数。因为反向边也有相同的权重, 故将由同一个节点发散出的边的权重归一化, 并把节点看作状态, 边的权重则视为转移概率, 构建马尔科夫链。马尔科夫链平衡分布反映出不断行走的随机游走者经过每个节点(状态)停留的时间间隔, 因为其经过与周围节点差异度高的节点的可能性更大, 这些节点的平衡分布值自然就会较高。平衡分布状态图就可作为图像的显著图  $S_A$ , 之后将  $S_A$  归一化。计算各通道的尺度归一化显著图, 经叠加后得到综合显著图  $S$ 。  $S$  的大小与原始图像一致, 其值则反映原始图像中对应像素处的显著值。因此, 可选择合适的阈值对综合显著图进行阈值分割以提取目标。提出的使用自适应阈值  $T$  对其进行分割, 即

$$T = \frac{3}{WH} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} S(x, y), \quad (4)$$

式中  $S(x, y)$  是像素  $(x, y)$  处的显著值;  $W$  和  $H$  分别表示图像的宽度和高度。阈值分割后得到的白色区域为经过上述 GBVS 模型提取出的包含目标的近似椭圆形区域, 这一区域的边缘并非实际目标的边缘, 如图 2(a) 所示。基于二维 Tsallis 熵提取出来的边缘是自然的边缘但包含非目标的边缘, 故将基于二维 Tsallis 熵分割所得二值图像和基于 GBVS 分割所得二值图像取交集, 从而得到最终更精确的实际目标区域, 如图 2(b) 所示。

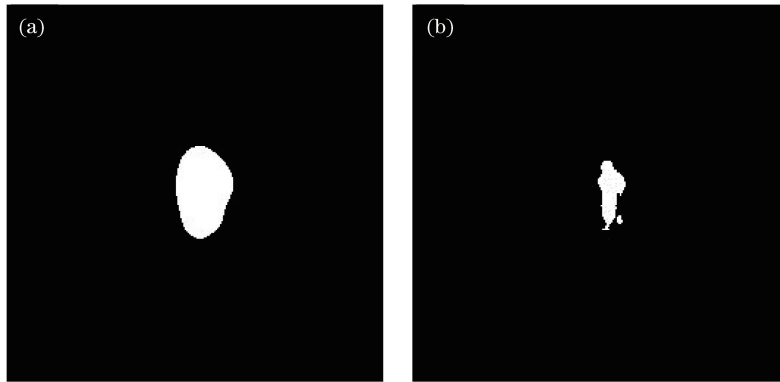


图 2 精确目标提取示意图。(a) 显著图阈值分割结果; (b) 最终提取的目标区域

Fig. 2 Schematic of accurate target extraction. (a) Threshold segmentation result of significant map;

(b) final extracted target region

## 4 NSST 域低频分量及高频分量的融合

红外图像反映热目标信息, 可见光图像体现场景的背景细节信息。图像融合的最终目的是使目标突出、背景细节清晰, 故在图像融合前对红外图像进行了目标提取。在此基础上, 将红外图像与可见光图像分别进行 NSST 变换, 得到各自对应的低频分量与高频分量; 然后对低频分量的目标区域、背景区域分别进行融合, 高频分量则根据方向子带信息和取大来确定; 最后进行 NSST 反变换得到红外与可见光融合图像。

### 4.1 基于引导滤波增强与目标提取的 NSST 域低频分量融合

一般来说, 低频分量包含图像的绝大部分能量。可见光图像具有背景信息丰富的特点, 但其对比度往往不高, 隐藏了较多细节信息, 因此对可见光图像进行增强可有效发掘其细节特征。与此同时, 对红外图像进行增强可突出热目标信息, 使其能够在背景增强的前提下保持目标与背景的对比度。引导滤波<sup>[18]</sup>作为一种新型的图像滤波算法, 在滤掉噪声的同时, 能更好地保护图像边缘。为了增强可见光图像中的细节信息、突出红外图像的热目标, 在图像融合前分别对可见光图像和红外图像进行引导滤波增强。引导滤波增强算法的原理是源图像与其经引导滤波后输出的图像相减得到细节图像, 将此细节图像与滤波后图像经加权得到增强图像。若在空域上直接采用引导滤波增强算法增强图像, 较易引起图像噪声放大, 且考虑到绝大部分噪声集中在图像的高频分量, 故仅对两幅源图像 NSST 域的低频分量利用引导滤波进行增强。

