**Fusion of mis-registered GFP and phase contrast images with convolutional sparse representation and adaptive region energy rule**

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Microscopy Research and Technique</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>MRT-19-212</td>
</tr>
<tr>
<td>Wiley - Manuscript type:</td>
<td>Research Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>31-May-2019</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Qiu, Chenhui; Zhejiang University, College of Biomedical Engineering and Instrument Science  
Peng, Wenxian; Hangzhou Medical College, Radiology department  
Wang, Yuanyuan  
Hong, Jiang; Hangzhou Hospital Zhejiang Armed Police Corps  
Xia, Shunren; Zhejiang University College of Biomedical Engineering and Instrument Science, Department of Biomedical Engineering |
| Classifications: | image analysis < IMAGING, bioimaging < IMAGING |
| Keywords:       | Image fusion, convolutional sparse representation (CSR), adaptive region energy (ARE) rule, discussion mechanism-based brain storm optimization (DMBSO) |
Fusion of mis-registered GFP and phase contrast images with convolutional sparse representation and adaptive region energy rule

Chenhui Qiu\textsuperscript{1,2}, Wenxian Peng\textsuperscript{4}, Yuanyuan Wang\textsuperscript{1,2}, Jiang Hong\textsuperscript{5} and Shunren Xia\textsuperscript{1,2,3,*}

\textsuperscript{1}Key Laboratory of Biomedical Engineering of Ministry of Education, Zhejiang University, Hangzhou, China
\textsuperscript{2}Zhejiang Provincial Key Laboratory of Cardio-Cerebral Vascular Detection Technology and Medicinal Effectiveness Appraisal, Zhejiang University, Hangzhou, China
\textsuperscript{3}School of Information & Electrical Engineering, Zhejiang University City College, Hangzhou, China
\textsuperscript{4}Radiology department of Hangzhou Medical College, Hangzhou, China
\textsuperscript{5}Hangzhou Hospital Zhejiang Armed Police Corps, Hangzhou, China

Chenhui Qiu, Ph. D, email: qch19881114@126.com
Wenxian Peng, Ph. D, email: wenxianpeng@hmc.edu.cn
Yuanyuan Wang, Ph. D, email: yuanyuan_wang1988@163.com
Jiang Hong, Ph. D, email: huanghj1@163.com

\*Shunren Xia, Ph. D, prof., (corresponding author), email: shunren_xia@163.com
Corresponding address: Building of Zhouyiqing, College of Biomedical Engineering & Instrument Science, Zhejiang University, No. 38, Zheda Road, Hangzhou City, Zhejiang Province, China

The authors declare no conflict of interest.

The abstract word count: 210,
The text word count: 4592,
The number of pages: 15,
The number of tables: 3,
The number of figures: 6,
The short running title: Mis-registered GFP and Phase Contrast Image Fusion with CSR

Keywords
\begin{itemize}
\item Image fusion, convolutional sparse representation (CSR), adaptive region energy (ARE) rule, discussion mechanism-based brain storm optimization (DMBSO)
\end{itemize}
**Research Highlights**

- A novel image fusion method based on convolutional sparse representation (CSR) is proposed.
- An adaptive region energy (ARE) rule which is calculated by discussion mechanism based brain storm optimization (DMBSO) algorithm is proposed to fuse the coefficients of base layers.
- The sum modified Laplacian (SML) rule is explored to fuse the coefficients of detail layers.
- The method aims at fusing the mis-registered phase contrast and GFP images.

**Abstract**

Biomedical image fusion is the process of combining the information from different imaging modalities to get a synthetic image. Fusion of phase contrast and green fluorescent protein (GFP) images is significant to predict the role of unknown proteins, analyze the function of proteins, locate the subcellular structure, etc. Generally, the fusion performance largely depends on the registration of GFP and phase contrast images. However, high precise image registration is a very challenging task. Hence we propose a new fusion method based on convolutional sparse representation (CSR) to fuse the mis-registered GFP and phase contrast images. At first the GFP and phase contrast images are decomposed by CSR to get the coefficients of base layers and detail layers. Secondly the coefficients of detail layers are fused by the sum modified Laplacian (SML) rule while the coefficients of base layers are fused by the proposed adaptive region energy (ARE) rule. ARE rule is calculated by discussion mechanism based brain storm optimization (DMBSO) algorithm. Finally the fused image is achieved by carrying out the inverse CSR. The proposed fusion method is tested on one hundred of mis-registered GFP and phase contrast images. The experimental results reveal that our proposed fusion method exhibits better fusion results and superior robustness than several existing fusion methods.
1 INTRODUCTION

Green fluorescent protein (GFP) is a protein which is capable of exhibiting bright green fluorescence when exposed to the blue light in ultraviolet range. GFP acts as an excellent tool in biology because it is able to constitute the internal chromophore only with molecular oxygen. It has been expressed in lots of species, such as bacteria, yeasts, fungi, fish and mammals so far (Tsien, 1998). Green fluorescence imaging is a representative optical imaging technique and GFP gene is frequently utilized as a reporter of expression in molecular biology. The GFP image is of low spatial resolution. It mainly offers the functional information that is relevant to the molecular distribution in biological living cells (Majlesara et al., 2017).

Phase contrast microscopy is an important optical technique in biological science. It can reveal cellular structure which is invisible with simple bright field microscope. It can convert the phase shifts in light passing through a transparent specimen to the changes of the image brightness. Although phase shifts are invisible, they can be observed by vision when they are displayed as brightness variations. Phase contrast imaging is an advanced optical microscopy technique which can help the biologists to study living cells and cells’ proliferating (David, Nohammer, Solak, & Ziegler, 2002). The structural information of a transparent specimen without the stain is provided in phase contrast image that has high spatial resolution (Koroleva, Tomlinson, Leader, Shaw, & Doonan, 2005).

