RESEARCH   **Open Access**

Compressive SAR raw data with principal component analysis

Wei Wang, Baoju Zhang* and Jiasong Mu

Abstract

The theory of synthetic aperture radar (SAR) signal model is briefly introduced which is constructed with a series of echo signals in range direction. The procedure of principal component analysis (PCA) is presented which is used as transformation basis to sparsify the signals. The joint compressive sensing (CS) and PCA algorithm is devised to realize SAR raw data compressive measurement. A SAR raw data for a point target is simulated and used to verify the performance of the joint CS and PCA algorithm. The numerical experimental results demonstrate that the PCA method has good sparse performance and the joint CS and PCA algorithm is possible to online compressively measure the SAR raw data.

Keywords: Synthetic aperture radar, Compressive sensing, Principal component analysis, Sparsity

Introduction

Synthetic aperture radar (SAR) is active and coherent microwave radar, which produces high spatial resolution images from a moving platform—an airplane or a satellite. SAR systems take advantage of the long-range propagation characteristics of radar signals and the complex information processing capability of modern digital electronics to provide high-resolution imagery. SAR technology has provided terrain structural information to geologists for mineral exploration, oil spill boundaries on water to environmentalists, sea state and ice hazard maps to navigators, and reconnaissance and targeting information to military operations [1]. And there are many other applications or potential applications.

A SAR receiver consists of a high-rate A/D converter followed by pulse compression. The resolution of SAR has been limited by the sampling rate of the fastest available A/D. And the performance of SAR system is further degraded by the trade-off that exists between the rate of sampling and the number of quantization level of an A/D. As the development of SAR imaging technology, the dimension of SAR imagery will become larger and larger and the resolution will become higher and higher [2]. The traditional SAR data acquisition follows the Nyquist sampling theorem and it increases the scale of SAR system for obtaining more and more data. Thus some problem such as data processing and transmission in real time would occur. Hence, how to compress the imagery data effectively during the acquiring process and reconstruct them accurately is a hot spot [3].

A new theory of compressive sensing (also known under the terminology of compressed sensing, compressive sampling, CS) in signal processing provides a fundamentally new approach to data acquisition which overwhelm the common Nyquist sampling theory. Consider a signal that is sparse in some basis (often using a wavelet-based transform coding scheme), the basic idea of CS is projecting the high-dimensional signal onto a measurement matrix, which is incoherent with the sparsifying basis, resulting to a low-dimensional sensed sequence. Then with a relatively small number of appropriately designed projection measurements, the underlying signal may be recovered exactly [4,5]. In contrast to the common framework of first collecting as much data as possible and then discarding the redundant data by digital compression techniques, CS seeks to minimize the collection of redundant data in the acquisition step. Because of the special advantages of CS, many references about SAR imagery raw data compressing and reconstruction based on CS have been researched.

Pruente [6] introduced CS to the field of ground moving target indication using multichannel stripmap SAR-data. And the results showed that the approach

* Correspondence: wdxyzbj@163.com
College of Physical and Electronic Information, Tianjin Normal University, Tianjin, China

was stable regarding noise and missing data. Although the statistics of SAR images imply that there was no basis or dictionary where the data could be assumed sparse, Rilling et al. [7] investigated two simple techniques, post processing and hybrid compressed sensing, to enhance the quality of the reconstructed SAR image in the non-compressible areas which could not be well described by a sparse approximation. Baraniuk [2] used CS ideas for SAR imaging and proposed two potential improvements: (i) eliminated the matched filter (MF) in the radar receiver and (ii) reduced the required sampling rate of the receiver A/D converter. And the author also indicated that there were a number of challengers to be overcome before an actual CS-based radar system would become a reality. Jiao et al. [8] proposed a new pursuit algorithm called Bayesian evolutionary pursuit algorithm for the CS reconstruction problem of SAR images with weak sparsity. To reduce the amount of stored SAR data, Sujit [1] proposed a method based on CS theory. According to the three steps in CS theory, in [3], DWT was utilized to make SAR imagery sparse and the random Gauss matrix after approximate QR decomposition was employed to complete the low-dimension measurement for sparse results based on the original SAR imagery divided into several sub-imageries. Furthermore, a modified OMP algorithm was proposed to reconstruct SAR imagery. In [9], the echo model of SAR imagery with multiple ships was converted into a CS model with orthogonal basis and used the CS-based method to achieve higher azimuth resolution. In [10], a sparse reconstruction method for SAR imaging based on CS theory was presented and compared with the traditional MF-based RD imaging method, the proposed method significantly suppressed the sidelobe and greatly improved the imaging performance of SAR when the target space was sparse.