令待滤波图像为  $p$ , 引导图像和滤波输出图像  $q$  是局部线性变换关系。在以像素点  $k$  为中心、大小为  $(2r+1) \times (2r+1)$  的局部方形窗口  $w_k$  中,

$$q_i = a_k I_i + b_k, \quad \forall i \in w_k, \quad (5)$$

式中  $q_i$  代表像素点  $i$  的滤波输出,  $I_i$  表示引导图像中像素  $i$  的值,  $w_k$  中的线性系数  $a_k$ 、 $b_k$  为常系数。为求出  $a_k$  和  $b_k$ , 需要对待滤波图像  $p$  施加约束条件。滤波输出图像  $q$  可视作输入图像  $p$  去除了一些噪声  $n_i$ , 即

$$q_i = p_i - n_i, \quad (6)$$

约束条件为在满足(5)式线性变换关系的同时, 最小化输入输出图像之间差的平方为

$$E(a_k, b_k) = \sum_{i \in w_k} [(a_k I_i + b_k - p_i)^2 + \epsilon a_k^2], \quad (7)$$

式中  $\epsilon$  是正则化参数, 用以防止出现较大的  $a_k$ 。系数  $a_k$ 、 $b_k$  可按下式求出:

$$a_k = \frac{\frac{1}{n_w} \sum_{i \in w_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \epsilon}, \quad (8)$$

$$b_k = \bar{p}_k - a_k \mu_k, \quad (9)$$

式中  $\mu_k$  和  $\sigma_k^2$  分别为引导图像  $I$  在窗口  $w_k$  中的均值和方差;  $n_w$  为窗口  $w_k$  中的像素总数;  $\bar{p}_k = (1/n_w) \sum_{i \in w_k} p_i$  是输入图像  $p$  在  $w_k$  中的均值。求出  $a_k$ 、 $b_k$  后,  $q_i$  可通过(10)式计算。然而像素点  $i$  被所有包含  $i$  的重叠窗口同时覆盖, 故在每个窗口中求出的  $q_i$  值并不唯一, 于是通过求均值获得最终的  $q_i$  值为

$$q_i = (1/n_w) \sum_{k | i \in w_k} (a_k I_i + b_k), \quad (10)$$

将  $\sum_{k | i \in w_k} a_k = \sum_{k \in w_i} a_k$ ,  $\sum_{k | i \in w_k} b_k = \sum_{k \in w_i} b_k$  代入(10)式, 得:

$$q_i = \bar{a}_i I_i + \bar{b}_i, \quad (11)$$

式中  $\bar{a}_i = (1/n_w) \sum_{k \in w_i} a_k$ ,  $\bar{b}_i = (1/n_w) \sum_{k \in w_i} b_k$  为所有包含像素点  $i$  的窗口中常系数均值。对源图像进行引导滤波后, 通过(12)式得到增强图像为

$$E = (p - q) \times \lambda + q, \quad (12)$$

式中  $\lambda$  是增强因子, 其值越大, 图像中的细节越清晰, 但同时噪声也会被放大。通过大量实验将  $\lambda$  设为  $1.7 - \bar{\omega}/255$ ,  $\bar{\omega}$  为可见光图像的全局灰度均值。

鉴于红外图像主要反映目标信息, 而可见光图像的背景信息相对更为丰富, 利用二维 Tsallis 熵和 GBVS 模型对红外图像进行目标提取, 得到目标区域和背景区域。并分别对可见光图像及红外图像 NSST 域上的低频分量进行引导滤波增强, 以发掘更多背景细节、凸显目标和保持目标与背景的对比度。融合后图像低频分量的目标区域系数选自增强后红外图像目标区域低频分量, 背景区域系数选择对应区域的增强后可见光图像低频分量, 这样得到的融合图像低频分量可以完整地保留红外图像的显著目标和可见光图像的背景细节。

#### 4.2 基于方向子带信息和的 NSST 域高频分量融合

图像经 NSST 分解得到的高频分量反映了图像的细节信息。对高频分量的处理直接影响了融合图像的清晰程度以及边缘细节信息的保持程度。NSST 系数模值大的地方, 说明该处存在明显的细节。由于高频带通分量的稀疏性, 显著特征在同一尺度所有方向子带上系数模值都较大, 而其他非显著特征系数模值接近于零。所以, 根据每个像素同一尺度上的方向子带信息足以区分并提取该像素上的细节。此外, 高频分量数据量较大, 基于区域特征的融合规则所需时间较长, 而所提出的高频融合规则既可较好地融合可见光图像与红外图像互补的边缘及纹理细节, 又可加快高频分量的融合速度。源图像在尺度  $l$  上的方向子带信息和为

$$C_l(x, y) = \sum_{1 \leq k \leq 2^n} |C_{k,l}(i, j)|, \quad (13)$$

式中  $2^n$  为尺度  $l$  上的高频方向子带数。据此, 对高频分量采用如下融合规则进行融合。若红外图像  $H$  在尺度  $l$  上像素  $(i, j)$  位置处的方向子带信息和  $C_l^H(i, j)$  大于或等于可见光图像  $V$  的方向子带信息和  $C_l^V(i, j)$ , 那么将红外图像  $H$  在尺度  $l$  上像素位置  $(i, j)$  处的高频分量系数作为融合图像在相应位置上的 NSST 高频分