As shown in Fig. 1 (a) and (b), there is one typical pair of GFP and phase contrast images of arabidopsis thaliana cell. Obviously, the difference of the two kinds of biomedical images is clear. GFP image provides a great deal of functional information and a spot of structural information. Phase contrast image mainly provides structural information but no functional information. Generally, GFP image like Fig. 1 (a) is not accurately registered with Fig. 1 (b) while Fig. 1 (c) is the accurately registered GFP image corresponding to Fig. 1 (b). In order to exhibit the difference between Fig. 1 (a) and Fig. 1 (c), we implement the subtractive operation to obtain Fig. 1 (d) which is the residual image of Fig. 1 (a) and (c).

The type of genetically incorporated fluorescent proteins and the phenotypic noise strength which is highly correlated to fluorescence signal both have influence on the spatial resolution, feature distribution and the target location of the GFP image. Besides, edge illumination setup and the resample operation of reconstruction algorithms affect the texture and corner structures in the phase contrast image (Stockmar, Feddersen, Cramer, Gruber, Jung, Bramkamp, & Shin, 2018; Overkamp, Beilharz, Weme, Solopova, Karsens, Kovacs, Kok, Kuipers, & Veening, 2013; Zhang, Pan, Shang, & Li, 2018; Zabler, 2018). These factors have impact on the misaligned problem of GFP and phase contrast images. Registration of GFP and phase contrast images is one kind of multi-modal image registration subjects. Its performance also depends on pixel-level or feature-level image registration algorithms. The accurate registration of GFP and phase contrast images is a difficult and challenging task (Sotiras, Davatzikos, & Paragios, 2013; Li, Kang, Fang, Hu, & Yin, 2017).

Image fusion which combines the different but complementary information to get a synthetic and informative image is widely used for better visual perception and further image processing such as segmentation, delineation and recognition (Goshtasby & Nikolov, 2007; Du, Li, Lu, & Xiao, 2016; James & Dasarathy, 2014). Fusion of GFP and phase contrast images is a good solution for predicting the role of unknown proteins, analyzing the function of proteins, precisely locating the subcellular structure (Borgeaud, Metzger, Scrignari, & Blokesch, 2015; Koroleva,
Multi-scale transform (MST) theories which are powerful tools of image representation are widely used in various image fusion tasks. Classical MST tools are originally proposed to represent one-dimensional signal, hence even their improved versions cannot capture edges, curves and textures of images very well. Multi-scale geometric analysis (MGA) tools are current MST ones. Curvelet transform (CVT), sharp frequency localization contourlet transform (SFL-CT), nonsubsampled contourlet transform (NSCT) and complex shearlet transform (CST) are the representative and popular MGA tools. They can extract the geometric structure information much more effectively because they are the “true” two-dimensional transforms (Li, Yang & Hu, 2011; Li, Kang, Fang, Hu, & Yin, 2017). The principles and characteristics of them have been expounded in our previous article (Qiu, Wang, Guo, & Xia, 2018). These MGA tools with proper fusion rules are extensively used in multi-modal image fusion subjects (Nencini, Garzelli, Baronti, & Alparone, 2007; Mehta & Marakarkandy, 2013; Feng, Wang, Wei, & Mi, 2013; Wang, Peng, Feng, He, Wu, & Yan, 2013; Ganasala & Kumar, 2014; Yin, Liu, Zhao, Yin, & Guo, 2014; Qiu, Wang, Guo, & Xia, 2018).

Sparse representation (SR) theory especially based on dictionary learning is another powerful tool of multi-modal image fusion research. SR theory is first applied by Yang and Li to solving image fusion problem, they exploit orthogonal matching pursuit (OMP) algorithm and sliding window technique in their fusion method (Yang & Li, 2010). They also explore the simultaneous orthogonal matching pursuit (SOMP) algorithm for sparse coding to improve the performance of fusion results (Yang & Li, 2012). A general framework for multi-modal and multi-focus image fusion that associates SR theory with MST tools is explored and discussed in detail by Liu et al. The advantages of SR and MST are retained and their defects are avoided (Liu, Liu, & Wang, 2015). Lately convolutional sparse representation (CSR) is first introduced by Liu et al. to fuse multi-focus and multi-modal images (Liu, Chen, Ward, & Wang, 2016). CSR is viewed as a new research direction for image fusion subject (Zhang, Liu, Blum, Han, & Tao, 2018).

For fusing phase contrast and GFP images, traditional stationary wavelet transform and auxiliary variable-transparency strategy are used by Li & Wang (2009). A fusion method for phase contrast and GFP images with SFL-CT is proposed by Feng et al. Maximum region energy rule, absolute-value-maximum rule and neighborhood consistency measurement rule are used in their method (Feng, Wang, Wei, & Mi, 2013). Qiu et al. propose a fusion approach that aims at fusing phase contrast and GFP images by exploiting CST and Haar wavelet based energy rule (Qiu, Wang, Guo, & Xia, 2018).

To a great extent, all of the existing fusion methods depend on high precise image registration. And this precondition has great influence on fusion performance. But it is a very challenging task (Du, Li, Lu, & Xiao, 2016; Goshtasby & Nikolov, 2007; James & Dasarathy, 2014; Sotiras, Davatzikos, & Paragios, 2013; Li, Kang, Fang, Hu, & Yin, 2017; Li, Yang & Hu, 2011). Liu et al. claim that the CSR based fusion method is robust to mis-registration, but they have not validated it in their paper (Liu, Chen, Ward, & Wang, 2016).