There are three main procedures should be executed in CS theory which is the sparse expression of original signal, low-dimensional sampling, and precisely reconstruction of original signals. According to the above research achievements in SAR raw data compressing based on CS theory, we can see that the SAR imagery data usually have poor sparsity feature and looking for suitable sparse transformation basis for SAR images is extremely significant. The two-dimensional SAR image formed is interpreted in the directions of range and azimuth. The range is the direction of signal propagation and the azimuth is the direction parallel to the flight path. In this article, the SAR signals are thought as composed with distributed signals of the range samples and principal component analysis (PCA) can be used to find transformations that sparsify the range samples. The remainder of the article is organized as follows. In the following section, the SAR raw data model is analyzed. In

Section "Joint PCA and CS for SAR raw data compression", we summarized a mathematic description of CS and PCA theory and the sparse transform procedure for SAR raw data was proposed. Simulation results are presented in Section "Simulation results and analysis". Finally, the article is ended with some conclusion.

SAR signal model and imaging theory
SAR is a high-resolution imaging radar technology. For the conventional standard SAR imaging, the high-resolution of range direction is obtained by the pulse compression of chirp signal and the high resolution of azimuth direction is received through the synthetic aperture formed by platform movement [11].

SAR signal model
Suppose SAR platform works in stripmap mode. The SAR platform flights along a track parallel to azimuth direction at altitude H with a constant velocity. Assuming the radar transmits a linear frequency modulated (LFM) pulse train as

$$f(t) = \sum_{n=-\infty}^{+\infty} \boldsymbol{rect}(t - nT_r) \exp \qquad (1)$$
$$\cdot \left\{ j2\pi[f_0(t - nT_r)] + \frac{1}{2}\alpha(t - nT_r)^2 \right\}$$

where α is the LFM, f_0 denotes carrier frequency, T_r is pulse repetition time and pulse width is τ, and $\mathrm{rec}(t)$ denotes the unit rectangular function.

So, the SAR echo signal at time t can be expressed as

$$s(t) = \iint_D \sigma\left(x', r'\right) W_\alpha\left(x - x'\right) W_r\left(r - r'\right) \qquad (2)$$
$$\cdot f\left(t - \frac{2r(x', r')}{c}\right) dx' dr'$$

where D denotes the radar beam irradiation area, $\sigma(x, r)$ represents the backscattering coefficients of target point (x, r), $W_\alpha(.)$, $W_r(.)$ are antenna gains of azimuth direction and range direction, respectively, and $r(x, r)$ states the distance between radar platform with target point (x, r).

The actual received echo signals are still pulse train and they have two kinds of time concept. The first one is fast-time which is included in the complex amplitude of the echo signal and reflects the echo signal varied trend. The fast-time signal is the range dimension signal actually and determines the radar range resolution characteristics. The other one is slow-time which is included in signal detention and reflects position changing of radar platform. The slow-time signal is also called azimuth dimension signal and determines the radar azimuth resolution characteristics.

