The multiscale directional bilateral filter and its application to multisensor image fusion

Jianwen Hu, Shutao Li*

College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

A R T I C L E   I N F O

Article history:
Received 16 September 2010
Received in revised form 7 December 2010
Accepted 17 January 2011
Available online 26 January 2011

Keywords:
Multisensor image fusion
Directional filter bank
Multiscale transform

A B S T R A C T

In this paper, a novel multiscale geometrical analysis called the multiscale directional bilateral filter (MDBF) which introduces the nonsubsampled directional filter bank into the multiscale bilateral filter is proposed. Through combining the characteristic of preserving edge of the bilateral filter with the ability of capturing directional information of the directional filter bank, the MDBF can better represent the intrinsic geometrical structure of images. The MDBF, which is a multiscale, multidirectional and shift-invariant image decomposition scheme, is used to fuse multisensor images in this paper. The source images are first decomposed into the directional detail subbands and the approximation subbands via the MDBF. Then, the directional detail subbands and the approximation subbands are fused according to the given fusion rule, respectively. Finally, the inverse MDBF is applied to the fused subbands to obtain the fused image. Experimental results over visible and infrared images and medical images demonstrate the superiority of our method compared with conventional methods in terms of visual inspection and objective measures.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

With the development of imaging techniques, multiple imaging devices can be used to acquire several different images of the same scene or object. Multiple sensor modalities can enhance the system performance and robustness in a wide range of modern military and civilian imaging applications. However, the increase of sensor modalities results in the “information overload” problem. In addition, viewing multiple sensor modalities simultaneously leads to an unnecessary load on the observer and combining information across a group of observers becomes almost impossible [1]. Some researcher found that multisensor image fusion is an effective technique to solve this problem [2]. Image fusion technique integrates the information from two or more images of the same scene or object into a composite image which is more informative and suitable for human visual perception or computer processing [3]. The benefits of image fusion include wider spatial and temporal coverage, decreased uncertainty and improved reliability [4]. Nowadays, image fusion has been applied to many fields such as military [5], remote sensing [6] and medical aid diagnosis [7].

So far, researchers have developed many image fusion algorithms [8–11], which can be categorized into pixel, feature, and decision levels according to the representation format at which image information is processed [3]. Pixel-level methods merge multiple input images into a single fused image in raw image representation. Compared with feature- and decision-level methods, pixel-level image fusion can preserve more original information. We focus on pixel-level image fusion in this paper.

Since multiscale transforms can effectively extract the important information of images such as lines and details, they are the most commonly used methods to fuse images [12–16]. The classical multiscale transforms include the Laplacian pyramid transform [17], the discrete wavelet transform (DWT) [18], the stationary wavelet transform (SWT) [19] and the dual-tree complex wavelet transform (DTCWT) [20]. To better capture the intrinsic geometrical structure of images, various multiscale geometrical analysis (MGA) methods, including curvelet [6], contourlet [21] and so on [22–24], have been developed. In this paper, we propose a novel MGA method based on the bilateral filter.

The bilateral filter proposed by Tomasi and Manduchi in [25] is a technique to smooth images while preserving edges. At present, the bilateral filter has been applied to various image processing tasks such as denoising, text editing and relighting, and tone management [26]. The kernel of the bilateral filter is space-variant, and it uses the difference between neighbor pixel values which are correlated with edges and details. Therefore, the bilateral filter can smooth images while preserving edges. It is well known that multiscale representation is important for image processing. Fattal et al. proposed the multiscale bilateral filter (MBF) and used it to enhance shape and detail [27]. However, edges and lines in natural images reflect important information of images, and they may...
present various directions. Therefore, the directional filter is very important for image processing to effectively extract the significant information of images. Thus, we propose the multiscale directional bilateral filter (MDBF), which introduces the directional filter bank (DFB) to the MBF to more effectively extract the information of images. The MDBF has proven to be effective in image fusion as shown in this paper.

The remainder of this paper is organized as follows. In Section 2, we review the principle of the bilateral filter and its multiscale extension version, and develop the MDBF through combining the DFB with the MBF. Section 3 proposes the multisensor image fusion method via the MDBF. The experimental results and discussions are presented in Section 4. Finally, concluding remarks and future works are given in Section 5.

2. Bilateral filter and multiscale directional bilateral filter

In this section, we briefly review the basic theory of the bilateral filter and the multiscale bilateral filter, and propose the multiscale directional bilateral filter.