In this paper, we propose a novel and robust fusion method which is based on CSR and specially designed fusion rules to fuse the mis-registered phase contrast and GFP images. CSR is an ideal image representation tool due to its shift-invariant property and its capability of coding an entire image rather than local image patches (Liu, Chen, Ward, & Wang, 2016). What’s more, fusion rules are also vitally important to the quality of fusion results (Kang, Fang, Hu, & Yin,
In our fusion method, we propose an adaptive region energy (ARE) rule which is calculated by discussion mechanism-based brain storm optimization (DMBSO) algorithm to merge the coefficients of base layers of CSR decomposition while the sum modified Laplacian (SML) rule is explored to merge the coefficients of detail layers of CSR decomposition. The experimental results indicate that the proposed fusion method achieves superior fusion results in terms of both subjective and objective assessments. Most importantly, the experimental results validate that our proposed method has satisfactory robustness to the inaccurate registration of phase contrast and GFP images.

The organization of remaining selections of this paper consists of subsequent three parts. CSR theory, DMBSO algorithm, proposed fusion rules and the framework of our fusion method are reported in Section 2. Afterward, Section 3 presents the experimental setting and fusion results. Finally the comparison and analysis of experimental results and the conclusion of this article are deployed in Section 4.

2 MATERIALS AND METHODS

2.1 CSR Theory

CSR also known as convolutional sparse coding (CSC) is studied under both signal processing community and machine learning community. It is the basic model of deconvolutional networks for signal or image feature extraction (Zeiler, Krishnan, Taylor, & Fergus, 2010; Liu, Chen, Ward, & Wang, 2016; Liu, Chen, Wang, Wang, Ward, & Wang, 2018). CSR models an entire image as the sum over a set of convolutions of coefficient maps and their corresponding dictionary filters, i.e.

$$\arg \min_{\{x_m\}} \frac{1}{2} \left\| \sum_m d_m \ast x_m - s \right\|^2_2 + \lambda \sum_m \|x_m\|_1$$

(1)

where $s$ is the image for sparse coding, $\{d_m\}$ is a set of $M$ dictionary filters, $\{x_m\}$ is a set of coefficient maps, $\lambda$ is the regularization parameter, and $\ast$ donotes convolutional operator. $s$ and each of the $\{x_m\}$ are considered to be $N$ dimensional vectors, where $N$ is the number of pixels in an image. This type of representation can be efficiently addressed by alternating direction method of multipliers (ADMM). Rewriting Eq. (1) in an appropriate form by introducing an auxiliary variable $y_m$ with a constraint, we obtain

$$\arg \min_{\{x_m\},\{y_m\}} \frac{1}{2} \left\| \sum_m d_m \ast x_m - s \right\|^2_2 + \lambda \sum_m \|x_m\|_1$$

$$s.t. \quad x_m - y_m = 0$$

(2)

Then the ADMM iterations for Eq. (2) are

$$\{x_m^{(j+1)}\} = \arg \min_{\{x_m\}} \frac{1}{2} \left\| \sum_m d_m \ast x_m - s \right\|^2_2$$

$$+ \frac{\rho}{2} \sum_m \|x_m - y_m^{(j)} + u_m^{(j)}\|^2_2$$

(3)
\[
\{y_m^{(j+1)}\} = \arg \min_{\{y_m\}} \lambda \sum_m \|y_m\|^2 \\
+ \frac{\rho}{2} \sum_m \|x_m^{(j+1)} - y_m + u_m^{(j)}\|_2^2
\]

(4)

\[
u_m^{(j+1)} = u_m^{(j)} + x_m^{(j+1)} - y_m^{(j+1)}
\]

(5)

where \(\rho\) is the step size which is equal to the augmented Lagrangian parameter (Wohlberg, 2016). The stopping criteria of ADMM and more details are discussed by Boyd, Parikh, Chu, Peleato, & Eckstein (2010).

2.2 DMBSO algorithm

Human being is the most intelligent organism in the world and the brainstorming process is widely utilized by people to create great ideas for solving problems. BSO introduces the idea of brainstorming process of human being to the design of optimization algorithm. Generally it executes the grouping, replacing and creating operators in order to approximate the global optimum of specific problem generation by generation (Cheng, Qin, Chen, & Shi, 2016; Shi, 2011). The steps of brainstorming process and the procedure of BSO algorithm are described in detail by Shi. (2011).

The new individual generation is described as follows.

\[
X_{new}^d = X_{selected}^d + \xi \cdot N(\mu, \sigma)
\]

(6)

where \(X_{new}^d\) is the \(d^{th}\) dimension of the individual which is newly generated, \(X_{selected}^d\) is the \(d^{th}\) dimension of the individual which is chosen to produce new individual, \(N(\mu, \sigma)\) is the Gaussian random function with mean \(\mu\) and variance \(\sigma\). \(\xi\) is a parameter that measures the contribution of the Gaussian random value and is calculated as follows.

\[
\xi = \logsig((0.5 \cdot N_{max\_gen} - N_{cur\_gen}) / k) \cdot \text{rand}( )
\]

(7)

where \(\logsig( )\) is a logarithmic sigmoid transfer function, \(k\) is used to adjust the slope of \(\logsig( )\) function, \(\text{rand}( )\) is a random value within (0,1). \(N_{max\_gen}\) and \(N_{cur\_gen}\) are the maximum number and current number of iterations respectively (Shi, 2011).

