Multi-focus Image Fusion Based on Sparse Representation with Adaptive Sparse Domain Selection

Yu Liu[1] Zengfu Wang[1, 2]
[1]University of Science and Technology of China, Hefei, Anhui, P.R. China
[2]Institute of Intelligent Machine, Chinese Academy of Sciences, Hefei, Anhui, P.R. China
liuyu1@mail.ustc.edu.cn zfwang@ustc.edu.cn

Abstract

Sparse representation (SR) has been widely used in many image processing applications including image fusion. As the contents vary significantly across different images, a highly redundant dictionary is always required in the sparse model, which reduces the algorithm stability and efficiency. This paper proposes a multi-focus image fusion method based on SR with adaptive sparse domain selection (SR-ASDS). Under SR-ASDS, numerous high-quality image patches are first classified into several categories according to their gradient information, and each category is applied into training a compact sub-dictionary. At the fusion process, a corresponding sub-dictionary is adaptively selected for a given pair of source image patches. Moreover, we present a general optimization framework for the merging rule design of the SR based image fusion. Numerous experiments on both clear images and the noisy ones demonstrate that the proposed method outperforms the fusion methods which use a single dictionary, in terms of several popular objective evaluation criteria.

1. Introduction

As different objects have different distances to the camera in a scene, it is difficult to obtain an all-focused image. Fortunately, Multi-focus image fusion technique offers an important low-cost approach to solve this problem. It utilizes the complementary information offered by several part-focused source images to compose an all-focused fused image, which is more suitable for further machine processing or human visual perception. In the past two decades, a variety of image fusion methods have been developed. In general, they can be divided into two categories: the spatial domain based methods and the transform domain based methods.

As the name suggests, the spatial domain based methods do the fusion task directly in the spatial domain. Specifically, the basic principle of this scheme is selecting the clearer image regions from the source images to construct the fused image with a certain focus measure criterion such as variance or spatial frequency. Both the block based and the image segmentation based methods belong to this category.

On the other hand, the transform domain based methods construct a more active branch of image fusion research. Many multi-scale analysis theories have been used as the framework of image fusion since the 1990s, such as various pyramids, discrete wavelet transform, bandlet, curvelet, contourlet and nonsubsampled contourlet (NSCT). In general, the transform domain based methods consist of the following three steps [1]. First, the source images are converted into a transform domain. Second, the transform coefficients are merged according to a certain fusion rule. Finally, the fused image is obtained by performing the inverse transform on the fused coefficients.

Sparse representation (SR) is a powerful statistical signal modeling technique which has been widely used in many image processing applications including image denosing [2], super-resolution [3, 4] and image fusion [5, 6, 7, 8] in the past few years. Actually, SR based image fusion can be viewed as a special transform domain based fusion method with some characteristics of the spatial domain based methods. It has been proved to be an excellent fusion method that can obtain state-of-the-art results. In this paper, we propose a novel multi-focus image fusion method based on SR with adaptive sparse domain selection (SR-ASDS).

The rest of this paper is organized as follows. Section 2 introduces the related works about SR and its applications in image fusion. Section 3 presents a general optimization framework for the merging rule design of the SR based image fusion. Section 4 describes the proposed fusion algorithm in detail. Experimental results and discussions are presented in Section 5, and Section 6 concludes the paper.

2. Related Works and Contributions

2.1. Basic Theory of Sparse Representation

Sparse modeling is a recently proposed image modeling technique which uses the sparsity prior of natural image signals. The basic idea of sparse representation is assuming that a natural signal can be well approximated by a linear combination of a small set of atoms from an over-complete dictionary, i.e., $x \approx D\alpha$, where $x \in R^n$ is the signal,

978-0-7695-5050-3/13 $26.00 © 2013 IEEE
DOI 10.1109/ICIG.2013.123

$D \in R^{n \times m}$ is the dictionary which contains m prototype signals be referred to atoms, for an redundant dictionary, it is usually $n < m$, and $\alpha \in R^m$ is the sparse coefficients. The goal of SR is to calculate the sparsest coefficients α for an input signal x with a given dictionary D, and the sparsity is often measured by the $l_0 - norm$ of vector α, which counts the nonzero coefficients in α. Therefore, this problem can be written as one of the optimization forms:

$$\alpha = \arg\min_{\alpha} \|x - D\alpha\|_2 \quad subject\ to \quad \|\alpha\|_0 < T \tag{1}$$

$$\alpha = \arg\min_{\alpha} \|\alpha\|_0 \quad subject\ to \quad \|x - D\alpha\|_2 < \varepsilon \tag{2}$$

Actually, the above two optimization problems have the same essence and we mainly use the second form in this paper. Some pursuit algorithms have been proposed to solve this non-convex problem, and the most popular one is orthogonal matching pursuit (OMP), which is a greedy algorithm also being used in many other applications.

