Image Fusion With Cosparse Analysis Operator

Rui Gao, *Student Member, IEEE*, Sergiy A. Vorobyov, *Senior Member, IEEE*, and Hong Zhao

***Abstract*—The letter addresses the image fusion problem, where multiple images captured with different focus distances are to be combined into a higher quality all-in-focus image. Most current approaches for image fusion strongly rely on the unrealistic noise-free assumption used during the image acquisition, and then yield limited fusion robustness. In our approach, we formulate the multifocus image fusion problem in terms of an analysis sparse model, and simultaneously perform the restoration and fusion of multifocus images. Based on this model, we propose an analysis operator learning, and define a novel fusion function to generate an all-in-focus image. Experimental evaluations confirm the effectiveness of the proposed fusion approach both visually and quantitatively, and show that our approach outperforms the state-of-the-art fusion methods.**

***Index Terms*—Alternating direction method of multiplier (ADMM), analysis K-singular value decomposition (K-SVD), analysis operator learning, cosparse representation, multifocus image fusion.**

I. Introduction

FUSION of multifocus images is a popular approach for generating an all-in-focus image of higher quality with less artifacts [1], [2]. It relies on the idea of combining a captured sequence of multifocus images with different focal settings. It plays a crucial role in many fundamental fields such as machine vision, remote sensing, and medical imaging [3]–[5]. During the last decade, two types of fusion approaches have been developed: the transform domain-based approaches and the spatial frequency-based approaches. Most existing transform domain-based methods [6]–[9] work with a limited basis, and make fusion excessively dependent on the choice of a basis. The latter approaches [10]–[13] require highly accurate subpixel or subregion estimates, and thus, fail to perform well in elimination of undesirable artifacts.

A prevalent approach for image fusion is based on the *synthesis sparse model*. Various manifold fusion methods have been proposed to explore the synthesis sparse model [3], [14]–[18]. The core ideas here are to describe source images as a linear combination of a few columns from a prespecified dictionary, merge sparse coefficients by a fusion function, and then generate an all-in-focus image using reconstructed sparse coefficients. While there has been extensive research on the synthesis sparse model, the *analysis sparse model* [19]–[22] is a recent construction that stands as a powerful alternative. It represents a signal $\boldsymbol{x}$ by multiplying it with the so-called analysis operator $\boldsymbol{\Omega}$ and emphasizes zero elements of the resulting analyzed vector $\boldsymbol{\Omega x}$ that describes the subspace containing the signal $\boldsymbol{x}$. It promotes strong linear dependencies between the rows of $\boldsymbol{\Omega}$, leads to a much richer union-of-subspaces, and shows the promise to be superior in various applications [23]–[25].

In this letter, we develop a novel fusion algorithm based on the analysis sparse model, which allows for simultaneous restoration and fusion of multifocus images. Specifically, we formulate the fusion problem as a regularized inverse problem of estimating an all-in-focus image given its reconstructed form, and take advantages of the correlations among multiple captured images for fusion using a cosparsity prior. The corresponding algorithm exploits a combination of variable splitting and alternating direction method of multipliers (ADMM) to learn the analysis operator $\boldsymbol{\Omega}$ that promotes cosparsity. Furthermore, a fusion function generates the cosparse representation of an all-in-focus image. Advantages of this approach are more flexible cosparse representation compared to the synthesis sparse approaches and better image restoration and fusion performance. The proposed approach also widens the applicability of the analysis sparse model.

II. Problem Formulation

Let $\boldsymbol{I}_1, \boldsymbol{I}_2, \ldots, \boldsymbol{I}_K$ be a sequence of multifocus images of the same scene acquired with different focal parameters. Our goal is to recover an all-in-focus image $\boldsymbol{I}_{\mathrm{F}}$ from K captured images $\{\boldsymbol{I}_k\}_{k=1}^K$. The model for describing the relationship between the sequence of multifocus images $\{\boldsymbol{I}_k\}_{k=1}^K$ and an unknown all-in-focus image $\boldsymbol{I}_{\mathrm{F}}$ can be formally expressed as

$$\boldsymbol{I}_k = F_k(\boldsymbol{I}_{\mathrm{F}}) + \boldsymbol{V}, \qquad k = 1, \ldots, K \tag{1}$$

where $F_k(\cdot)$ is a blurring operator that denotes the physical process of capturing k-th multifocus image [26], [27] and $\boldsymbol{V}$ is an additive zero-mean white Gaussian noise matrix with entries drawn at random from the normal distribution $\mathcal{N}(0, \sigma^2)$. The blurring operator $F_k(\cdot)$ in most cases is unknown and irreversible; therefore, it is too complex or impossible to find a sequence $\{F_k(\cdot)\}_{k=1}^K$ out of all possible operators. Instead, it is more favorable to seek a compromise between the physical modeling of image capture and signal approximation.