2.1. The basic theory of bilateral filter and multiscale bilateral filter

Gaussian filtering is one of the most used methods for image smoothing. It is a weighted average of the intensity of the adjacent positions with a weight decreasing with the spatial distance to the center position. The underlying assumption of Gaussian filtering is that images vary slowly over space. Nevertheless, this assumption fails at edges. To overcome this problem, the bilateral filter was developed [25]. It is a nonlinear, noniterative and local technique, and can smooth image while preserving edges, by means of a non-linear combination of nearby pixel values. So far, due to the advantage of preserving edges, the bilateral filter has obtained extensive applications [26]. The bilateral filter combines intensity values based on both their geometric closeness and their photometric similarity, and prefers near values to distant values in both spatial and range domain. Closeness refers to vicinity in the spatial domain, and similarity to vicinity in the range. The bilateral filter (BF) can be viewed as the combination of the spatial filter and the range filter, and it is defined by

$$\text{BF}|I_p = \frac{1}{W_p} \sum_{q \in \mathcal{N}} G_r(\|p - q\|)G_s(l_p - l_q)$$

with $W_p = \sum_{q \in \mathcal{N}} G_r(\|p - q\|)G_s(l_p - l_q)$.

where $p$ and $q$ are two-dimensional vectors representing the spatial location of the image, $l_p$ and $l_q$ are the intensity values of the location $p$ and $q$ respectively, $\mathcal{N}$ denotes the neighborhood of the pixel $p$, $W_p$ is a normalization factor, $G_r = \exp\left(-x^2/\sigma^2\right)$, $G_s$ is a spatial Gaussian function that decreases symmetrically as distance from the center increases, and $G_{r, s}$ are the standard deviations of Gaussian functions $G_r$ and $G_s$, respectively, determine the amount of filtering for the image $I$.

The bilateral filter is often used to decompose an image into an approximation subband and a detail subband. The approximation subband is obtained by applying Eq. (1) to the original image, and the detail subband is the difference between the original image and the approximation subband. However, generally, one level of the decomposition is not enough to extract the important information of images since images may contain information of various resolutions. Fattal et al. extended the original bilateral filter into multiscale form [27]. An idea of the dyadic wavelet transform known as the à trous algorithm is used in the multiscale bilateral filter (MBF), which is defined as

$$I_p^{j+1} = \frac{1}{W_p} \sum_{q \in \mathcal{N}} W^\prime(p - q)G_r(l_p - l_q/\sqrt{2^j})$$

with $W^\prime(x) = \begin{cases} G_s(\|x/\sqrt{2^j}\|) & \text{if } x \leq 2^j \text{ and } \|x/\sqrt{2^j}\| < m \\ 0 & \text{otherwise} \end{cases}$

where $j$ denotes the $j$th level of the decomposition, $W_p$ is still a normalization factor and similar to Eq. (2). The neighborhood $\mathcal{N}$ is modified so that the pixel location $q$ addresses only the points where $W^\prime$ is non-zero. The MBF is iterated over approximation subbands according to Eq. (3). After $J$ levels multiscale bilateral filtering, the image is decomposed into an approximation subband and $J$ detail subbands. The MBF is a shift-invariant scheme because it uses a technique based on the à trous algorithm which is shift-invariant.

2.2. Multiscale directional bilateral filter

The edges and lines in images may have various directions due to the intrinsic geometrical structure of typical natural images. For
example, the line discontinuities which extensively exist in images correspond to directional information of images [28]. Recent developments on new two-dimensional multiscale representation such as curvelet [29] and contourlet [21] indicate that the directionality is a crucial feature for an efficient image representation. In addition, the research on visual perception indicates that the cells having directional selectivity are found in the retinas and visual cortices of the entire major vertebrate classes [30]. Therefore, obtaining the direction information is important to effectively extract the significant information of images. In this paper, we combine the MBF with the directional filter, and propose the multiscale directional bilateral filter (MDBF).