量系数。反之,则选择可见光图像  $V$  在尺度  $l$  上像素位置  $(i, j)$  处的高频分量系数作为融合图像相应位置的 NSST 域高频分量系数。

#### 4.3 本文融合方法的流程图及具体步骤

基于目标提取与引导滤波增强的红外与可见光图像融合方法的流程如图 3 所示。

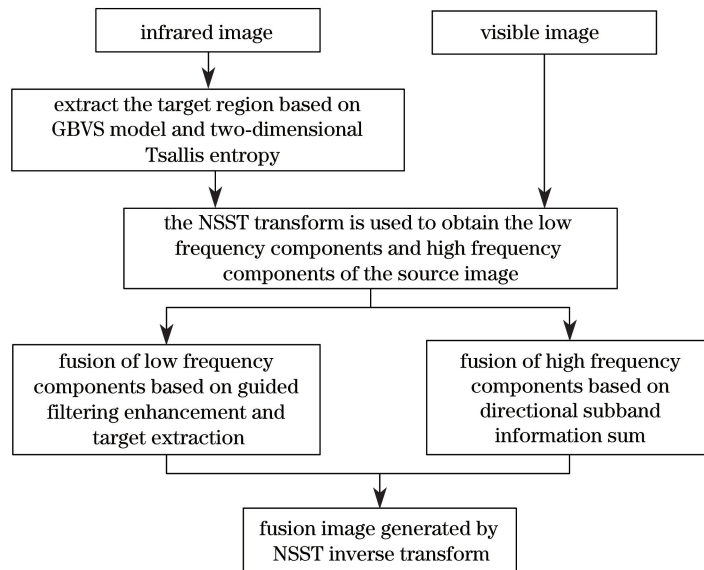


图 3 本文融合方法流程图

Fig. 3 Flow chart of the proposed fusion method

融合方法的具体步骤如下:

- 1) 对红外图像通过二维 Tsallis 熵和基于 GBVS 的视觉注意模型提取目标区域;
- 2) 对红外图像及可见光图像分别进行 NSST 分解;
- 3) 对于两幅源图像 NSST 域的低频分量和高频分量,使用不同的融合规则进行处理:先进行引导滤波增强,对增强后的低频分量采用基于引导滤波增强与目标提取的融合规则进行融合,而高频分量利用方向子带信息和取大的方法进行融合;
- 4) 对处理后的低频分量和高频分量进行 NSST 逆变换,获得最终的融合结果。

## 5 实验结果与分析

为了验证本文方法的有效性,针对大量红外图像和可见光图像进行了融合实验,并与基于 LP 变换、基于 WT、基于 SWT、基于 NSCT<sup>[15]</sup>和文献[13]中基于目标提取与 NSCT 变换等 5 种方法的融合结果进行了比较。其中,WT 方法、SWT 方法的小波基函数为“bior3.7”,分解层数为 2。为了更为合理地比较融合结果,对前 3 种方法,即 LP 方法、WT 方法、SWT 方法均采用常用的低频求平均、高频取大的融合规则。其中的实验图像均来源于常用红外与可见光图像融合测试数据集“UN Camp”。由于篇幅限制,现以 3 组大小均为  $256 \times 256$  且已进行过配准的红外图像和可见光图像融合为例进行说明。图 4 给出了 3 组红外图像和可见光图像,第 1 组的可见光图像中目标人物与背景的灰度值相近,很难看出有人物存在;第 2 组可见光图在光照不足的环境下拍摄,处于黑暗中的人无法很好地识别出来;第 3 幅可见光图像中的目标由于被花草遮挡,几乎完全看不见。从图 4 可以看出,3 幅红外图像的目标都较为明显,而可见光图像包括了不同景物的边缘细节,如何有效地保留红外图像和可见光图像中有效信息将直接影响到融合后图像效果的好坏。图 5 是分别采用本文方法和文献[13]方法提取的红外图像目标区域,为了清楚地显示细节,对尺寸进行了调整。图 6~8 分别给出了针对上述两组图像利用 6 种融合方法的实验结果。实验参数设置如下:NSST 中,采用“maxflat”滤波器进行多尺度分解,分解层数为 3,各层带通方向子带数目取[2,2,2]。实验运行环境:中央处理器为 Intel core i3、2.1 GHz,内存为 2G 的 PC 机,32 位 Win7 操作系统,Matlab 版本为 Matlab7.11.0.(R2010b)。