Assuming the cosparsity prior, each image patch $\boldsymbol{i}_{\mathrm{F}} \in \mathbb{R}^m$ of the image $\boldsymbol{I}_{\mathrm{F}}$ is said to have a cosparse representation over the known analysis operator $\boldsymbol{\Omega} \in \mathbb{R}^{h \times m}$ with $h \geqslant m$, if there exists a sparse analyzed vector $\boldsymbol{\Omega i}_{\mathrm{F}} \in \mathbb{R}^h$. The model emphasizes on l zero coefficients of $\boldsymbol{\Omega i}_{\mathrm{F}}$ and defines $\boldsymbol{\Omega}_\Lambda$ as a submatrix of

Manuscript received April 13, 2017; accepted April 13, 2017. Date of publication April 24, 2017; date of current version May 10, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. James E. Fowler. *(Corresponding author: Sergiy A. Vorobyov.)*

R. Gao is with Department Signal Processing and Acoustics, Aalto University, Espoo 02150, Finland, and also with Department Computer Application Technology, Northeastern University, Shenyang 110819, China (e-mail: rui.gao@aalto.fi).

S. A. Vorobyov is with Department Signal Processing and Acoustics, Aalto University, Espoo 02150, Finland (e-mail: svor@ieee.org).

H. Zhao is with Department Computer Application Technology, Northeastern University, Shenyang 110819, China (e-mail: zhaoh@neusoft.com).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2017.2696055

$\boldsymbol{\Omega}$ with rows that belong to the cosupport set Λ. The cosupport set Λ consists of the row indices, which determine the subspace that $\boldsymbol{i}_{\mathrm{F}}$ is orthogonal to. Then, every image patch $\boldsymbol{i}_{\mathrm{F}}$ can be characterized by its cosupport and estimated by solving the following optimization problem:

$$\min_{\hat{\boldsymbol{i}}_{\mathrm{F}}} \left\|\hat{\boldsymbol{i}}_{\mathrm{F}} - \boldsymbol{i}_{\mathrm{F}}\right\|_2^2 \quad \text{s.t. } \left\|\boldsymbol{\Omega}\hat{\boldsymbol{i}}_{\mathrm{F}}\right\|_1 \leqslant \epsilon,\ \boldsymbol{\Omega}_{\Lambda}\hat{\boldsymbol{i}}_{\mathrm{F}} = 0 \tag{2}$$

where $\hat{\boldsymbol{i}}_{\mathrm{F}}$ is a denoised estimation of $\boldsymbol{i}_{\mathrm{F}}$, ϵ is a tolerance error, and $\|\cdot\|_1$ and $\|\cdot\|_2$ are respectively the l_1- and l_2-norms of a vector. Since $\boldsymbol{i}_{\mathrm{F}}$ is unknown, we choose to use the cosparse representation vector $\boldsymbol{\Omega}\hat{\boldsymbol{i}}_{\mathrm{F}}$ for recovering $\hat{\boldsymbol{i}}_{\mathrm{F}}$ through an optimal fusion of cosparse coefficients $\{\boldsymbol{\Omega}\boldsymbol{i}_k\}_{k=1}^K$. Using correlations among multiple images, the proposed approach defines the fusion function as $T(\cdot)$ that generates an optimal cosparse representation and returns the corresponding indices of image patches. Thus, a natural generalization of problem (2) for recovering a clean all-in-focus image patch $\hat{\boldsymbol{i}}_{\mathrm{F}}$ is given as

$$\min_{k,\hat{\boldsymbol{i}}_{\mathrm{F}}} \left\|\hat{\boldsymbol{i}}_{\mathrm{F}} - \boldsymbol{i}_k\right\|_2^2 \quad \text{s.t. } \left\|T\left(\{\boldsymbol{\Omega}\boldsymbol{i}_k\}_{k=1}^K\right)\right\|_1 \leqslant \epsilon, \boldsymbol{\Omega}_{\Lambda}\boldsymbol{i}_k = 0. \tag{3}$$

The role of $T(\cdot)$ in (3) is to provide a meaningful constraint on how closely the optimal patch $\boldsymbol{i}_k$ approximates $\boldsymbol{i}_{\mathrm{F}}$. We replace the cosparse representation $\boldsymbol{\Omega}\hat{\boldsymbol{i}}_{\mathrm{F}}$ by the corresponding optimal $\boldsymbol{i}_k$ of the input image $\boldsymbol{I}_k$, with respect to the cosparse analysis operator $\boldsymbol{\Omega}$. Therefore, the major subproblems here are learning the analysis operator $\boldsymbol{\Omega}$ and defining the fusion function T.