SAR imaging

Assuming N_r and N_a denote sample size of range and azimuth dimension signal, respectively, then the matrix of SAR echo signals can be written as

$$S = [s(1,1), s(1,2), \ldots, s(1,N_r); s(2,1), s(2,2), \ldots, \quad (3)$$

$$\cdot s(2,N_r); \ldots; s(N_a,1), s(N_a,2), \ldots, s(N_a,N_r)]$$

where $s(i,1), s(i,2), \ldots, s(i,N_r)$ denotes the echo signal at time i which are sampled following Shannon's celebrated theorem. SAR imaging is to extract the two-dimensional distribution of target scattering coefficient $\sigma(x, r)$ from echo signal train S. The most universal methods for SAR imaging are based on MF theory to deal with rang and azimuth dimension signal separately, such as Rang-Doppler algorithm, Chirp-Scaling algorithm, Spectral Analysis algorithm, etc. [10]. However, there are many obvious defects in these traditional SAR imaging methods. For instance, according to the Nyquist sampling theory, a large number of echo samplers must be gathered in order to raise SAR range resolution which will lead to an excessive burden of SAR acquisition and storage system. Therefore, how to realized high-resolution SAR from low-dimension samplers is extremely significant. The emergence of CS theory makes it possible.

Joint PCA and CS for SAR raw data compression

In this section, we first review the basic theory of CS and PCA and we subsequently illustrate the sparse transformation procedure for SAR raw data.

Basic theory of CS

Recently, CS has received considerable attention. The theory of CS states that a signal may be captured with a small set of non-adaptive, linear measurements as long as the signal is sparse in some basis (such as DCT, wavelet) [12,13]. We can write in matrix notation:

$$y = \Phi x = \Phi \Psi \theta = \Theta \theta \quad (4)$$

where Φ is an $M \times N$ measurement matrix with $M<N$, $y \in R^M$ is the vector of samples observed. θ is the transform coefficient vector of signal x under orthonormal basis ψ, i.e., $\theta = \Psi^T x$.

Known the observed vector y and the measurement matrix Φ, sparse transformation basis ψ, the recovery of the unknown signal x could be achieved by searching for the l_0-sparsest representation that agrees with the measurements [14]:

$$\hat{\theta} = \arg\min \|\theta\|_0$$

Subject to

$$y = \Phi \Psi \theta \quad (5)$$

where the l_0 pseudo-norm $\| \bullet \|_0$ corresponds to the number of non-zero elements. As it is well known,

this is a daunting NP-complete combinatorial optimization problem which cannot be solved directly in practice. Instead of solving the l_0-minimization problem, non-adaptive CS theory seeks to solve the 'closest possible' tractable minimization problem, i.e., the l_1-minimization:

$$\hat{\theta} = \arg\min \|\theta\|_1$$

Subject to

$$y = \Phi \Psi \theta \quad (6)$$

This modification leads to a much simpler convex problem, but it involves expensive computations when applied to large-scale signals. Therefore, a second approach using iterative greedy methods has been proposed, such as matching pursuit (MP), orthogonal matching pursuit (OMP), StOMP, Subspace Pursuit (SP), and CoSaMP.

PCA

From the recovery algorithm of CS, it can be shown obviously that the appropriate sparse transformation basis ψ can improve recovery precision and reduce the computations. In this article, the PCA is the technique providing the transformation matrix ψ.

PCA is a mathematical procedure that uses an orthogonal transform to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables.

Assuming $x_i \in R^N$ is the samples at discrete time i=1,2,...,K with a fixed sampling rate. The measurements mean vector and covariance matrix of x_i can be defined as

$$\bar{x} = \frac{1}{K} \sum_{i=1}^{K} x_i$$

$$C = \frac{1}{K} \sum_{i=1}^{K} (x_i - \bar{x})(x_i - \bar{x})^T \quad (7)$$

Considered the above equations, the orthonormal matrix P can be constructed whose columns are the unitary eigenvectors of covariance matrix C, placed according to the decreasing order of the corresponding eigen values. Then we can define

$$z_i = P^T (x_i - \bar{x}) \quad (8)$$

where P^T is the inverse matrix of P. Due to the construction of P, we have the entries of z_i are ordered as follows: $z_i(1) \geq z_i(2) \geq \cdots \geq z_i(N)$ If the instances $x_1, x_2, \ldots, x_K$ are correlated, there exists an $M < N$ such that $j > M$ we have

