Remote-sensing Images Fusion by Compressed Sensing in Contourlet Transform Domain

Yang Senlin, Li Yuanyuan
School of Physics and Mechatronic Engineering
Xi'an University
Xi'an 710065, P. R. China
linsenyang@stu.xjtu.edu.cn

Wan Guobin
School of Electronics and Informatics
Northwestern Polytechnical University
Xi'an 710072, P. R. China
gbwan@nwpu.edu.cn

***Abstract*—The compressive fusion by block-based compressed sensing (BCS) produces a low computation cost, but suffers from low quality in recovery since BCS sampling lacks in global feature. Therefore, a new compressive fusion with the contourlet-transform-based BCS (CTBCS) is proposed. Firstly, the CTBCS is implemented with block-size and subrate dependent on decomposition-level. Then the compressive samplings are fused by rule of linear weighting. Finally, the fused image is reconstructed by the iterative thresholding projection (ITP) algorithm with removal of blocking artifacts. Field test shows the CTBCS achieves better compressive fusion than that by BCS does. With better consideration of global feature, the CTBCS fusion simplifies fusion decision and provides better compressive fusion for big images of remote-sensing.**

Keywords-compressed sensing; remote sensing; fusion; reconstruction; contourlet transform

I. INTRODUCTION

Remote-sensing fusion is widely applied in military surveillance, environmental protection, traffic monitoring, and disaster forecast. The classical fusion techniques are the principle component analysis (PCA) and the intensity-hue-saturation (IHS) transform. In the last decade, multiresolution decomposition schemes based on wavelets, curvelets, and contourlets have become popular in image fusion [1]. However, these methods use all samples of input images for fusion decision, which is extremely inconvenient when handling large images. Hence, the question arises whether the fusion decision can be made based upon partial information retrieved from the input images without impacts on the final art? This is exactly the functionality that compressed sensing (CS) theory [2] is offering, or more particularly compressive fusion since it allows for reconstructing the fused image from a limited set of samples [3-6]. Nonetheless, most of the solutions adhere to the global compressive sampling approach, which results in huge memory requirements and an expensive computational cost for reconstruction of large images. Block-based compressed sensing (BCS) [7] contributes to fast computation and small memory requirements compared to global compressive observation, but suffers from a reduced quality in reconstruction since the block-diagonal measurement lacks in consideration of global features. Additionally, blocking artifacts deteriorate the quality of fused images. To address these problems, deploying compressive sampling in transform domain that allow for retaining global features as well is most likely an interesting research track to exploit. Hence, we propose a novel compressive fusion method based BCS and multiscale transform to alleviate the drawbacks of classical BCS solutions.

The remainder of this paper is organized as follows. First, in Section 2, we make a brief description for the BCS sampling in contourlet transform (CT) domain. Then, in Section 3, we fuse the compressive samplings of CT coefficients with rule of linear weighting. In Section 4, we describe the iterative thresholding projection (ITP) algorithm to reconstruct the fused image, and, in Section 5, we present the experimental results of proposed method. Finally, we draw the concluding remarks in Section 6.

II. COMPRESSIVE SENSING OF CONTOURLET SUBBANDS

Consider a vectorized image $x \in \mathbb{R}^{N\times 1}$ and suppose that an orthogonal basis $\psi \in \mathbb{R}^{N\times N}$ provides a K -sparse representation for x. In terms of matrix notation, we have $x = \psi\upsilon$, in which $\upsilon \in \mathbb{R}^{N\times 1}$ can be well-approximated using only $K << N$ non-zero entries. The CS theory states that such a signal x can be reconstructed by taking only $M = O(K \log N)$ linear, non-adaptive projections with $y = \Phi x$, in which $y \in \mathbb{R}^{M\times 1}$ represents the compressive observation and $\Phi \in \mathbb{R}^{M\times N}$ $(M << N)$ is the measurement matrix. For precise reconstruction, measurement matrix and sparse basis should be satisfied with the restricted isometric property (RIP) condition [2]. Generally, the aforementioned CS projection uses a global sampling, which requires a significant amount of memory space for the storage of measurement matrix and represents a high computational cost for image reconstruction.

For computational efficiency, the block-based compressed sensing (BCS) divides the image to be analyzed into many blocks with size of $B\times B$-pixel and then samples with an identical operator [7]. Letting $x_i \in \mathbb{R}^{B^2\times 1}$ represent the vectorized signal of the block i of input image x, we get the BCS measurement, i.e., $y_i = \Phi_B x_i$, in which $y_i \in \mathbb{R}^{M_B\times 1}$ and $\Phi_B \in \mathbb{R}^{M_B\times B^2}$ $(M_B << B^2)$ is the block-based

measurement operator. Note that BCS applied block-by-block to an image is equivalent to a constrained measurement with a block-diagonal matrix to the whole-image. Ideally, the sampling operator should be global such that entire image should contribute to all measurements. However, an equivalent block-diagonal measurement operator is short of such holistic sampling. As a result, the BCS enables fast reconstruction, but simultaneously suffers from reduced reconstruction quality and blocking artifacts due to their reliance on block-based observation.