III. Robust Fusion via Analysis Sparse Model

Analysis operator learning aims at constructing an operator $\boldsymbol{\Omega}$ suitable for a family of signals of interest. We first propose a practical approach for learning the analysis operator by variable splitting and ADMM. Then, with the analysis operator fixed, we define the optimal fusion function.

A. Analysis Operator Learning

Suppose a training set $\boldsymbol{Y} = [\boldsymbol{y}_1\ \boldsymbol{y}_2\ \ldots\ \boldsymbol{y}_N] \in \mathbb{R}^{m\times N}$ is formed from a set of N clean vectorized images $\boldsymbol{X} = [\boldsymbol{x}_1\ \boldsymbol{x}_2\ \ldots\ \boldsymbol{x}_N] \in \mathbb{R}^{m\times N}$ contaminated by an additive zero-mean white Gaussian noise $\boldsymbol{V} \in \mathbb{R}^{m\times N}$. Our task is to find $\boldsymbol{\Omega} \in \mathbb{R}^{h\times m}$, which enforces the coefficient vector $\boldsymbol{\Omega}\boldsymbol{x}_i$ to be sparse for each $\boldsymbol{x}_i$. This problem for $\boldsymbol{\Omega}$ can be cast as

$$f(\boldsymbol{\Omega}, \boldsymbol{x}_i) \triangleq \min_{\boldsymbol{\Omega},\boldsymbol{x}_i} \frac{1}{2}\|\boldsymbol{x}_i - \boldsymbol{y}_i\|_2^2 + \lambda\|\boldsymbol{\Omega}\boldsymbol{x}_i\|_1 \tag{4}$$

where λ is a regularization parameter. To prevent $\boldsymbol{\Omega}$ from being degenerate, it is common to normalize its rows $\{\boldsymbol{\omega}_j\}_{j=1}^h$ so that they have their l_2-norms equal to 1. Then, the constraint set can be described as

$$\mathcal{C} \triangleq \begin{cases} \boldsymbol{\Omega}_{\Lambda_i}\boldsymbol{x}_i = 0,\ \mathrm{rank}(\boldsymbol{\Omega}_{\Lambda_i}) = m - r,\ 1 \leq i \leq N \\ \|\boldsymbol{\omega}_j\|_2^2 = 1,\ 1 \leq j \leq h \end{cases} \tag{5}$$

where Λ_i is the index set of the rows in $\boldsymbol{\Omega}$ corresponding to zero elements in $\boldsymbol{x}_i$, $\mathrm{rank}(\cdot)$ denotes the rank of a matrix, and $m - r$ elements of $\boldsymbol{x}_i$ are zeros.

Problem (4)–(5) is nonconvex in variables $\boldsymbol{\Omega}$ and $\boldsymbol{x}_i$. A fundamental approach to addressing it is to alternate between the two sets of variables $\boldsymbol{\Omega}$ and $\boldsymbol{x}_i$, i.e., minimize over one set of variables while keeping the other fixed. Motivated by the first-order surrogate (FOS) approach [28] and ADMM [29], [30], we propose the following FOS-ADMM algorithm for cosparse coding.

Cosparse coding: With $\boldsymbol{\Omega}$ fixed, we update each column of $\boldsymbol{X}$. For notation simplicity, we drop the column index in $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$. We observe that the objective function (4) for fixed $\boldsymbol{\Omega}$, i.e., $f(\boldsymbol{x})$, is the sum of two functions: $f_1(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_2^2$ and $f_2(\boldsymbol{v}) = \lambda\|\boldsymbol{v}\|_1$ where $\boldsymbol{v} = \boldsymbol{\Omega}\boldsymbol{x}$. It enables us to transform the problem of minimizing $f(\boldsymbol{x})$ into the following constraint optimization problem

$$\min_{\boldsymbol{x},\boldsymbol{v}} f_1(\boldsymbol{x}) + f_2(\boldsymbol{v}) \quad \text{s.t. } \boldsymbol{v} = \boldsymbol{\Omega}\boldsymbol{x}. \tag{6}$$

The iterative algorithm to solve (6) can be then expressed in the following ADMM form:

$$\boldsymbol{x}^{(t+1)} = \mathrm{argmin}_{\boldsymbol{x}} f_1(\boldsymbol{x}) + \frac{\mu}{2}\left\|\boldsymbol{\Omega}\boldsymbol{x} - \boldsymbol{v}^{(t)} - \boldsymbol{d}^{(t)}\right\|_2^2 \tag{7a}$$

$$\boldsymbol{v}^{(t+1)} = \mathrm{argmin}_{\boldsymbol{v}} f_2(\boldsymbol{v}) + \frac{\mu}{2}\left\|\boldsymbol{\Omega}\boldsymbol{x}^{(t+1)} - \boldsymbol{v} - \boldsymbol{d}^{(t)}\right\|_2^2 \tag{7b}$$