$z_i(j)$ is negligible. In other words, the N-dimensional vector z_i can be considered as an M sparse vector and the value of M depends on the level of correlation among the gathered samples $x_1, x_2, \ldots, x_K$ [15,16].

Joint CS and PCA for compressive SAR imaging

Given the SAR imaging theory and Equation (3), we can rewrite the echo signals as $s_i \in C^{N_r}, i = 1, 2, \ldots, N_a$, $s_i = [s(i,1), s(i,2), \ldots, s(i,N_r)]$. Due to the echo signals reflect the backscatter coefficients of target, the instances $s_1, s_2, \ldots, s_{N_a}$ are temporally correlated, that is, the PCA algorithm can be used to sparsify the echo signal. Considered that the echo signals are complex, we can write the compressive measurement of SAR echo signal through CS and PCA as follows.

$$y_i^{re} = \Phi[\text{Re}(s_i) - \text{Re}(\bar{s})] = \Phi P_{re} \theta_i^{re} \qquad (9)$$

$$y_i^{im} = \Phi[\text{Im}(s_i) - \text{Im}(\bar{s})] = \Phi P_{im} \theta_i^{im} \qquad (10)$$

The real and imaginary part of echo signal is calculated, respectively. Where $\bar{s}$ denotes the mean vector of S_i, y^{re}, y^{im} denote the low-dimensional measurements, P_{re}, P_{im} represent the orthonormal transformation basis constructed based on Equation (7), θ_i^{re} and θ_i^{im} are the sparse coefficients.

As above, we can obtain the recovery sparse vector $\hat{\theta}_i re$ and $\hat{\theta}_i im$ through acquired low-dimensional measurements and recovery algorithms such as OMP. Then the estimations of s_i can be written as follows.

$$\hat{s}_i = \text{Re}(\hat{s}_i) + j\text{Im}(\hat{s}_i)$$
$$= \left[\text{Re}(\bar{s}) + P_{re}\hat{\theta}_i^{re} \right] + j \left[\text{Im}(\bar{s}) + P_{im}\hat{\theta}_i^{im} \right] \qquad (11)$$

In short, we can achieve high-resolution SAR from low-dimension samplers based on joint CS and PCA schemes. Furthermore, it is noted that real SAR echo signals are in general non-stationary. This means that the statistics that we need to use in the PCA must be acquired at runtime and it makes that an online estimation of these parameters possible.

