

Article

Distributed Compressed Sensing Based Ground Moving Target Indication for Dual-Channel SAR System

Jing Liu *, Xiaoqing Tian^{ID}, Jiayuan Jiang and Kaiyu Huang^{ID}

Ministry of Education Key Laboratory of Intelligent and Network Security, School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China;
tianxiaoqing2017@stu.xjtu.edu.cn (X.T.); jyjiang@stu.xjtu.edu.cn (J.J.); huangkaiyu@stu.xjtu.edu.cn (K.H.)

* Correspondence: elelj20080730@gmail.com

Received: 2 June 2018; Accepted: 18 July 2018; Published: 21 July 2018



Abstract: The dual-channel synthetic aperture radar (SAR) system is widely applied in the field of ground moving-target indication (GMTI). With the increase of the imaging resolution, the resulting substantial raw data samples increase the transmission and storage burden. We tackle the problem by adopting the joint sparsity model 1 (JSM-1) in distributed compressed sensing (DCS) to exploit the correlation between the two channels of the dual-channel SAR system. We propose a novel algorithm, namely the hierarchical variational Bayesian based distributed compressed sensing (HVB-DCS) algorithm for the JSM-1 model, which decouples the common component from the innovation components by applying variational Bayesian approximation. Using the proposed HVB-DCS algorithm in the dual-channel SAR based GMTI (SAR-GMTI) system, we can jointly reconstruct the dual-channel signals, and simultaneously detect the moving targets and stationary clutter, which enables sampling at a further lower rate in azimuth as well as improves the reconstruction accuracy. The simulation and experimental results show that the proposed HVB-DCS algorithm is capable of detecting multiple moving targets while suppressing the clutter at a much lower data rate in azimuth compared with the compressed sensing (CS) and range-Doppler (RD) algorithms.

Keywords: distributed compressed sensing; Variational Bayesian; dual-channel SAR; ground moving-target indication

1. Introduction

Synthetic aperture radar (SAR) is a kind of high resolution imaging radar, which is capable of working at long distances, all-weather, and day and night. In the SAR system, the scene is observed by an antenna at different positions in the azimuth direction. The coherent information recorded at the different positions is used to synthesize a very long antenna to improve the azimuth resolution. The dual-channel SAR system is widely applied in the field of ground moving-target indication (GMTI). In practice, the Doppler shift of the stationary clutter spreads the Doppler bandwidth of the clutter returns due to the motion of the platform. Accordingly, the Doppler bandwidth of the clutter will mask that of moving targets [1]. Dual-channel SAR can effectively suppress the clutter, which is propitious to slowly moving targets detection [2]. Some traditional approaches, such as the displaced phase center antenna (DPCA) approach, and along-track interferometry (ATI) detection [3], have achieved good results. However, relying on the Nyquist sampling theorem, dual-channel SAR is obliged to transmit and store substantial raw data samples, which limits its application in practice.

Compressed sensing (CS) [4–7], as a novel sparse reconstruction technique, is able to recover the signals that have been sampled below the traditional Nyquist sampling rate. It provides a good solution

for the transmission and storage of substantial raw data samples. CS aims at minimizing the number of measurements and attempts to recover the original signal from these necessary measurements by solving an ℓ_0 -norm optimization problem, or equivalently a convex ℓ_1 -norm optimization problem under the restricted isometry property (RIP) condition [8].

A lot of valuable literature on CS-based SAR imaging and the GMTI system has been published in recent years. A CS-based SAR imaging algorithm is proposed in [9], which demonstrates that CS has the potential to eliminate the need for the pulse compression matched filter at the receiver, and reduce the required receiver analog-to-digital conversion bandwidth. In [10], CS is applied to randomly select lines in azimuth from the signals after range compression, which allows the implementation of wide-swath modes without reducing the azimuth resolution. Sun et al. [11] propose a CS-based method for joint sparse recovery of all channel or sub-aperture images. In [12], a fast CS-based SAR imaging algorithm is proposed to save the computational cost both in time and memory. In [13], a phase error correction method is proposed for the CS-based radar imaging system based on approximated observation, which improves the defocus caused by the phase error. In [14], CS is utilized in tomographic SAR system, extended to the SAR elevation direction for 3-D imaging. In the field of GMTI, in [15] CS is used to estimate the velocities and positions of moving targets. In [16,17], a CS-based space-time adaptive processing (STAP) algorithm is proposed to detect the moving targets. The above CS-based algorithms can efficiently reduce data rate and improve image quality, as well as improve anti-jamming performance. However, for the dual-channel SAR based GMTI (SAR-GMTI) system, with the increase of the imaging resolution, the resulting substantial raw data samples aggravate transmission and storage load. Thus it is required to develop an efficient algorithm that can further reduce the data rate.

