Medical Image Fusion via Convolutional Sparsity Based Morphological Component Analysis

Yu Liu, Xun Chen, Rabab K. Ward, *Fellow, IEEE*, and Z. Jane Wang, *Fellow, IEEE*

***Abstract*—In this letter, a sparse representation (SR) model named convolutional sparsity based morphological component analysis (CS-MCA) is introduced for pixel-level medical image fusion. Unlike the standard SR model, which is based on single image component and overlapping patches, the CS-MCA model can simultaneously achieve multi-component and global SRs of source images, by integrating MCA and convolutional sparse representation (CSR) into a unified optimization framework. For each source image, in the proposed fusion method, the CSRs of its cartoon and texture components are first obtained by the CS-MCA model using pre-learned dictionaries. Then, for each image component, the sparse coefficients of all the source images are merged and the fused component is accordingly reconstructed using the corresponding dictionary. Finally, the fused image is calculated as the superposition of the fused cartoon and texture components. Experimental results demonstrate that the proposed method can outperform some benchmarking and state-of-the-art SR-based fusion methods in terms of both visual perception and objective assessment.**

***Index Terms*—Medical image fusion, sparse representation (SR), morphological component analysis (MCA), convolutional sparse representation (CSR), dictionary learning.**

I. Introduction

DUE to the diversity of image capturing mechanisms, medical images with different modalities may reflect very different categories of organ/tissue information. For example, the computed tomography (CT) images can clearly exhibit dense structures like bones and implants, while the magnetic resonance (MR) images provide high-resolution anatomical information for soft tissues. The aim of pixel-level medical image fusion technique is to combine the complementary information in multi-modality medical images by generating a composite image, which is hoped to be more suitable for physician observation or machine perception [1].

A variety of image fusion methods have been proposed in the last few decades. According to a recent survey [2], these methods can be generally grouped into four categories based on the image transform strategy adopted: multi-scale decomposition (MSD)-based methods [3]–[12], sparse representation (SR)-based methods [13]–[25], methods performed in other domains [26]–[34] and methods based on combination of different transforms [35], [36]. It is noted that the SR-based methods have recently emerged as an influential branch in image fusion, although they have only appeared within the last decade since the pioneering work by Yang and Li [13]. A comprehensive review on this topic can be found in [37]. In this letter, we mainly focus on SR-based image fusion.

One of the most crucial issues in SR-based image fusion is the adoption of an SR model [37]. The early SR-based fusion methods [13], [15] employ the standard sparse coding model [38] which is based on single image component and local patches. The source images are divided into a set of overlapping patches in the original spatial domain for sparse coding. This practice is followed by most existing SR-based fusion methods, which generally attempt to the promote model performance by adding elaborate constraints into the model [19], [22], designing more effective dictionary learning strategies [20], employing multiple sub-dictionaries for representation [16], [18], etc.

To address the issue of single-component representation, Jiang and Wang [17] proposed a novel multi-component SR-based fusion method via morphological component analysis (MCA) [39], which can obtain the sparse representations of cartoon and texture components of each source image. This component separation process can significantly improve the flexibility for designing more effective fusion strategies. On the other hand, to overcome the drawbacks caused by patch-based coding, Liu *et al.* [21] proposed a global SR-based image fusion method by introducing the convolutional sparse representation (CSR) model [40], where sparse coding is performed over the entire image rather than on overlapping patches to achieve better representations.

In this letter, we study SR-based image fusion that simultaneously addresses the above two issues: multi-component and global sparse representations. The contributions are two-fold:

1) We introduce an SR model, the convolutional sparsity based morphological component analysis (CS-MCA), into the field of image fusion. By integrating MCA and CSR into a unified optimization framework, this model can achieve multi-component and global sparse representations of source images at the same time.
2) We propose a new medical image fusion method based on the CS-MCA model. Experimental results show that the

Manuscript received November 22, 2018; revised January 22, 2019; accepted January 24, 2019. Date of publication January 28, 2019; date of current version February 11, 2019. This work was supported in part by the National Natural Science Foundation of China under Grants 61701160 and 81571760, in part by the Provincial Natural Science Foundation of Anhui under Grant 1808085QF186, in part by the Fundamental Research Funds for the Central Universities under Grant JZ2018HGTB0228, and in part by the SenseTime Research Fund. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Charles Kervrann. *(Corresponding author: Xun Chen.)*

Y. Liu is with the Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China (e-mail: yuliu@hfut.edu.cn).

X. Chen is with Hefei National Laboratory for Physical Sciences at the Microscale and Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230026, China (e-mail: xunchen@ustc.edu.cn).