The dictionary D plays a very important role in a sparse model. There are two main approaches to obtain a dictionary. The first one is the pre-constructed designed dictionary based on analytical methods, such as DCT, wavelets and curvlets, which share the advantages of fast implementation. However, this category of dictionaries is restricted to signals of a certain type and cannot be used for an arbitrary family of signals of interest. The second one is the dictionary learned from a large number of example image patches, using a certain training algorithm such as MOD or K-SVD [9]. This type of dictionaries has a better representative ability and always outperforms the pre-constructed ones, leading to state-of-the-art results in image restoration and reconstruction.

2.2. SR Based Image Fusion

Sparse representation has been applied into image fusion tasks in the past three or four years [5, 6, 7, 8], and the fusion results reach or even exceed the state-of-the-art level. Yang and Li [5] have presented a basic framework for SR based image fusion, where they use the sparse coefficients as the local salient feature and the coefficients are merged with the popular "choose-max" fusion rule using their $l_1 - norm$. The fused image is reconstructed using the merged coefficients with a given dictionary. The whole image is processed by using the "sliding window" technique, which is the most widely-used technique in various SR based image processing applications. Yang and Li [6] replaced the basic OMP used in their previous work with SOMP to calculate the sparse coefficients to guarantee that the different source images are decomposed into the same subset of dictionary atoms. Yu et al. [7] proposed an image fusion algorithm based on joint sparse representation (JSR). In their method, the source image is represented with the common and innovation sparse coefficients separately by JSR. Moreover, they use a special "weighted average" fusion rule instead of the "choose-max". However, their main target is multi-modal image fusion, such as the fusion of CT and MR images in medical image processing. For multi-focus image fusion task, the source images are acquired from the same type of image sensors, so the JSR model and "weighted average" rule are not very suitable.

It should be noted that all the above mentioned SR based image fusion methods only use a single dictionary, either a pre-constructed DCT dictionary or a dictionary learned by K-SVD. Actually, the contents vary significantly across different images or even different patches in a single image, so a single highly redundant dictionary is always required in order to represent various image structures well. However, sparse decomposition over a highly redundant dictionary is potentially unstable and time consuming. Furthermore, when the source images are corrupted by noise, the disadvantages of a single dictionary become more obvious for its limited representative ability.

2.3. Contributions

Here we list two main contributions of this paper. The first one is we propose a SR based multi-focus image fusion scheme with adaptive sparse domain selection (ASDS-SR) to overcome the aforementioned disadvantages of using a single dictionary. Motivated by a recent work [4] of Dong et al., we use several compact sub-dictionaries instead of a single highly redundant dictionary to represent signals. In our method, numerous high-quality natural image patches are first classified into several categories according to their local gradient information, and each category is applied into training a compact sub-dictionary. And then, at the fusion process, for a given pair of source image patches, a corresponding sub-dictionary is adaptively selected based on the patches' gradient information.

The second contribution is that we present a general optimization framework for the merging rule design of the SR based image fusion, which indicates that the popular "choose-max" fusion rule is a special version of it. Considering the consecution, we will first present it in the next section.

3. A General Optimization Framework

There are two important characteristics a satisfactory fused image needs to have. The first is the high gradient information should be kept in the fusion result from the source images. In general, the more gradient information the fusion result has, the better it is. The second one is the difference between the fused and source images should be as small as possible. Yang et al. [8] have proposed an object function to evaluate the fusion performance. However, the generality of its penalty factors is not sufficient. Considering these two criteria, we propose a more general optimization framework:

$$\gamma^* = \arg\min_{\gamma} f_1(s_1,s_2,\alpha,\beta)\| D(\gamma-\alpha)\|_2^2 + f_2(s_1,s_2,\alpha,\beta)\| D(\gamma-\beta)\|_2^2$$
$$subject\ to:\ \|\gamma\|_0 < L \quad (3)$$