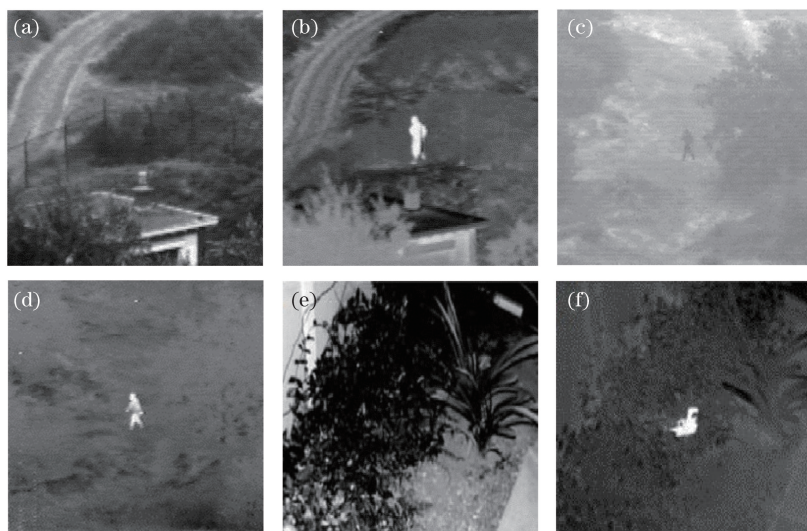


图 4 3 组可见光图像和红外图像。(a)第 1 组可见光图像;(b)第 1 组红外图像;(c)第 2 组可见光图像;  
(d)第 2 组红外图像;(e)第 3 组可见光图像;(f)第 3 组红外图像

Fig. 4 Three groups of visible images and infrared images. (a) Visible image of group 1;  
(b) infrared image of group 1; (c) visible image of group 2; (d) infrared image of group 2;  
(e) visible image of group 3; (f) infrared image of group 3

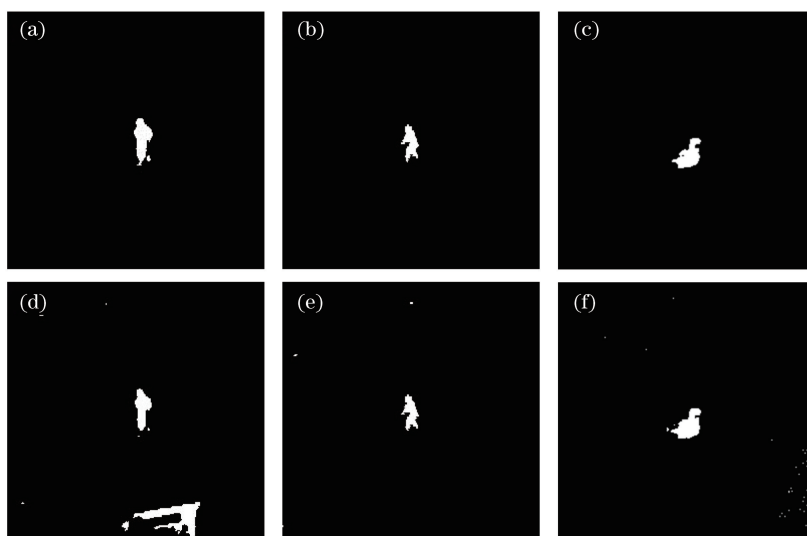


图 5 本文方法和文献[13]方法从 2 幅红外图像中提取的目标。(a)本文方法从图 4(b)提取的目标;  
(b)本文方法从图 4(d)提取的目标;(c)本文方法从图 4(f)提取的目标;(d)文献[13]方法从图 4(b)提取的目标;  
(e)文献[13]方法从图 4(d)提取的目标;(f)文献[13]方法从图 4(f)提取的目标

Fig. 5 Targets extracted from two infrared images by the proposed method and the method in Ref. [13].  
(a) Targets extracted from the Fig. 4(b) by the proposed method; (b) targets extracted from the Fig. 4(d) by  
the proposed method; (c) targets extracted from the Fig. 4(f) by the proposed method;  
(d) targets extracted from the Fig. 4(b) by the method in Ref. [13]; (e) targets extracted from the Fig. 4(d) by  
the method in Ref. [13]; (f) targets extracted from the Fig. 4(f) by the method in Ref. [13]