R. K. Ward and Z. J. Wang are with Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: rababw@ece.ubc.ca; zjanew@ece.ubc.ca).

This letter has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/LSP.2019.2895749

proposed method outperforms several benchmarking and state-of-the-art SR-based fusion methods in terms of both visual quality and objective assessment.

II. Related Work and Motivation

A. Standard SR Model Based Image Fusion

In [13], Yang and Li first introduced SR into image fusion. In their method, a source images is divided into a set of overlapping patches, and the standard sparse coding model [38] is applied independently on each image patch. Mathematically, the applied SR model can be expressed as

$$\min_{\mathbf{x}} ||\mathbf{x}||_0 \quad \text{s.t. } ||\mathbf{y}-\mathbf{D}\mathbf{x}||_2 < \varepsilon, \tag{1}$$

where $\mathbf{y} \in \mathrm{R}^n$ means a stacked vector version of an image patch of size $\sqrt{n} \times \sqrt{n}$. $\mathbf{D} \in \mathrm{R}^{n \times m}$ means an over-complete dictionary. $\mathbf{x} \in \mathrm{R}^m$ is the sparse vector to be calculated and the sparsity is measured by its l_0-norm which counts the number of non-zero entries. ε represents the tolerance of reconstruction error. The orthogonal matching pursuit (OMP) algorithm is employed to solve this optimization problem. In [15], Yang and Li improved this method by applying the simultaneous OMP (SOMP) algorithm, which can ensure that the source image patches at the same location are decomposed by identical dictionary atoms. For these patch-based SR methods, the target image is finally obtained by aggregating all reconstructed patches and averaging the overlapping pixels.

B. MCA-Based Image Fusion

In [17], Jiang and Wang proposed a multi-component SR-based image fusion method by adopting the MCA model [39], which can be formulated as

$$\min_{\mathbf{x}_c,\mathbf{x}_t} \frac{1}{2}\left\|\mathbf{y}-\mathbf{D}_c\mathbf{x}_c-\mathbf{D}_t\mathbf{x}_t\right\|_2^2 + \lambda_c||\mathbf{x}_c||_1 + \lambda_t||\mathbf{x}_t||_1, \tag{2}$$

where $\mathbf{x}_c$ and $\mathbf{x}_t$ denote the sparse representations of the cartoon and texture components using dictionaries $\mathbf{D}_c$ and $\mathbf{D}_t$, respectively. The l_1-norm is used to constrain the sparsity, while λ_c and λ_t are regularization parameters.[1] Clearly, the main advantage of the MCA model is that it can separate the cartoon and texture components from the original image for individual fusion. Since the cartoon and texture components focus on distinct image contents (i.e., the cartoon component mainly contains piecewise smooth contents such as geometric structures in large scales, while the texture component hold repeated/oscillating patterns and fine details in small scales), this separation can allow better flexibility in designing more effective fusion strategies.

C. CSR-Based Image Fusion

In [21], Liu *et al.* introduced a CSR-based image fusion approach to address the defects caused by patch-based representation manner. The CSR model [40] is defined as

$$\min_{\mathbf{X}_m} \frac{1}{2}\left\|\mathbf{Y}-\sum_{m=1}^{M}\mathbf{d}_m * \mathbf{X}_m\right\|_2^2 + \lambda\sum_{m=1}^{M}\left\|\mathbf{X}_m\right\|_1, \tag{3}$$

where $\mathbf{Y}$ is an entire image instead of a local image patch in Eq. (1). It is modeled as the sum over a set of M convolutions between local dictionary filters $\{\mathbf{d}_m\}$ and global sparse coefficient maps $\{\mathbf{X}_m\}$. To make a clear distinction with the patch-based SR, the image and sparse maps are written in capitals. We continue to use the l_2-norm and l_1-norm since the contents within them can be stacked as column vectors. As opposed to the standard SR model using independent patch-based coding, the result of CSR is single-valued and optimized over the entire image, leading to better capability in detail preservation for image fusion [21]. Moreover, CSR has the property of shift invariance [40], which is of great significance for an image fusion method in ensuring its robustness to mis-registration.