In this work, we tackle the problem by exploiting the correlation between two channels of the dual-channel SAR system. The antennas of dual-channel SAR are configured in the along-track direction. Thus the SAR images acquired by each channel are highly correlated during a flight along the track. Using distributed compressed sensing (DCS) based algorithms, we can jointly reconstruct the dual-channel signals, which enables sampling at a further lower rate in azimuth as well as improves the reconstruction accuracy.

DCS [18] exploits both intra-signal and inter-signal correlations. Three different joint sparsity models have been proposed, i.e., joint sparsity model 1 (JSM-1), joint sparsity model 2 (JSM-2) and joint sparsity model 3 (JSM-3). In JSM-1, a sparse signal consists of a common component and an innovation component. Such signals may exist when large-scale phenomena affects all signal sources and local phenomena affects specific individual source. In JSM-2, all signals are constructed from the same sparse set of basis vectors, but with different coefficient values. JSM-3 extends JSM-1 so that the common component needs no longer to be sparse in any basis. In this work, the echo signal of dual-channel SAR is the combination of the signals from both stationary clutter and moving targets. After time delay and phase compensation, the signal from stationary clutter is common to all channels, while the signals from moving targets are specific and different across different channels. Thus the JSM-1 model is adopted, considering that the signal from stationary clutter can be treated as the common component, while those from moving targets can be treated as the innovation components.

Many reconstruction algorithms have been proposed for the JSM-1 model, either in a central manner or a distributed manner. The central recovery methods assume the presence of a fusion center where all information is gathered. The joint recovery strategy (JRS) [18,19] is to stack the common component and all innovation components together as one single sparse signal, and reconstruct the single sparse signal via ℓ_0 -norm minimization or ℓ_1 -norm minimization. There are well developed algorithms such as subspace pursuit (SP), basic pursuit (BP), iterative hard thresholding (IHT), etc., to solve the minimization problem. However, JRS is computationally expensive since the dimension of the sparse signal is multiplied by stacking. As another branch of the methods, the distributed recovery performs the reconstruction in the network, with no fusion center. The Texas Hold 'Em algorithm [20] separates the estimation of the common component and innovation components by the

measurement decomposition process. Valsesia et al. [21] exploit the side information from one of the signals to reconstruct the innovation components. Recently Sundman et al. [22] propose the parallel pursuit with side information (SIPP) algorithm, which extends SP by using the side information of common support set at each iteration. SIPP can reconstruct the signals in a distributed manner, thus achieving much less computing time than JRS. This implies that the information of support set can be utilized to accelerate the reconstruction process. In [23], an alternating direction method of multipliers (ADMM) based algorithm is proposed for the JSM-1 model in a decentralized framework, which exploits the local cooperation between nodes.

In this work, we consider dealing with the JSM-1 model in a different way, i.e., in a Bayesian framework. Bayesian methods essentially approximate the posterior distributions of sparse signals according to the prior knowledge and data, which is effective in dealing with uncertain models and large noises. We consider using the variational Bayesian algorithm for the JSM-1 model. Variational Bayesian (VB) inference [24] obtains the optimal solution by iterating over a set of mutually dependent equations, providing a good approximation to the exact posterior with a readily monitored convergence.