$$\boldsymbol{d}^{(t+1)} = \boldsymbol{d}^{(t)} - \left(\boldsymbol{\Omega}\boldsymbol{x}^{(t+1)} - \boldsymbol{v}^{(t+1)}\right) \tag{7c}$$

where μ is the augmented Lagrangian penalty and $\boldsymbol{d}^{(t)}$ is the vector of Lagrange multipliers at t-th iteration. Note that the updates for $\boldsymbol{x}$ and $\boldsymbol{v}$ are separated into (7a) and (7b). Subproblem (7a) is a convex quadratic problem, and it can be addressed by the FOS approach [28] that consists of solving iteratively the optimization problem

$$\boldsymbol{x} = \mathrm{argmin}_{\boldsymbol{x}}\, \mathcal{G}(\boldsymbol{x}) \tag{8}$$

where $\mathcal{G}(\boldsymbol{x}) = (\nabla_{\boldsymbol{x}} f_1(\boldsymbol{x}))^T(\boldsymbol{x}^{(t+1)} - \boldsymbol{x}) + \frac{1}{2}L\left\|\boldsymbol{x}^{(t+1)} - \boldsymbol{x}\right\|_2^2 + f_2(\boldsymbol{x})$, $\nabla_{\boldsymbol{x}} f_1(\boldsymbol{x})$ is the gradient of $f_1(\boldsymbol{x})$, L is the Lipschitz constant of $f_1(\boldsymbol{x})$, and $(\cdot)^T$ stands for the transpose.

As for problem (7b) for updating $\boldsymbol{v}$, it turns out to be a simple shrinkage problem. Thus, we just employ the soft-thresholding operator $\mathrm{soft}(\cdot)$, and the update (7b) becomes

$$\begin{aligned}\boldsymbol{v}^{(t+1)} &= \mathrm{soft}\left(\boldsymbol{x}^{(t+1)} - \boldsymbol{d}^{(t)}, \frac{\lambda}{\mu}\right) \\ &= \mathrm{sign}\left(\boldsymbol{x}^{(t+1)} - \boldsymbol{d}^{(t)}\right) \odot \max\left\{\left|\boldsymbol{x}^{(t+1)} - \boldsymbol{d}^{(t)}\right| - \frac{\lambda}{\mu}, 0\right\}\end{aligned} \tag{9}$$

where $\mathrm{sign}(\cdot)$ is the sign function and $\odot$ stands for the component-wise product.

Analysis operator update: With $\boldsymbol{X}$ fixed, we turn to updating $\boldsymbol{\Omega}$ that amounts to obtaining each row $\boldsymbol{\omega}_j$ of $\boldsymbol{\Omega}$. The update of $\boldsymbol{\omega}_j$ should be affected only by those columns of $\boldsymbol{X}$ that are orthogonal to $\boldsymbol{\omega}_j$ [23]. Denote by J a set of indices of those columns, then the corresponding optimization problem can be written as

$$\begin{aligned}\min_{\boldsymbol{\omega}_j} \sum_{i\in J} \|\boldsymbol{x}_i - \boldsymbol{y}_i\|_2^2 \quad &\text{s.t. } \boldsymbol{\Omega}_{\Lambda_i}\boldsymbol{x}_i = 0,\ \|\boldsymbol{\omega}_j\|_2^2 = 1, \\ &\mathrm{rank}(\boldsymbol{\Omega}_{\Lambda_i}) = m - r\end{aligned} \tag{10}$$

where $\{\boldsymbol{x}_i\}_{i\in J}$ and $\{\boldsymbol{y}_i\}_{i\in J}$ form respectively the submatrices of $\boldsymbol{X}$ and $\boldsymbol{Y}$, which contain columns found to be orthogonal to $\boldsymbol{\omega}_j$.

Problem (10) leads to updates of the cosupport sets Λ_i in each iteration. Motivated by [23], we simplify the updates in Λ_i by

using the following approximation:

$$\min_{\boldsymbol{\omega}_j} \sum_{i\in J} \left\|\boldsymbol{\omega}_j^T \boldsymbol{y}_i\right\|_2^2 \quad \text{s.t. } \|\boldsymbol{\omega}_j\|_2^2 = 1 \tag{11}$$

as an alternative that can be solved based on the singular value decomposition (SVD) of the submatrix of $\boldsymbol{Y}$ formed by $\{\boldsymbol{y}_i\}_{i\in J}$.

B. Local Optimal Fusion

When the analysis operator $\boldsymbol{\Omega}$ is learned, we yet cannot directly compute the cosparse representation of $\boldsymbol{\Omega}\boldsymbol{i}_{\mathrm{F}}$. Instead, we work with the collection of the cosparse representations $\{\boldsymbol{\Omega}\boldsymbol{i}_k\}_{k=1}^K$, and then seek the optimal one to recover the corresponding all-in-focus patch.