[1]The standard SR model can also be expressed using an unconstrained form as $\min_{\mathbf{x}} \frac{1}{2}\left\|\mathbf{y}-\mathbf{D}\mathbf{x}\right\|_2^2 + \lambda||\mathbf{x}||_1$.

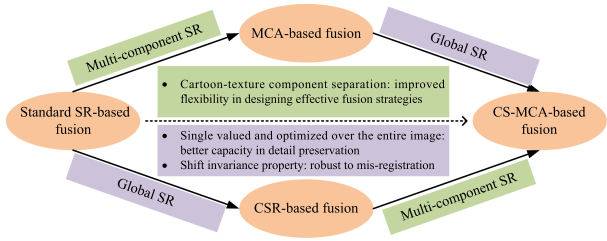

Fig. 1. The relation of this work to prior works.

D. Motivation of This Work

As mentioned above, the MCA-based method [17] and CSR-based method [21] provide two different fashions to promote SR-based image fusion, namely, multi-component representation and global representation, leading to their respective advantages as summarized in Fig. 1. However, each of them neglects the issues addressed by the other one. In this work, inspired by some recent advances in convolutional sparse coding (CSC) [40], [41], the CS-MCA model (presented in Section III-A) is introduced to combine these advantages by simultaneously achieving multi-component and global sparse representations of source images. The relation of this work to prior works is shown in Fig. 1.

III. The Proposed Method

A. The CS-MCA Model

By referring to the above standard SR, MCA and CSR models, the CS-MCA model is defined as

$$\min_{\mathbf{x}_{c,m},\mathbf{x}_{t,m}} \frac{1}{2}\left\|\mathbf{Y}-\sum_{m=1}^{M_c}\mathbf{d}_{c,m} * \mathbf{X}_{c,m} - \sum_{m=1}^{M_t}\mathbf{d}_{t,m} * \mathbf{X}_{t,m}\right\|_2^2 + \lambda_c\sum_{m=1}^{M_c}\left\|\mathbf{X}_{c,m}\right\|_1 + \lambda_t\sum_{m=1}^{M_t}\left\|\mathbf{X}_{t,m}\right\|_1, \tag{4}$$

where $\{\mathbf{d}_{c,m}\}_{m=1}^{M_c}$ and $\{\mathbf{d}_{t,m}\}_{m=1}^{M_t}$ denote two sets of dictionary filters for the SR of the cartoon and texture components, respectively. In our method, they are independently pre-learned from cartoon and texture images using the CSR dictionary learning approach presented in [40]. $\{\mathbf{X}_{c,m}\}_{m=1}^{M_c}$ and $\{\mathbf{X}_{t,m}\}_{m=1}^{M_t}$ are the corresponding sparse coefficient maps. It is clear that the CS-MCA model can be viewed as the CSR-based version of the MCA model given in Eq. (2) as well as the MCA-based version of the CSR model given in Eq. (3), leading to multi-component and global sparse representations at the same time.

The CS-MCA model can be solved iteratively over $\mathbf{X}_{c,m}$ and $\mathbf{X}_{t,m}$ by the following two sub-problems:

Algorithm 1: Algorithm for Solving the CS-MCA model.

Input: $\mathbf{Y}$, $\{\mathbf{d}_{c,m}\}_{m=1}^{M_c}$, $\{\mathbf{d}_{t,m}\}_{m=1}^{M_t}$, λ_c, λ_t, L
  for $i = 1$: L
    Calculate $\mathbf{X}_{c,m}$ by solving Eq. (5);
    Calculate $\mathbf{X}_{t,m}$ by solving Eq. (6);
  end for
Output: $\{\mathbf{X}_{c,m}\}_{m=1}^{M_c}$ and $\{\mathbf{X}_{t,m}\}_{m=1}^{M_t}$.