Bamberger and Smith constructed a two-dimensional directional filter bank (DFB) through combining critically-sampled two-channel fan filter banks with resampling operations [31]. This DFB is efficiently implemented via an l-level binary tree decomposition that results in $2^l$ subbands with directional wedge-shaped frequency partitioning. One disadvantage of the DFB in [31] is that it does not have lowpass or highpass subbands. Therefore, some developers devoted to multiresolution critically-sampled DFB. In [32], the directional filters at multiple scales are obtained by cascading fan filter banks at the $2^l$ outputs of the DFB of [31]. In [33], two-channel parallelogram filter banks are cascaded at the output of a four-band DFB to provide a different multiscale DFB. Nguyen and Oraintara proposed a class of multiresolution directional filter banks, and each DFB except for the uniform DFB in this class can be used to decompose an image yielding 12 directions [28]. The above DFBs are shift-variant because they are critically-sampled. Nevertheless, the property of shift-invariance is important for many image processing applications such as image fusion. A shift-invariant directional extension can be obtained by a non-subsampled directional filter bank (NSDFB), which is constructed through removing the downsamplers and upsamplers in each two-channel filter bank in the DFB tree structure and upsampling the filters accordingly [22]. Similarly with the critically-sampled DFB, the filter banks in the NSDFB tree structure are obtained from

<table>
<thead>
<tr>
<th>Algorithm 1</th>
<th>Image fusion method based on the MDBF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
<td>$N$ source images $I_n$, the number of decomposition levels $J$, the standard deviations $\sigma_r$ and $\sigma_s$.</td>
</tr>
<tr>
<td><strong>Output:</strong></td>
<td>the fused image $I_f$.</td>
</tr>
<tr>
<td><strong>Step 1: Decomposition</strong></td>
<td>for $n=1, \ldots, N$ do</td>
</tr>
<tr>
<td></td>
<td>$I_n^0 = I_n$</td>
</tr>
<tr>
<td></td>
<td>for $j=1, \ldots, J$ do</td>
</tr>
<tr>
<td></td>
<td>$I_n^{j+1}(p) = \frac{1}{W_{pq}} \sum_{p,q} w(p-q) G_{pq}(I_n^j(p) - I_n^j(q)) I_n^j(q)$</td>
</tr>
<tr>
<td></td>
<td>$C_n^j = I_n^{j+1} - I_n^j$</td>
</tr>
<tr>
<td></td>
<td>end for</td>
</tr>
<tr>
<td></td>
<td>$C_n^{j+1} = I_n^{j+1}$</td>
</tr>
<tr>
<td></td>
<td>end for</td>
</tr>
<tr>
<td><strong>Step 2: Fusion</strong></td>
<td>$C_f^{j+1}(p) = C_n^{j+1}(p)$, $k = \arg \max { C_n^{j+1}(p) }$, $j=1, \ldots, J$</td>
</tr>
<tr>
<td></td>
<td>$C_f^{j+1} = \frac{1}{N} \sum_{n=1}^{N} C_n^{j+1}$</td>
</tr>
<tr>
<td><strong>Step 3: Reconstruction</strong></td>
<td>for $j=1, \ldots, J$ do</td>
</tr>
<tr>
<td></td>
<td>Apply the NSDFB reconstruction to $C_n^j$ to obtain $C_f^j$</td>
</tr>
<tr>
<td></td>
<td>end for</td>
</tr>
<tr>
<td></td>
<td>$I_f = \sum_{j=1}^{J} C_f^j$</td>
</tr>
</tbody>
</table>

**Fig. 2.** The pseudo-code of the proposed fusion method.
a single nonsampled filter bank with fan filters. To obtain the multidirectional decomposition, the NSDFB are iteratively used. Fig. 1 presents the four-channel nonsampled directional filter bank.

The multiscale directional bilateral filter is constructed through combining the MBF with the NSDFB. The MBF is firstly applied to the original image to obtain the detail subbands and the approximation subband. Then, the detail subbands are fed into a NSDFB so that the direction information is captured. It is worth noting that the combined scheme is shift-invariant since both the MBF and the DFB used in this paper are nonsampled. In addition, the nonsampled contourlet transform (NSCT) [22], which has been extensively used for various image processing applications such as image fusion, is similar to the MBF. However, compared with the NSCT, our method has the following advantages. Firstly, due to that the weight of the MBF changes alongside pixel locations, the multiscale transform of the MBF is adaptive. This enables that the MBF can better capture edges and details. Secondly, for the multiscale transform of the MBF is simpler, there is no need to design the pyramid filter. Thirdly, the MDBF can better capture edges and details. Secondly, for the multiscale transform of the NSCT, our method has the following advantages. Firstly, due to that the source images

3. Multisensor image fusion using multiscale directional bilateral filter

In the section, we will discuss in details the multisensor image fusion algorithm using the MBF. To simplify the discussion, it is firstly assumed that the fusion process only involves two registered source images denoted by A and B to generate a more informative fused image F. The fusion procedure is described as follows:

1. Firstly, the source images A and B are decomposed by the MBF to obtain an approximation subband and a series of directional detail subbands respectively:

\[
\{C^d_A(p), C^{d+1}_A(p) \mid 1 \leq j \leq J, d = 1, 2, \ldots, 2^d\}
\]

and

\[
\{C^d_B(p), C^{d+1}_B(p) \mid 1 \leq j \leq J, d = 1, 2, \ldots, 2^d\},
\]

where \(p\) is the spatial location of subband coefficients, \(J\) is the number of decomposition levels, \(C^{d+1}(p)\) represents the approximation subband coefficients at the coarsest scale, and \(C^d(p)\) denotes the directional detail subband coefficients at the \(j\)th scale and the \(d\)th direction.

