Feature-Based Image Fusion with a Uniform Discrete Curvelet Transform

Regular Paper

Liang Xu¹, Junping Du¹,*, Qian Hu¹ and Qingping Li¹

¹ Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia, School of Computer Science, Beijing University of Posts and Telecommunications, Beijing, China

* Corresponding author E-mail: junpingdu@126.com

Received 16 May 2012; Accepted 8 Mar 2013

DOI: 10.5772/56345

© 2013 Xu et al.; licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract The uniform discrete curvelet transform (UDCT) is a novel tool for multiscale representations with several desirable properties compared to previous representation methods. A novel algorithm based on UDCT is proposed for the fusion of multi-source images. A novel fusion rule for different subband coefficients obtained by UDCT decomposition is discussed in detail. Low-pass subband coefficients are merged to develop a fusion rule based on a feature similarity (FSIM) index. High-pass directional subband coefficients are merged for a fusion rule based on a complex coefficients feature similarity (CCFSIM) index. Experimental results demonstrate that the proposed algorithm fuses all of the useful information from source images without introducing artefacts. Compared with several state-of-the-art fusion methods, it yields a better performance and achieves higher efficiency.

Keywords Image Fusion, UDCT, FSIM Index, CCFSIM Index

1. Introduction

With the application of image sensors in many fields, multi-source image fusion techniques are increasingly important. Images of a scene can be captured using different sensors, times, angles and distances. These images may contain a large amount of different content that can provide complementary and redundant information. Image fusion approaches can transform all the important information from each source image into a fused image while eliminating superfluous data. The fused image can provide a better description of a scene than any of the individual source images [1, 2]. In many image-based application fields, image fusion is widely regarded as an important and promising research area. So far, image fusion has been successfully used in many real-world fields, such as defence surveillance, medical imaging, remote sensing and computer vision [3–5].

During the last decade, there has been much research into image fusion methods and numerous tools have been developed to solve different problems. These can be categorized into spatial domain and transform domain techniques [6]. However, fusion methods based on multiscale decomposition [7] in the transform domain are increasingly popular because of their better robustness and reliability. Pyramid-based [8–11] and discrete wavelet transform (DWT) [12–14] approaches are typically used in image fusion. Of these, DWT methods have some advantages, such as: temporal–frequency
localization, increased directional information and low redundancy [15, 16]. However, DWT approaches also have some drawbacks in practical applications. A 2D DWT is directly constructed as the tensor product of two 1D transforms, so it has only limited directions and is isotropic for each scale and location. In addition, DWT methods cannot effectively represent a signal that has features along smooth curves. To overcome these drawbacks of DWT in image analysis, a large number of new multiscale transforms have been developed in recent years. Examples include ridgelets [17], curvelets [18], contourlets [19] and NSCT [20]. Compared to traditional transforms, these are true 2D image representation tools with multiscale, multi-direction and anisotropy features.

The principle for selecting coefficients is another key step in image fusion. A variety of fusion strategies have been discussed in the literature and these can mainly be divided into three categories: pixel-based, window-based and area-based [21–23]. Window-based and area-based fusion rules make full use of the local characteristics of neighbourhood pixels and thus are superior to pixel-based rules [24].

In 2010, Truong and Chauris proposed a uniform discrete curvelet transform (UDCT), for which the forward and inverse transforms form a tight and self-dual frame [25]. This means that input images can be reconstructed perfectly. As a novel tool for multiscale representation, UDCT has higher approximation accuracy for geometric shapes and optimal sparsity. UDCT has several desirable properties for image analysis, such as a lower redundancy ratio, a hierarchical data structure and easy implementation. In addition, UDCT runs rapidly and fully satisfies image fusion processes in practice. Moreover, UDCT is shift-invariant in an energy sense for each complex band. Therefore, we applied UDCT to the field of multi-source image fusion for the first time.

The major contribution of this paper is the proposal of a novel fusion algorithm for multi-source images based on UDCT and a feature similarity (FSIM) index [26]. The input images are decomposed into subbands at different scales and directions using UDCT. Low-pass subband coefficients are merged to develop an FSIM-based fusion rule. The gradient magnitude component in the FSIM index is obtained by considering horizontal, vertical and two diagonal directions. In this way, the local features of an image can be better represented than when only horizontal and vertical directions are considered. The high-pass directional subband coefficients are merged to develop a CCFSIM-based fusion rule. Redundant and complementary regions can easily be distinguished according to FSIM and CCFSIM index values. A weighted average process is used for the redundant region and a selection process is applied for the complementary region [27]. The local energy is used as a saliency measure in the low-pass subbands. Feature magnitude (FM) is used as a saliency measure in the high-pass subbands. The proposed fusion rule improves the performance of fusion systems to yield better quality fused images.