Therefore, it is advisable to apply BCS in multiscale transform domain, which allows for accounting of global features. Consider that the CT has good properties in locality, shift-invariance and directional selection, and enables fusion of image features separately at individual scales. Therefore, better reconstruction and fast computation may be achieved simultaneously by BCS sampling from CT decomposition. Assume that θ is the CT coefficients of input image $x \in \mathbb{R}^{N\times 1}$ with L levels, i.e., $\theta = \Omega x$, in which $\Omega \in \mathbb{R}^{N\times N}$ is the CT decomposition operator. Suppose that the CT coefficients of subband j at level $l \subset (1,\cdots,L)$ is divided into $B_l \times B_l$ blocks and then sampled by a level-reliance measurement operator $\Phi_l \in \mathbb{R}^{M_l \times B_l^2}$ ($M_l << B_l^2$). For convenience, letting subscript Λ represent the index set $\{l, j, i\}$, and $\theta_\Lambda \in \mathbb{R}^{B_l^2 \times 1}$ denote vectorized coefficients of block i in subband j at level l, we get

$$y_\Lambda = \Phi_l \theta_\Lambda \tag{1}$$

where $y_\Lambda \in \mathbb{R}^{M_l \times 1}$ is BCS sampling for CT coeffcients at index-Λ, and M_l is the number of sampling at level l. For convenience, the ratio between M_l and B_l^2 is defined as the sampling ratio, i.e., subrate.

Notwithstanding many measurement operators have been proposed in the literature [8], quite few of them can be simultaneously equipped with universality, computational efficiency and optimal reconstruction. The structural random matric (SRM) is universal with a variety of sparse signals. With SRM, the number of measurements required for exact reconstruction is nearly optimal. Moreover, it requires very low complexity and fast computation based on block processing. Therefore, the SRM is employed here as the measurement matrix Φ_l.

Since coefficients from various decomposition levels contribute differently to the quality of recovered image, recovery may be enhanced by distinctive sampling rate (i.e., subrate) c_l for each level [7]. In such way, a larger subrate can be attributed to the subband contributing more significantly to the recovery quality, and vice versa. For convenience, let the baseband of CT decomposition be represented by level $l = 0$, the target subrate c of the whole image can be expressed as

$$c = \frac{1}{N_{ct}} \sum_{l=0}^{L} c_l \sum_{j=1}^{J_l} n_{l,j} \tag{2}$$

where L represent the number of scales, N_{ct} is the total number of CT coefficients, J_l is the number of subbands at level l, and $n_{l,j}$ is the number of coefficients in subband j at level l. Therefore, once the target subrate of the input image is given, the subrate at each decomposition level can be level-by-level determined by equation (2). Usually, full sampling may be used for the subband coefficients of baseband and level $l = 1$ given their importance for signal recovery. For other levels, subrates may be scattered as $c_l = \alpha c_{l+1}$ ($l = 1,\cdots,L$), and the constant $\alpha > 1$.

Clearly, such kind of compressive sampling uses different block-sizes and subrates for CT subbands. For convenience, we name it as CT-based BCS (CTBCS) in the paper.

III. Fusion by Linear Weighting

After input images have been sampled by CTBCS, a strategy is necessary to be adopted for efficient fusion. In image fusion, the linear weighting is a simple and intuitive rule, which assigns different weights to achieve good enough fusion according to the characteristics of data. For CTBCS, linear weighting fusion can better preserve spectral characteristics, and it will not bring any additional blocking noise. Therefore, the linear weighting is employed here to fuse the CTBCS of input images. Let $y_\Lambda^{(A)} = \left\{\hat{y}_\Lambda^{(A)}(m)\right\}^T$ and $y_\Lambda^{(B)} = \left\{\hat{y}_\Lambda^{(B)}(m)\right\}^T$ ($m = 1,\cdots,M_l$) denote the CTBCS of input images A and B on Λ, the linear weighting fusion gives

$$\hat{Y}_\Lambda(m) = a\,\hat{y}_\Lambda^{(A)}(m) + b\,\hat{y}_\Lambda^{(B)}(m),\ m = 1,\cdots,M_l \tag{3}$$

where $\hat{Y}_\Lambda(m)$ is the fused sample-m on Λ, a and b are the weighting coefficients with $a + b = 1$. The CTBCS fusion on Λ can be expressed as $Y_\Lambda = \left\{\hat{Y}_\Lambda(m)\right\}^T$.