2. Then, according to the given fusion rule, the approximation subbands and the directional detail subbands are fused respectively to obtain the fused subbands

\[
\{C^d_F(p), C^{d+1}_F(p) \mid 1 \leq j \leq J, d = 1, 2, \ldots, 2^d\}.
\]

3. Finally, the fused image \(F\) is obtained via the inverse MDBF from the fused approximation and directional detail subbands.

The aim of image fusion is to integrate important features from the source images to the fused image. Three components of image fusion algorithm are identifying, comparing and transferring the significant feature such as edges, details and directions. As a result, a method which can capture these features is highly necessary. The MBF possesses the abilities of preserving edges and capturing directional information. Therefore, the MBF can be used to identify the salient features. The sharp brightness changes which correspond to the edge and contrast of images result in large variation of range weight of bilateral filter. The sharp brightness changes lead to large transform values. Moreover, if the edge or object presents a certain direction, the directional filter with the same direction gives large response. Thus, the transform values of the MBF can be directly used to measure the salience of features, and the largest of absolute transform values represents the most salient features. In this paper, the choose-max fusion rule is used for directional detail subbands:

![Fig. 4. The variation of the metric Q_ABF on different \(\sigma_s\) and \(\sigma_r\).](image)

![Fig. 3. The pseudo-code of the NSDFB decomposition.](image)
\[ C_j^{d}(p) = \begin{cases} C_A^{d}(p) & \text{if } \left| C_A^{d}(p) \right| > \left| C_B^{d}(p) \right| \\ C_B^{d}(p) & \text{else} \end{cases} \] (5)

The approximation subbands represent the outline of the original image. In this paper, the approximation subbands are fused by the average method, which is commonly used in multiscale transform based image fusion methods [13,14] and is defined as follows:

\[ C_{j+1}^{d}(p) = \frac{C_{j}^{d}(p) + C_{j+1}^{d}(p)}{2} \] (6)

We can notice that the above fusion method can be easily extended to the case that the number of source images is more than two. The pseudo-codes of the proposed fusion method for \( N \) source images and the NSDFB decomposition are shown in Figs. 2 and 3, respectively. The NSDFB reconstruction is similar to the NSDFB decomposition, and the difference between them is that the DFB are different. To implement the perfect reconstruction, the decomposition DFB and the reconstruction DFB must satisfy Bezout identity. The symbol "\( \ast \)" in Fig. 3 represents convolution with upsampled filters. The number of decomposition direction is \( 2^L \) for \( L \) levels of binary tree decomposition.
4. Experimental results and discussions

4.1. Parameter setting and quality evaluation metrics

The proposed method is tested over five pairs of multisensor images. As the difference of the fused images is sometimes too small to be distinguished by human eyes, six quality metrics is used to objectively evaluate the performance of various methods, i.e., mutual information (MI), $Q_0$, $Q_W$, $Q_E$, $Q_{AB/F}$, and visual information fidelity (VIF). Mutual information evaluates the fusion result from information theory viewpoint and reflects the amount of information transferred from source images to fused image [34]. The metric $Q_0$ models image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion [35]. The

Table 1
The objective evaluation of various methods for the “UN camp” images.