The remainder of the paper is organized as follows. Section 2 reviews basic UDCT theory in brief. Section 3 describes the proposed image fusion algorithm in detail. Section 4 presents and discusses the experimental results. Section 5 concludes.

2. Uniform discrete curvelet transform

In this section, we briefly review UDCT theory and the properties [25] used in subsequent sections.

UDCT is a new version of the discrete curvelet transform that is based on multirate filter bank (FB) theory. UDCT is implemented in the Fourier domain and is designed as a multiresolution FB consisting of a set of discrete filters and decimation and up sampling blocks. This takes advantage of both an FFT-based discrete curvelet transform and an FB-based contourlet transform.

To illustrate the structure of multiscale UDCT decomposition, a three-level UDCT FB is displayed in Figure 1(a) for \( J=3 \) (\( 1 \leq j \leq 5 \)) scales and \( 2N_j=3\times2^j \) \( (n\geq0) \) directional bands for the \( j \)th scale. The number of directional subbands satisfies the parabolic scaling rule.

First, a set of 2N 2D directional filters \( F_j(\omega) , l=1, \ldots, 2N_j \) and a low-pass filter \( F_0(\omega) \) are constructed, for which the directional subbands and the low-pass subband can be decimated without aliasing. Since the directional filters have one-sided support in the frequency domain, they have complex subband coefficients. The 2D signal \( x(n) \) is filtered using \( F_j(\omega) \) and \( F_0(\omega) \). Second, the filtered signals are down sampled with three decimation ratios for the \( 2N \) directional bands and the low-pass band for the \( (2N+1) \)th band FB. \( D_1^{(N)} \) is the decimation ratio for the low-pass band. \( D_1^{(N)} \) and \( D_2^{(N)} \) are the decimation ratios for the first and last \( 3\times2^j \) directional bands, respectively. The three decimation ratios are:

\[
D_0^{(N)} = 2I, \quad D_1^{(N)} = \text{diag} \left( \frac{2N}{3} \right), \quad D_2^{(N)} = \text{diag} \left( \frac{2N}{3} \right). \quad (1)
\]

Finally, a multiscale UDCT is constructed by cascading the same FB at a lower band, i.e., the output of \( D_0^{(N)} \) in Figure 1(a).

However, in practical implementations, UDCT does not need to follow an iterative structure as in Figure 1(a). It can be implemented directly, as in Figure 1(b). All the curvelet functions are estimated at once according to the number of scales and directions of the transform. In
Figure 1(b), $\hat{F}_{j,l}(\omega)$ denotes the directional filters for scale $j$ and direction $l$ corresponding to $F_{j}(\omega)$. The equivalent low-pass filter is $\hat{F}_0(\omega)$, corresponding to $F_0(\omega)$. With the equivalent filters, the decimation ratios $D_{0,0}$, $D_{j,0}$ and $D_{j,1}$ are defined as:

\[
D_{0,0} = \text{diag}\left\{2^j, 2^j\right\}, \\
D_{j,0} = \text{diag}\left\{2, \frac{2N}{3}\right\} \cdot 2^{j-1}, \\
D_{j,1} = \text{diag}\left\{\frac{2N}{3}, 2\right\} \cdot 2^{j-1}.
\]

(2)

---

**Figure 1.** Structure of the multiscale UDCT. (a) Iterative multiple level of UDCT. (b) Equivalent FB.

The UDCT inherits the advantages of both curvelet and contourlet transforms. Moreover, compared to existing transforms, it has several additional properties such as a lower redundancy ratio, a hierarchical data structure, easy implementation and shift invariance for each complex band in the energy sense. The lower redundancy ratio of UDCT is very practical in industrial applications. A more detailed description is available elsewhere [21].

### 3. Feature-based image fusion with UDCT

In this section, a novel fusion algorithm based on UDCT and the FSIM index is discussed in detail.