where α and β are the sparse coefficients of two corresponding source image patches: patch1 and patch2. Without loss of generality, we only consider the fusion task of two source images here. γ is the optimization variable, which is the unknown fused coefficients. $\| D(\gamma-\alpha)\|_2^2$ and $\| D(\gamma-\beta)\|_2^2$ characterize the difference between the fused image and two source images. The functions $f_1(s_1,s_2,\alpha,\beta)$ and $f_2(s_1,s_2,\alpha,\beta)$ are the penalty factors, where s_1 and s_2 represent the gradient information of the two patches, such as spatial frequency. When patch1 contains more information than patch2, then $f_1(s_1,s_2,\alpha,\beta) > f_2(s_1,s_2,\alpha,\beta)$, i.e. the penalty of the first term is heavier than the second one, so the fused image will be more similar to patch1. When the situation is the opposite, the conclusion is also contrary.

In SR based image fusion scheme, the "choose-max" fusion rule using $l_1 - norm$ of the sparse coefficients is very popular, which is a special situation of Eq. (3). Actually, when the penalty factors are $f_1 = +\infty$ when $\|\alpha\|_1 > \|\beta\|_1$ and $f_2 = +\infty$ when $\|\alpha\|_1 < \|\beta\|_1$, the optimal solution of Eq. (3) is: if $\|\alpha\|_1 > \|\beta\|_1$, then $\gamma = \alpha$; if $\|\alpha\|_1 < \|\beta\|_1$, then $\gamma = \beta$, which is the "choose-max" fusion rule.

Considering the task in this paper is multi-focus image fusion, we use the "choose-max" rule here. The further study of this general framework will be the future work.

4. Fusion Based on ASDS-SR

4.1. Learning the Sub-dictionaries

As K-SVD algorithm has been widely approved in training dictionaries, all sub-dictionaries are learned with K-SVD in this paper. At first, a large number of image patches should be randomly sampled from a set of high-quality natural images, and then the "plain patches" whose intensity variance is less than a certain threshold should be eliminated, so we can obtain M meaningful patches, noted by $S = \{p_1, p_2, ..., p_M\}$.

In order to learn a series of sub-dictionaries, we need to divide all patches in S into several categories and the gradient dominant direction is used for classification. Motivated by the orientation assignment method of SIFT descriptor [10], we calculate the dominant direction by using an orientation histogram, which is formed by the gradient orientations of all pixels in a patch. We obtain the horizontal gradient $G_x(x,y)$ and vertical gradient $G_y(x,y)$ of an image patch by *Sobel* operator, so the gradient magnitude $m(x,y)$ and gradient orientation $\theta(x,y)$ are:

$$m(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \quad (4)$$
$$\theta(x,y) = \tan^{-1}(G_y(x,y)/G_x(x,y)) \quad (5)$$

The orientation histogram has 6 bins covering the 360 degree range of orientations. Each sample added to the histogram is weighted by its gradient magnitude. We construct 7 subsets $S_0, S_1, ..., S_6$ to hold the corresponding patches. If the differences among the 6 bins are small enough, the corresponding patch will be viewed as an irregular structure patch and classified into S_0. We use the ratio between the peak's height and the sum of all bins as the criterion. A patch will be classified into S_0 when its ratio is smaller than 2/6. Otherwise, the patch will be classified into a corresponding set from $S_1, ..., S_6$ according to the peak of the histogram.

Therefore, there are 7 sub-dictionaries in total, where $D_1, ..., D_6$ are learned from $S_1, ..., S_6$ and the last one D_0 is learned from all patches in S. In our experiments, the patch size is set to 8×8 which has been proved to be an appropriate setting for many image processing applications [2, 5]. The training data consists of 100000 8×8 image patches that randomly sampled from a database of 40 high-quality natural images. All the sub-dictionaries sizes are 128, and three of them (D_2, D_4, D_0) are shown in Figure 1. We can see that the atoms' gradient directions in D_2 and D_4 are consistent while the situation is different in D_0.