Simulation results and analysis

To illustrate the effectiveness of the proposed joint CS and PCA algorithm, we test its performances on simulated SAR imaging of point target. In the experiments, we chose random Gaussian matrix for measurement matrix Φ and OMP for recovery algorithm. Due to the randomness of measurement matrix Φ, all the experiment results are the mean value of ten times computations. All the computations are run on a 2.60-GHz Inter Core i5 laptop with 4-GB RAM.
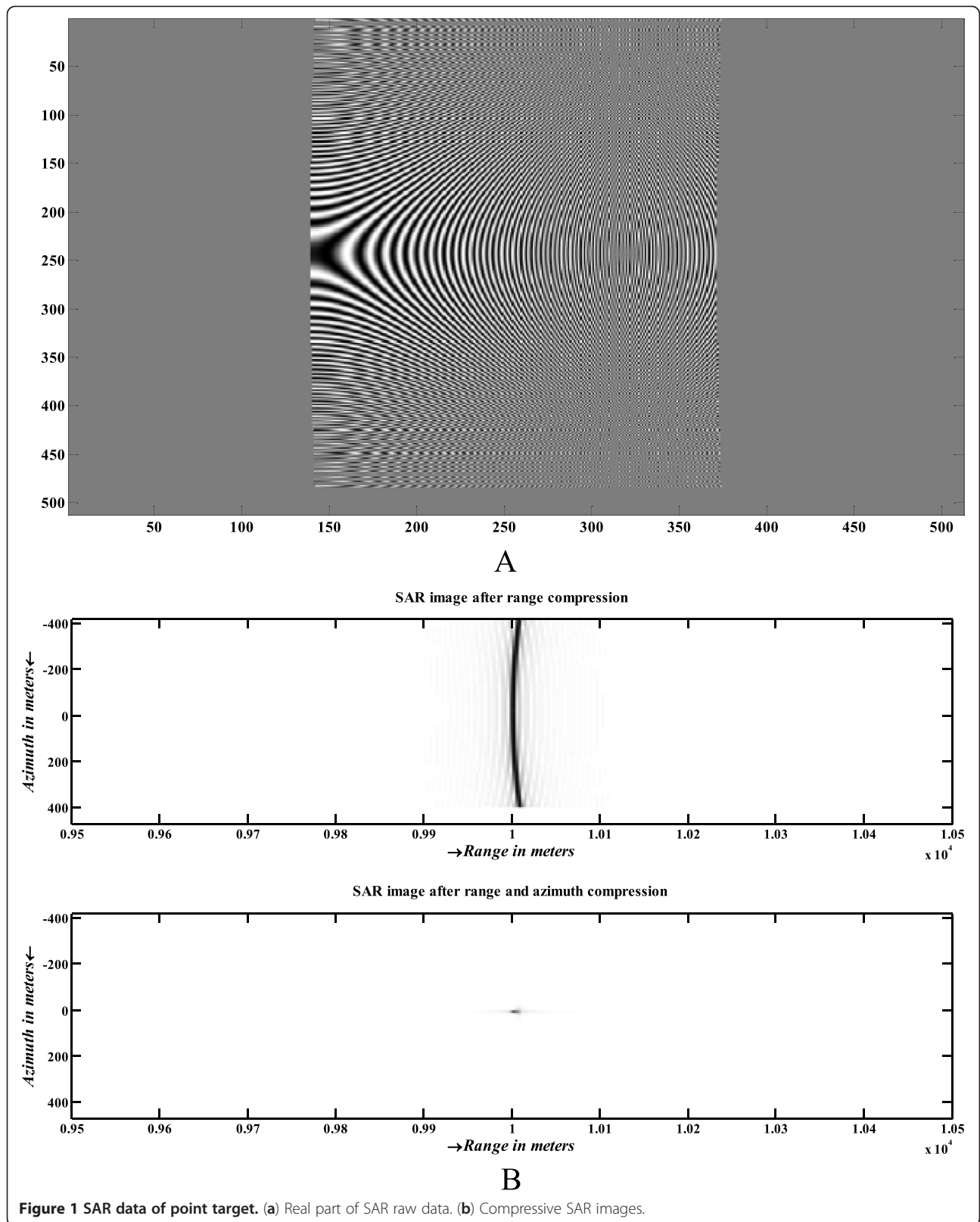
In this section, we simulate the SAR echo signals from a point target with carrier frequency $f_0 = 1 \ \textbf{GHz}$, pulse repetition time $T_r = 0.02 \ s$, and pulse width $\tau = 5 \ \boldsymbol{\mu s}$. In Figure 1, we show the real part of SAR raw data with $N_r = N_a = 512$, SAR image after range compression and SAR image after range and azimuth compression.