Figure 2 illustrates the block diagram of the proposed image fusion algorithm. To simplify the discussion, we only consider a pair of source images (A and B) that are merged into a composite image (F). Again, it is assumed that the source images have been registered. The key idea in Figure 2 is that a pair of input images is decomposed into different subbands using UDCT and the FSIM index and CCFSIM index are then used to combine the subband coefficients. Finally, the fused image is reconstructed by applying the inverse UDCT to the merged coefficients. The proposed image fusion approach consists of the following steps:

**Step 1:** Input images $A$ and $B$ are decomposed into different scale and direction subbands using UDCT. The coefficients $\{C^A_{j}(x,y), C^B_{j}(x,y)\}$ are obtained, where $C^a(x,y)$ denotes the low-pass subband coefficients of the input images at the coarsest scale and $C_{j}(x,y)$ denotes the high-pass directional subband coefficients at the $j$th scale and in the $l$th direction.

**Step 2:** Different fusion rules are applied to merge the low-pass and high-pass subband coefficients. The low-pass subband coefficients are merged using an FSIM-based fusion rule. The high-pass directional subband coefficients are merged using a CCFSIM-based fusion rule. The corresponding subband coefficients of the fused image $\{F_{j}(x,y), \hat{F}_{j,l}(x,y)\}$ are then obtained.

**Step 3:** Apply the inverse UDCT to the fused coefficients $\{F_{j}(x,y), \hat{F}_{j,l}(x,y)\}$ and then obtain the fused image $F$.

As discussed in Section 1, besides multiscale transform methods, fusion rules are also key factors in image fusion schemes. Existing fusion rules have been described in detail elsewhere [7]. Considering the characteristics of the subband coefficients decomposed by UDCT, the FSIM index [26] is used as an additional tool to discriminate complementary and redundant regions between the source images. The FSIM index is a measure of feature similarity among images. Phase congruency (PC) and image gradient magnitude (GM) are two components of the FSIM index. As complementary components, PC and GM reflect different aspects of the human visual system.

First, convolution between 2D log-Gabor filters and the input image $f(x,y)$ yields a set of orthogonal vectors $[e_{n,o}(x,y), o_{n,o}(x,y)]$ for scale $n$ and orientation $o$. The local amplitude is defined as:

\[
A_{n,o} = \sqrt{e_{n,o}(x,y)^2 + o_{n,o}(x,y)^2}.
\]

(3)

The PC at position $(x,y)$ is defined as:

\[
\text{PC}(x,y) = \frac{\sum_n \left(\sum_o e_{n,o}(x,y)^2 + \sum_o o_{n,o}(x,y)^2\right)}{\varepsilon + \sum_n \sum_o A_{n,o}(x,y)},
\]

(4)

where $\varepsilon$ is a small positive constant and the value of PC lies between 0 and 1. The closer PC is to 1, the more salient the feature.
The image gradient can be computed using convolution masks. Sobel [28], Prewitt [28] and Scharr [29] are commonly used gradient operators. The Scharr operator can be used to obtain horizontal and vertical image gradients [26]. However, GM is obtained by considering horizontal, vertical and two diagonal directions in the present study. In this way, the local features of an image can be better represented than when only horizontal and vertical gradients are considered. Horizontal, vertical and diagonal Sobel operators are applied to the image and four directional gradients \( (G_x, G_y, G_{x+y} \text{ and } G_{x-y}) \) are obtained. The horizontal, vertical and diagonal Sobel operators are written as:

\[
\begin{bmatrix}
0 & 0 & 0 \\
1 & 2 & 1 \\
-1 & -2 & -1
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 1 \\
2 & 0 & 2 \\
-1 & 0 & 1
\end{bmatrix}
\]

The GM of input image \( f(x, y) \) is then defined as:

\[
G = \frac{1}{4} \left(G_x^2 + G_y^2 + G_{x+y}^2 + G_{x-y}^2 \right).
\]

The similarity \( S_t(x, y) \) for input signals \( f_1(x, y) \) and \( f_2(x, y) \) is defined as:

\[
S_t(x, y) = \frac{2PC_t(x, y)PC_s(x, y) + T_1}{PC_t^2(x, y) + PC_s^2(x, y) + T_2} \cdot \frac{2G_t(x, y)G_s(x, y) + T_3}{G_t^2(x, y) + G_s^2(x, y) + T_4}
\]

where \( T_1 \) and \( T_2 \) are positive constants and \( S_t(x, y) \) is a real number between 0 and 1.