<table>
<thead>
<tr>
<th>Fusion methods</th>
<th>VIF</th>
<th>MI</th>
<th>$Q_{AB/F}$</th>
<th>$Q_W$</th>
<th>$Q_E$</th>
<th>$Q_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWT</td>
<td>0.5731</td>
<td>2.0678</td>
<td>0.4063</td>
<td>0.6429</td>
<td>0.4627</td>
<td>0.5891</td>
</tr>
<tr>
<td>SWT</td>
<td>0.6115</td>
<td>2.0699</td>
<td>0.4246</td>
<td>0.6432</td>
<td>0.4627</td>
<td>0.6031</td>
</tr>
<tr>
<td>DTCWT</td>
<td>0.6161</td>
<td>2.0728</td>
<td>0.4271</td>
<td>0.6343</td>
<td>0.4591</td>
<td>0.6068</td>
</tr>
<tr>
<td>NSCT</td>
<td>0.6363</td>
<td>2.0728</td>
<td>0.4458</td>
<td>0.6668</td>
<td>0.4881</td>
<td>0.6312</td>
</tr>
<tr>
<td>MBF</td>
<td>0.6119</td>
<td>2.0732</td>
<td>0.4346</td>
<td>0.7015</td>
<td>0.5010</td>
<td>0.6183</td>
</tr>
<tr>
<td>MDBF</td>
<td>0.6377</td>
<td>2.0744</td>
<td>0.4450</td>
<td>0.7093</td>
<td>0.5135</td>
<td>0.6347</td>
</tr>
</tbody>
</table>

Table 2
The objective evaluation of various methods for the “Gun” images.

<table>
<thead>
<tr>
<th>Fusion methods</th>
<th>VIF</th>
<th>MI</th>
<th>$Q_{AB/F}$</th>
<th>$Q_W$</th>
<th>$Q_E$</th>
<th>$Q_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWT</td>
<td>0.5730</td>
<td>2.3185</td>
<td>0.6566</td>
<td>0.8545</td>
<td>0.7286</td>
<td>0.7304</td>
</tr>
<tr>
<td>SWT</td>
<td>0.6491</td>
<td>2.3233</td>
<td>0.6758</td>
<td>0.8565</td>
<td>0.7414</td>
<td>0.7176</td>
</tr>
<tr>
<td>DTCWT</td>
<td>0.6563</td>
<td>2.3250</td>
<td>0.6882</td>
<td>0.8564</td>
<td>0.7361</td>
<td>0.7373</td>
</tr>
<tr>
<td>NSCT</td>
<td>0.6714</td>
<td>2.3269</td>
<td>0.6982</td>
<td>0.8656</td>
<td>0.7485</td>
<td>0.7470</td>
</tr>
<tr>
<td>MBF</td>
<td>0.6994</td>
<td>2.3277</td>
<td>0.7066</td>
<td>0.8861</td>
<td>0.7731</td>
<td>0.7572</td>
</tr>
<tr>
<td>MDBF</td>
<td>0.7201</td>
<td>2.3294</td>
<td>0.7103</td>
<td>0.8882</td>
<td>0.7773</td>
<td>0.7599</td>
</tr>
</tbody>
</table>

Fig. 6. The “Gun” source images and fused images of various fusion algorithms.
metric \( Q_W \) reflects that how much salient information in the source images is transferred into the fused image [36]. Both the metrics \( Q_E \) [36] and \( Q^{AB/F} \) [37] can evaluate the edge information in the fused image. Nevertheless, their emphases are different. The metric \( Q_E \) evaluates the quality of the fused image by combining \( Q_W \) of the original images with \( Q_W \) of the corresponding edge images. Therefore, the metric \( Q_E \) contains the information of both the original images and the edge images. However, \( Q_E \) does not contain the information of the edge orientation, which is an important attribute of edges. The metric \( Q^{AB/F} \) reflects the quality of visual information obtained from the fusion of input images through combining the edge strength and orientation preserving. The VIF proposed by Sheikh and Bovik [38] assesses image quality by modeling human visual system, natural scenes and image distortion. It can quantify the distortions including additive noises, blurs, and global or local contrast changes. Note that the larger the value of all the metrics mentioned above, the better the fused image.

The proposed method is compared with some well known multisensor image fusion methods based on multiscale transforms including the DWT, SWT, DTCWT, NSCT, and MBF. For all compared methods, the fusion rule is the same as that of the proposed method. The wavelet basis Biorthogonal “bior(2,2)’’ is used for the DWT and SWT based methods. For the DTCWT-based method, the Legall 5–3 tap filter and the quarter sample shift orthogonal 10–10 tap filter are used as the first-level and other-levels filters respectively. For the NSCT-based method, the “pyrexc” filter derived from one-

![Fig. 7. The “Tropical” source images and fused images of various fusion algorithms.](image-url)
dimension using maximally flat mapping function with 2 vanishing moments is used as the pyramid filter, and the “vk” filter derived from the McClellan transform is used as the directional filter. These filters present the best results for the above methods [16]. The directional filter of our method is the same as that of the NSCT-based method. The number of decomposition levels of the DWT, SWT, DTCWT, and MBF-based methods is set to 3. Three levels of decomposition, with 8, 8, 16 directions from coarser scale to finer scale, are used in our method and the NSCT-based method.