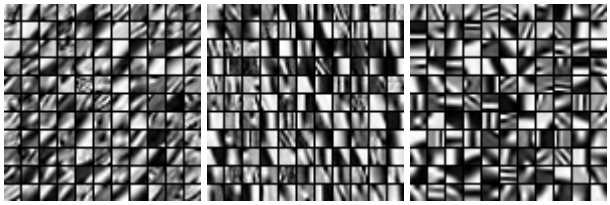


Figure 1. Sub-dictionaries (D_2, D_4, D_0)

4.2. Detailed Fusion Scheme

We also use the SR based fusion framework proposed in [5]. Assuming that there are N registered input source images $I_1, I_2, ..., I_N$. After learning the sub-dictionaries, $I_1, I_2, ..., I_N$ can be merged by taking the following steps.

S1. Construct an empty image I_F. Use the "sliding window" technique to divide each source image into 8×8 patches, from left-top to right-bottom with a certain step length (1 pixel is used in our algorithm). Suppose the source image size is $H\times W$, then there are $(H-8+1)\times(W-8+1)$ groups of patches and each group consists of N patches sampled from N source images. For each group, repeat **S2** to **S4**.

S2. For a given group of N corresponding patches $p_1, p_2, ..., p_N$, choose p_m whose variance is the largest. Then

classify p_m into a category from $S_0, S_1, ..., S_6$ by using the same approach proposed in the sub-dictionaries learning section. Therefore, a corresponding sub-dictionary D_k can be selected from $D_0, D_1, ..., D_6$.

S3. Transform $p_1, p_2, ..., p_N$ into vectors $v_1, v_2, ..., v_N$ via lexicographic ordering. Then use OMP algorithm to calculate the corresponding sparse coefficients $\alpha_1, \alpha_2, ..., \alpha_N$ with the selected dictionary D_k in **S2**.

S4. Apply the "choose-max" fusion rule. The fused coefficients $\alpha_F = \alpha_{k^*}$, where $k^* = \arg\max_k \| \alpha_k \|_1, k = 1, ..., N$. The fused vector $v_F = D_k \alpha_F$, and the fused patch p_F can be obtained by reshaping v_F into an 8×8 block. Then add p_F into its corresponding positions in I_F and update the accumulation times of each pixel.

S5. Finally, for each pixel in I_F, the value is divided by the accumulation times at its position to obtain the fused result.

5. Experimental Results

5.1. Experiment Setups

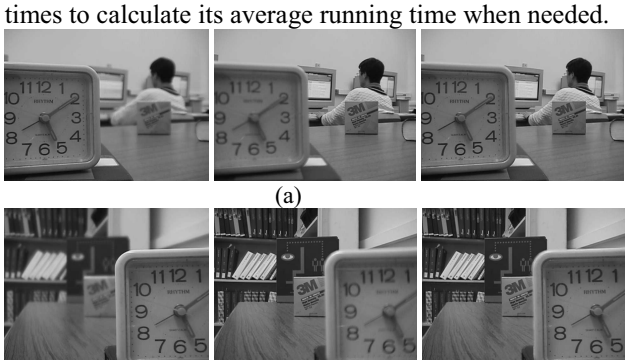
In [5], the SR based multi-focus image fusion method has been compared with many popular fusion methods, such as DWT, MWT, SWT, CVT and NSCT. Their experimental results indicate that the advantages of SR based method are very clear on both clear image datasets and noisy ones. In our experiments, we also confirm it, but it is not our focal point. Our main target is to compare the proposed method with the basic SR based method [5].

In our experiments, there are two basic SR based methods being used to compare. One uses a single dictionary with size of 128 while the other's dictionary size is 256. The two dictionaries are both learned with K-SVD under the same condition. Besides, Pulse coupled neural networks (PCNN) is a novel biological neural network which has been successfully used in image fusion. In our experiments, a PCNN based fusion method named NSCT-SF-PCNN [11] is also used for comparison. Therefore, there are four fusion methods in our experiments: NSCT-SF-PCNN, SR-128, SR-256 and SR-ASDS-128.

Our experiments have been tested on two common-used multi-focus image datasets, and we use the all-focus results offered by website [12] as the reference images. Figure 2 shows the source and reference images. However, the reference image usually does not exist in image fusion tasks, so we also test the proposed method on many other datasets to verify its effectiveness. Furthermore, both clear and noisy source images are tested in our experiments.