The FSIM index between \( f_1(x, y) \) and \( f_2(x, y) \) is defined as:

\[
FSIM = \sum_{x,y} S_t(x, y) PC_{\omega}(x, y)
\]

where \( PC_{\omega}(x, y) = \max(PC(x, y), PC_s(x, y)) \) is used to weight the importance of \( S_t(x, y) \) in the overall similarity measure. \( \Omega \) denotes the image region.

The high-pass directional subband coefficients provide detail-rich information. They can effectively express salient features of images such as edges, lines and contours. The residual low-pass subband coefficients represent the main energy of source images and provide rich structural information. Here, the FSIM index is applied to the low-pass coefficients to distinguish complementary and redundant regions. According to the FSIM score, a weighting or selecting rule is used to merge coefficients. The high-pass subband coefficients of the UDCCT decomposition are complex. Accordingly, a CCFSIM index was developed by considering phase changes for the complex coefficients.

3.1 Fusion rule for low-pass subband coefficients

A fusion rule for the low-pass subband was developed based on the local region defined around centre point \((x, y)\). The size, \( M \times N \), is \( 3 \times 3 \) or \( 5 \times 5 \). Using (4) and (6), PC and GM maps are first obtained using a sliding window for the overall low-pass subband. The FSIM index between coefficients \( C_{\alpha}^A(x, y) \) and \( C_{\beta}^B(x, y) \) for the local region is then computed according to (8).

The FSIM index reflects the similarity of low-pass subband coefficients between input images. The FSIM value is used to distinguish redundant and complementary regions. A threshold \( T \) is defined between 0 and 1. Here, we take \( T=0.7 \). Regions with FSIM\(>T\) have high similarity and there is redundant information between coefficients \( C_{\alpha}^A(x, y) \) and \( C_{\beta}^B(x, y) \). A weighted method can preserve important information in the input images and decrease noise and redundant information. The low-pass subband coefficients of the fused image are defined as:

\[
F_{\beta}(x, y) = \omega_{\alpha}(x, y) \cdot C_{\alpha}^A(x, y) + \omega_{\beta}(x, y) \cdot C_{\beta}^B(x, y)
\]

where the weights \( \omega_{\alpha}(x, y) \) and \( \omega_{\beta}(x, y) \) depend on the local energy of the coefficients \( C_{\alpha}^A(x, y) \) and \( C_{\beta}^B(x, y) \). The low-pass coefficients reflect the coarsest image scale, which contains the main energy and provides abundant structural features. Consequently, the local energy can
effectively represent the saliency of the low-pass coefficients. The weights are defined as:

$$
\begin{align*}
\omega_A(x,y) &= \frac{E_A(x,y)+\varepsilon}{E_A(x,y)+E_B(x,y)+\varepsilon}, \\
\omega_B(x,y) &= 1 - \omega_A(x,y)
\end{align*}
$$

where $\varepsilon$ is a small positive constant used to avoid a denominator of zero. The local energy of the low-pass coefficients is defined as:

$$
E(x,y) = \sum_{x,y\in\Omega} w(x,y) |C(x,y)|^2
$$

where $w(x,y)$ is a Gaussian template of size $W\times N$ with a standard deviation of 0.5. The sum of the coefficients in the Gaussian template is 1, in order to enhance the robustness of the algorithm.

For regions with $FSIM_{<T}$, the low-pass subband coefficients $C_A^A(x,y)$ and $C_B^B(x,y)$ are significantly different and represent complementary information. A selection rule is applied to the coefficients and thus coefficients for the fused image are obtained from regions with greater salience between the source images. The salience criterion uses the local energy $E(x,y)$ in (11). The low-pass subband coefficients for the fused image are calculated as:

$$
F(x,y) = \begin{cases} 
C_A^A(x,y), & \text{for } E_A(x,y) > E_B(x,y) \\
C_B^B(x,y), & \text{for } E_A(x,y) < E_B(x,y)
\end{cases}
$$

**Remark 1:** The similarity threshold $T$ among regions was set to $T=0.7$ in a large number of studies because experimental results demonstrated that this threshold yields an optimal performance. We also used an adaptive approach to set a similarity threshold of $T=\max_{x,y}(FSIM(x,y))$, where 1-1 denotes the absolute value and constant $k$ is set to 0.7. However, experiments revealed that this adaptive threshold is inferior to the fixed value of $T=0.7$ in image fusion performance.