The size of the neighborhood $r$, $s$, and $r_s$ are three important parameters in our method and the MBF-based method. The bilateral filter can be viewed as the combination of the spatial filter and the range filter. The spatial filter reflects the geometric closeness, while the range filter reflects the intensity similarity. The parameters $r_s$ and $r$, specify the amount of filtering for the image. When the parameter $r_s$ is over-large, the spatial filter is close to the mean filter in a given support domain, and the bilateral filter is approximate to the range filter. This causes that the distant pixels and the close pixels have similar effect to results, which is not appropriate. When the parameter $r_s$ is under-small, the spatial filter is close to the impulse filter. The surrounding pixels will be underweighted and therefore the filter takes no effect. The parameter $r$, has similar effect to the range filter. The effective support domain of the Gaussian function $G_r$, constrains the size of the neighborhood $r$. Therefore, the size of the neighborhood $r$ is dependent on the parameter $r_s$. An experiment is performed over the medical images Fig. 8a and b through varying $r_s$ and $r$. We firstly fix $r_s$ to 2 and alter $r$, then fix $r$ to 1 (the range of pixel intensity is $[0 1]$) and alter $r_s$. The values of the metric $Q_{\text{AB/F}}$ of the above experiment are presented in Fig. 4. It can be seen that the metric $Q_{\text{AB/F}}$ first increases, then decreases, finally tends to remain steady. The metric $Q_{\text{AB/F}}$ reaches the largest value when $r_s$ and $r$ equal to 2 and 1 respectively. Therefore, the parameter $r_s$ and $r$ are set to 2 and 1. And the size of the neighborhood $r$ is set to $5 \times 5$. From Fig. 4, it can be seen that high values of $r_s$ and $r$ would be also good since the optimal settings are used for the fixed standard deviations. In addition, the directional filter is used to further decompose detail subbands after bilateral filtering. This improves the performance of image fusion, which makes that high values would be also good.

4.2. Experimental results on visible and infrared images

Since the spectral bands of visible and infrared are different, they can reflect different features of scenes. Usually, visible images provide spatial details, while infrared images can show some especial objects such as concealed creatures and weapons which have different temperature with their backgrounds. Therefore, visible images and infrared images present complementary features. These two images can be fused to obtain a composite image which can simultaneously show spatial details and especial objects. We present the experimental results over three pairs of visible and infrared images using our method and the compared methods in the following.

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image440x297 to 536x392)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image440x412 to 536x508)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image440x526 to 536x622)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image454x68 to 517x110)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image454x183 to 522x224)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image536x297 to 632x392)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image536x412 to 632x508)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image536x526 to 632x622)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image545x68 to 608x110)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image545x183 to 608x224)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image632x297 to 728x392)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image632x412 to 728x508)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image632x526 to 728x622)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image645x68 to 698x110)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image645x183 to 701x224)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image728x297 to 824x392)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image728x412 to 824x508)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image728x526 to 824x622)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image738x68 to 801x110)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image738x183 to 795x224)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image824x297 to 920x392)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image824x412 to 920x508)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image824x526 to 920x622)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image838x68 to 901x110)

![Fig. 8. The “Medical-1” images and fused images of various fusion algorithms.](image838x183 to 901x224)