IV. Reconstruction

After CTBCS are fused, the final art can be derived by CS reconstruction. Many technologies have been proposed in literatures for CS reconstruction [2-7], e.g., linear programming, gradient projection sparse reconstruction (GPSR), iterative greedy methods, and iterative thresholding methods. However, both linear programming method and min-TV optimization method are of high computational cost. The GPSR and iterative greedy methods suffer lower reconstruction quality for same number of iteration. The ITP [7, 9] is capable of fast reconstruction with good quality and benefits for eliminating the blocking artifacts. For iteration n, we define the iterative thresholding function as

$$S(\Omega, \Theta, \theta^{(n)}) = \Omega\Theta^{-1}\tilde{w}^{(n)} \tag{4}$$

where $\theta^{(n)}$ is the CT coefficients at iterative n, $\Theta \in \mathbb{R}^{N\times N}$ is an directional transform (DT) operator for thresholding filtering. For convenience, let Ω^{-1} and Θ^{-1} be the inverse operators of CT and DT, respectively. In equation (4), $\tilde{w}^{(n)}$ is the thresholding estimation of DT coeffcients $w^{(n)} = \Theta\Omega^{-1}\tilde{\theta}^{(n)}$, in which $\tilde{\theta}^{(n)}$ is the blocking iterative projection of $\theta^{(n)}$. For an arbitrary index-Λ, the projection of $\theta_v^{(n)}$ is defined by $\tilde{\theta}_\Lambda^{(n)} = \theta_\Lambda^{(n)} + \Phi_l^T\left(Y_\Lambda - \Phi_l\theta_\Lambda^{(n)}\right)$. The Θ used here is the complex-valued wavelet transform. For DT coefficients $w^{(n)}$, assume that there are strong dependencies between an arbitrary coefficient ω_l at scale l and its parent ω_{l+1} at scale $l+1$. Further, suppose that these coefficients are corrupted by i.i.d. additive white Gaussian noise, model the statistics of these coefficients with a bivariate shrinkage function, and subsequently estimate the coefficients with a soft-threshold scheme [10]. In this way, the thresholding filtering gives the thresholding estimation $\tilde{w}^{(n)}$ of DT coefficients $w^{(n)}$. The thesholding estimation benefits for removal of blocking artifacts. Moreover, for each iteration, the Wiener filtering is also applied to impose smoothness in addition to the sparsity inherent to iteration.

V. RESULTS

The proposed method is tested with field remote sensing images. Two source images, as shown in Fig.1a-1b, are the PAN and MS images, respectively. The PAN image comes from the Ziyuan Satellite while the MS image is from the Landsat-7 Satellite. These images are after geometric adjustment and with 1024×1024 pixels. The algorithm is implemented by Matlab-7 on a computer of 2.8-GHz CPU with 2G memory. The MS images are firstly converted into 3-components by the IHS transform. After matching processing, the PAN image and I-component are fused for substituting the original I-component, and the final fusion is derived by inverse IHS transform. The images to be fused are decomposed by CT with 3-level at a subrate of 0.25, and the block-size for 3-level are 16, 32 and 64, respectively. Fig.2a shows the fused image by rule of linear weighting. Also, Fig.2b is the BCS fusion, and Fig.2c is the CTBCS fusion, which are all fused by linear weighting. Additionally, the quantitative analysis of fused images is given in Table.1, which includes the mean value (MV), standard deviation (STD), information entropy (IE), average gradient (AG), bias index (BI) and correlation coefficient (CC) [2]. Clearly, these fusions illustrate very small difference in values such as MV, STD, IE, BI, and CC. However, the compressive fusion by CTBCS produces the better values in AG than that by BCS, which are more close to that of CT fusion. Moreover, for different subrates, term the traditional CT fusion as the ground truth, the peak signal-to-noise ratio (PSNR) between CT fusion and compressive fusions with CTBCS and BCS are calculated and shown in Fig.3. The result shows that the fusion by CTBCS have better values in PSNR than that by BCS. Therefore, the CTBCS fusion achieves better performance than BCS fusion since the CTBCS takes better global features into account.

VI. CONCLUSION

We presents a compressive fusion for remote-sensing images based on the CTBCS and ITP algorithm. The CTBCS sampling takes better global features into account via compressive measurement in the CT domain. The field tests show that CTBCS achieves better compressive fusion than BCS does. With better consideration for global feature, the CTBCS fusion may provide a better solution for the compressive fusion of big images in remote sensing.

ACKNOWLEDGMENT

Thanks for supports of the National Natural Science Fund (under Grant 61401356), and Natural Science Foundation Fund of Shaanxi Province (under Grant 2012jq5006).

REFERENCES

[1] X. Yang and L. Jiao, "Fusion algorithm for remote sensing images based on non-subsampled contourlet transform," Acta Automatica Sinica, vol.34, no.3, pp.274-281, Mar.2008.