Although BSO algorithm proposed by Shi (2011) is a promising and effective for optimizing search, it encounters premature convergence and falls into the local optimum in some searching problem. DMBSO algorithm introduces the discussion mechanism which uses separate inter-group and intra-group discussions into the individual updating of BSO algorithm to address the defects. Specifically, the global searching ability is enhanced by linearly decreasing times of inter-cluster discussion at the initial phase, while the local searching ability is strengthened by linearly increasing times of intra-cluster discussion at the later phase (Yang, Shi, & Xia, 2013). They are expressed as Eq. (8) and (9) respectively.
\[ N_{\text{times\_inter\_cluster}} = N_{\text{max\_time}} \cdot (1 - \frac{N_{\text{cur\_gen}}}{N_{\text{max\_gen}}}) \]  
\[ N_{\text{times\_intra\_cluster}} = N_{\text{max\_time}} \cdot (\frac{N_{\text{cur\_gen}}}{N_{\text{max\_gen}}}) \]

where \( N_{\text{times\_inter\_cluster}} \) denotes the times of current inter-group discussion and \( N_{\text{times\_intra\_cluster}} \) denotes the times of current intra-group discussion, \( N_{\text{cur\_gen}} \) and \( N_{\text{max\_gen}} \) are the numbers of current and maximal individual updating separately. \( N_{\text{max\_time}} \) is set as the maximal time of cluster discussions.

2.3 Fusion rule

The detail layers and base layers of the input images can be achieved by two-scale decomposition with CSR mentioned above. Then the coefficients of detail layers and base layers should be merged with appropriate fusion rule, which also plays vitally important role in the performance of fusion results.

2.3.1 ARE rule

The general fusion rule for low frequency subbands or base layers is the point-based rule or stationary region-based rule (Li, Yang & Hu, 2011; Li, Kang, Fang, Hu, & Yin, 2017). These rules largely rely on high precise image registration and are not suitable to the mis-registered GFP and phase contrast images. In order to obtain excellent fusion performance of this research, we propose ARE rule which is calculated by DMBSO algorithm to merge the coefficients of base layers. Fig. 2 visualizes the base layer of each subfigure in Fig. 1. There is slight difference between the base layers of mis-registered GFP and registered GFP images which is vividly exhibited in Fig. 2 (4). Actually the difference of different GFP or phase contrast images is also different. If an adaptive region rule can cover the difference and select the important coefficients obeying pre-designed objective function, the mis-registration occurred in spatial domain can be overcome in the base layers. Since there are a great deal of energy and some salient features in the base layers, ARE rule which retains the energy effectively and adjusts the region adaptively is utilized to fuse the base layers.

The base layers of GFP image, phase contrast image and the fused image are denoted as \( B_G \), \( B_P \) and \( B_F \) respectively. In this paper, the region energy of a local region of \( B_G \) or \( B_P \) is defined as

\[ E^X(i,j) = \sum_{(s,t) \in \Omega(i,j)} \left[ \sqrt{X(s,t) - u^X(i,j)} \right]^2, \quad X = G \text{ or } P \]  

where \( \Omega(i,j) \) represents an adjustable local region with the size of \( m \times n \) whose center is located at the position \( (i,j) \). \( r^X(s,t) \) is the coefficient within the region \( \Omega(i,j) \) of \( B_G \) or \( B_P \). \( u(i,j) \) is the mean value of the coefficients within the region \( \Omega(i,j) \).
Essentially \( B_F \) is viewed as a matrix composed of the fused coefficients \( r^F(i, j) \) which can be computed as follows.

\[
r^F(i, j) = \begin{cases} 
  r^G(i, j), & E^G(i, j) \geq E^P(i, j) \\
  r^P(i, j), & \text{otherwise}
\end{cases}
\]  

(11)

As \( B_G \), \( B_P \) and \( B_F \) have the same size of the input images which is large enough, they contain a great deal of energy of the corresponding images. Besides, the region with fixed size is not appropriate since the GFP and phase contrast images are not registered precisely. In this paper, the size of \( \Omega(i, j) \) that can be adjusted according to \( B_G \) and \( B_P \) is calculated as follows.

\[
\arg \max_{(m,n)} SF(m,n,i,j)
\]

s.t. \( m \in [3, M], n \in [3, N] \), and \( m, n \) is odd numbers  

(12)

where the objective function \( SF(m,n,i,j) \) is denoted as the spatial frequency of \( B_F \) which is formulated as follows.

\[
SF(m,n,i,j) = \sqrt{F_x^2 + F_y^2}
\]  

(13)

where \( F_x \) and \( F_y \) are the horizontal and vertical spatial frequencies respectively and expressed as follows (Huang, & Jing, 2007).

\[
F_x = \sqrt{\frac{1}{M \times N} \sum_{i=2}^{M} \sum_{j=2}^{N} [r^F(i, j) - r^F(i, j-1)]^2}
\]  

(14)

\[
F_y = \sqrt{\frac{1}{M \times N} \sum_{i=2}^{M} \sum_{j=2}^{N} [r^F(i, j) - r^F(i-1, j)]^2}
\]  

(15)

where \( M \times N \) is the size of \( B_F \) which is also the size of input images.

Eq. (12) can be viewed as an optimizing search problem and efficiently solved by DMBSO algorithm.

2.3.2 SML rule

The detail layers of mis-registered GFP and phase contrast images are acquired by the decomposition with CSR. They are similar to the high frequency subbands of multi-scale transform. SML rule is explored in this paper to merge the coefficients of detail layers.