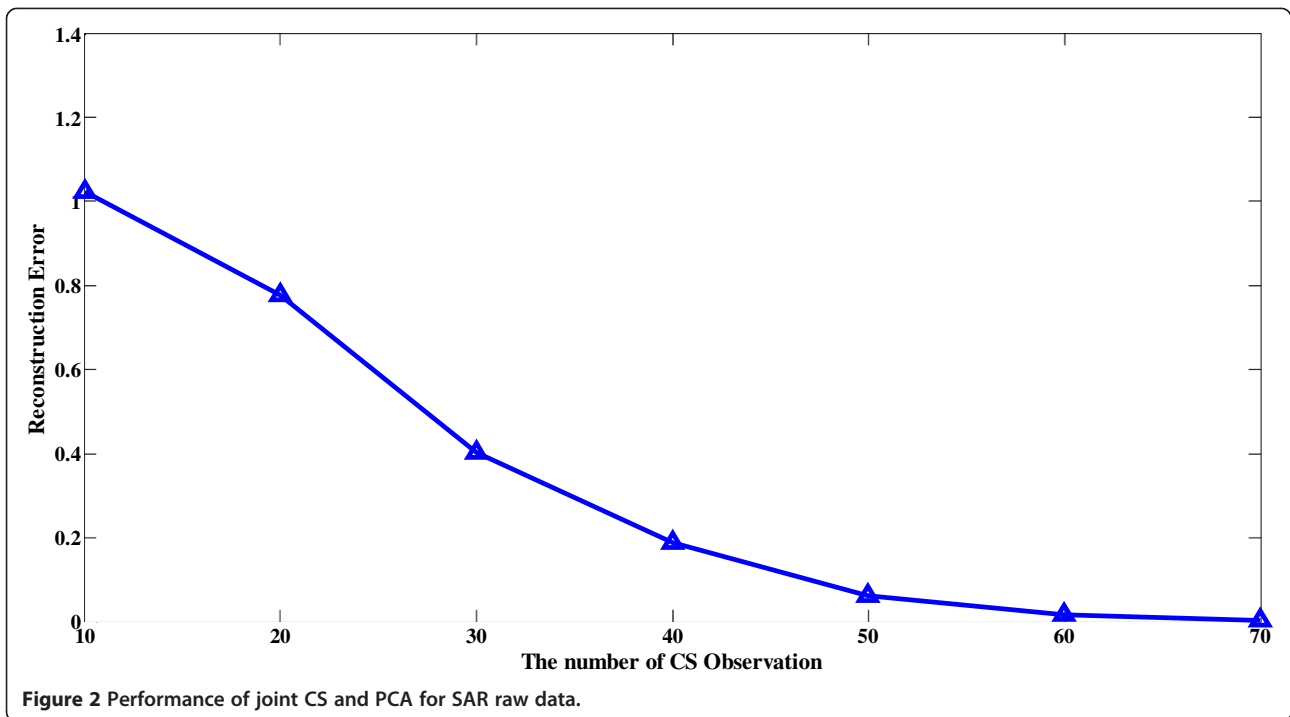
To test the feasibility of the joint CS and PCA scheme, we test the performance of the algorithm based on the above real part of SAR raw data. In Figure 2, it is showed that the trend of reconstruction error versus the number of CS observations. The reconstruction error is defined as $\frac{\epsilon = \|(\text{Re}(S) - \text{Re}(\hat{S}))\|_2}{\|\text{Re}(S)\|_2}$, where S is the original matrix of SAR echo signals according to Equation (3) and $\hat{S}$ is the reconstructed matrix of SAR echo signals on the basis of Equations (3) and (11). The number of CS observations is the amount of random measuring with CS measurement matrix Φ for per echo signal s_i. From Figure 2, it can be seen clearly that we can get high compression effect with CS and PCA and the PCA algorithms have good performance to sparsify the SAR raw data.

