Fusion method for infrared and other-type images based on the multi-scale Gaussian filtering and morphological transform

LI Zhi-Jian[1,2],   YANG Feng-Bao[1*],   GAO Yu-Bin[3],   JI Lin-Na[1],   HU Peng[1]

（1. School of Information and Communication Engineering, North University of China, Taiyuan 030051, China;
2. North University of China, Shuozhou, Shuozhou 036000, China;
3. Department of Mathematics, North University of China, Taiyuan 030051, China）

Abstract：To ensure the fusion quality and efficiency simultaneously, a novel image fusion method based on multi-scale Gaussian filtering and morphological transform is proposed. The multi-scale Gaussian filtering is designed to decompose the source images into a series of detail images and approximation images. The multi-scale top- and bottom-hat decompositions are used respectively to fully extract the bright and dark details of different scales in each approximation image. The multi-scale morphological inner- and outer-boundary decompositions are constructed to fully extract boundary information in each detail image. Experimental results demonstrate that the proposed method is comparable to or even better in comparison with typical multi-scale decomposition-based fusion methods. Additionally, the method operates much faster than some advanced multi-scale decomposition-based methods like NSCT and NSST.

Key words：image fusion, multi-scale decomposition, multi-scale Gaussian filtering, morphological transform
PACS：07. 57. −C, 07. 57. Ty, 85. 25. Pb, 85. 60. Gz

基于多尺度高斯滤波和形态学变换的红外与其他类型图像融合方法

李志坚[1,2],   杨风暴[1*],   高玉斌[3],   吉琳娜[1],   胡　鹏[1]
（1. 中北大学 信息与通信工程学院，山西 太原　030051;
2. 中北大学 朔州校区，山西 朔州　036000;
3. 中北大学 数学系，山西 太原　030051）

摘要：为了同时保证融合质量和效率，提出了一种基于多尺度高斯滤波和形态学变换的图像融合方法。设计了多尺度高斯滤波，将源图像分解为一系列细节图像和近似图像，并使用多尺度顶帽和底帽分解来完全提取近似图像中不同尺度的明暗细节。构造了多尺度形态学内外边缘分解，以充分提取细节图像的边缘信息。实验结果表明，该方法与典型的基于多尺度分解的融合方法相当甚至更好，同时比一些先进的基于多尺度分解的方法(如 NSCT 和 NSST)运算速度快得多。

关　键　词：图像融合；多尺度图像分解；多尺度高斯滤波；形态学变换
中图分类号：TN219；TP391.4　　**文献标识码**：A

Introduction

Visible, infrared, and infrared polarization images individually captured by different sensors present complementary information of the same scene, and they could be combined by the image fusion technology to obtain a new, more accurate, comprehensive, and reliable image description of the scene [1]. The fusion methods are varied with image sources, fusion requirements, or purposes [2-4]. In general, the fusion methods can be classified into pixel-, feature- and decision-level. Compared with the latter two fusion levels, the first fusion level can

maintain the source image data as much as possible，so it plays an important role in most image processing tasks. Major pixel-level image fusion methods can be put into four groups according to their adopted theories [5]，namely multi-scale decomposition-based methods，the sparse representation-based methods，methods in other domains，and methods combining different transforms. For the multi-scale decomposition-based methods the decomposition schemes and fusion rules are two aspects that affect fusion quality and efficiency.

For the decomposition schemes，various methods are proposed，like the discrete wavelet transform（DWT）[6]，dual-tree complex wavelet transform（DTCWT）[7]，stationary wavelet transform（SWT）[8]，wavelet packet transform（WPT）[9]，non-subsampled contourlet transform（NSCT）[10-11]，and non-subsampled shearlet transform（NSST）[12]. And many practices proved that NSCT and NSST usually outperform other multi-scale decomposition-based methods in representing 2-D singular signals contained in digital images [13]. But the design of multi-directional filter banks for NSCT and NSST is relatively complex and computational time-consuming，which greatly reduces the efficiency of image fusion.

Fusion rules generally include low and high frequency coefficients fusion rules. The AVG-ABS rule is a simple fusion rule which uses the average rule to combine low-frequency coefficients and uses the absolute maximum rule to combine the high-frequency coefficients. The AVG-ABS fusion rule is computed easily and implemented simply，however，it always causes distortions and artifacts [14-15]. To overcome these shortcomings and improve the fusion quality，a large number of rules have been proposed [15-20]. These rules in Refs. 15-20 have achieved satisfactory results，but they have the disadvantage of high computational complexity.

To ensure both the fusion quality and computational efficiency simultaneously，a novel multi-scale decomposition-based fusion method with dual decomposition structures is proposed. Our method is dedicated to improving the image fusion quality and efficiency from the aspect of image decomposition scheme，while for the rule aspect，our method only uses simple AVG-ABS rule. Firstly，inspired by the idea of constructing octaves in SIFT[21] and SURF[22] algorithms，the source images are decomposed into a series of detail and approximation images by multi-scale Gaussian filters to construct the undecimated pyramid structures. The multi-scale Gaussian filters have increasing standard deviation as well as up-scaling size. Secondly，for the approximation images，i. e.，the top layers of the undecimated pyramid structures，multi-scale morphology top- and bottom-hat decompositions [23-24] are used to fully extract bright and dark details of different scales on the background，and then the contrast of the fused layer is improved by the absolute maximum rule. Thirdly，the multi-scale morphological inner- and outer-boundary decompositions are especially constructed based the idea of constructing multi-scale top- and bottom-hat decompositions. For each detail image，these two morphology decompositions are implemented to ex-

tract the boundary information. And then the decomposed coefficients are combined by the approach of choosing absolute maximum. At last，the fused image is reconstructed through taking the inverse transforms corresponding to the decompositions mentioned.