从图 5 可以看出,本文方法提取的目标区域完整、准确,且不包含某些背景区域或者孤立的噪声点,而文献[13]提取的目标区域不够准确,包含一些背景区域或者孤立的微小噪声点。从图 6~8 的主观视觉效果上来看,本文方法的融合图像充分反映了预先进行目标提取所起的作用,无论从热目标的显著程度还是在图像背景细节表现方面均明显优于其他 5 种融合方法,图像对比度、清晰度高。LP 方法所得融合图像的地物轮廓较模糊、对比度不高、目标也不太明显;WT 方法所得融合图像中引入了较明显的虚影,景像不自然,图像

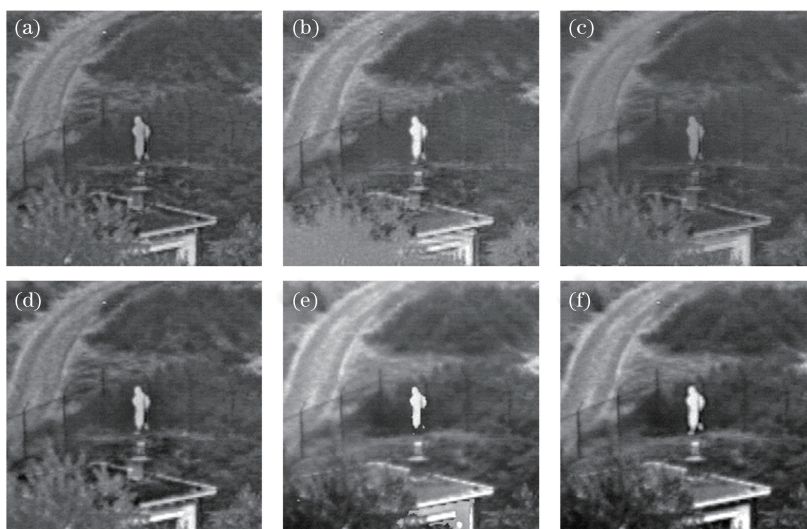


图 6 第 1 组图像的融合结果。(a) LP 方法;(b) WT 方法;(c) SWT 方法;  
(d) NSCT 方法;(e) 文献[13]方法;(f) 本文方法

Fig. 6 Fusion results on the images of group 1. (a) LP method; (b) WT method; (c) SWT method;  
(d) NSCT method; (e) method in Ref. [13]; (f) proposed method

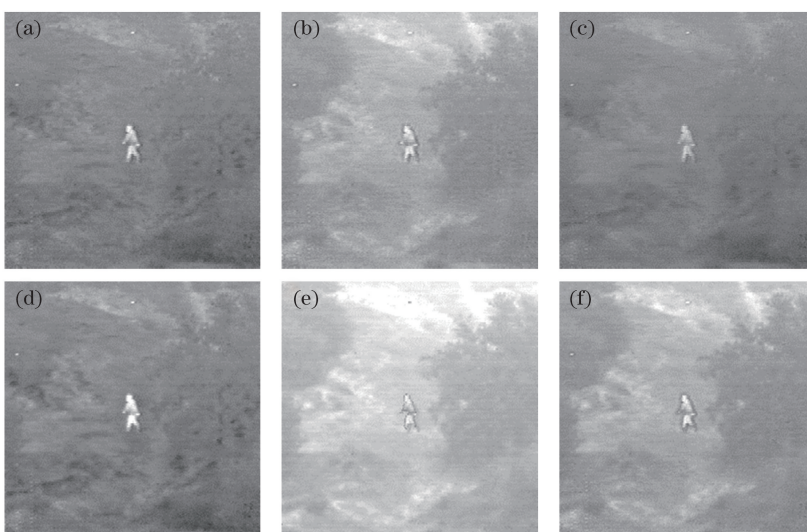


图 7 第 2 组图像的融合结果。(a) LP 方法;(b) WT 方法;(c) SWT 方法;  
(d) NSCT 方法;(e) 文献[13]方法;(f) 本文方法

Fig. 7 Fusion results on the images of group 2. (a) LP method; (b) WT method; (c) SWT method;  
(d) NSCT method; (e) method in Ref. [13]; (f) proposed method