Sub-problem $\mathbf{X}_{c,m}$ (given the fixed $\mathbf{X}_{t,m}$):

$$\min_{\mathbf{X}_{c,m}} \frac{1}{2}\left\|\mathbf{Y}' - \sum_{m=1}^{M_c} \mathbf{d}_{c,m} * \mathbf{X}_{c,m}\right\|_2^2 + \lambda_c \sum_{m=1}^{M_c} \|\mathbf{X}_{c,m}\|_1, \quad (5)$$

where $\mathbf{Y}' = \mathbf{Y} - \sum_{m=1}^{M_t} \mathbf{d}_{t,m} * \mathbf{X}_{t,m}$.

Sub-problem $\mathbf{X}_{t,m}$ (given the fixed $\mathbf{X}_{c,m}$):

$$\min_{\mathbf{X}_{t,m}} \frac{1}{2}\left\|\mathbf{Y}'' - \sum_{m=1}^{M_t} \mathbf{d}_{t,m} * \mathbf{X}_{t,m}\right\|_2^2 + \lambda_t \sum_{m=1}^{M_t} \|\mathbf{X}_{t,m}\|_1, \quad (6)$$

where $\mathbf{Y}'' = \mathbf{Y} - \sum_{m=1}^{M_c} \mathbf{d}_{c,m} * \mathbf{X}_{c,m}$.

These two sub-problems can be efficiently solved by the ADMM-based CSC algorithm presented in [40]. The overall algorithm for solving the CS-MCA model is summarized in Algorithm 1.

B. Detailed Fusion Scheme

Let $\mathbf{I}^k, k \in \{1, \ldots, K\}$ denote a set of K pre-registered source images, the proposed CS-MCA-based medical image fusion method consists of the following steps.

1) Sparse Coding Based on the CS-MCA Model: Apply the CS-MCA model to each source image $\mathbf{I}^k$. Let $\{\mathbf{X}_{c,m}^k\}_{m=1}^{M_c}$ and $\{\mathbf{X}_{t,m}^k\}_{m=1}^{M_t}$ denote the obtained sparse coefficient maps of the cartoon and texture components, respectively.

2) Fusion of Sparse Coefficient Maps: In accordance with the related SR-based fusion methods [13], [15], [17], [21], the l_1-norm of sparse coefficient vectors is employed as the activity level measurement of source images. For either $\mathbf{X}_{c,m}^k$ or $\mathbf{X}_{t,m}^k$, we use the same fusion strategy (the strategy used in the CSR-based fusion method [21] is adopted in this work to make a fair comparison) but with independent setting of free parameters. Thus, for notational simplicity, we adopt the character n ($n \in \{c, t\}$) to universally denote the cartoon and texture components when applicable. Specifically, let $\mathbf{X}_{n,1:M_n}^k(x, y)$ denote the M_n-dimensional vector containing the coefficients of $\mathbf{X}_{n,m}^k$ at pixel (x, y). The initial activity level map $\mathbf{A}_n^k(x, y)$ is defined as

$$\mathbf{A}_n^k(x, y) = \left\|\mathbf{X}_{n,1:M_n}^k(x, y)\right\|_1, \ n \in \{c, t\}. \quad (7)$$

Further, a window-based strategy is applied to improve the robustness to mis-registration and noise. The final activity level map is calculated as

$$\bar{\mathbf{A}}_n^k(x, y) = \frac{\sum_{p=-r_n}^{r_n} \sum_{q=-r_n}^{r_n} \mathbf{A}_n^k(x+p, y+q)}{(2r_n+1)^2}, \ n \in \{c, t\}, \quad (8)$$

where the parameters r_c and r_t are the window radius for the cartoon and texture components, respectively. Finally, the fused coefficient maps $\{\mathbf{X}_{n,m}^f\}_{m=1}^{M_n}$, with $n \in \{c, t\}$, are obtained using the "choose-max" rule:

$$\mathbf{X}_{n,1:M_n}^f(x, y) = \mathbf{X}_{n,1:M_n}^{k^*}(x, y), \ k^* = \arg\max_k(\bar{\mathbf{A}}_n^k(x, y)). \quad (9)$$