Using the sliding window technique, each image $\boldsymbol{I}_k$ can be divided into small $n \times n$ patches, from left-top to right-bottom. For convenience, we introduce a matrix $\boldsymbol{W}_{i,j}$ to extract (i, j)-th block from the image $\boldsymbol{I}_k$. Visible artifacts may occur on block boundaries, and we also introduce an overlapping patch of length p for each small patch, and demand that the reconstructed all-in-focus patches would agree to each other on the overlapping areas. According to [16], the block or, equivalently, the set of indices $\{i,\ j,\ k\}$ corresponding to the biggest value in the set $\left\{\|\boldsymbol{\Omega}\boldsymbol{W}_{i,j}\boldsymbol{I}_k\|_1\right\}_{k=1}^K$ is chosen to reconstruct the fused image. Thus, the problem of finding the optimal cosparse representation can be formulated as

$$T(\{\boldsymbol{\Omega}\boldsymbol{W}_{i,j}\boldsymbol{I}_k\}_{k=1}^K) = \boldsymbol{\Omega}\boldsymbol{W}_{i,j}\boldsymbol{I}_k$$
$$\{i,\ j,\ k\} = \underset{i,j,k}{\operatorname{argmax}} \left(\left\{\|\boldsymbol{\Omega}\boldsymbol{W}_{i,j}\boldsymbol{I}_k\|_1\right\}_{k=1}^K\right). \tag{12}$$

Given such optimal fusion function (12), the fusion problem can be cast as the basis pursuit problem with the cosparse regularization term $\|\boldsymbol{\Omega}\boldsymbol{W}_{i,j}\boldsymbol{I}_k\|_1$. Thus, problem (3) can be replaced by the following problem of finding the initial estimate $\hat{\boldsymbol{i}}_{\mathrm{F}_0}$ of the fused patch $\boldsymbol{i}_{\mathrm{F}}$

$$\min_{\hat{\boldsymbol{i}}_{\mathrm{F}_0}} \sum_{i,j,k} \left\|\hat{\boldsymbol{i}}_{\mathrm{F}_0} - \boldsymbol{W}_{i,j}\boldsymbol{I}_k\right\|_2^2$$
$$\text{s.t. } \boldsymbol{\Omega}_\Lambda \boldsymbol{W}_{i,j}\boldsymbol{I}_k = 0,\ \sum_{i,j,k} \|\boldsymbol{\Omega}\boldsymbol{W}_{i,j}\boldsymbol{I}_k\|_1 \le \epsilon. \tag{13}$$

Since problem (13) is convex, its solution can be efficiently computed by using many existing algorithms [20], [23]–[25].

C. Global Reconstruction

The above-explained local optimal fusion is used to recover local details for each all-in-focus patch, respecting spatial compatibility between neighboring patches. In order to remove possible artifacts and improve spatial smoothness, the global reconstruction constraint between the initial image estimate $\hat{\boldsymbol{I}}_{\mathrm{F}_0}$, formed from all $\hat{\boldsymbol{i}}_{\mathrm{F}_0}$'s, and the final estimate $\hat{\boldsymbol{I}}_{\mathrm{F}}$ can be applied to make a further improvement.

The size of $\boldsymbol{\Omega}$ is suitable to represent a small image patch, and it is too small to apply for the entire image. Therefore, we expand the size of $\boldsymbol{\Omega}$ and define

$$\boldsymbol{\Omega}_{\mathrm{F}} \triangleq \begin{bmatrix} \boldsymbol{\Omega}\boldsymbol{W}_{1,1} & \boldsymbol{\Omega}\boldsymbol{W}_{1,2} & \dots & \boldsymbol{\Omega}\boldsymbol{W}_{1,u} \\ \boldsymbol{\Omega}\boldsymbol{W}_{2,1} & \boldsymbol{\Omega}\boldsymbol{W}_{2,2} & \dots & \boldsymbol{\Omega}\boldsymbol{W}_{2,u} \\ \vdots & \vdots & \dots & \vdots \\ \boldsymbol{\Omega}\boldsymbol{W}_{v,1} & \boldsymbol{\Omega}\boldsymbol{W}_{v,2} & \dots & \boldsymbol{\Omega}\boldsymbol{W}_{v,u} \end{bmatrix} \tag{14}$$

Algorithm 1: Fusion.