1　Related theories and work

1. 1　The pyramid transforms

The theory and mathematical representation for constructing multiresolution pyramid transform scheme are presented by Ref. 25 and extended by Ref. 26. A domain V_j of signals is assigned at each level，the analysis operators $\psi_j^{\uparrow}$ maps an image to a higher level in the pyramid，while the synthesis operator $\psi_j^{\downarrow}$ maps an image to a lower level in the pyramid，i. e. $\psi_j^{\uparrow}:V_j:V_{j+1}$ and $\psi_j^{\downarrow}:V_{j+1}:V_j$. The detail signal $y = x - \hat{x}$ contains information of x which does not exist in $\hat{x}$，where $\hat{x} = \psi_{j,i}^{\downarrow}\psi_{i,j}^{\uparrow}(x)$ and $-$ is a subtraction operator $(x,\hat{x}) \mapsto x - \hat{x}$ mapping $\widehat{V}_j \times Y_j$ into the set V_j. The decomposition process of an input image f is expressed as Eq. 1：

$$f^{(0)} \rightarrow \left\{y^{(1)}f^{(2)}\right\} \rightarrow \cdots \rightarrow \left\{y^{(1)},y^{(2)},\cdots,y^{(j)}f^{(j+1)}\right\} \rightarrow \cdots,(1)$$

where

$$\begin{cases} f^{(0)} = f \in V_0 \\ f^{(j+1)} = \psi_j^{\uparrow}\left(f^{(j)}\right) \in V_{j+1}, & j \geqslant 0 \\ y^{(j)} = f^{(j)} - \psi_j^{\downarrow}\left(f^{(j+1)}\right) \in Y_j \end{cases} \quad . \quad (2)$$

And the reconstruction process through the backward recursion is expressed as Eq. 3：

$$f = f^{(0)},f^{(j)} = \psi_j^{\downarrow}\left(f^{(j+1)}\right) + y^{(j)},j \geqslant 0 \quad , \quad (3)$$

Eq. 1 and Eq. 3 are called the pyramid transform and the inverse pyramid transform respectively.

1. 2　Scale space representation and multi-scale Gaussian filtering

The scale space of an image can be generated through convolving the image with Gaussian filters，and it has been successfully applied in SIFT [21] to detect key points which are invariant to scales. In Ref. 26 the scale space is divided into octaves. For each octave，the initial image is iteratively convolved with Gaussians with increasing standard deviation to generate a set of scale space images（Gaussian images），and one of the Gaussian images is downsampled to obtain the initial image of the next octave. Then，the Difference of Gaussians（DoG）images are obtained by subtracting adjacent Gaussian images. In SURF [22]，in order to omit the down-sampling step，the scale space is obtained by increasing the size of filter.

Inspired by the above algorithms，we have the source image repeatedly convolved with Gaussian filters whose standard deviation and size increase simultaneously to construct undecimated pyramid structure. Then，the DoG images are produced by subtracting adjacent Gaussian images. Accordingly，the transform scheme of such pyramid is given by Eq. 4：

$$\begin{cases} f^{(0)} = f \in V_0 \\ f^{(j+1)} = G_j * f^{(j)} \in V_{j+1}, \ j \geqslant 0 \\ y^{(j)} = f^{(j)} - f^{(j+1)} \in Y_j \end{cases} , \quad (4)$$

where

$$G_j = \left(1 / 2\pi\sigma_j^2\right) \cdot \exp\left[-\left(x^2 + y^2\right) / 2\sigma_j^2 \right] . \quad (5)$$

G_j is the Gaussian kernels (filters) with a size of $(6\sigma_j + 1) \times (6\sigma_j + 1)$ in this paper, and * denotes the convolution operation. The parameter σ_j is the standard deviation, which is increasing with j, and in this paper $\sigma_{j+1} = k\sigma_j, k > 1$. Then the source image f can be decomposed into an approximation image and a set of detail images as shown in scheme 1, and it can also be exactly reconstructed through the following recursion:

$$f = f^{(0)}, f^{(j)} = f^{(j+1)} + y^{(j)}, \ j \geqslant 0 . \quad (6)$$

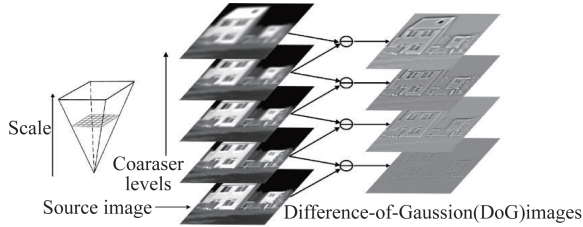The four-level decomposition scheme is illustrated in Fig. 1.