[2] Y.C. Eldar and G Kutynio. Compressed sensing: theory and applications [M]. New York: Cambridge University Press, 2012.

[3] T. Wan and Z. Qin, "An application of compressive sensing for image fusion," International Journal of Computer Mathematics, vol.8, no.18, pp.3915-3930, 2011.

[4] T. Wan, N. Canagarajah, and A. Achim, "Compressive image fusion," in Proc. of 16th ICIP, Oct.12-Oct.15, 2008, pp.1308-1311.

[5] A. Divekar and O. Ersoy, "Image fusion by compressive sensing," in Proc. of 17th International Conference on Geoinformatics, Aug.12-Aug.14, 2009, pp.1-6.

[6] S. Yang, G. Wan, B. Zhang, and X. Chong, "Remote sensing images fusion based on block compressed sensing," in Proc. of SPIE, vol.8910, Aug. 30, 2013, pp. 891017-8.

[7] J.E. Fowler, M. Sungkwang, E.W Tramel, "Block-Based Compressed Sensing of Images and Video," Foundations and Trends in Signal Processing, vol. 4, no. 4, pp. 297-416, 2010.

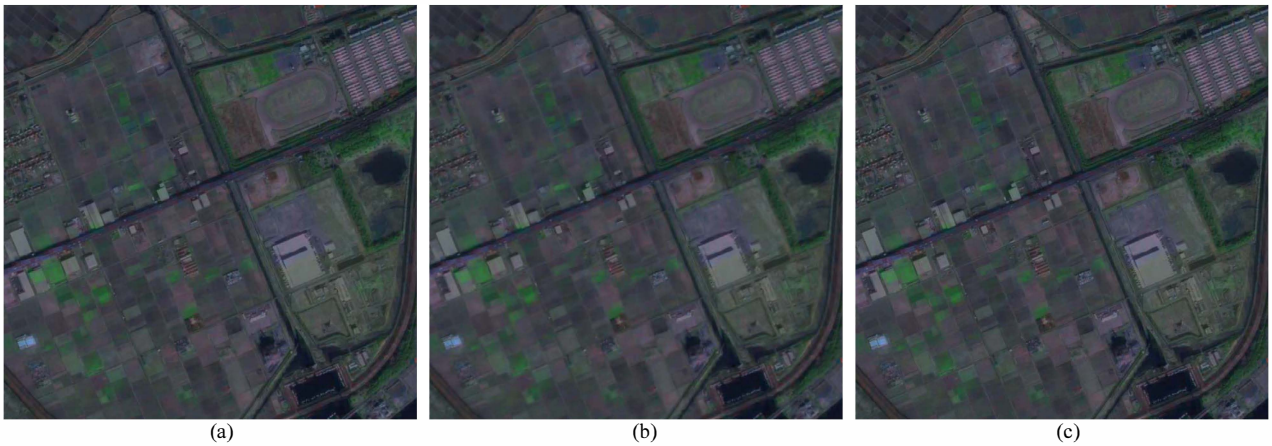
[8] T.T. Do, L. Gan, N.H. Nguyen, and T.D. Trans, "Fast and efficient compressive sensing using structurally random matrix," IEEE transaction on Signal Processing, vol.60, no.1, pp. 139-154, Jan. 2012.

[9] J. Ma, "Single-pixel remote sensing," IEEE Geoscience and Remote Sensing Letters, vol.6, no.2, pp. 199-203, Apr. 2009.

[10] L. Sendur and I.W. Selesnick, "bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," IEEE Transaction on Signal Processing, vol. 50, no. 11, pp. 2744-2756, Nov. 2002.

(a) (b)

Figure.1. The original PAN (a) and MS (b) images



(a) (b) (c)

Figure.2. The fusion images from CTBCS (a), BCS (b), and traditional CT (c) with rule of linear weighting.

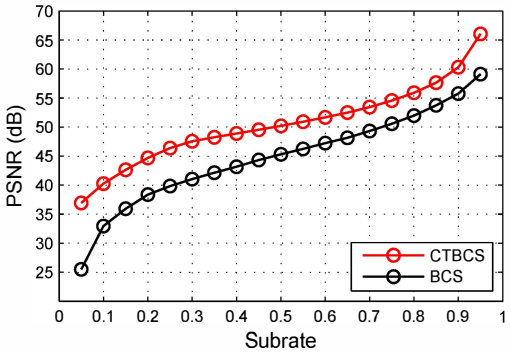


Figure 3. PSNR between CT fusion and compressive fusions with CTBCS (red) and BCS (black)

Table 1. Quantitative analysis for fusion images

Fusion	MV	STD	AG	IE	CC	BI
CTBCS	65.37	17.01	2.07	1.83	0.75	0.36
BCS	65.31	16.50	1.71	1.82	0.74	0.37
CT	65.42	17.06	2.19	1.83	0.75	0.36