Input: Multifocus images $\{\boldsymbol{I}_k\}_{k=1}^K$, analysis operator $\boldsymbol{\Omega}$ obtained by analysis operator learning

Output: the estimate of the all-in-focus image $\hat{\boldsymbol{I}}_{\mathrm{F}}$

1: **for** each patch $\boldsymbol{i}_k$ from $\boldsymbol{I}_k$ **do**
2: Compute the fusing coefficients and generate the optimal set of K indices of image patches, using (12);
3: Find the initial estimates $\hat{\boldsymbol{i}}_{\mathrm{F}_0}$, using (13);
4: **end for**
5: Define the global analysis operator $\boldsymbol{\Omega}_{\mathrm{F}}$, using (14);
6: Reconstruct the all-in-focus image $\hat{\boldsymbol{I}}_{\mathrm{F}}$, using (15).

as the global analysis operator where u and v are indices of patches' boundaries. Using the result from the local optimal fusion, the entire image can be redefined using the reconstruction constraint by solving the problem

$$\min_{\hat{\boldsymbol{I}}_{\mathrm{F}}} \left\|\hat{\boldsymbol{I}}_{\mathrm{F}} - \hat{\boldsymbol{I}}_{\mathrm{F}_0}\right\|_2^2 + \lambda' \left\|\boldsymbol{\Omega}_{\mathrm{F}}\hat{\boldsymbol{I}}_{\mathrm{F}}\right\|_1 \tag{15}$$

where λ' is the parameter controlling the sparsity penalty and representation fidelity. Hence, the entire process of the optimal fusion (Subsections III-B and III-C) is summarized in Algorithm 1.

IV. Experimental Results

We verify the restoration and fusion performance of the proposed approach by the visual comparisons, and then discuss the quantitative assessments. We have tested our approach for a number of images, and here one representative example is shown. Specifically, fusion experiments over the standard multifocus dataset [31] are conducted. Throughout the experiments, the tolerance error in the proposed approach is set as $\epsilon = 0.1$, the maximum number of iterations is 1000, the patch size is $n = 7$, and the overlapping length is $p = 1$. During the analysis operator learning, the generated training set consists of 10 000 two-dimensional normalized samples of size 7×7 extracted at random from the natural images. Considering the tradeoff between fusion quality and computations, the analysis operator size is fixed to 64×49. All the experiments are performed on a PC running Inter(R) Xeon(R) 3.40 GHz CPU.

In the noise-free ($\sigma = 0$) and noisy ($\sigma = 15$) cases, the proposed approach is compared with well-known fusion approaches, including the image fusion approach based on spatial frequency in discrete cosine transform (SF-DCT) [5] and the sparse representation K-SVD-based image fusion approach (SR-KSVD) [16]. The fusion results of the noise-free images "Dog" are shown in Fig. 1, including the magnified details in the lower-right corners of the images. There are noticeable differences in the edge of the wall. The SF-DCT method (see Fig. 1(c)) produces blocking artifacts, and the SR-KSVD method (see Fig. 1(d)) introduces undesired smoothing. Our proposed method (see Fig. 1(e)) eliminates some artificial distortions, and gives better visual result. To test the robustness of our approach, we add Gaussian noise to the multifocus images. In Fig. 2, the results for the approaches tested are shown. Note that the SF-DCT method needs the denoising preprocessing, and then fuses multifocus images. Fig. 2(c) shows the circle blurring effect around strong boundaries. The image (see Fig. 2(d)) also shows the blurring effect for the SR-KSVD method. Our approach is capable of providing restoration and fusion

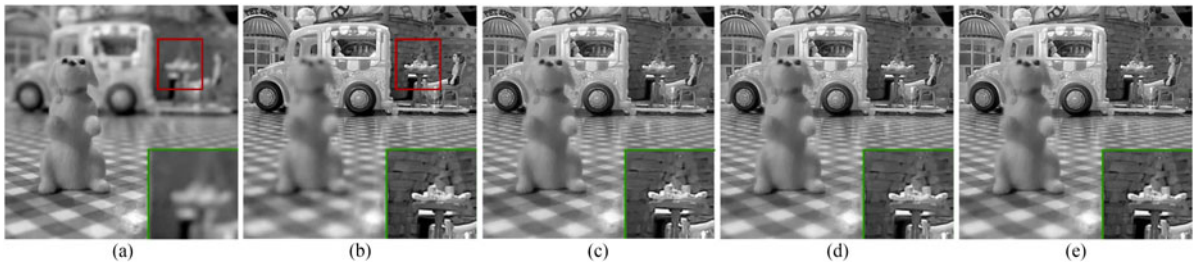

Fig. 1. Source images and the fusion results, with $\sigma = 0$. (a) First source image with focus on the front. (b) Second source image with focus on the background. Fused images obtained by (c) SF-DCT, (d) SR-KSVD, and (e) Proposed method.