Fig. 1　Example of four-level decomposition by multi-scale Gaussian filtering
图1　用多尺度高斯滤波进行4层分解的示例

1.3　Multi−scale morphological transforms

The multi-scale top-hat transform using structuring elements with up-scaling size can extract the light and dark details at different image scales in image fusion[24]. Based the idea of constructing multi-scale top-hat transform, the multi-scale morphological inner-boundary transform is constructed. These two kinds of morphological transforms can be expressed as Eq. 4, with the Gaussian kernel G_j being replaced by morphological opening operation Open_{b_j} and erosion operation Erosion_{b_j}, respectively. For purpose of extracting details of different scales, the scale of structuring element b_j increases with j. The inverse transforms can be expressed as Eq. 6.

The multi-scale morphological bottom-hat transform and its inverse are shown as follows

$$\begin{cases} f^{(0)} = f \in V_0 \\ f^{(j+1)} = \text{Close}_{b_j}(f^{(j)}) \in V_{j+1}, \ j \geqslant 0 \\ y^{(j)} = f^{(j+1)} - f^{(j)} \in Y_j \end{cases} , \quad (7)$$

$$f = f^{(0)}, f^{(j)} = f^{(j+1)} - y^{(j)}, \ j \geqslant 0 . \quad (8)$$

where the analysis operator is morphological closing operation Close_{b_j}, with b_j also increasing with j. The morphological outer-boundary transform and its inverse are similar to the bottom-hat transform and its inverse, with

Close_{b_j} being replaced by dilation operation Dilation_{b_j}.

2　Proposed method framework

The proposed fusion method comprises three processes that are multi-scale decomposition, fusion, and reconstruction.

2.1　Multi−scale decomposition process

The K-level decomposition of a given source image f by the scheme (4) has the form

$$f \rightarrow \left\{ y^{(1)}, y^{(2)}, \cdots y^{(k)}, \cdots y^{(K)}, f^{(K+1)} \right\} , \quad (9)$$

where $y^{(k)}$ represents the detail image at level k and $f^{(K+1)}$ denotes the approximation image of this multi-scale structure.

$f^{(K+1)}$ is a coarse representation of f and usually inherit a few bright and dark details, thus the multi-scale top- and bottom-hat decompositions are used to extract bright objects on a dark background and dark objects on a bright background of different scales, respectively. Henceforth, $f^{(K+1)}$ can be decomposed by schemes mentioned in subsection 1.3 as

$$f^{(K+1)} \rightarrow \begin{cases} \left\{ y_t^{(1)}, y_t^{(2)}, \cdots y_t^{(l)}, \cdots y_t^{(M)}, f_t^{(M+1)} \right\} \\ \left\{ y_b^{(1)}, y_b^{(2)}, \cdots y_b^{(l)}, \cdots y_b^{(M)}, f_b^{(M+1)} \right\} \end{cases} , \quad (10)$$

where $y_t^{(l)}$ and $y_b^{(l)}$ represent the detail images at level l obtained by the top- and bottom-hat decomposition process, respectively. And $f_t^{(M+1)}$ and $f_b^{(M+1)}$ denote the approximation images of the multi-scale top- and bottom-hat structure, respectively. Figure 2 is given as an example of three-level top- and bottom-hat decompositions.
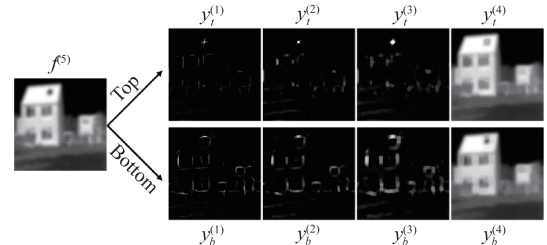

Fig. 2　Example of three-level top- and bottom-hat decompositions of the input image $f^{(5)}$
图2　对输入图像 $f^{(5)}$ 进行3层顶、底帽分解的示例

The detail image $y^{(k)}$ in scheme 9 comprises various details like edges and lines, thus the multi-scale inner- and outer-boundary transforms mentioned in subsection 1.3 are used to extract inner-boundary as well as outer-boundary information of different scales. Hence, $y^{(k)}$ can be decomposed as

$$y^{(k)} \rightarrow \begin{cases} \left\{ y_i^{(k,1)}, y_i^{(k,2)}, \cdots y_i^{(k,l)}, \cdots y_i^{(k,N_k)}, f_i^{(k,N_k+1)} \right\} \\ \left\{ y_o^{(k,1)}, y_o^{(k,2)}, \cdots y_o^{(k,l)}, \cdots y_o^{(k,N_k)}, f_o^{(k,N_k+1)} \right\} \end{cases} , \quad (11)$$

where $y_i^{(k,l)}$ and $y_o^{(k,l)}$ represent the detail images at level l of $y^{(k)}$ that are obtained by the inner- and outer-boundary decomposition process, respectively. $f_i^{(k,N_k+1)}$ and $f_o^{(k,N_k+1)}$ are the approximation images of $y^{(k)}$ at the highest level

LI Zhi-Jian *et al*: Fusion method for infrared and other-type images based on the multi-scale Gaussian filtering and morphological transform

6 期

813

of the multi-scale inner- and outer-boundary structure, respectively. Figure 3 gives an example of three-level inner- and outer-boundary decompositions.
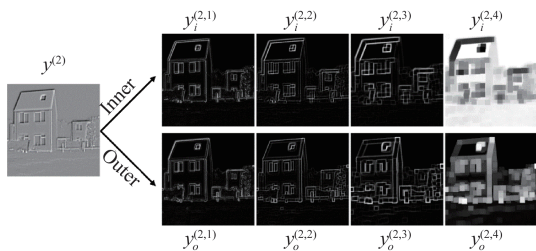

Fig. 3 Example of three-level inner and outer-boundary decompositions of the input image $y^{(2)}$.
图 3 对输入图像 $y^{(2)}$ 进行 3 层内、外边缘分解的示例

2.2 Fusion process

In this paper, the composite approximation coefficients of the approximation image in the multi-scale top- and bottom-hat structures take the average of the approximation of the sources. For the composite detail coefficients of the detail images, the absolute maximum selection rule is used.