### 3.2 Fusion rule for high-pass subband coefficients

This section describes the fusion process for high-pass directional subband coefficients $C_A^A(x,y)$ and $C_B^B(x,y)$. The local region around the centre point $(x,y)$ is first defined for a size of $M\times N$. The high-pass directional subband coefficients $C_A^A(x,y)$ and $C_B^B(x,y)$ are complex numbers, so a CCFSIM index was developed based on the FSIM index. The idea was inspired by CW-SSIM [30], which considers how a phase change for complex coefficients impacts image features. Image features are mainly reflected by the relative phase patterns for the complex coefficients. Moreover, shifting of the phase of all coefficients by a constant value will not change the image features. Thus, the CCFSIM index is defined as:

$$
CCFSIM = \frac{\sum_{x,y\in\Omega} S_0(x,y) |PC(x,y) + k|}{\sum_{x,y\in\Omega} |PC(x,y)|} + \frac{\sum_{x,y\in\Omega} C_0^A(x,y) |C_0^A(x,y) + k|}{\sum_{x,y\in\Omega} |C_0^A(x,y)|},
$$

where $\Omega$ denotes a local region of size $M\times N$, $k$ is a small positive constant that improves the robustness of the CCFSIM index and $C_0^A(x,y)$ denotes the complex conjugate of $C_0^A(x,y)$. The value of the CCFSIM index lies within 0–1; a value close to 1 indicates strong similarity between $C_0^A(x,y)$ and $C_0^B(x,y)$.

When the high-pass coefficients for the jth scale and lth orientation are merged, the threshold is first defined as $T=0.7$. For regions with $CCFSIM_{<T}$, there is more shared information and more redundancy among source images, so a weighted method is selected. For regions with $CCFSIM_{<T}$, little information is shared and the source images are complementary. In this case a selection method is used to preserve detail information in the source images. The proposed fusion scheme is written as:

$$
F_M(x,y) = \begin{cases} 
C_A^A(x,y), & \text{for } CCFSIM_{<T} \text{ and } FM_A(x,y) > FM_B(x,y) \\
C_B^B(x,y), & \text{for } CCFSIM_{<T} \text{ and } FM_A(x,y) < FM_B(x,y) \\
\alpha_f(x,y) C_A^A(x,y) + \alpha_f(x,y) C_B^B(x,y), & \text{for } CCFSIM_{>T}
\end{cases}
$$

where $FM(x,y)$ is the feature magnitude of the region, defined as:

$$
FM(x,y) = \frac{\sum_{x,y\in\Omega} w(x,y) |C_j(x,y)|^2}{\sum_{x,y\in\Omega} w(x,y)}
$$

$PC(x,y)$ and $G(x,y)$ can be extracted from the PC and GM maps obtained when computing the CCFSIM index. $w(x,y)$ is a Gaussian template of size $W\times N$ with a standard deviation of 0.5. The sum of the coefficients in the Gaussian template is 1. $\varepsilon$ is a small positive constant; here, we set $\varepsilon=1$. In addition, $\alpha$ and $\beta$ are applied to adjust the relative importance of PC and GM. Here, we use $\alpha=1$ and $\beta=2$. FM represents local features and the amount of information contained in the image. FM can effectively represent salient features in the high-pass subbands of source images.

The weights $\alpha_f(x,y)$ and $\beta_f(x,y)$ depend on $FM_A(x,y)$ and $FM_B(x,y)$ and are defined as:

$$
\alpha_f(x,y) = \frac{FM_A(x,y)}{FM_A(x,y) + FM_B(x,y)}, \quad \beta_f(x,y) = 1 - \alpha_f(x,y).
$$

Finally, all the UDCT coefficients are merged and the inverse UDCT is applied to the coefficients of the fused image for reconstruction.

**Remark 2:** In (15), $\alpha$ and $\beta$ are used to tune the weights of PC and GM in the FM, $\alpha$ and $\beta$ are positive constants, and are commonly set to 1, 2 or 4. The constant $\varepsilon$ is used to avoid the emergence of zero and enhance the reliability of the algorithm.
4. Experiments and analysis

In this section, the proposed algorithm is tested on several sets of images. The results are compared with those for different fusion algorithms to validate the performance. For comparison, we use the discrete wavelet transform (DWT), contourlet transform (CNT), nonsampled contourlet transform (NSCT), shiftable complex directional pyramid transform (SCDPT) [31] and UDCT-simple. All of these use averaging and absolute maximum selection schemes for merging low- and high-pass subband coefficients, respectively.