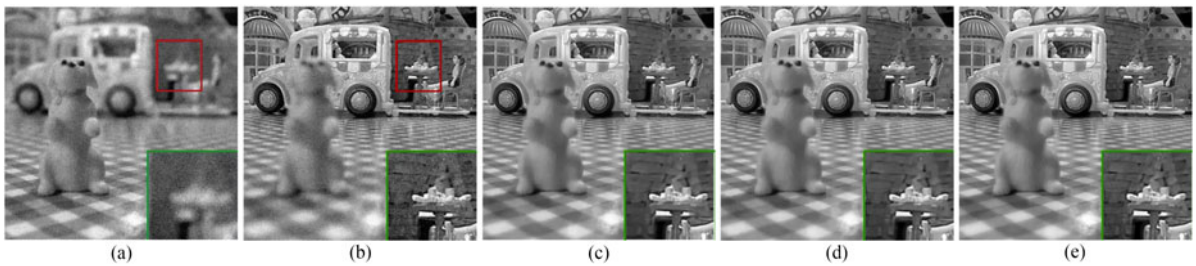

Fig. 2. Source images and the fusion results, with $\sigma = 15$. Same order as in the Fig. 1.

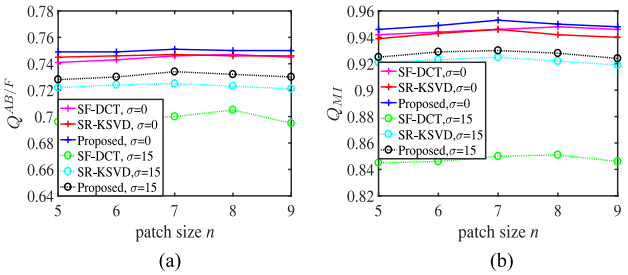

Fig. 3. Fusion performances versus patch size n.

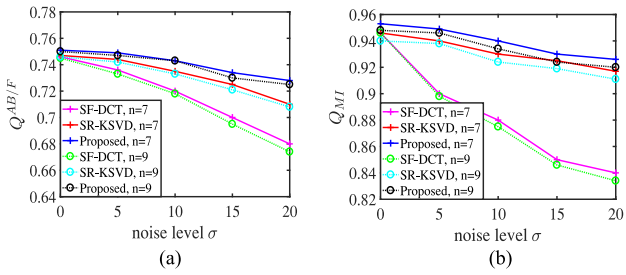

Fig. 4. Fusion performances versus noise level σ.

simultaneously, and it performs the best as it visually appears in Fig. 2(e).

More objectively, we test the impact of different parameters' selection on the proposed approach. The objective evaluation is based on the following two state-of-the-art fusion performance metrics: Q_{MI} [32], which measures how well the mutual information from the source images is preserved in the fused image; and $Q^{AB/F}$ [33], which evaluates how well the edge information transfers from the source images to the fused image. The values of Q_{MI} and $Q^{AB/F}$ range from 0 to 1, with 1 representing the ideal fusion. First, we conduct several experiments for different patch sizes, and compare the performance in the noise-free ($\sigma = 0$) and noisy ($\sigma = 15$) cases for the aforementioned methods in Fig. 3. The employed patch sizes are $n = \{5, 6, 7, 8, 9\}$. In either case $\sigma = \{0, 15\}$, the values of $Q^{AB/F}$ (see Fig. 3(a)) and Q_{MI} (see Fig. 3(b)) for the proposed approach are always larger than those for the SF-DCT and SR-KSVD methods. It means that our approach preserves well the mutual information and transfers efficiently the edge information from source images. When patch size is 7×7, the values of $Q^{AB/F}$ and Q_{MI} are the best. Thus, we set the patch size $n = 7$, and also conduct fusion experiments with different noise levels $\sigma = \{0, 5, 10, 15, 20\}$. The results are shown in Fig. 4. It can be seen that all the methods tested show larger values of Q_{MI} and $Q^{AB/F}$ when σ is equal to zero. With the increase of the noise level, the values of Q_{MI} and $Q^{AB/F}$ gradually decrease, while the proposed method performs the best. Table I presents the average running time of the aforementioned methods. As expected, the proposed approach achieves the restoration and fusion with high quality in reasonable time.

TABLE I
TIME MEASURE OF FUSION METHODS

Measure	Methods in noise-free case			Methods in noisy case ($\sigma = 15$)		
	SF-DCT	SR-KSVD	Ours	SF-DCT	SR-KSVD	Ours
Times(s)	1.073	4.342	4.567	2.179	5.281	5.981

V. CONCLUSION

A novel fusion approach for combining multifocus noisy images into a higher quality all-in-focus image based on the analysis sparse model has been presented. Using the cosparsity prior assumption, we have proposed an analysis operator learning approach based on ADMM. Furthermore, an efficient fusion processing via the learned analysis operator has been presented. Extensive experiments have demonstrated that the proposed approach can fuse images with remarkably high quality, and have confirmed the highly competitive performance of our proposed algorithm. As a future work, a more flexible penalty function can be employed in the fusion problem, which can possibly lead to even better results.