2.2.1 Fusion rules for the multi−scale top− and bottom−hat structures

The vector coordinate $\boldsymbol{n} = (n, m)$ is used here to denote the location of an image. For instance, $y_t^{(l)}(\boldsymbol{n}|A)$ represents the detail coefficient for the multi-scale top-hat structure at location $\boldsymbol{n}$ within level l of source image A. And the notation $(\cdot)$ will be used to denote an image, e. g., $y_t^{(l)}(\cdot|A)$ refers to the detail image.

The arbitrary fused detail coefficient $y_t^{(l)}(\boldsymbol{n}|F)$ and the fused approximation coefficient $y_t^{(l)}(\boldsymbol{n}|F)$ of the multi-scale top-hat structure are obtained through

$$\begin{cases} y_t^{(l)}(\boldsymbol{n}|F) = \max\left\{\left|y_t^{(l)}(\boldsymbol{n}|A)\right|, \left|y_t^{(l)}(\boldsymbol{n}|B)\right|\right\} \\ f_t^{(M+1)}(\boldsymbol{n}|F) = \alpha_t f_t^{(M+1)}(\boldsymbol{n}|A) + \beta_t f_t^{(M+1)}(\boldsymbol{n}|B) \end{cases} . (12)$$

The weights α_t and β_t take 0. 5, which preserves the mean intensity of the two source images. Likewise, $y_b^{(l)}(\boldsymbol{n}|F)$ and $f_b^{(M+1)}(\boldsymbol{n}|F)$ of the multi-scale bottom hat structure are obtained through

$$\begin{cases} y_b^{(l)}(\boldsymbol{n}|F) = \max\left\{\left|y_b^{(l)}(\boldsymbol{n}|A)\right|, \left|y_b^{(l)}(\boldsymbol{n}|B)\right|\right\} \\ f_b^{(M+1)}(\boldsymbol{n}|F) = \alpha_b f_b^{(M+1)}(\boldsymbol{n}|A) + \beta_b f_b^{(M+1)}(\boldsymbol{n}|B) \end{cases}, (13)$$

with $\alpha_b = \beta_b = 0.5$.

The selective rule in Eq. 12 means that we choose the brighter ones in the bright details, and the selective rule in Eq. 13 means that we choose the darker ones in the dark details. In this way, the bright and dark details of different scales can be fully extracted and hence the contrast at each layer can be improved.

2.2.2 Fusion rules for the multi−scale inner− and outer−boundary structures

For an arbitrary fused detail coefficient $y_i^{(k,l)}(\boldsymbol{n}|F)$ of the multi-scale inner-boundary structures, we only use the absolute maximum selection rule:

$$y_i^{(k,l)}(\boldsymbol{n}|F) = \begin{cases} y_i^{(k,l)}(\boldsymbol{n}|A) & \text{if } \left|y_i^{(k,l)}(\boldsymbol{n}|A)\right| > \left|y_i^{(k,l)}(\boldsymbol{n}|B)\right| \\ y_i^{(k,l)}(\boldsymbol{n}|B) & \text{otherwise} \end{cases} . (14)$$

So is the fused approximation coefficient $f_i^{(k, N_k+1)}(\boldsymbol{n}|F)$. In such way, the boundary information such as edges and lines of different scales can be well preserved. Likewise, arbitrary $y_o^{(k,l)}(\boldsymbol{n}|F)$ and $f_o^{(k, N_k+1)}(\boldsymbol{n}|F)$ of the multi-scale outer-boundary structures are also obtained by the absolute maximum selection rule.

2.3 Reconstruction process

According to Eqs. 6 and 8, the reconstruction of the approximation image $f^{(K+1)}(\cdot|F)$ can be obtained through the multi-scale top- and bottom-hat inverse transforms as

$$f^{(K+1)}(\cdot|F) = \left(\gamma_{M+1} f_t^{(M+1)}(\cdot|F) + \sum_{l=0}^{M} \gamma_l y_t^{(l)}(\cdot|F)\right)\Big/2 + \left(\kappa_{M+1} f_b^{(M+1)}(\cdot|F) - \sum_{l=0}^{M} \kappa_l y_b^{(l)}(\cdot|F)\right)\Big/2 \quad, \quad (15)$$

which means both bright and dark information are of equal importance to the source image. In addition, we attach equal importance to the features of different scale levels, thus the weights in Eq. 15 are set to be $\gamma_p = \kappa_p = 0.5$, $p = 1, 2, \cdots, (M+1)$.