There is one base layer and a series of detail layers obtained when an image is decomposed by CSR. The coefficients in each detail layer are very sparse ones. Thus the mis-registration occurred in spatial domain has little influence on fusing the coefficients of detail layers when a region-based rule is explored. SML rule with fixed window size proposed by Huang et al. (2007) is suitable to merge the coefficients of detail layers in this research.

The detail layers of GFP image, phase contrast image and the fused image are denoted as \( D_G \).
Dp and Df respectively. The coefficients of DG, DP and DF are denoted as hG(i, j), hP(i, j) and hF(i, j) respectively. The modified Laplacian (ML) operator is expressed as follows.

\[
ML^X(i, j) = \left| 2h^X(i, j) - h^X(i - 1, j) - h^X(i + 1, j) \right| + \left| 2h^X(i, j) - h^X(i, j - 1) - h^X(i, j + 1) \right| \quad X = G \text{ or } P
\]  

Then the SML operator is

\[
SML^X(i, j) = \sum_{(i,j) \in W} ML^X(i, j)
\]  

where \( W \) is a local region or window of DG and DP. It is worthwhile to note that CSR is shift-invariant for the entire image representation. Besides, \( h^G(i, j), h^P(i, j) \) and \( h^F(i, j) \) are sparse coefficients of detail layers and indicate geometric structure information of input images. Thus, the size of \( W \) almost has no influence on the fusion performance. It is set to \( 4 \times 4 \) (Huang, & Jing, 2007).

Finally, the SML rule which based on SML operator is obtained as follows (Qu, Yan, & Yang, 2009).

\[
h^F(i, j) = \begin{cases} 
  h^G(i, j), & SML^G(i, j) \geq SML^P(i, j) \\
  h^P(i, j), & \text{otherwise} 
\end{cases}
\]

### 2.4 Framework of the proposed fusion method

As shown in our previous article (Qiu, Wang, Guo, & Xia, 2018), the GFP image should be converted from RGB model to IHS model before the fusion procedure. Firstly the phase contrast image and the intensity component of GFP image are both decomposed by CSR. Secondly the coefficients of base layers and detail layers are merged with ARE rule and SML rule respectively. Then the fused intensity component is achieved by performing the inverse CSR on both fused base layers and detail layers. The schematic diagram of the proposed fusion procedure is shown in Fig. 3. Finally the fusion result is obtained by conducting the IHS-to-RGB conversion.

### 3 RESULTS

#### 3.1 Experimental setting

The GFP and phase contrast images used in our experiments have the same size \( 256 \times 256 \), and they are not accurately registered in pairs. All the experiments are programmed in MATLAB 7.14 and operated on a PC with 4.00-GHz CPU, 32-GB RAM and 64-bit OS.

In order to evaluate the performance of our proposed method, CST-Haar (Qiu, Wang, Guo, & Xia, 2018), SFL-CT (Feng, Wang, & Wei et al., 2013), NSCT-SR (Liu, Liu & Wang, 2015), SOMP (Yang & Li, 2012) and CSR (Liu, Chen & Ward et al, 2016) based fusion methods are also carried out for comparisons.

The absolute-value-maximum rule is used to merge high frequency subbands and the Haar
wavelet energy-based rule is used to merge low frequency subbands in the CST-Haar based method (Qiu, Wang, Guo, & Xia, 2018). For SFL-CT based method, maximum region energy rule is utilized to merge the coefficients of the approximate subbands, absolute-value-maximum rule is exploited to fuse the coefficients of the finest detailed subbands, and neighborhood consistency measurement rule is designed to fuse the coefficients of other detailed subbands (Feng, Wang, Wei, & Mi, 2013). NSCT-SR based method uses absolute-value-maximum rule to merge the high frequency subbands and designs a SR based approach with offline trained dictionary to merge the low frequency subbands (Liu, Liu, & Wang, 2015). SOMP based method employs norm-1 maximum rule to fuse the sparse coefficients of multi-modal images (Yang & Li, 2012). Because phase contrast and GFP images are achieved by different imaging modalities, the maximum rule is selected to fuse the coefficients of detail layers while the weighted averaging rule is exploited to fuse the coefficients of base layers (Liu, Chen, Ward, & Wang, 2016). The pyramidal filter is set as ‘pyrex’ and the directional filter is set as ‘7-9’ respectively in NSCT-SR based method (Li, Yang & Hu, 2011; Liu, Liu, & Wang, 2015). The decomposition levels are all set to 4 while the directional numbers form coarser to finer scales are set as $2^1$, $2^2$, $2^3$ and $2^4$ in both SFL-CT and NSCT-SR based methods (Li, Yang & Hu, 2011).

Subjective assessments and objective evaluations should be utilized to measure the performances of image fusion results simultaneously. $Q_0$ (Wang & Bovik, 2002), $Q_{psr}$ (Piella & Heijmans, 2003), $Q_E$ (Piella & Heijmans, 2003), $Q^{AB/F}$ (Xydeas & Petrovic, 2000), sum of the correlations of differences (SCD) (Aslantas & Bendes, 2015) and visual information fidelity for fusion (VIFF) (Han, Cai, Cao, & Xu, 2013) are exploited to measure the fusion results objectively in our experiments. Qiu et al. have explained these objective evaluation metrics in detail (Qiu, Wang, Guo, & Xia, 2018).