As is clear from the PCA procedure, the statistics (sample mean and covariance matrix) must be generated at runtime. Hence, in order to implement joint CS and PCA for real signals, we should implement our algorithm at two phases: training phase and subsequent monitoring phase. In training phase, we collect a series of signals to compute samples mean and orthonormal transformation matrix. Then, in the monitoring phase, it is compressive measurement of subsequent received signals using the statistics computed in the previous phase [15]. To test the effective of the algorithm of the two phases, we simulate the trend of reconstruction error ϵ versus the number of training signals. In Figure 3, with the constant number of CS observation, the simulation results show that the tendency of reconstruction error is decreasing with the increasing amount of training signals. And we can also see that the fluctuation is smaller when a proper amount of training signals is used to compute the statistics. From the experimental results, we can stress that the joint CS and PCA algorithm can readily be applied to online compressively measuring SAR raw data.

Conclusion

In this section, we analyze the structure of SAR imaging data which can be processed through range direction and azimuth direction, separately. And the PCA method is used to sparsify the range direction raw data of SAR and the data are compressed by CS theory. From the simulation results, it demonstrates that the PCA method

Figure 1 SAR data of point target. (**a**) Real part of SAR raw data. (**b**) Compressive SAR images.

**Figure 2 Performance of joint CS and PCA for SAR raw data.**

has good sparse performance for SAR raw data which are temporal correlated in range dimension and the joint CS and PCA algorithm is possible to online compressive measure the SAR raw data.

However, there are some challenges to be overcome before a high efficient joint CS and PCA algorithm for real-time SAR image compression become a reality. First, how to decide an appropriate amount of echo signals to determine the statistics should be analyzed in order to lessen computation burden. Second, for a large number of SAR echo signals, we should decide which part of echo signals is temporal

Figure 3 The outline performance of joint CS and PCA algorithm.
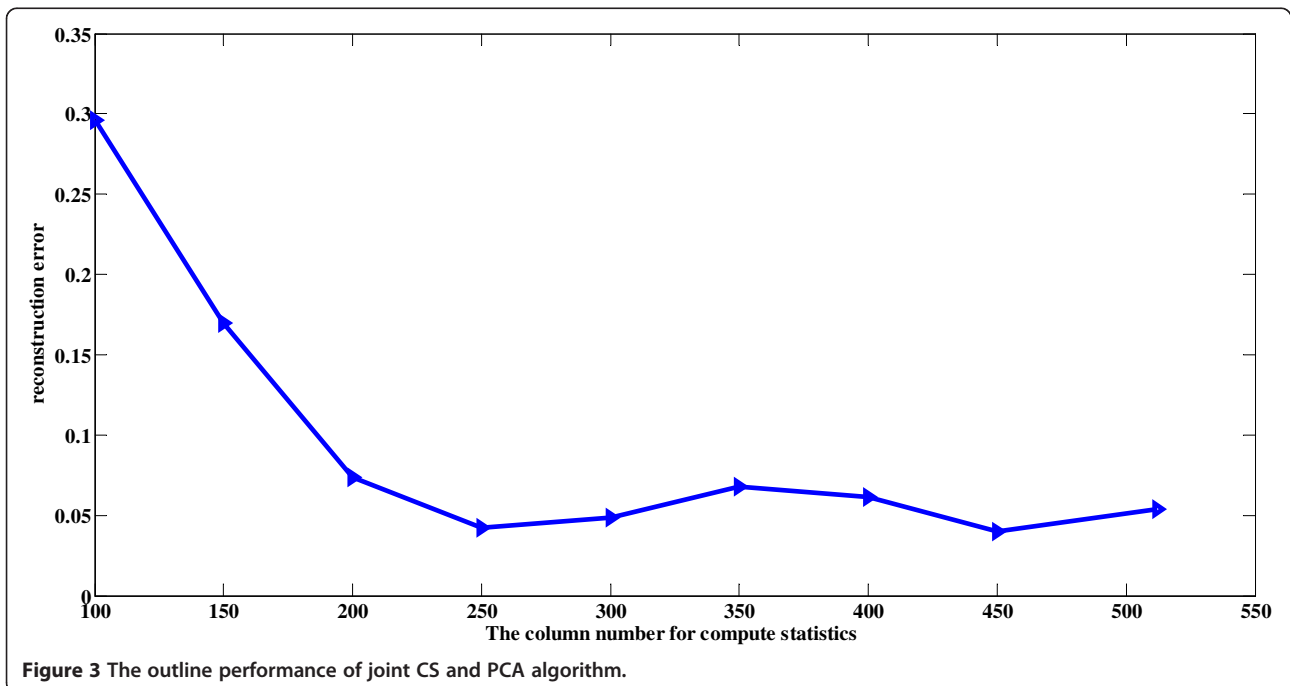
correlated, and it is the premise to use joint CS and PCA algorithm.