In addition to the subjective visual perception, several objective criteria are used to evaluate the fusion results. When the reference image does not exist, Q, Q_W, Q_E [13] and $Q^{AB/F}$ [14] are used, which have been widely used in fusion performance evaluation, and we set $a = 0.5$ to calculate the Q_E. While the reference image exists, in addition to the above criteria, the PSNR and Structural Similarity (SSIM) [15] are also used as evaluation criteria.

Our experiments are implemented in Matlab 7.10 on an AMD 2.8GHz PC with 4G RAM. Each algorithm will run 5 times to calculate its average running time when needed.

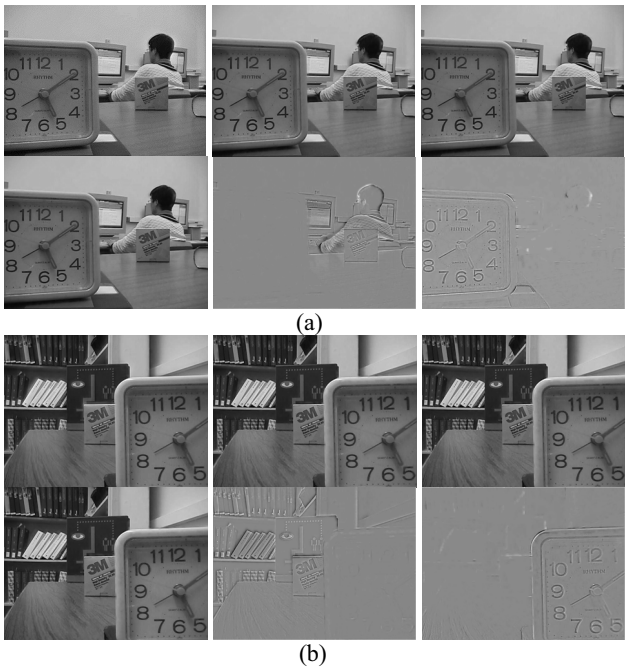


(a) (b)

Figure 2. Multi-focus image datasets. For (a) and (b), from left to right: the first source image, the second source image and the reference image.

5.2. Fusion Results of Clear Images

In this part, the parameter ε in Eq. (2) is set to 1 according to the analysis in [5]. Figure 3 shows the fusion results of the source images in Figure 2. We can see that the results of our method are satisfactory. The objective evaluation criteria are listed in Table 1.



(a) (b)

Figure 3. Fusion results of clear images. For datasets (a) and (b), the first row: NSCT-SF-PCNN, SR-128, SR-256; the second row: ASDS-SR-128 and its differences with the two source images.

Table 1. Comparison of objective criteria of different methods on the two clear image datasets

	PSNR/db	*SSIM*	Q	Q_W	Q_E	$Q^{AB/F}$	*Time/s*
			image set (a)				
NSCT-SF-PCNN	36.2720	0.9837	0.7563	0.9162	0.9349	0.7189	1236.6
SR-128	39.1089	0.9881	0.7596	0.9194	0.9362	0.7361	284.18
SR-256	**39.7528**	**0.9886**	0.7602	0.9199	**0.9376**	**0.7418**	437.57
ASDS-SR-128	39.3803	0.9883	**0.7610**	**0.9202**	0.9366	0.7388	374.34
			image set (b)				
NSCT-SF-PCNN	33.4214	0.9628	0.7487	0.9101	0.9284	0.6879	1284.2
SR-128	36.2738	0.9775	0.7783	0.9099	0.9280	0.7009	330.53
SR-256	36.3693	0.9778	**0.7819**	0.9105	**0.9293**	**0.7053**	518.30
ASDS-SR-128	**36.6008**	**0.9788**	0.7814	**0.9108**	0.9288	0.7028	430.22

Table 2. Comparison of objective criteria of different methods on noisy image datasets at $\sigma = 5$

	PSNR/db	*SSIM*	Q	Q_W	Q_E	$Q^{AB/F}$	*Time/s*
			image set (a)				
SR-128	30.5079	0.8188	0.4393	0.8617	0.8549	0.4939	436.80
SR-256	30.6815	0.8185	0.4399	0.8644	0.8589	0.4963	692.41
ASDS-SR-128	**31.2226**	**0.8265**	**0.4480**	**0.8731**	**0.8690**	**0.5113**	500.88
			image set (b)				
SR-128	29.4234	0.8154	0.5398	0.8518	0.8368	0.4543	445.42
SR-256	29.6439	0.8156	0.5395	0.8590	0.8479	0.4637	696.21
ASDS-SR-128	**30.0165**	**0.8214**	**0.5439**	**0.8654**	**0.8535**	**0.4682**	507.58