3) Reconstruction of the Fused Image: Having $\{\mathbf{X}_{c,m}^f\}_{m=1}^{M_c}$ and $\{\mathbf{X}_{t,m}^f\}_{m=1}^{M_t}$, the fused image $\mathbf{I}^f$ is reconstructed as

$$\mathbf{I}^f = \sum_{m=1}^{M_c} \mathbf{d}_{c,m} * \mathbf{X}_{c,m}^f + \sum_{m=1}^{M_t} \mathbf{d}_{t,m} * \mathbf{X}_{t,m}^f. \quad (10)$$

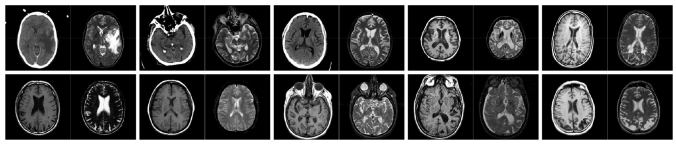

Fig. 2. Ten pairs of source images used in our experiments.

IV. Experiments

A. Experimental Settings

1) Source Images: In our experiments, ten pairs of multi-modal medical images are used for testing, as shown in Fig. 2. They are all from the Whole Brain Atlas [42], a popular medical image database created by Harvard Medical School.

2) Compared Methods: Four existing SR-based image fusion methods are employed for performance comparison: the SR-SOMP-based method [15], the MCA-based method [17], the CSR-based method [21], and the cartoon-texture decomposition and sparse representation (CTD-SR)-based method [24]. The first three can be viewed as the benchmarking methods with regard to our method, as presented in Section II. The last one is a state-of-the-art SR-based fusion method, in which CTD and SR are adopted but with a totally different pipeline in contrast to our method. In the CTD-SR-based method [24], the CTD is performed using a total variation (TV)-based algorithm and the texture components are fused with a patch-based SR approach. The source codes of the SR-SOMP, MCA and CTD-SR based methods are all provided by their original authors. The code of the CSR-based method is publicly available at [43]. All the parameters in these methods are set to the default values for unbiased comparison. In our method, we experimentally fix $\lambda_c = \lambda_t = \max(0.6 - 0.1 \times i, 0.005)$ during the i-th iteration and the number of iterations L is set to 6.

3) Objective Metrics: In [44], the objective evaluation metrics for image fusion are summarized into four categories: information theory based metrics, image feature based metrics, image structure similarity based metrics and human perception inspired metrics. In our experiments, for each category, one widely used metric is adopted. Specifically, the applied metrics include the localized mutual information (LMI) [45], the gradient-based metric $Q^{AB/F}$ [46], the structure similarity based metric Q_E [47] and the human visual system (HVS)-based metric Q_{HVS} [48]. For each of these four metrics, a higher value indicates a better performance. Please refer to [45]–[48] for more details of these metrics.

4) Learning Dictionary Filters: In our experiments, using the dictionary learning approach for CSR presented in [40], 60 cartoon images and 60 texture images collected from the Internet are applied to learn $\{\mathbf{d}_{c,m}\}_{m=1}^{M_c}$ and $\{\mathbf{d}_{t,m}\}_{m=1}^{M_t}$, respectively.