REFERENCES

[1] A. A. Goshtasby and S. Nikolov, "Image fusion: Advances in the state of the art," *Inf. Fusion*, vol. 8, no. 2, pp. 114–118, Apr. 2007.

[2] T. Wan, C. Zhu, and Z. Qin, "Multifocus image fusion based on robust principal component analysis," *Pattern Recognit. Lett.*, vol. 34, no. 9, pp. 1001–1008, Jul. 2013.

[3] Q. Zhang and M. D. 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.

[4] V. N. Gangapure, S. Banerjee, and A. S. Chowdhury, "Steerable local frequency based multispectral multifocus image fusion," *Inf. Fusion*, vol. 23, pp. 99–115, May 2015.

[5] 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 Signal Process. Lett.*, vol. 22, no. 2, pp. 220–224, Feb. 2015.

[6] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.

[7] V. Aslantas and R. Kurban, "Fusion of multi-focus images using differential evolution algorithm," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8861–8870, Dec. 2010.

[8] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Inf. Fusion*, vol. 26, no. 7, pp. 971–979, Jul. 2008.

[9] Z. Zhou, S. Li, and B. Wang, "Multi-scale weighted gradient-based fusion for multi-focus images," *Image Vis. Comput.*, vol. 20, pp. 60–72, Nov. 2014.

[10] J. Tian and L. Chen, "Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure," *Signal Process.*, vol. 92, no. 9, pp. 2137–2146, Sep. 2012.

[11] A. L. Da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[12] O. Rockinger, "Image sequence fusion using a shift-invariant wavelet transform," in *Proc. IEEE Int. Image Process.*, Santa Barbara, CA, Oct. 1997, vol. 3, pp. 288–291.

[13] Q. Zhang and B. Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Process.*, vol. 89, no. 7, pp. 1334–1346, Jul. 2009.

[14] S. Ambat, S. Chatterjee, and K. Hari, "Fusion of algorithms for compressed sensing," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3699–3704, May 2010.

[15] H. Li, L. Li, and J. Zhang, "Multi-focus image fusion based on sparse feature matrix decomposition and morphological filtering," *Opt. Commun.*, vol. 342, pp. 1–11, May 2015.

[16] 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.

[17] M. Nejati, S. Samavi, and S. hirani, "Multi-focus image fusion using dictionary-based sparse representation," *Inf. Fusion*, vol. 25, pp. 72–84, Sep. 2015.

[18] R. Gao, S. A. Vorobyov, and H. Zhao, "Multi-focus image fusion via coupled dictionary training," in *Proc. IEEE 41st Int. Conf. Acoust., Speech Signal Process.*, Shanghai, China, 2016, pp. 1666–1670.

[19] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, Jun. 2007.

[20] J. Dong, W. Wang, W. Dai, M. D. Plumbley, Z. Han, and J. Chambers, "Analysis SimCO algorithms for sparse analysis model based dictionary learning," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 417–431, Jan. 2016.

[21] M. Seibert, J. Wörmann, R. Gribonval, and M. Kleinsteuber, "Learning cosparse analysis operators with separable structures," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 120–130, Jan. 2016.

[22] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "The cosparse analysis model and algorithms," *Appl. Comput. Harmon. Anal.*, vol. 34, no. 1, pp. 30–56, Jan. 2013.

[23] R. Rubinstein, T. Peleg, and M. Elad, "Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 661–677, Feb. 2013.

[24] M. Yaghoobi, S. Nam, R. Gribonval, and M. E. Davies, "Constrained overcomplete analysis operator learning for cosparse signal modelling," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2341–2355, May 2013.

[25] S. Hawe, M. Kleinsteuber, and K. Diepold, "Analysis operator learning and its application to image reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2138–2150, Feb. 2013.

[26] S. Pertuz, D. Puig, M. A. Garcia, and A. Fusiello, "Generation of all-in-focus images by noise-robust selective fusion of limited depth-of-field images," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 1242–1251, Mar. 2013.

[27] M. Subbarao, T. Choi, and A. Nikzad, "Focusing techniques," *J. Opt. Eng.*, vol. 32, pp. 2824–2836, Mar. 1993.

[28] J. Mairal, "Incremental majorization-minimization optimization with application to large-scale machine learning," *SIAM J. Optim.*, vol. 25, no. 2, pp. 829–855, Apr. 2015.

[29] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 3, pp. 293–318, Nov. 1992.

[30] S. Xie and S. Rahardja, "Alternating direction method for balanced image restoration," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4557–4567, Nov. 2012.

[31] M. Nejati, S. Samavi, and S. Shirani, "Multi-focus image fusion using dictionary-based sparse representation," *Inf. Fusion*, vol. 25, pp. 72–84, Sep. 2015.

[32] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on information measure for performance of image fusion," *Electron. Lett.*, vol. 44, no. 18, pp. 1066–1067, Aug. 2008.

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