左下部树叶与中部灌木丛处的纹理细节较模糊; SWT 方法所得融合图像的边缘较模糊, 目标与背景对比度不强、目标不清晰; NSCT 方法较好地保留了可见光图像的地物轮廓, 但仍存在图像对比度不高、背景较模糊的问题。与前面 4 种方法相比, 文献[13]方法所得融合图像对比度有所提高, 包含的细节信息也更丰富, 但图 6(e) 中因为提取的目标不够准确导致图像左下角屋檐处存在严重的不连续性失真现象, 同时目标有所缺失, 如图 6(e) 中的人物脚部没有提取出来。且部分区域存在过增强的现象, 导致某些背景和区域过亮, 一些细节被掩盖, 如图 8(e) 的部分区域。而本文方法所得融合图像视觉效果较好, 人物目标提取完整、准确, 且较为明显; 如图 6(f) 左下角的树叶与屋顶处的纹理细节更为连续、清晰。图像对比度高, 更好地保留了源图像中地物的轮廓、纹理信息, 虚影也得到了较好的抑制。

为了进一步对上述 6 种不同融合方法所得到的实验结果进行客观定量评价, 选取信息熵、平均梯度、相关系数、空间频率、扭曲度和标准差等作为客观定量评价指标<sup>[10]</sup>。信息熵用来衡量图像所包含的信息量, 其



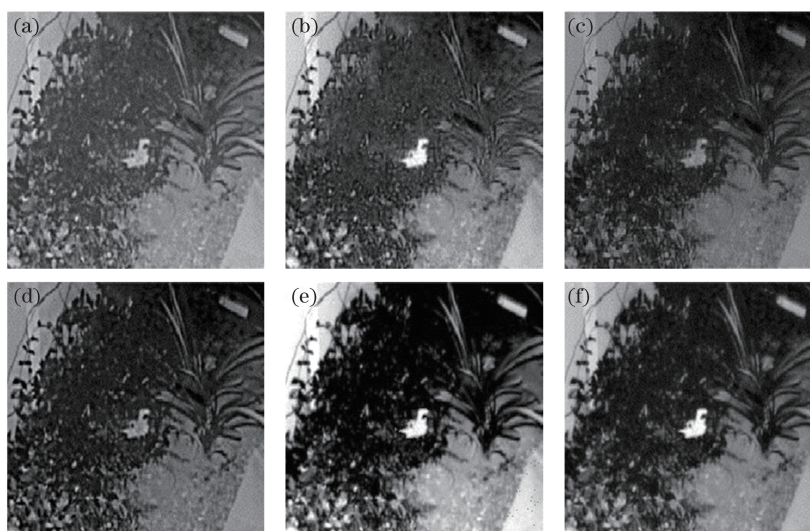


图 8 第 3 组图像的融合结果。(a) LP 方法;(b) WT 方法;(c) SWT 方法;(d) NSCT 方法;(e)文献[13]方法;(f)本文方法  
Fig. 8 Fusion results on the images of group 3. (a) LP method; (b) WT method; (c) SWT method;  
(d) NSCT method; (e) method in Ref. [13]; (f) proposed method

值越大,融合图像所含的信息越多;平均梯度可以反映图像微小细节反差及纹理变化,用来评价图像的清晰度,其值越高,图像越清晰;相关系数反映了融合图像与可见光图像的相关程度,其值越大,则融合图像保持可见光图像的结构信息越好;空间频率是反映图像的总活程度,其值越大,图像越活跃;扭曲度反映了图像的失真程度,其值越小,则说明图像失真越轻微;标准差反映了图像灰度相对于灰度平均值的离散情况,图像的标准差越大,则图像的灰度级越分散,图像反差越大,所含的信息更丰富。表 1 给出了上述 6 种融合方法的 6 个客观指标定量评价结果。

表 1 6 种图像融合方法的定量评价结果  
Table 1 Quantitative evaluation results of six image fusion methods