Table 3. Comparison of objective criteria of different methods on noisy image datasets at $\sigma = 10$

	PSNR/db	*SSIM*	Q	Q_W	Q_E	$Q^{AB/F}$	*Time/s*
			image set (a)				
SR-128	27.5012	0.5960	0.3233	0.7979	0.7848	0.3851	429.90
SR-256	27.5462	0.5983	0.3269	0.8056	0.7933	0.3944	687.14
ASDS-SR-128	**27.7482**	**0.6056**	**0.3318**	**0.8149**	**0.8014**	**0.4029**	506.19
			image set (b)				
SR-128	26.6483	0.6101	0.4008	0.7864	0.7586	0.3479	436.78
SR-256	26.7398	0.6108	0.4014	0.7937	0.7688	0.3524	694.69
ASDS-SR-128	**26.9277**	**0.6191**	**0.4087**	**0.8051**	**0.7788**	**0.3638**	511.84

In Table 1, the values in bold indicate the highest quality measure obtained over all fusion methods. We can see that the SR based methods have an obvious advantage over NSCT-SF-PCNN on all criteria including the efficiency. As expected, both ASDS-SR-128 and SR-256 outperform SR-128 slightly. The perform difference between SR-256 and ASDS-SR-128 is not very obvious except for the running time, which indicates ASDS-SR-128 is more efficient than SR-256 because of the less dictionary size.

5.3. Fusion Results of Noisy Images

Sparse model has been successfully used in image denoising application, leading to state-of-the-art performance [2]. When the source images are corrupted by noise, the SR based fusion methods can accomplish the fusion and restoration tasks simultaneously [5].

In this part, we will compare the fusion results of the above three SR based fusion methods on noisy source images. The source images in Figure 2 are corrupted by adding zero-mean additive Gaussian white noise with different standard deviations σ of 5 and 10. The parameter ε in Eq. (2) is set to 1.15σ according to [2]. As the lack of space, only the fusion results of ASDS-SR-128 at $\sigma = 5$ are shown in Figure 4. The original clear source images are used to evaluate the fusion performance. The objective criteria of $\sigma = 5$ and $\sigma = 10$ are listed in Table 2 and Table 3. Since the noise is random, each experiment is carried out 5 times to get a mean evaluation value.

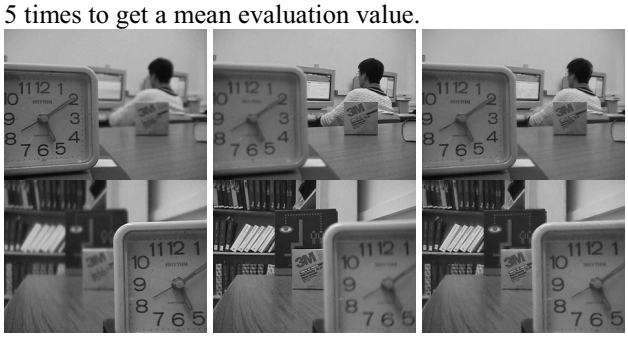


Figure 4. Fusion results of noisy images. For each dataset, from left to right: the first noisy source image, the second noisy source image and the fused image with ASDS-SR-128 at $\sigma = 5$.

5.4. More Experimental Results

To further confirm the effectiveness of the proposed method, 10 pairs of multi-focus images shown in Figure 5 are fused by SR-128, SR-256 and ASDS-SR-128 to make a comparison. Both the clear images and their noisy versions

are used to test. The average objective criteria of 10 pairs of images using different fusion methods are listed in Table 4.