Similarly, inner- and outer-boundary information are considered to be equally important to the source image, and so are the features of different scale levels. Thus, according to Eqs. 6 and 8, the reconstruction of an arbitrary detail image $y^{(k)}(\cdot|F)$ through the multi-scale inner- and outer-boundary inverse transforms can be obtained as

$$y^{(k)}(\cdot|F) = \left(f_i^{(k, N_k+1)}(\cdot|F) + \sum_{l=1}^{N_k} y_i^{(k,l)}(\cdot|F)\right)\Big/2 + \left(f_o^{(k, N_k+1)}(\cdot|F) - \sum_{l=1}^{N_k} y_o^{(k,l)}(\cdot|F)\right)\Big/2 \quad . \quad (16)$$

At last, the fused image $f(\cdot|F)$ can be reconstructed by

$$f(\cdot|F) = f^{(0)}(\cdot|F) = f^{(K+1)}(\cdot|F) + \sum_{k=1}^{K} y^{(k)}(\cdot|F). (17)$$

3 Experiments

3.1 Experimental setups

In order to validate the performance of the proposed method, experiments are conducted on two categories of source images including ten pairs of infrared-visible images (Fig. 4(a)) and eight pairs of infrared intensity-polarization images (Fig. 4(b)). The two source images in each pair are pre-registered and the size of each image is set to 256×256 pixels. The experiments in this paper are programmed by Matlab 2016b and run on an Intel(R) Core(TM) i5-6500 CPU @ 3. 20GHz Desktop with 16. 0 GB RAM.

Various pixel-level multi-scale decomposition-based methods including DWT, DTCWT, SWT, WPT, NSCT, and NSST are compared with the proposed method. All the compared methods adopt the simple AVG-ABS rule. According to Ref. 13, most of the methods mentioned above perform well when the decomposition levels for them are set to 3. Thus, for purpose of making reliable and persuasive comparisons, the decomposition levels
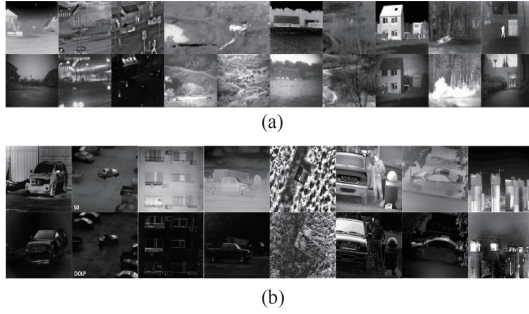
Fig. 4   The two kinds of source images (a) infrared-visible images, (b) infrared intensity-polarization images
图 4   两类源图像(a)红外与可见光图像,(b)红外光强与偏振图像

for the methods mentioned above are all set to 3. And to make each method achieve a good performance the other parameters are also suggested by Ref. 13, some of which are listed in Table 1.

Table 1   The parameters set in the compared methods. 'Filter' represents the Orientation filter; 'Levels' denotes the decomposition levels and the corresponding number of orientations for each level.
表1   对比方法中参数的设置。"Filter"表示方向滤波器;"Levels"表示分解层数和每个层对应的方向数

Methods	Pyramid filter	Filter	Levels
DWT	rbio1. 3		3
DTCWT	5−7	q−6	3
SWT	bior1. 3		3
WPT	bior1. 3		3
NSCT	maxflat	dmaxflat5	4,8,16
NSST	maxflat		4,8,16

For NSST, the size of the local support of shear filter at each level are selected as 8, 16, 32. As for the proposed method, the parameters σ_0 and k for the multi-scale Gaussian filtering process in Eq. 5 are selected experimentally. In this experiment, the source images are decomposed by 3-layer multi-scale Gaussian decomposition, and different fused images are obtained by changing the parameters σ_0 and k. During the fusion process, the AVG-ABS rule is also adopted. When σ_0 and k are in certain value, every fusion image will be evaluated by seven objective assessment metrics (mentioned in subsection 3.2). For each metric, its mean value is obtained by averaging the evaluation results of the fusion images. Then, seven mean values are summed to get the sum values of objective metrics. Figure 5 gives three surface plots which show variations of the sum of the seven metrics with σ_0 and k. As shown in Fig. 5, the optimal values of σ_0 and k for the four kinds of images are obtained. The structuring elements in the multi-scale inner- and outer-boundary decompositions are selected as square, and in the multi-scale top- and bottom-hat decompositions they are chosen to be disk. σ_0 and k in Eq. 5 and the parameters K, M, N_1, N_2, and N_3 in schemes

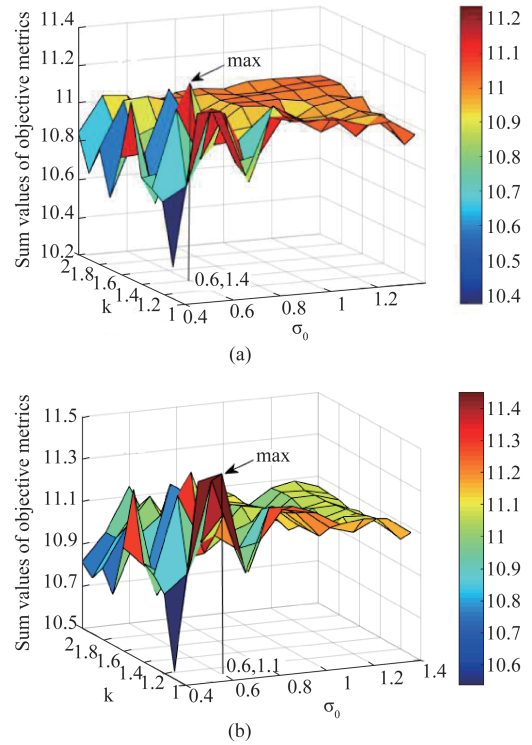Fig. 5   Estimation of the parameters σ_0 and k for (a) infrared-visible images, (b) infrared intensity-polarization images
图 5   参数 σ_0 和 k 的估算(a)红外与可见光图像,(b)红外光强与偏振图像

9, 10, and 11 are set as shown in Table 2 to make the proposed method achieve a good performance.