Image group	Fusion method	Information entropy	Average gradient	Correlation coefficient	Spatial frequency	Distortion degree	Mutual information	Standard deviation
Images of group 1	LP method	6.4951	7.7574	0.7525	11.0577	19.4206	11.0401	26.5996
	WT method	6.7708	7.8434	0.5763	11.0364	23.2909	11.7224	32.8766
	SWT method	6.3207	6.1602	0.7573	8.6306	19.6157	10.6074	24.0802
	NSCT method	6.7655	7.8662	0.7163	11.3051	19.9156	11.5751	31.0908
	Method in Ref. [13]	7.3105	7.4442	0.9321	12.9534	20.0626	<b>13.3453</b>	45.4183
	Proposed method	<b>7.3687</b>	<b>8.9181</b>	<b>0.9383</b>	<b>13.3877</b>	<b>8.8702</b>	12.7921	<b>46.6459</b>
Images of group 2	LP method	6.0510	7.7796	0.7692	<b>9.6541</b>	35.8572	8.9382	16.4286
	WT method	6.3190	<b>7.8598</b>	0.9617	9.6030	3.5831	9.4615	20.9288
	SWT method	5.9878	6.9860	0.7974	8.5302	35.0970	8.9300	15.5457
	NSCT method	6.1282	7.4556	0.7374	9.3233	36.1309	9.0837	9.5244
	Method in Ref. [13]	6.3027	4.4741	0.9868	6.3052	33.4371	<b>10.5666</b>	23.0787
	Proposed method	<b>6.4606</b>	7.6566	<b>0.9637</b>	9.5117	<b>4.5828</b>	9.6762	<b>23.3445</b>
Images of group 3	LP method	7.1235	16.5607	0.9428	21.5501	61.5165	12.0966	41.1500
	WT method	7.2539	<b>17.4572</b>	0.9148	22.4568	24.9870	12.7090	53.5894
	SWT method	7.0456	14.0310	0.9540	17.6974	26.0376	11.7899	39.6084
	NSCT method	7.1799	17.0119	0.9437	22.3278	24.2580	12.1845	43.4219
	Method in Ref. [13]	7.2641	16.8610	0.9578	<b>24.2308</b>	16.9047	<b>14.5343</b>	54.0441
	Proposed method	<b>7.3404</b>	13.5783	<b>0.9788</b>	18.1892	<b>16.5072</b>	12.5676	<b>55.4027</b>

从表 1 中可以看出,本文方法的 6 个定量评价指标中,除了平均梯度略低于 NSCT 方法或者 WT 方法,其他所有 5 个指标都高于另外 5 种方法,尤其在信息熵、相关系数、标准差、扭曲度这 4 个指标上都明显优于其他 5 种方法,表明边缘保持度高、图像细节信息较为丰富。这说明本文融合方法的综合性能最佳,与主观视觉效果相符。值得一提的是,本文方法融合图像的扭曲度相对其他方法明显低得多,而文献[13]方法的扭曲度明显较高,说明引导滤波对可见光图像进行细节增强的同时,有效保护了源可见光图像的有用信息,而文献[13]中的增强方法在对图像增强的同时无法保留好可见光图像的有用信息。NSST 能够更稀疏地表示图像,低频分量采用基于引导滤波增强与目标提取的融合规则进行融合,高频分量利用方向子带信息和取大的规则加以融合,使得最终的融合图像能够更加突出热目标和增强背景的边缘、纹理细节特征,融合效果最好。综合来看,本文方法无论在主观视觉效果还是客观定量评价指标上,都要优于其他 5 种方法。

## 6 结 论

提出了一种基于目标提取与引导滤波增强的红外与可见光图像融合方法。该方法充分利用了 NSST 的多分辨率分析等性质,通过目标提取与引导滤波增强算法处理原始红外图像与可见光图像,从而对经 NSST 分解后的高低频分量设计不同的融合规则,更好地完成了红外和可见光图像的融合。大量实验结果表明,该图像融合方法在主观视觉效果和客观定量评价指标上均优于基于 LP 变换、基于 WT 变换、基于 SWT 变换、基于 NSCT 变换、基于目标提取与 NSCT 变换等方法,不仅能够有效突出目标,而且发掘了更多背景细节信息,是一种行之有效的红外和可见光图像融合方法。