### 3.2 Experimental results

Fig. 4 shows the fusion results of Fig.1 (a) and (b) acquired by different fusion methods mentioned above. The yellow square block at the top-left part of each sub-figure is the enlarged view of the image that is located in the lower-right yellow square block. Clearly, Fig. 4(b) and (c) contain some obvious artifacts and they are inclined to preserve a large amount of redundant information from the GFP image. And Fig.4 (b) is much better than Fig.4 (c) in this aspect. The brightness of Fig. 4 (a) tends to dark as a whole. It implies that the brightness and contrast of Fig. 4 (a) are distinctly reduced. Fig. 4 (d) and (e) relatively preserve the information from input images in acceptable sense. However, the edge in Fig. 4 (d) is blurred if we observe the enlarged yellow square block carefully. Besides, there are some mismatches with different degrees in Fig. 4 (a)-(e). The mismatches are distinct in Fig. 4 (b), (d) and (e). Fig. 4 (f) obtained by proposed fusion method acquires the satisfactory fusion performance especially compared with other sub-figures. It is noteworthy that there is little mismatch in Fig. 4 (f).

The objective evaluation metrics of different fusion results of Fig.1 (a) and (b) is shown in Table 1. The larger value of each metric mentioned above indicates better fusion performance. The largest value for each metric is labeled in bold. The second largest value is labeled with underline. The proposed method acquires the best results for four metrics and the second best result for one metric because it utilizes CSR to decompose the input images and uses ARE rule.
and SML rule to merge the corresponding coefficients.

As shown in Fig. 5, there is another example of phase contrast and GFP images of arabidopsis thaliana cell. Fig. 5 (a) and (b) are one pair of mis-registered images while Fig. 5 (c) is the accurately registered GFP image corresponding to Fig. 5 (b). Fig. 5 (d) is the residual image of Fig. 5 (a) and (c).

The fusion results of Fig. 5 (a) and (b) acquired by different fusion methods are exhibited in Fig. 6. The visual analysis of Fig. 6 is very similar to that of Fig. 4. We can find that Fig. 6 (f) exhibits the better fusion result than Fig. 6 (a)-(e). It can be observed distinctly by the enlarged yellow square block at bottom-right.

The objective evaluation metrics of different fusion results of Fig. 5 (a) and (b) is shown in Table 2. Our proposed approach acquires the best results for three metrics and the second best result for one metric.

CSR decomposition obtains the sparse coefficients of the entire image. It is a shift-invariant tool for sparse coding and less sensitive to mis-registration. So the mismatch in Fig. 4 (a) is less serious than that in Fig. 4 (b)-(e). Similarly, the mismatch in Fig. 6 (a) is less distinct than that in Fig. 6 (b)-(e). ARE rule measures the activity-level and region energy of the coefficients of base layers of mis-registered GFP and phase contrast images decomposed by CSR, and merges them. SML rule highlights discontinuities and larger values of coefficients, deemphasizes the slowly varying coefficients of detail layers of mis-registered GFP and phase contrast images decomposed by CSR, and merges them. To a great extent, the two rules combining with CSR can preserve valuable and salient information of input images and overcome the mis-registration.

The objective evaluation metrics of the fusion results of one hundred pairs of mis-registered phase contrast and GFP images among different fusion methods are depicted in Table 3. The means and the standard deviations (SD) of different metrics are used to statistically measure the whole performances of these methods. Obviously, our proposed method acquires the best results for four metrics and the second best result for one metric. Moreover, the subjective observations of these fused images also indicate that the proposed method exhibits satisfactory visual performances including the robustness to inaccurate registration.

4 DISCUSSION

4.1 Comparison and analysis

In this paper, a novel fusion method is proposed to fuse mis-registered phase contrast and GFP images based on CSR decomposition. The base layers have plenty of energy and the sizes of them are as large as the input images. Besides, the phase contrast and GFP images are not accurately registered. So ARE rule which can choose the salient coefficients in adaptive local region and be insensitive to mis-registration is proposed in this paper to more effectively merge the base layers. The detail layers are merged by SML rule which is an effective region based fusion rule. As shown in Fig. 4 and Fig. 6, the proposed method exhibits the best fusion results in visual sense. From Table 1 to Table 3, we can find that the excellent objective evaluation metrics is achieved by the proposed method.

NSCT and SFL-CT are the modified version of contourlet transform (CT) while CST is the complex-valued extended visions of shearlet transform (ST). CT has the defect of shift-variance. NSCT has shift-invariant property, but it is the simplified version of CT essentially. Thus, its ability of image representation is weaker than CT’s (Cunha & Zhou, 2006). SFL-CT is developed
to solve the problem of the frequency non-localization of CT, but it lacks shift-invariance (Qu, Yan & Yang, 2009). CST possesses shift-invariant property only in its magnitude response (Reisenhofer, 2014). As two-dimensional signal representation tool for fusing mis-registered images, the ability of extracting geometric structure information and shift-invariant property are both important factors. Hence, the MGA based fusion methods with general fusion rules mentioned above cannot deal with the mis-registered condition.

The sliding window technique which divides input images into patches is exploited in the SOMP and NSCT-SR based methods (Yang & Li, 2012; Liu, Liu, & Wang, 2015). Moreover, these methods exploit OMP or SOMP algorithm to get the very sparse coefficients (Tropp, 2004; Tropp, Gilbert, & Strauss, 2006). So they are much easy to preserve redundant information and produce distinct artifacts in multi-modal image fusion. Besides, since they are patch-wise methods, they are not robust to mis-registration (Liu, Chen, Ward, & Wang, 2016). VIFF metric is also calculated by sliding window technique and it measures the quantity of visual information (Han, Cai, Cao, & Xu, 2013), so the SOMP and NSCT-SR based methods perform very well in terms of VIFF metric.