Table 2   The parameters of the proposed method for the four kinds of source images.
表2   所提出的方法对四种源图像的参数

Source images	Parameters		
	σ_0	k	$[K, M, N_1, N_2, N_3]$
Infrared−visible	0. 6	1. 4	$[3,2,0,1,2]$
Infrared intensity−polarization	0. 6	1. 1	$[3,2,1,1,2]$

3. 2   Objective assessment metrics

Seven representative metrics, i. e. , Q_0 [27], Q_E [28], $Q^{AB/F}$ [29], information entropy (IE) [30], mutual information (MI) [31], Tamura contrast (TC) [32], and visual information fidelity (VIF) [33] are employed to evaluate the proposed method comprehensively. The variable n in TC is chosen to be 4.

3. 3   Experimental results
3. 3. 1   Subjective assessment

In this section, the subjective assessment of the fusion methods is done by comparing the visual results obtained from the above and proposed methods. One sample pair in each type of source images are selected for visual comparison as shown in Figs. 6 and 7.

In Fig. 6, both the DWT and WPT methods distort the edges of the roof, which was shown clearly in magnified squares. The DTCWT, SWT, NSCT and NSST methods produce artificial edges in the sky around the

LI Zhi-Jian *et al*：Fusion method for infrared and other-type images based on the multi-scale Gaussian filter-
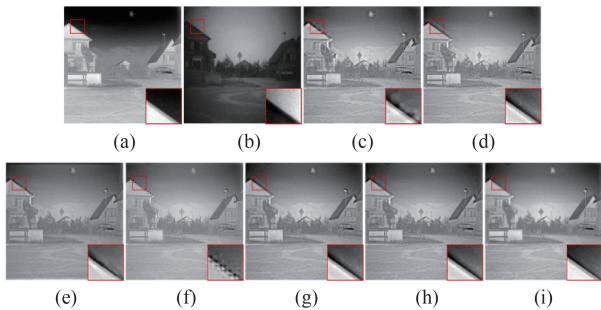ing and morphological transform

6 期

815

Fig. 6　Fusion results of one pair of the infrared-visible images (a) infrared image, (b) visible image, (c)-(i) the fusion results of the DWT, DTCWT, SWT, WPT, NSCT, NSST, and the proposed methods.
图6　红外与可见光图像中一对图像的融合结果(a)红外图像，(b)可见光图像，(c)-(i)依次为 DWT、DTCWT、SWT、WPT、NSCT、NSST 和所提方法的融合结果
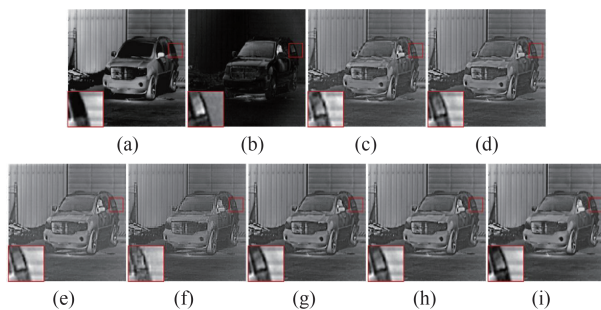

Fig. 7　Fusion results of one pair of the infrared intensity-polarization images (a) Infrared intensity image, (b) Infrared polarization image, (c)-(i) the fusion results of the DWT, DTCWT, SWT, WPT, NSCT, NSST, and the proposed methods.
图7　红外光强与偏振图像中一对图像的融合结果(a)红外光强图像，(b)红外偏振图像，(c)-(i)依次为 DWT、DTCWT、SWT、WPT、NSCT、NSST 和所提方法的融合结果

roof，while the result obtained by the proposed method is free from such artifacts or brightness distortions. In addition，the walls and the clouds in the sky in Fig. 6(i) are brighter those in Fig. 6(g) and (h)，which means that the fused image of the proposed method has better contrast.

The edges of the car are distorted heavily in Fig. 7 (f)，and slightly distorted in Figs. 7(c-e) which is shown more clearly in the corresponding regions in magnified square. And Figs. 7(c-h) show some artifacts around the edges of the car. However，in Figs. 7(i) there are no distortions or certain artifacts. In addition，the car in magnified square of Fig. 7(i) is the darker than those in Figs. 7 (h) and (i)，which demonstrate that the proposed method has better contrast.

The above experiments confirm that the proposed method performs better in visual effect for the two categories of source images. Although adopting the simple AVG-ABS rule，the proposed method does not generate certain artifacts or distortions and simultaneously preserves the detail information of source images as much as possible.

3. 3. 2　Objective assessment

The objective assessment of the seven multi-scale decomposition-based methods are shown in Tables 3. For the infrared-visible images，the proposed method performs the best on all the seven metrics. For the infrared intensity-polarization images，the proposed method performs the best on the other five metrics except Q_0 and Q_E on which it performs the second best. It can also be obtained from Tables 3 that compared with the seven methods，the proposed method always has the best assessment on metrics $Q^{AB/F}$，IE，MI，TC，and VIF. It means that the proposed method can transfer the original information of source image including the edges and brightness details to the fused image sufficiently，and improve the contrast of the fused image.