TABLE I
IMPACTS OF FREE PARAMETERS ON OBJECTIVE PERFORMANCE

		$r_t = 2$	$r_t = 3$	$r_t = 4$	$r_t = 5$
LMI	$r_c = 2$	0.8517	0.8508	0.8497	0.8501
	$r_c = 3$	0.8544	0.8540	0.8533	0.8521
	$r_c = 4$	0.8562	0.8560	0.8558	0.8555
	$r_c = 5$	0.8575	0.8571	0.8562	0.8559
	$r_c = 6$	**0.8576**	0.8573	0.8568	0.8564
	$r_c = 7$	0.8575	0.8568	0.8564	0.8563
$Q^{AB/F}$	$r_c = 2$	0.5685	0.5683	0.5682	0.5682
	$r_c = 3$	0.5834	0.5830	0.5830	0.5829
	$r_c = 4$	0.5963	0.5960	0.5957	0.5959
	$r_c = 5$	0.6052	0.6048	0.6043	0.6042
	$r_c = 6$	0.6081	0.6078	0.6072	0.6068
	$r_c = 7$	**0.6095**	0.6091	0.6086	0.6084
Q_E	$r_c = 2$	0.6313	0.6313	0.6312	0.6312
	$r_c = 3$	0.6337	0.6336	0.6335	0.6333
	$r_c = 4$	0.6376	0.6375	0.6373	0.6372
	$r_c = 5$	**0.6397**	0.6396	0.6394	0.6392
	$r_c = 6$	0.6385	0.6383	0.6380	0.6379
	$r_c = 7$	0.6360	0.6358	0.6356	0.6354
Q_{HVS}	$r_c = 2$	0.6770	0.6762	0.6743	0.6744
	$r_c = 3$	0.6805	0.6794	0.6777	0.6778
	$r_c = 4$	0.6837	0.6828	0.6812	0.6810
	$r_c = 5$	**0.6866**	0.6856	0.6840	0.6839
	$r_c = 6$	0.6860	0.6852	0.6839	0.6839
	$r_c = 7$	0.6857	0.6849	0.6837	0.6836

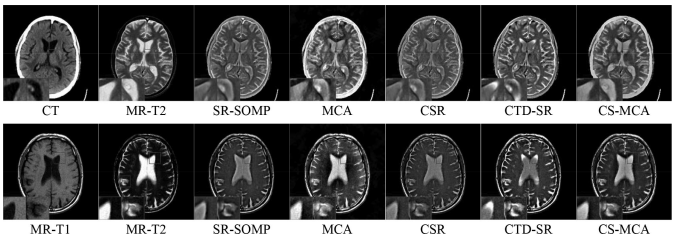

Fig. 3. Two sets of image fusion results obtained by different methods.

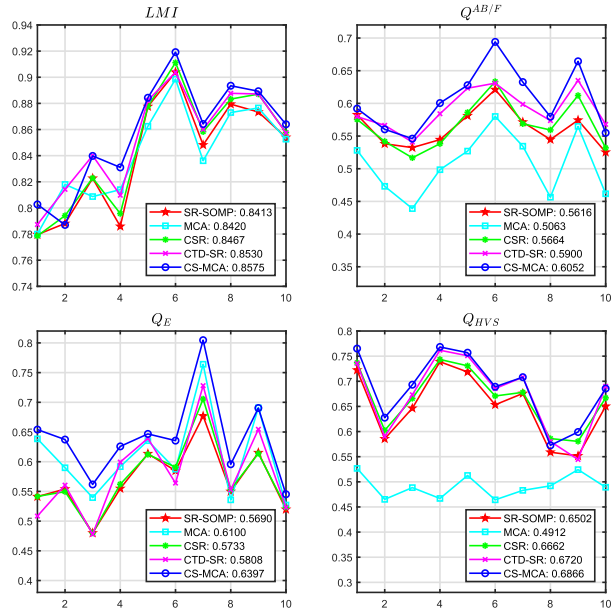

Fig. 4. Objective performance of different fusion methods on four metrics.

TABLE II
THE AVERAGE RUNNING TIME OF DIFFERENT METHODS WHEN FUSING TWO IMAGES OF SIZE 256×256 PIXELS

Methods	SR-SOMP	MCA	CSR	CTD-SR	CS-MCA
Time/s	958	262	30.6	23.8	78.1

It is noted that only the SR-SOMP, CSR and CS-MCA are implemented in pure MATLAB.

In accord with the settings used in the CSR-based fusion method [21], we fix $M_c = M_t = 32$ and the spatial size of each filter is set to 8×8. Please refer to the supplementary material for more details about dictionary learning.

B. Analysis of Free Parameters

Table I reports the objective performance of the proposed method using different values of r_c and r_t in Eq. (8). The average scores over all ten testing examples are calculated, and the highest score for each metric is indicated in bold. It can be seen that the effects of r_c and r_t are distinctly different, which actually reflects the benefits of applying multi-component SR. A smaller value of r_t tends to produce better results, but the impact is relatively slight. The impact of r_c is much more significant and a clear increasing trend from 2 to 5 exists for all the metrics. However, when it exceeds 5, the scores of some metrics start to be stable or even decrease. Based on the above observations, we set $r_c = 5$ and $r_t = 2$ in our method.