Four sets of different image types were tested to evaluate the performance of the proposed algorithm: a set of out-of-focus images, a set of multimodal medical images, a set of images of navigation aids for helicopter pilots and a set of remote sensing images. The image data were evaluated using subjective visual inspection and objective assessment tools. The parameters for the different fusion algorithms are shown in Table 1.

| Method     | Scale decomposition level | Directions from coarser to finer scale | Scale filter | Directional filter | Local region size |
|------------|---------------------------|---------------------------------------|--------------|--------------------|-------------------|
| DWT        | 3                         | –                                     | db4          | –                  | –                 |
| CNT        | 3                         | 4, 8, 8                               | pkva         | pkva               | –                 |
| NSCT       | 3                         | 2, 4, 8                               | maxflat      | dmaxflat7          | –                 |
| SCDPT      | 3                         | 4, 8, 8                               | nalias       | pkva               | –                 |
| UDCT-simple| 3                         | 6, 12, 24                             | nalias       | nalias             | –                 |
| UDCT-FSIM  | 3                         | 6, 12, 24                             | nalias       | nalias             | 5x5               |

Table 1. Parameters for the different fusion methods

Figure 3. Out-of-focus source images (256 level, size 256x256) and fused images: (a) focus on the right-hand clock; (b) focus on the left-hand clock; and fused images using (c) DWT, (d) CNT, (e) NSCT, (f) SCDPT, (g) UDCT-simple and (h) UDCT-FSIM methods.

Figure 4. Magnified regions from the fused images in Figure 3(c)-(h) using (a) DWT, (b) CNT, (c) NSCT, (d) SCDPT, (e) UDCT-simple and (f) UDCT-FSIM methods.
4.1 Visual analysis

The first experiment was performed on a pair of out-of-focus clock images with perfect registration, as shown in Figure 3. Comparison of the source (Figure 3(a), (b)) and fused images (Figure 3(c)–(h)) shows that important information in the source images is well integrated. However, the images fused using DWT and CNT (Figure 3(c),(d)) are not clear enough and have lower contrast; artefacts were also introduced. The images fused using the other approaches (Figure 3(e)–(h)) are obviously clearer and have stronger contrast than the DWT and CNT results. The differences among the images in Figure 3(e)–(h) are very slight, so it is difficult to evaluate the image quality by direct visual inspection. To observe the image quality in more detail, one area in the images was magnified.

Figure 4(a)–(f) shows magnified images of the region marked by the boxes in Figure 3(c)–(h). The performance of the different fusion algorithms can be observed from these magnified images. The DWT-based fused image has edge bends and serious deformation (Figure 4(a)). The wavelet transform has limited directions and cannot characterize smooth curves. Thus, aliasing easily occurs and leads to image distortion. The CNT-based fused image in Figure 4(b) has a similar problem to that of Figure 4(a). The quality of the images fused using the NSCT, SCDPT and UDCT-simple methods is significantly better and the edges are smoother (Figure 4(c)–(e)). However, slight image distortion is still evident. NSCT, SCDPT and UDCT are multiscale and multi-directional tools for image representation. They have several desirable features, such as high approximation accuracy for geometric shapes, good sparsity representation and an optimal frequency response. Consequently, the fused images in Figure 4(c)–(e) have better visual quality than those in Figure 4(a),(b). However, the NSCT, SCDPT and UDCT-simple fusion schemes are pixel-based simple fusion rules that do not consider neighbourhood pixels. Thus, they are sensitive to noise and artefacts can easily be introduced. Compared to the other fused images, the UDCT-FSIM image in Figure 4(f) shows optimal quality, with the best visual effect and smoother and sharper edges. This comparison reveals that the UDCT-FSIM-based approach effectively determines complementary or redundant information between source images. It can preserve all the important information of the source images while avoiding artefacts. In addition, UDCT-FSIM has greater robustness. In conclusion, the proposed fusion algorithm has optimal performance.

Figures 5–7 show source images and images fused using different fusion algorithms for different applications. The visual effects for the image sets were the same as for Figure 3. Image fusion by the UDCT-FSIM based approach perfectly preserved useful information and the fused image is close to the source images. For the other fusion methods, loss of information and distortion are evident for the fused images. Thus, the proposed fusion algorithm yields better performance for both multifocus and multimodal images.