3. 3. 3　Comparison of computational efficiency

To verify the efficiency of the proposed method，an experiment is conducted on the image sequences named as "Nato_camp"，"Tree"，and "Dune" from the TNO Im-

Table 3　Objective assessment of all methods (the best result of each metric is highlighted in bold).
表3　所有方法的客观评价值(每个指标的最佳结果用粗体突出显示)

Images	Methods	Q_0	$Q^{AB/F}$	Q_E	IE	MI	TC	VIF
Infrared-visible	DWT	0. 439 1	0. 485 8	0. 226 8	6. 660 1	2. 165 8	0. 258 8	0. 293 6
	DTCWT	0. 444 6	0. 517 3	0. 257 9	6. 683 0	2. 223 5	0. 293 7	0. 294 9
	SWT	0. 445 2	0. 509 7	0. 245 7	6. 615 5	2. 187 2	0. 220 3	0. 278 4
	WPT	0. 407 9	0. 395 2	0. 161 4	6. 638 5	2. 194 9	0. 274 5	0. 273 8
	NSCT	0. 466 9	0. 528 1	0. 259 5	6. 696 1	2. 263 3	0. 294 0	0. 314 5
	NSST	0. 465 3	0. 523 1	0. 257 0	6. 685 8	2. 257 5	0. 290 2	0. 310 3
	Proposed	**0. 475 7**	**0. 535 6**	**0. 268 9**	**6. 735 9**	**2. 470 7**	**0. 317 7**	**0. 362 6**
Infrared intensity−polarization	DWT	0. 385 3	0. 420 6	0. 167 6	6. 478 2	2. 266 4	0. 347 6	0. 219 6
	DTCWT	0. 394 4	0. 458 5	**0. 208 9**	6. 570 7	2. 341 5	0. 468 4	0. 243 7
	SWT	0. 387 5	0. 439 1	0. 193 1	6. 473 0	2. 342 9	0. 330 8	0. 230 0
	WPT	0. 346 9	0. 343 9	0. 119 8	6. 405 2	2. 291 7	0. 443 7	0. 197 2
	NSCT	0. 413 3	0. 467 5	0. 197 7	6. 564 6	2. 391 7	0. 458 5	0. 257 4
	NSST	**0. 413 8**	0. 464 1	0. 199 5	6. 574 0	2. 389 8	0. 459 7	0. 259 2
	Proposed	0. 413 4	**0. 469 0**	0. 201 3	**6. 658 0**	**2. 624 1**	**0. 547 8**	**0. 313 7**

Table 4　Average processing time (unit: sec.) comparison of eight methods. Each value represents the average run time of a frame in a certain sequence.
表4　8种方法的平均处理时间(单位:秒)的比较。每个值表示特定序列中帧的平均运行时间

Image sequences	DWT	DTCWT	SWT	WPT	NSCT	NSST	Proposed
Nato_camp	0. 018 0	0. 036 2	0. 064 7	0. 140 1	24. 517 3	2. 307 2	0. 141 9
Tree	0. 016 5	0. 035 7	0. 064 3	0. 139 8	24. 821 5	2. 292 3	0. 141 1
Duine	0. 017 1	0. 036 1	0. 064 1	0. 140 6	24. 584 1	2. 288 1	0. 141 2

age Fusion Dataset[34]. Table 4 shows the average processing time of all methods for a frame. Compared with the DWT, DTCWT, SWT, and WPT methods, the proposed method is more time-consuming because these four methods contain one types of multi-scale decomposition while the proposed method contains two, i. e. , the multi-scale decomposition using multi-scale Gaussian filtering and the multi-scale morphological decomposition, as mentioned in Sec. 2. Compared with the NSCT and NSST methods which also contain two kinds of multi-scale decomposition, the proposed method is far more efficient mainly because the design of the multi-directional filter banks for NSCT and NSST is relatively complex and the processing speed of multi-directional filtering is much lower than that of multi-scale morphological operations.

4　Conclusions

Experiments on both visual quality and objective assessment demonstrate that although adopting the simple AVG-ABS rule, the proposed method does not generate certain artifacts or distortions and performs very well in aspects like information preservation and contrast improvement. Under the premise of ensuring image fusion quality, the proposed method is also proved computationally efficient. The proposed method provides an option for the fusion situations needing both high quality and particularly computational efficiency, such as fast high-resolution images fusion and video fusion.

Acknowledgment

References

[1] Ma J, Ma Y, Li C. Infrared and visible image fusion methods and applications: A survey [J]. *Information Fusion*, 2018, **45** (S15662553517307972).

[2] Xin J, Qian J, Yao S, et al. A Survey of infrared and visual image fusion methods [J]. *Infrared Physics & Technology*, 2017, **85**: 478–501.

[3] Yang F, Wei H. Fusion of infrared polarization and intensity images using support value transform and fuzzy combination rules [J]. *Infrared Physics & Technology*, 2013, **60**: 235–243.

[4] Hu P, Yang F, Wei H, et al. Research on constructing difference-features to guide the fusion of dual-modal infrared images [J]. *Infrared Physics & Technology*, 2019, **102**: 102994.