C. Comparison With Other Methods

Fig. 3 shows two sets of image fusion results and two close-ups are provided in each set. The source images in the first set are CT and MR-T2 (T2-weighted MR) images, while in the second set are MR-T1 (T1-weighted MR) and MR-T2 images. It can be seen that the SR-SOMP and CSR methods tend to lose some spatial details and image energy. The MCA and CTD-SR methods suffer from obvious undesirable visual artifacts in many fused regions. Overall, the proposed method obtains more competitive visual quality in consideration of detail extraction, energy preservation and the prevention of visual artifacts. Fig. 4 shows the objective performance of different methods. For each metric, the ten scores of a method are connected to generate a curve and the average score is provided in the legend. It is clear that the proposed CS-MCA method outperforms other methods on all the four metrics. Table II lists the average running time of different methods when fusing two images of size 256×256 pixels with a 4.0 GHz CPU and 32 GB RAM. It is noted that the SR-SOMP, CSR and our CS-MCA methods are implemented in *pure* MATLAB, while MATLAB and C/C++ mixed-programming is applied in the MCA and CTD-SR methods. It is believed that the efficiency of our method can be greatly improved by applying more efficient implementation techniques such as C/C++ programming and GPU acceleration.

V. CONCLUSION

In this letter, we introduce a novel SR model CS-MCA for image fusion. By integrating MCA and CSR into a unified optimization framework, the CS-MCA model can simultaneously achieve multi-component and global sparse representations of source images. A new medical image fusion method based on the CS-MCA model is proposed. Experimental results verify the effectiveness of the proposed method.

REFERENCES

[1] A. James and B. Dasarathy, "Medical image fusion: A survey of the state of the art," *Inf. Fusion*, vol. 19, pp. 4–19, 2014.

[2] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, "Pixel-level image fusion: A survey of the state of the art," *Inf. Fusion*, vol. 33, pp. 100–112, 2017.

[3] H. Li, B. Manjunath, and S. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995.

[4] X. Qu, J. Yan, H. Xiao, and Z. Zhu, "Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain," *Acta Autom. Sin.*, vol. 34, no. 12, pp. 1508–1514, 2008.

[5] L. Yang, B. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing*, vol. 72, pp. 203–211, 2008.

[6] Z. L. G. Bhatnagar and Q. Wu, "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1014–1024, Aug. 2013.

[7] S. Das and M. K. Kundu, "A neuro-fuzzy approach for medical image fusion," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3347–3353, Dec. 2013.

[8] R. Shen, I. Cheng, and A. Basu, "Cross-scale coefficient selection for volumetric medical image fusion," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 4, pp. 1069–1079, Apr. 2013.

[9] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.

[10] Z. Zhou, B. Wang, S. Li, and M. Dong, "Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters," *Inf. Fusion*, vol. 30, pp. 15–26, 2016.

[11] P. Hill, M. Al-Mualla, and D. Bull, "Perceptual image fusion using wavelets," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1076–1088, Mar. 2017.

[12] J. Du, W. Li, and B. Xiao, "Anatomical-functional image fusion by information of interest in local Laplacian filtering domain," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5855–5866, Dec. 2017.

[13] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 884–892, Apr. 2010.

[14] N. Yu, T. Qiu, F. Bi, and A. Wang, "Image features extraction and fusion based on joint sparse representation," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 1074–1082, Sep. 2011.

[15] B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Inf. Fusion*, vol. 13, no. 1, pp. 10–19, 2012.

[16] S. Li, H. Yin, and L. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 12, pp. 3450–3459, Dec. 2012.

[17] Y. Jiang and M. Wang, "Image fusion with morphological component analysis," *Inf. Fusion*, vol. 18, no. 1, pp. 107–118, 2014.

[18] Y. Liu and Z. Wang, "Simultaneous image fusion and denosing with adaptive sparse representation," *IET Image Process.*, vol. 9, no. 5, pp. 347–357, 2015.

[19] Q. Zhang and M. Levine, "Robust multi-focus image fusion using multi-task sparse representation and spatial context," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2045–2058, May 2016.