CSR based fusion method obtains the sparse coefficients of the entire image rather than the image patches. It is the sparsest representation tool with respect to the entire image (Liu, Chen, Ward, & Wang, 2016; Wohlberg, 2016). However, Liu et al. uses general fusion rules to fuse base layers and detail layers (Liu, Chen, Ward, & Wang, 2016). Though their method obtains good fusion performances in pre-registered multi-focus images, infrared and visible images and CT and MR images, it cannot obtain better fusion results and superior robustness than the proposed method in this fusion subject. Actually it also verifies the significance of the ARE and SML fusion rules we have explored.

4.2 Conclusion and future work

Our proposed fusion method decomposes the input images by CSR. Moreover, ARE rule which can adaptively adjust the region size by DMBSO algorithm is proposed to merge the coefficients in base layers of CSR. SML rule is also explored to merge the coefficients in detail layers of CSR. The experiments verify that our proposed method acquires the excellent fusion results in fusing mis-registered phase contrast and GFP images.

Because GFP image is of low spatial resolution while phase contrast image is of high spatial resolution, the registration of phase contrast and GFP images is very challenging task and the accurate registration is hard to implement. Thus, our proposed method is much more practical and effective as it do not rely on precise registration. However, this method is specially designed to fuse the GFP and phase contrast images. It may not be suitable for other kinds of images such as SPECT/PET and MR/CT images. Besides, decomposition and reconstruction of CSR and the DMBSO algorithm searching optimal region size need to take more time than other fusion methods. The time consumption of pre-registration is repaid in our fusion method. Fortunately biomedical image fusion is not a real-time work and the hardware such as GPU and TPU can reduce the running time of the proposed method.

The satisfactory fusion results acquired by our proposed method will greatly contribute to the research of predicting the role of unknown proteins, analyzing the function of proteins, precisely locating the subcellular structure, etc.

ACKNOWLEDGMENTS
The authors sincerely thank the editors and anonymous reviewers for their constructive comments. The authors also thank GFP database of John Innes Center for supplying GFP and phase contrast images (http://data.jic.bbsrc.ac.uk/cgi-bin/gfp). This work is supported by National Key Research and Development Program of China (No. 2016YFC1306600).

REFERENCES
Li T., & Wang Y. (2009). Fusion of GFP fluorescent and phase contrast images using stationary wavelet transform based variable-transparency method. Optics and Precision Engineering, 17(11),
2871-2879.


Table 1
Objective evaluation metrics of the fusion results of Fig. 1 (a) and (b)

Table 2
Objective evaluation metrics of the fusion results of Fig. 5 (a) and (b)

Table 3
Objective evaluation metrics of the fusion results of one hundred pairs of mis-registered GFP and phase contrast images

Fig. 1. One typical example of mis-registered GFP and phase contrast images of arabidopsis thaliana cell. (a) mis-registered GFP image (b) corresponding phase contrast image (c) accurately registered GFP image (d) the residual image of Fig. 1 (a) and (c).

Fig. 2. The visualization of base layers of the images in Fig. 1 decomposed by CSR (a) the base layer of the mis-registered GFP image (b) the base layer of the phase contrast image (c) the base layer of registered GFP image (d) the base layer of the residual image of Fig. 1 (a) and (c).
Fig. 3. The schematic diagram of the proposed fusion method.

Fig. 4. The fusion results of Fig.1 (a) and (b). (a) CSR based method (b) SOMP based method (c) NSCT-SR based method (d) SFL-CT based method (e) CST-Haar based method (f) proposed method.

Fig. 5. Another typical example of mis-registered GFP and phase contrast images of Arabidopsis thaliana cell. (a) mis-registered GFP image (b) corresponding phase contrast image (c) accurately registered GFP image (d) the residual image of Fig. 5 (a) and (c).

Fig. 6. The fusion results of Fig.5 (a) and (b). (a) CSR based method (b) SOMP based method (c) NSCT-SR based method (d) SFL-CT based method (e) CST-Haar based method (f) proposed method.
Fusion of GFP and phase contrast images based on convolutional sparse representation (CSR) decomposition.

The coefficients of base layers are merged by the proposed adaptive region energy (ARE) rule. The coefficients of detail layers are merged by sum modified Laplacian (SML) rule.

379x226mm (300 x 300 DPI)
Fig. 1. One typical example of mis-registered GFP and phase contrast images of arabidopsis thaliana cell. (a) mis-registered GFP image (b) corresponding phase contrast image (c) accurately registered GFP image (d) the residual image of Fig. 1 (a) and (c).

180x199mm (300 x 300 DPI)
Fig. 2. The visualization of base layers of the images in Fig. 1 decomposed by CSR (a) the base layer of the mis-registered GFP image (b) the base layer of the phase contrast image (c) the base layer of registered GFP image (d) the base layer of the residual image of Fig. 1 (a) and (c).

313x249mm (300 x 300 DPI)
Fig. 3. The schematic diagram of the proposed fusion method.
Fig. 4. The fusion results of Fig.1 (a) and (b). (a) CSR based method (b) SOMP based method (c) NSCT-SR based method (d) SFL-CT based method (e) CST-Haar based method (f) proposed method.
Fig. 5. Another typical example of mis-registered GFP and phase contrast images of Arabidopsis thaliana cell. (a) mis-registered GFP image (b) corresponding phase contrast image (c) accurately registered GFP image (d) the residual image of Fig. 5 (a) and (c).