## 参 考 文 献

- [1] Wang R, Du L. Infrared and visible image fusion based on random projection and sparse representation[J]. *International Journal of Remote Sensing*, 2014, 35(5): 1640-1652.
- [2] Zhao J, Gao X, Chen Y, *et al.* Multi-window visual saliency extraction for fusion of visible and infrared images[J]. *Infrared Physics & Technology*, 2016, 76: 295-302.
- [3] Xia Yu, Qu Shiru. Color image fusion framework based on improved  $(2D)^2$  PCA [J]. *Acta Optica Sinica*, 2014, 34(10): 1010001.  
夏 余, 曲仕茹. 改进  $(2D)^2$  PCA 的彩色图像融合框架[J]. *光学学报*, 2014, 34(10): 1010001.
- [4] Zhang Su, Fu Qiang, Duan Jin, *et al.* Low contrast target polarization recognition technology based on lifting wavelet[J]. *Acta Optica Sinica*, 2015, 35(2): 0211002.  
张 肃, 付 强, 段 锦, 等. 基于提升小波的低对比度目标偏振识别技术[J]. *光学学报*, 2015, 35(2): 0211002.
- [5] Liu Kun, Guo Lei, Chang Weiwei. Regional feature self-adaptive image fusion algorithm based on contourlet transform[J]. *Acta Optica Sinica*, 2008, 28 (4): 681-686.  
刘 坤, 郭 雷, 常威威. 基于 Contourlet 变换的区域特征自适应图像融合算法[J]. *光学学报*, 2008, 28(4): 681-686.
- [6] Cunha A L, Zhou J P, Do M N. The non-subsampled contourlet transform: theory, design, and applications[J]. *IEEE Transactions on Image Processing*, 2006, 15(10): 3089-3101.
- [7] Guo K, Labate D. Labate D. Optimally sparse multidimensional representation using shearlets[J]. *SIAM Journal on Mathematical Analysis*, 2007, 39(1): 298-318.
- [8] Easley G, Labate D, Lim W Q. Sparse directional image representations using the discrete shearlet transform[J]. *Applied & Computational Harmonic Analysis*, 2008, 25(1): 25-46.
- [9] Wu Yiquan, Tao Feixiang. Multispectral and panchromatic image fusion based on improved projected gradient NMF in NSST domain[J]. *Acta Optica Sinica*, 2015, 35(4): 0410005.  
吴一全, 陶飞翔. 改进投影梯度 NMF 的 NSST 域多光谱与全色图像融合[J]. *光学学报*, 2015, 35(4): 0410005.
- [10] Kong W W. Technique for gray-scale visual light and infrared image fusion based on non-subsampled shearlet transform[J]. *Infrared Physics & Technology*, 2014, 63(11): 110-118.
- [11] Guo Ming, Wang Shuman. Image fusion based on region and directional variance weighted entropy[J]. *Systems Engineering & Electronics*, 2013, 35(4): 720-724.  
郭 明, 王书满. 基于区域和方向方差加权信息熵的图像融合[J]. *系统工程与电子技术*, 2013, 35(4): 720-724.
- [12] Gao Ying, Wang Amin, Zhi Pengfei, *et al.* Image fusion algorithm based on region segmentation and lifting wavelet

- transform[J]. Journal of Northwestern Polytechnical University, 2014, 32(4): 569-575.
- 高 颖, 王阿敏, 支朋飞, 等. 基于区域分割与提升小波变换的图像融合算法[J]. 西北工业大学学报, 2014, 32(4): 569-575.
- [13] Xing Suxia, Xiao Hongbing, Chen Tianhua, *et al.* Study of image fusion technology based on object extraction and NSCT[J]. Journal of Optoelectronics•Laser, 2013, 24(3): 583-588.
- 邢素霞, 肖洪兵, 陈天华, 等. 基于目标提取与 NSCT 的图像融合技术研究[J]. 光电子•激光, 2013, 24(3): 583-588.
- [14] Kong W, Lei Y, Ni X. Fusion technique for grey-scale visible light and infrared images based on non-subsampled contourlet transform and intensity-hue-saturation transform[J]. IET Signal Processing, 2011, 5(1): 75-80.
- [15] Guo Ming, Fu Zheng, Xi Xiaoliang. Novel fusion algorithm for infrared and visible images based on local energy in NSCT domain[J]. Infrared & Laser Engineering, 2012, 41(8): 2229-2235.
- 郭 明, 符 拯, 奚晓梁. 基于局部能量的 NSCT 域红外与可见光图像融合算法[J]. 红外与激光工程, 2012, 41(8): 2229-2235.
- [16] Li Jiyong, Dang Jianwu, Wang Yangping. Medical image segmentation algorithm based on quantum clonal evolution and two-dimensional Tsallis entropy[J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(3): 465-471.
- 李积英, 党建武, 王阳萍. 融合量子克隆进化与 2 维 Tsallis 熵的医学图像分割算法[J]. 计算机辅助设计与图形学学报, 2014, 26(3): 465-471.
- [17] Harel J, Koch C, Perona P. Graph-based visual saliency[C]. Advances in Neural Information Processing Systems, 2007, 19: 545-552.
- [18] He K, Sun J, Tang X. Guided image filtering[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(6): 1397-1409.
- [19] Yin Wen, Li Yuanxiang, Zhou Zeming, *et al.* Remote sensing image fusion method based on sparse representation[J]. Acta Optica Sinica, 2013, 33(4): 0428003.
- 尹 雯, 李元祥, 周则明, 等. 基于稀疏表示的遥感图像融合方法[J]. 光学学报, 2013, 33(4): 0428003.