[5] Li S, Kang X, Fang L, et al. Pixel-level image fusion: A survey of the state of the art [J]. *Information Fusion*, 2017, **33**: 100–112.

[6] Amolins K, Zhang Y, Dare P. et al. Wavelet based image fusion techniques: An introduction, review and comparison [J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2007, **62**(4): 249–263.

[7] Selesnick I W, Baraniuk R G, Kingsbury N C. The dual-tree complex wavelet transform [J]. *IEEE Signal Processing Magazine*, 2005, **22**(6): 123–151.

[8] Singh D, Garg D, Pannu H S, Efficient landsat image fusion using fuzzy and stationary discrete wavelet transform [J]. *Journal of Photographic Science*, 2017, **65**(2): 108–114.

[9] Walczak B, Bogaert B V D, Massart D L. Application of wavelet packet transform in pattern recognition of near-IR data [J]. *Analytical Chemistry*, 1996, **68**(10): 1742–1747.

[10] Da Cunha A L, Zhou J, Do M N. The nonsubsampled contourlet transform: theory, design, and applications [J]. *IEEE Transactions on Image Processing*, 2006, **15**(10): 3089–3101.

[11] Zhu Z, Zheng M, Qi G, et al. A phase congruency and local laplacian energy based multi-modality medical image fusion method in NSCT domain [J]. *IEEE Access*, 2019, **7**: 20811 – 20824.

[12] Ming Y, Wei L, Xia Z, et al. A novel image fusion algorithm based on nonsubsampled shearlet transform [J]. *Optik – International Journal for Light and Electron Optics*, 2014, **125**(10): 2274–2282.

[13] Li S, Yang B, Hu J. Performance comparison of different multi-resolution transforms for image fusion [J]. *Information Fusion*, 2011, **12** (2): 74–84.

[14] Li S, Kang X, Hu J. Image fusion with guided filtering [J]. *IEEE Transactions on Image Processing*, 2013, **22**(7): 2864–2875.

[15] Du J, Li W, Xiao B. Anatomical-functional image fusion by information of interest in local laplacian filtering domain [J]. *IEEE Transactions on Image Processing*, 2017, **26**(12): 5855–5865.

[16] Bhatnagar G, Wu J, Liu Z. Directive contrast based multimodal medical image fusion in NSCT domain [J]. *IEEE Transactions on Multimedia*, 2013, **9**(6): 1014–1024.

[17] Gong J, Wang B, Lin Q, et al. Image fusion method based on improved NSCT transform and PCNN model [C]. 2016 9th International Symposium on Computational Intelligence and Design (ISCID), 2016.

[18] Ma T, Jie M, Bin F, et al. Multi-scale decomposition based fusion of infrared and visible image via total variation and saliency analysis [J]. *Infrared Physics & Technology*, 2018, **92**: 154–162.

[19] Yin M, Liu. X, Yu Liu, et al. Medical image fusion with parameter-adaptive pulse coupled-neural network in nonsubsampled shearlet transform domain [J]. *IEEE Transactions on Instrumentation & Measurement*, 2018, **68**(1): 1–16.

[20] Li Y, Sun Y, Huang X, et al. An image fusion method based on sparse representation and sum modified-laplacian in NSCT domain [J]. *Entropy*, 2018, **20**(7): 522.

[21] Lowe D G. Distinctive image features from scale-invariant key points [J]. *International Journal of Computer Vision*, 2004, **60**(2): 91–110.

[22] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF) [J]. *Computer Vision & Image Understanding*, 2008, **110** (3): 346–359.

[23] Mukhopadhyay S, Chanda B. Fusion of 2D grayscale images using multiscale morphology [J]. *Pattern Recognition*, 2001, **34** (10): 1939–1949.

[24] Bai X, Gu S, Zhou F, et al. Multiscale top-hat selection transform based infrared and visual image fusion with emphasis on extracting regions of interest [J]. *Infrared Physics & Technology*, 2013, **60** (5): 81–93.

[25] Goutsias J, Heijmans H M. Nonlinear multiresolution signal decomposition schemes--part I: morphological pyramids [J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2000, **9**(11): 1862–1876.

[26] Piella G. A general framework for multiresolution image fusion: from pixels to regions [J]. *Information Fusion*, 2003, **4**(4): 259–280.

[27] Wang Z, Bovik A C. A universal image quality index [J]. *IEEE Signal Processing Letters*, 2002, **9**(3): 81–84.

[28] Piella G, Heijmans H. A new quality metric for image fusion [C]. In-

ternational Conference on Image Processing, 2003.

[29] Hong R. Objective image fusion performance measure [J]. *Military Technical Courier*, 2000, **56**(2): 181-193.

[30] Roberts W J, Van J A A, Ahmed F. Assessment of image fusion procedures using entropy, image quality, and multispectral classification [J]. *Journal of Applied Remote Sensing*, 2008, **2**(1): 1-28.

[31] Qu G, Zhang D, Yan P. Information measure for performance of image fusion [J]. *Electronics Letters*, 2002, **38**(7): 313-315.

[32] Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception [J]. *IEEE Trans. syst. man. cybernet*, 1978, **8**(6): 460-473.

[33] Han Y, Cai Y, Cao Y, *et al*. A new image fusion performance metric based on visual information fidelity [J]. *Information Fusion*, 2013, **14**(2): 127-135.

[34] http://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029.