[20] M. Kim, D. Han, and H. Ko, "Joint patch clustering-based dictionary learning for multimodal image fusion," *Inf. Fusion*, vol. 27, pp. 198–214, 2016.

[21] Y. Liu, X. Chen, R. Ward, and Z. Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1882–1886, Dec. 2016.

[22] H. Li, X. He, D. Tao, Y. Tang, and R. Wang, "Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning," *Pattern Recognit.*, vol. 79, pp. 130–146, 2018.

[23] Q. Zhang, T. Shi, F. Wan, R. S. Blum, and J. Han, "Robust sparse representation based multi-focus image fusion with dictionary construction and local spatial consistency," *Pattern Recognit.*, vol. 83, pp. 299–313, 2018.

[24] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Inf. Sci.*, vol. 432, pp. 516–529, 2018.

[25] H. Yin, "Tensor sparse representation for 3-D medical image fusion using weighted average rule," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 11, pp. 2622–2633, Nov. 2018.

[26] S. Li, J. Kwok, and Y. Wang, "Combination of images with diverse focuses using the spatial frequency," *Inf. Fusion*, vol. 2, no. 3, pp. 169–176, 2001.

[27] S. Daneshvar and H. Ghassemian, "MRI and PET image fusion by combining IHS and retina-inspired models," *Inf. Fusion*, vol. 11, pp. 114–123, 2010.

[28] L. Cao, L. Jin, H. Tao, G. Li, Z. Zhuang, and Y. Zhang, "Multi-focus image fusion based on spatial frequency in discrete cosine transform domain," *IEEE Singal Process. Lett.*, vol. 22, no. 2, pp. 220–224, Feb. 2015.

[29] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Inf. Fusion*, vol. 31, pp. 100–109, 2016.

[30] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multiple visual features measurement with gradient domain guided filtering for multisensor image fusion," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 4, pp. 691–703, Apr. 2017.

[31] W. Zhao and H. Lu, "Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2283–2294, Sep. 2017.

[32] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Inf. Fusion*, vol. 36, pp. 191–207, 2017.

[33] Y. Zhang, X. Bai, and T. Wang, "Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure," *Inf. Fusion*, vol. 35, pp. 81–101, 2017.

[34] Y. Chen, J. Guan, and W. K. Cham, "Robust multi-focus image fusion using edge model and multi-matting," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1526–1541, Mar. 2018.

[35] S. Li and B. Yang, "Hybrid multiresolution method for multisensor multimodal image fusion," *IEEE Sens. J.*, vol. 10, no. 9, pp. 1519–1526, Sep. 2010.

[36] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Inf. Fusion*, vol. 24, no. 1, pp. 147–164, 2015.

[37] Q. Zhang, Y. Liu, R. Blum, J. Han, and D. Tao, "Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review," *Inf. Fusion*, vol. 40, pp. 57–75, 2018.

[38] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Berlin, Germany: Springer, 2010.

[39] J. Starck, M. Elad, and D. Donoho, "Redundant multiscale transforms and their application for morphological component separation," *Adv. Imag. Electron Phys.*, vol. 132, pp. 287–348, 2004.

[40] B. Wohlberg, "Efficient algorithms for convolutional sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 301–315, Jan. 2016.

[41] H. Zhang and V. M. Patel, "Convolutional sparse and low-rank coding-based image decomposition," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2121–2133, May 2018.

[42] Nov. 2, 2015. [Online]. Available: http://www.med.harvard.edu/AANLIB/

[43] Jun. 15, 2018. [Online]. Available: https://sites.google.com/site/yuliu316316/

[44] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.

[45] M. Hossny, S. Nahavandi, D. Creighton, and A. Bhatti, "Image fusion performance metric based on mutual information and entropy driven quadtree decomposition," *Electron. Lett.*, vol. 46, pp. 1266–1268, 2010.

[46] C. S. Xydeas and V. S. Petrovic, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, 2000.

[47] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. 10th IEEE Int. Conf. Image Process.*, 2003, pp. 173–176.

[48] Y. Chen and R. Blum, "A new automated quality assessment algorithm for image fusion," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1421–1432, 2009.