![Figure 5. Medical source images (256 level, size 256×256) and fused images: (a) CT image; (b) MR image; and fused images using (c) DWT, (d) CNT, (e) NSCT, (f) SCDPT, (g) UDCT-simple and (h) UDCT-FSIM methods.](image-url)
Figure 6. Source and fused images of navigation aids for helicopter pilots (256 level, size 256×256): (a) source low-light-television (LLTV) sensor image; (b) source thermal imaging forward-looking-infrared (FLIR) sensor image; and fused images using (c) DWT, (d) CNT, (e) NSCT, (f) SCDPT, (g) UDCT-simple and (h) UDCT-FSIM methods.

Figure 7. Remote sensing source images (256 level, size 256×256) and fused images: (a, b) input images; and fused images using (c) DWT, (d) CNT, (e) NSCT, (f) SCDPT, (g) UDCT-simple and (h) UDCT-FSIM methods.

4.2 Quantitative analysis

Visual analysis was used to evaluate four image sets, but this is very subjective. Observers may report different results for an image, depending on experience and perspective. Thus, visual assessment alone is not an accurate measure of algorithm performance. Therefore, objective quantitative analysis tools were also used to evaluate the performance of different fusion algorithms. Three metrics were used for evaluation: information entropy (IE), mutual information (MI) [32] and an objective image fusion performance measure (Q_{ABF}) [33]. IE quantifies the average information content in an image. MI indicates how much of the input information the fused image contains. Q_{ABF}, proposed by Xydeas and Petrović, reflects the preservation of input edge
information in the fused image. The larger the values for the three metrics, the better are the fusion results.

5. Conclusion

A novel image fusion algorithm based on UDCT is proposed. We applied UDCT, a novel tool for multiscale and multi-directional decomposition, to the field of multi-source image fusion for the first time and observed a considerable improvement in performance. Using the UDCT characteristics, coefficients are selected according to FSIM and a CCFSIM index for the low-pass and high-pass subbands. Depending on the FSIM and CCFSIM scores, complementary and redundant information between source images can be distinguished. According to the complementarity or redundancy, a weighting or selection rule is applied to merge the coefficients. The local energy is used as a saliency measure for the low-pass subbands. FM is used as a saliency measure for the high-pass subbands. Experiments confirmed that our algorithm yields encouraging performance in terms of both visual analysis and objective quality metrics.

Figure 8 shows the metric results for the different fusion methods for the images in Figures 3 and 5–7.

6. Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program) 2012CB821200 (2012CB821206), the National Natural Science Foundation of China (No. 91024001, No.61070142) and the Beijing Natural Science Foundation (No. 4111002).