Table 3

<table>
<thead>
<tr>
<th>Fusion methods</th>
<th>VIF</th>
<th>MI</th>
<th>Q_{\text{AB/F}}</th>
<th>Q_{\theta}</th>
<th>Q_{\gamma}</th>
<th>Q_{\alpha}</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWT</td>
<td>0.5589</td>
<td>2.2137</td>
<td>0.6059</td>
<td>0.8109</td>
<td>0.6758</td>
<td>0.7785</td>
</tr>
<tr>
<td>SWT</td>
<td>0.6329</td>
<td>2.2235</td>
<td>0.6436</td>
<td>0.8176</td>
<td>0.6918</td>
<td>0.7912</td>
</tr>
<tr>
<td>DTCWT</td>
<td>0.6172</td>
<td>2.2201</td>
<td>0.6309</td>
<td>0.8066</td>
<td>0.6725</td>
<td>0.7826</td>
</tr>
<tr>
<td>NSCT</td>
<td>0.6445</td>
<td>2.2255</td>
<td>0.6584</td>
<td>0.8304</td>
<td>0.7072</td>
<td>0.8037</td>
</tr>
<tr>
<td>MBF</td>
<td>0.6748</td>
<td>2.2364</td>
<td>0.6652</td>
<td>0.8577</td>
<td>0.7362</td>
<td>0.8200</td>
</tr>
<tr>
<td>MDBF</td>
<td>0.6999</td>
<td>2.2375</td>
<td>0.6728</td>
<td>0.8606</td>
<td>0.7425</td>
<td>0.8254</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Fusion methods</th>
<th>VIF</th>
<th>MI</th>
<th>Q_{\text{AB/F}}</th>
<th>Q_{\theta}</th>
<th>Q_{\gamma}</th>
<th>Q_{\alpha}</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWT</td>
<td>0.6255</td>
<td>2.2544</td>
<td>0.6208</td>
<td>0.6546</td>
<td>0.4513</td>
<td>0.6244</td>
</tr>
<tr>
<td>SWT</td>
<td>0.7169</td>
<td>2.2811</td>
<td>0.6878</td>
<td>0.6433</td>
<td>0.4450</td>
<td>0.6390</td>
</tr>
<tr>
<td>DTCWT</td>
<td>0.6800</td>
<td>2.2612</td>
<td>0.6539</td>
<td>0.6283</td>
<td>0.4246</td>
<td>0.6035</td>
</tr>
<tr>
<td>NSCT</td>
<td>0.7161</td>
<td>2.2802</td>
<td>0.6921</td>
<td>0.6581</td>
<td>0.4487</td>
<td>0.6448</td>
</tr>
<tr>
<td>MBF</td>
<td>0.7601</td>
<td>2.2809</td>
<td>0.7109</td>
<td>0.7379</td>
<td>0.5264</td>
<td>0.6604</td>
</tr>
<tr>
<td>MDBF</td>
<td>0.7705</td>
<td>2.2821</td>
<td>0.7195</td>
<td>0.7397</td>
<td>0.5328</td>
<td>0.6618</td>
</tr>
</tbody>
</table>
The first experiment is performed on a frame of “UN camp” sequence images provided by Toet et al. [39] in the ImageFusion.org website [40]. Fig. 5a is the visible image presenting spatial details, while Fig. 5b is the infrared image which clearly shows a person. The infrared image was captured by a Radiance HS IR camera (Raytheon), which is sensitive for 3–5 μm. Fig. 5c–h is the fused images of various methods. It can be seen that the person of Fig. 5c–f is darker than that of Fig. 5g and h. That is, the person of Fig. 5c–f is more easily confused with the background, whose intensity is lower.

The “Medical-2” images and fused images of various fusion algorithms are shown in Fig. 9.

### Table 5

<table>
<thead>
<tr>
<th>Fusion methods</th>
<th>VIF</th>
<th>MI</th>
<th>Q\textsubscript{PSNR}</th>
<th>Q\textsubscript{AV}</th>
<th>Q\textsubscript{S}</th>
<th>Q\textsubscript{E}</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWT</td>
<td>0.3099</td>
<td>2.4247</td>
<td>0.5578</td>
<td>0.6409</td>
<td>0.4525</td>
<td>0.5027</td>
</tr>
<tr>
<td>SWT</td>
<td>0.3772</td>
<td>2.4614</td>
<td>0.6305</td>
<td>0.6462</td>
<td>0.4701</td>
<td>0.5136</td>
</tr>
<tr>
<td>DTCWT</td>
<td>0.3960</td>
<td>2.4404</td>
<td>0.6209</td>
<td>0.6395</td>
<td>0.4707</td>
<td>0.4937</td>
</tr>
<tr>
<td>NSCT</td>
<td>0.3623</td>
<td>2.4666</td>
<td>0.6329</td>
<td>0.6535</td>
<td>0.4652</td>
<td>0.5366</td>
</tr>
<tr>
<td>MBF</td>
<td>0.3610</td>
<td>2.4678</td>
<td>0.6325</td>
<td>0.6907</td>
<td>0.4953</td>
<td>0.5454</td>
</tr>
<tr>
<td>MDBF</td>
<td>0.3753</td>
<td><strong>2.4721</strong></td>
<td><strong>0.6435</strong></td>
<td><strong>0.6926</strong></td>
<td><strong>0.5002</strong></td>
<td><strong>0.5458</strong></td>
</tr>
</tbody>
</table>

Fig. 9. The “Medical-2” images and fused images of various fusion algorithms.
low. Table 1 gives the numerical results of the objective evaluation, and the best results in terms of individual quality metric are labeled in bold. From this table, it can be found that our method is slightly inferior to the NSCT-based method for the metric $Q_{AB/F}$, and the proposed method outperforms all other methods for other metrics. On the whole, our method provides the best result for the “UN camp” images.