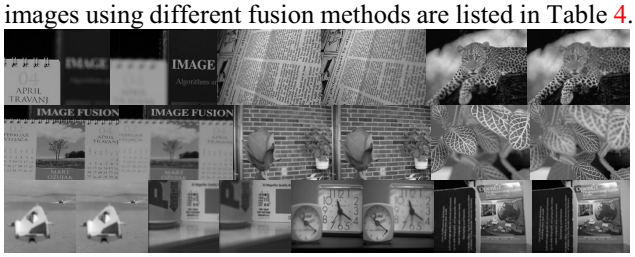


Figure 5. 10 pairs of multi-focus images

Table 4. Comparison of average criteria of 10 pairs of images

	SR-128	SR-256	ASDS-SR-128
	clear images		
Q	0.8285	0.8293	**0.8298**
Q_W	0.9091	0.9093	**0.9098**
Q_E	0.9084	**0.9087**	0.9086
$Q^{AB/F}$	0.7400	0.7402	**0.7403**
Time/s	107.97	175.87	130.99
	noisy images $\sigma = 10$		
Q	0.5476	0.5495	**0.5523**
Q_W	0.8512	0.8542	**0.8584**
Q_E	0.8258	0.8291	**0.8350**
$Q^{AB/F}$	0.5027	0.5056	**0.5105**
Time/s	111.47	178.17	131.75

As reference images do not exist, the PSNR and SSIM are not included in the criteria. For clear images, the SR-256 and ASDS-SR-128 have a slight advantage over SR-128, but ASDS-SR-128 is more efficient than SR-256. For noisy images, the advantage of ASDS-SR-128 is very obvious compared with the other two methods. Overall, the proposed method is more practical than the basic SR based fusion methods [5].

6. Conclusion

This paper has presented a novel multi-focus image fusion method based on sparse representation with adaptive sparse domain selection (ASDS-SR). In our method, a series of compact sub-dictionaries are learned and one of them is adaptively selected according to the patch gradient direction, which makes the proposed method more efficient and stable than the basic SR based methods. Experimental results demonstrate that ASDS-SR has advantages over the basic SR based methods on various objective image fusion evaluation criteria, especially when the source images are corrupted by noise. We have also proposed a general optimization framework for the fusion rule design. However, this paper has not gone deep into this optimization framework and its further study will be taken in the future.

Acknowledgments

This paper is supported by the Fundamental Research Funds for the Central Universities of China (No. WK2100100009), NSFC (No.61175033) and STP (No.11010202192) of Anhui.

References

[1] G. Piella. A general framework for multiresolution image fusion: from pixels to regions. Information Fusion, 4 (4):259-280, 2003. 1

[2] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. on Image Processing, 15(2):3736–3745, 2006. 1, 3, 5

[3] J. Yang, J. Wright, T. Huang and Y. Ma. Image super-resolution via sparse representation. IEEE Trans. on Image Processing, 19(11):2861-2873. 2010. 1

[4] W. Dong, L. Zhang, G. Shi, and X. Wu. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. IEEE Trans. on Image Processing, 20(7):1838–1857, 2011. 1, 2

[5] B. Yang and S. Li. Multifocus image fusion and restoration with sparse representation. IEEE Trans. on Instrumentation and Measurement, 59 (4):884–892, 2010. 1, 2, 3, 4, 5, 6

[6] B. Yang and S. Li. Pixel-level image fusion with simultaneous orthogonal matching pursuit. Information Fusion, 13(1):10–19, 2012. 1, 2

[7] N. Yu, T. Qiu, F. Bi and A. Wang. Image features extraction and fusion based on joint sparse representation. IEEE Journal of Selected Topics in Signal Processing, 5(5):1074-1082, 2011. 1, 2

[8] G. Yang, X. Xu and H. Man. Optimum image fusion via sparse representation. 20th Annual Wireless and Optical Communications Conference, 1-4, 2011. 1, 2

[9] M. Aharon, M. Elad and A. Bruckstein. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing, 54 (11):4311–4322, 2006. 2

[10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91-110, 2004. 3

[11] X. Qu, J. Yan, H Xiao and Z. Zhu. Image fusion algorithm based on spatial frequency-motivated plus coupled neural network in nonsubsampled contourlet transform domain. Acta Automatica Sinica, 34(12):1508-1514, 2008. 4

[12] Investigations of Image Fusion. Electrical Engineering and Computer Science Department, Lehigh University. http://www.ece.lehigh.edu/SPCRL/IF/image_fusion.htm. 4

[13] G. Piella and H. Heijmans. A new quality metric for image fusion. In Proceedings of the IEEE International Conference on Image Processing, 173-176, 2003. 4

[14] C.S. Xydeas and V. Petrovic. Objective image fusion performance measure. Electronics Letters 36(4):308–309, 2000. 4

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error measurement to structural similarity. IEEE Transactions on Image Processing, 3(4):600–612, 2004. 4