Competing interests

The authors declare that they have no competing interests.

References

1. B. Sujit, Fast encoding of synthetic aperture radar raw data using compressed sensing, in *IEEE/SP 14th Workshop on Statistical.* Signal Process **08**, 448–452 (2007)
2. R. Baraniuk, *Compressive radar imaging*, vol. 04 (IEEE Radar Conference, Houston, USA, 2007), pp. 128–133
3. F. Zhu, Q. Zhang, J.-B. Yan et al., *SAR imagery compressing and reconstruction method based on compressed sensing*, vol. 03 (Electromagnetic Research Symposium Proceedings, KL, Malaysia, 2012), pp. 235–239
4. Y. Rivenson, A. Stern, Compressed imaging with a separable sensing operator. IEEE Signal Process. Lett. **16**(6), 449–452 (2009)
5. L. He, L. Carin, Exploiting structure in wavelet-based bayesian compressive sensing. IEEE Trans Signal Process. **57**(9), 3488–3497 (2009)
6. L. Pruente, *Application of compressed sensing to SAR/GMTI-data*, vol. 06 (2010 8th European Conference on Synthetic Aperture Radar, 2010)
7. G. Rilling, M. Davies, B. Mulgrew, *Compressed sensing based compression of SAR raw data*, vol. 03 (Signal Processing with Adaptive Sparse Structured Representations, 2009), pp. 1–6
8. W. Jiao, L. Fang, L.C. Jiao et al., Compressive sensing SAR image reconstruction based on Bayesian framework and evolutionary computation. IEEE Trans. Image Process. **20**(7), 1904–1911 (2011)
9. L. Yabo, Q. Yinghui, L. Jun et al., *SAR imaging of multiple ships based on compressed sensing*, vol. 10 (2nd Asian-Pacific Conference on Synthetic Aperture Radar, 2009), pp. 112–115
10. S.J. Wei, X.L. Zhang, J. Sh et al., Sparse reconstruction for SAR imaging based on compressed sensing. Electromag. Res **109**, 63–81 (2010)
11. M. Soumekh, *Synthetic Aperture Radar Signal Processing with Matlab Algorithms* (Wiley, New York, 1999)
12. E. Candès, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies. IEEE Trans. Inf. Theory **52**, 5406–5425 (2006)
13. E. Candès, M. Wakin, An introduction to compressive sampling-a sensing/sampling paradigm that goes against the common knowledge in data acquisition. IEEE Signal Process. Mag. **25**(2), 21–30 (2008)
14. A. Averbuch, D. Shai, D. Shay, Adaptive compressed image sensing using dictionaries. SIAM J. Imag. Sci. **5**(1), 57–89 (2012)
15. R. Masiero, G. Quer, D. Munaretto et al., *Data acquisition through joint compressive sensing and principal component analysis* (IEEE Globecom 2009, Honolulu, Hawaii, 2009), pp. 1–6
16. R. Masiero, G. Quer, M. Rossi, and M. Zorzi, A Bayesian analysis of compressive sensing data recovery in wireless sensor networks, in *IEEE Ultra Modern Telecommunications and Control Systems ICUMT (SenseApp Workshop),* 2009 ICUMT'09 **10**, 1–6 (2009)