Next, the second experiment is conducted on the “Gun” images. Fig. 6a is the visible image clearly revealing the background, while Fig. 6b is the infrared image which displays a gun in the grass. Fig. 6c–h is the fused images of different methods. Similarly with Fig. 5, the guns in Fig. 6c–f are darker than those in Fig. 6g and h. The numerical results of the objective evaluation are shown in Table 2. From this table, it can be found that our method exceeds all other methods. For example, for the metric $Q_{AB/F}$, our method reaches 0.7103, while the second best result is 0.7066 obtained by the MBF-based method and the third best result is 0.6982 obtained by the NSCT-based method. The metric $Q_{EB/R}$ evaluates the amount of edge information. That is, compared with other methods, our method preserves better edges. This is because that the MDBF combines the characteristics of preserving edge with capturing directional information.

The third experiment is performed over “Tropical” images provided by University of Bristol and Waterfall Solutions Ltd. [41] in the ImageFusion.org website [40]. Fig. 7a, which presents clear background such as foliage, is a visible image. Fig. 7b, which shows a person occluded by foliage, is an infrared image. It is acquired by a Ratheon Thermal-eye 250D camera, whose range of spectral response is 7–14 μm. Fig. 7c–h is the fused images of various methods. It can be seen that all important features of the source images are integrated into the fused images. The contrast of Fig. 7g and h is slightly higher than that of Fig. 7c–f. Table 3 presents the objective evaluation of various methods. From this table, it can be seen that our method provides the best result for all the metrics. For example, for the metric $Q_{E}$ our method reaches 0.7425, while the second best result is 0.7362 and the third best result is only 0.7072. Since the metric $Q_{E}$ evaluates the amount of edge information, our fused image can preserve better edges.

### 4.4. Comparison of computation time

In this section, the comparison of computation time for various methods is presented. All experiments are performed on a PC with an AMD Athlon 2.2 GHz CPU and 1.75 GB RAM using Matlab 7.3. The elapsed time (s) of various methods and the size of the source images are presented in Table 6. From this table, it can be seen that the DWT-based method is the fastest, followed by the DTCWT, SWT, NSCT, MBF-based method and our method. Notice that some fast methods of bilateral filtering have been proposed [43,44]. Therefore, these methods can be used to speed up our method.

### 5. Conclusions

In this paper, we develop the MDBF through combining the multiscale bilateral filter with the nonsubsampled directional filter bank. The multiscale bilateral filter can smooth images while preserving edges and the nonsubsampled directional filter bank captures directional information. This combination enables the MDBF to possess the property of shift-invariance, multiscale and multidirection. Compared with the NSCT which is extensively used for various image processing applications, the multiscale transform of the MDBF is adaptive, which results in the characteristic of preserving edges.

To validate the effectiveness of the MDBF, the MDBF is applied to multisensor image fusion. Five experiments over visible and infrared images and medical images are performed. Six metrics of image fusion performance are used for the objective evaluation. These metrics reflect the quality of the fused images through evaluating edge preserving, contrast, the amount of information transferred, blur, etc. The experiments demonstrate that the proposed method presents superior fused images in term of quantitative evaluation metrics. Therefore, our method compared with traditional methods can better preserve edge and contrast and is more informative, which is consistent with subjective inspection.
However, the MDBF is more time-consuming than traditional methods because our method has many directions and is space-variant and nonsampled. Notice that some fast methods of bilateral filtering have been developed. So, a part of on-going work is extending the basic fast bilateral filter to the multiscale form to speed up the multiscale transform of the MDBF. In addition, another part of our future work is applying the MDBF to other fields of image processing.

Acknowledgements

The authors would like to thank the editor and anonymous reviewers for their detailed review, valuable comments and constructive suggestions. The authors also would like to thank the providers of the source images. This paper is supported by the National Natural Science Foundation of China (Nos. 60871096 and 60835004), the Ph.D. Programs Foundation of Ministry of Education of China (No. 200805320006), the Key Project of Chinese Ministry of Education (2009-120), and the Open Projects Program of National Laboratory of Pattern Recognition.

References

[23] T. Li, Y. Wang, Biological image fusion using a NSCT based variable-weight method, Information Fusion, in press. doi:10.1016/j.inffus.2010.03.007.