7. References

[1] Pajares G, de la Cruz J M (2004) A Wavelet-based Image Fusion Tutorial. Patt. Recogn. 37(9): 1855-1872.
[2] De I, Chanda B (2006) A Simple and Efficient Algorithm for Multifocus Image Fusion Using Morphological Wavelets. Signal Process. 86: 924-936.
[3] Yang L, Guo B L, Ni W (2008) Multimodality Medical Image Fusion Based on Multiscale Geometric Analysis of Contourlet Transform. Neurocomputing, 72: 203-211.
[4] Yang B, Li S T (2010) Multifocus Image Fusion and Restoration with Sparse Representation. IEEE Trans. Instrum. Meas. 59(4): 884-892.
[5] Liu Z, Forsyth D S, Saizadeh M S, Fahr A (2008) A Data-fusion Scheme for Quantitative Image Analysis by Using Locally Weighted Regression and Dempster-Shafer theory. IEEE Trans. Instrum. Meas. 57(11): 2554-2560.
[6] Mitianoudis N, Stathaki T (2007) Pixel-based and Region-based Image Fusion Schemes Using ICA Bases. Inf. Fusion. 8(2): 131-142.
[7] Piella G (2003) A General Framework for Multiresolution Image Fusion: From Pixels to Regions. Inf. Fusion. 4(4): 259-280.
[8] Burt P J, Adelson E H (1983) The Laplacian Pyramid as a Compact Image Code. IEEE Trans. Commun. 31(4): 532-540.
[9] Toet A (1989) A Morphological Pyramidal Image Decomposition. Patt. Recogn. Lett. 9(4): 255-261.
[10] Burt P J (1992) A Gradient Pyramid Basis for Pattern Selective Image Fusion. In: Proceedings of the Society for Information Display Conference. pp. 467-470.
[11] Petrović V S, Xydeas C S (2004) Gradient-based Multiresolution Image Fusion. IEEE Trans. Image Process. 13(2): 228-237.
[12] Ranchin T, Wald L (1993) The Wavelet Transform for the Analysis of Remotely Sensed Images. Int. J. Remote Sens. 14(3): 615-619.
[13] Kovačević J, Vetterli M (1995) Nonseparable Two- and Three-dimensional Wavelets. IEEE Trans. Signal Process. 43(5): 1269-1273.
[14] Amolins K, Zhang Y, Dare P (2007) Wavelet Based Image Fusion Techniques – An Introduction, Review and Comparison. ISPRS J. Photogram. Remote Sens. 62(4): 249-263.
[15] Lewis J J, O’Callaghan R J, Nikolov S G, et al. (2007) Pixel- and Region-based Image Fusion with Complex Wavelets. Inf. Fusion. 8(2): 119-130.
[16] Mallat S G (1989) Multifrequency Channel Decomposition of Images and Wavelet Models. IEEE Trans. Acoust. Speech Signal Process. 37(12): 2091-2110.
[17] Do M N, Vetterli M (2003) The Finite Ridgelet Transform for Image Representation. IEEE Trans. Image Process. 12(1): 16-28.
[18] Candès E, Demanet L, Donoho D, Ying L (2006) Fast Discrete Curvelet Transforms. Multiscale Model. Simulation. 5(3): 861-899.
[19] Do M N, Vetterli M (2005) The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. IEEE Trans. Image Process. 14(12): 2091-2106.
[20] Cunha L D, Zhou J P (2006) The Nonsubsampled Contourlet Transform: Theory, Design, and Applications. IEEE Trans. Image Process. 15(10): 3089-3101.
[21] Burt P J, Kolczynski R J (1993) Enhanced Image Capture Through Fusion. In: Proceedings of the 4th International Conference on Computer Vision. pp. 173-182.
[22] Zhang Z, Blum R S (1997) Region-based Image Fusion Scheme for Concealed Weapon Detection. In: Proceedings of the 31st Annual Conference on Information Science and Systems. pp. 168-173.
[23] Li M, Cai W, Tan Z (2006) A Region-based Multi-sensor Image Fusion Scheme Using Pulse-coupled Neural Network. Patt. Recogn. Lett. 27(16): 1948-1956.
[24] Zhang Z, Blum R S (1999) A Categorization of Multiscale Decomposition-based Image Fusion Schemes with a Performance Study for a Digital Camera Application. Proc. IEEE. 87(8): 1315-1326.
[25] Nguyen T T, Chauris H (2010) Uniform Discrete Curvelet Transform. IEEE Trans. Signal Process. 58(7): 3618-3634.
[26] Zhang L, Zhang L, Mou X Q, Zhang D (2011) FSIM: A Feature Similarity Index for Image Quality Assessment. IEEE Trans. Image Process. 20(8): 2378-2386.
[27] Zhang Q, Wang L, Li H J, Ma Z K (2011) Similarity-based Multimodality Image Fusion with Shiftable Complex Directional Pyramid. Patt. Recogn. Lett. 32(13): 1544-1553.
[28] Jain R, Kasturi R, Schunck B G (1995) Machine Vision. McGraw-Hill. pp. 140-185.
[29] Jähne B, Haubecker H, Geibler P (1999) Handbook of Computer Vision and Applications. Academic Press. pp. 125-151.
[30] Sampat M P, Wang Z, Gupta S, Bovik A C, Markey M K (2009) Complex Wavelet Structural Similarity: A New Image Similarity Index. IEEE Trans. Image Process. 18(11): 2385-2401.
[31] Nguyen T T, Oriantara S (2008) The Shiftable Complex Directional Pyramid-Part I: Theoretical Aspects. IEEE Trans. Signal Process. 56(10): 4651-4660.
[32] Qu G H, Zhang D L, Yan P F (2002) Information Measure for Performance of Image Fusion. Electron. Lett. 38(7): 313-315.
[33] Xydeas C S, Petrović V (2000) Objective Image Fusion Performance Measure. Electron. Lett. 36(4): 308-309.