179x199mm (300 x 300 DPI)
Fig. 6. The fusion results of Fig.5 (a) and (b). (a) CSR based method (b) SOMP based method (c) NSCT-SR based method (d) SFL-CT based method (e) CST-Haar based method (f) proposed method.
<table>
<thead>
<tr>
<th>Method</th>
<th>VIFF</th>
<th>SCD</th>
<th>$Q_b$</th>
<th>$Q_w$</th>
<th>$Q_e$</th>
<th>$Q_{AB/F}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>proposed</td>
<td>0.5853</td>
<td>1.5892</td>
<td>0.8691</td>
<td>0.9253</td>
<td>0.6689</td>
<td>0.7644</td>
</tr>
<tr>
<td>CST-Haar</td>
<td>0.7236</td>
<td>1.6873</td>
<td>0.8453</td>
<td>0.9039</td>
<td>0.6406</td>
<td>0.7360</td>
</tr>
<tr>
<td>SFL-CT</td>
<td>0.3329</td>
<td>1.2560</td>
<td>0.7542</td>
<td>0.8132</td>
<td>0.5505</td>
<td>0.6229</td>
</tr>
<tr>
<td>NSCT-SR</td>
<td>0.8398</td>
<td>0.7630</td>
<td>0.4418</td>
<td>0.6334</td>
<td>0.3930</td>
<td>0.4929</td>
</tr>
<tr>
<td>SOMP</td>
<td>0.7229</td>
<td>0.6369</td>
<td>0.7269</td>
<td>0.8949</td>
<td>0.5827</td>
<td>0.6865</td>
</tr>
<tr>
<td>CSR</td>
<td>0.4397</td>
<td>1.4827</td>
<td>0.6936</td>
<td>0.7503</td>
<td>0.5182</td>
<td>0.6736</td>
</tr>
<tr>
<td>Method</td>
<td>VIFF</td>
<td>SCD</td>
<td>Q₀</td>
<td>Qₘ</td>
<td>Qₜ</td>
<td>Q_{AB/F}</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>proposed</td>
<td>0.8166</td>
<td>1.4628</td>
<td>0.9537</td>
<td>0.9651</td>
<td>0.7236</td>
<td>0.8416</td>
</tr>
<tr>
<td>CST-Haar</td>
<td>0.7984</td>
<td><strong>1.5132</strong></td>
<td>0.9353</td>
<td>0.9438</td>
<td>0.7097</td>
<td><strong>0.8522</strong></td>
</tr>
<tr>
<td>SFL-CT</td>
<td>0.4040</td>
<td>1.1905</td>
<td>0.8393</td>
<td>0.8464</td>
<td>0.6257</td>
<td>0.7328</td>
</tr>
<tr>
<td>NSCT-SR</td>
<td><strong>0.8851</strong></td>
<td>0.5101</td>
<td>0.5544</td>
<td>0.7635</td>
<td>0.5095</td>
<td>0.5667</td>
</tr>
<tr>
<td>SOMP</td>
<td>0.8463</td>
<td>0.9410</td>
<td>0.7849</td>
<td>0.9372</td>
<td>0.6230</td>
<td>0.7548</td>
</tr>
<tr>
<td>CSR</td>
<td>0.4303</td>
<td>1.4117</td>
<td>0.7404</td>
<td>0.7631</td>
<td>0.5513</td>
<td>0.7170</td>
</tr>
</tbody>
</table>
Table 3

Objective evaluation metrics of the fusion results of one hundred pairs of mis-registered GFP and phase contrast images

<table>
<thead>
<tr>
<th>Method</th>
<th>VIFF</th>
<th>SCD</th>
<th>$Q_b$</th>
<th>$Q_w$</th>
<th>$Q_e$</th>
<th>$Q^{AB/F}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>proposed</td>
<td>0.5711/0.1765</td>
<td>1.5598/0.0989</td>
<td>0.9029/0.0280</td>
<td>0.9181/0.0349</td>
<td>0.6746/0.0568</td>
<td>0.7911/0.0360</td>
</tr>
<tr>
<td>CST-Haar</td>
<td>0.6937/0.0850</td>
<td>1.6432/0.1019</td>
<td>0.8882/0.0284</td>
<td>0.8990/0.0332</td>
<td>0.6488/0.0555</td>
<td>0.7798/0.0399</td>
</tr>
<tr>
<td>SFL-CT</td>
<td>0.3326/0.0497</td>
<td>1.2068/0.0755</td>
<td>0.7898/0.0272</td>
<td>0.8073/0.0317</td>
<td>0.5638/0.0506</td>
<td>0.6641/0.0431</td>
</tr>
<tr>
<td>NSCT-SR</td>
<td><strong>0.8519/0.1923</strong></td>
<td>0.8592/0.2423</td>
<td>0.3596/0.1953</td>
<td>0.5635/0.1833</td>
<td>0.3690/0.1126</td>
<td>0.5333/0.0846</td>
</tr>
<tr>
<td>SOMP</td>
<td>0.7229/0.2307</td>
<td>0.6381/0.1742</td>
<td>0.6247/0.1442</td>
<td>0.8607/0.0679</td>
<td>0.5341/0.1099</td>
<td>0.6545/0.0852</td>
</tr>
<tr>
<td>CSR</td>
<td>0.4323/0.0471</td>
<td>1.5075/0.1375</td>
<td>0.7010/0.0217</td>
<td>0.7430/0.0193</td>
<td>0.5234/0.0459</td>
<td>0.6820/0.0424</td>
</tr>
</tbody>
</table>

Note: The values before “/” are Statistical mean, and the values after “/” are the Statistical SD.