Variational Bayesian has been widely applied in sparse signal reconstruction [25,26]. In this work, we propose a novel algorithm, namely the hierarchical variational Bayesian based distributed compressed sensing (HVB-DCS) algorithm for the JSM-1 model, which decouples the common component from the innovation components by applying variational approximation. The proposed algorithm is capable of approximating the complex signal in a hierarchical framework. Moreover, we apply the proposed HVB-DCS algorithm for signal reconstruction in the dual-channel SAR based GMTI system. Based on the received signal from both the moving targets and the stationary clutter, we decouple the common component (stationary clutter) from the innovation components (moving targets). We derive separable PDF functions for common component and innovation components, respectively. We can then obtain the reconstructed signal (including the common and innovation components) by applying the maximum a posteriori (MAP) estimation. The performance of the algorithm is verified by both simulation on point targets, and experiment on real SAR data.

The main contribution of this paper is threefold. First, we model the received signals of the dual-channel SAR-GMTI system in the JSM-1 model of DCS, by exploiting the correlation between the two channels. Secondly, we propose a novel HVB-DCS algorithm for the JSM-1 model, which decouples the common component from the innovation components by applying variational approximation. The proposed HVB-DCS algorithm approximates the complex signal in a hierarchical framework. Thirdly, we apply the proposed HVB-DCS algorithm for signal reconstruction in the dual-channel SAR-GMTI system. In the framework of the JSM-1 model, the proposed algorithm can simultaneously generate the stationary clutter (common component) and the moving targets (innovation components), which omits the DPCA process adopted in the traditional GMTI system. The simulation and experimental results show that the proposed HVB-DCS algorithm is capable of detecting multiple moving targets while suppressing the clutter at a much lower data rate compared with the CS and range-Doppler (RD) algorithms.

The organization of this paper is as follows. First, we briefly introduce the theory of DCS in Section 2. We then present the signal model of dual-channel SAR system in Section 3. The DCS-based dual-channel SAR-GMTI system is introduced in Section 4, where the echo signals of dual-channel SAR-GMTI system are represented in the JSM-1 model. The HVB-DCS algorithm is proposed in Section 5, which decouples the common component from the innovation components by applying variational approximation, in a hierarchical framework. The simulation and experimental results are shown in Section 6, and the conclusions are summarized in Section 7.

2. Brief Introduction to Distributed Compressed Sensing

CS [4] is an emerging field for signal and image processing. CS aims to recover a sparse signal with a reduced number of linear measurements, which is significantly less than that required by the

traditional bandwidth constraint based on the Shannon-Nyquist sampling theorem. CS theory has been widely applied in various fields, such as magnetic resonance imaging [27], target tracking [28], biomedical monitoring [29], computer-generated holography [30], machine learning [31], signal acquisition in wireless sensor networks [32], image/video compression [33], etc.

As an important branch of CS, DCS [18] considers a scenario where correlated sparse signals are captured from multiple sources. DCS aims to exploit the inter-signal and intra-signal correlations of those sparse signals, so that the ensemble of signals can be reconstructed from fewer measurements than the standard CS approach requires. Baron et al. in [18] propose three various joint sparsity models (JSMs) for the correlated signals, namely JSM-1, JSM-2 and JSM-3. Among them, JSM-1 is capable of modeling a wide range of scenarios where large-scale phenomena affect all signal sources and local phenomena affect specific individual source. In this work, we focus on the JSM-1 model.

2.1. Joint Sparsity Model-1

In the JSM-1 framework, an ensemble of vectors (signals) $\{\mathbf{x}_j\}_{j=1}^J$ are jointly sparse if each vector is sparse and comprised of a common component and an innovation component, as

$$\mathbf{x}_j = \mathbf{z}_c + \mathbf{z}_j, j \in \{1, 2, \dots, J\}, \quad (1)$$

where $\mathbf{z}_c \in \mathbb{R}^N$ denotes the common component of sparse vector \mathbf{x}_j , and $\mathbf{z}_j \in \mathbb{R}^N$ denotes the innovation component of \mathbf{x}_j . Both \mathbf{z}_c and \mathbf{z}_j are supposed to be sparse.

The DCS problem consists in reconstructing jointly sparse vectors from their measurements, which are obtained by

$$\mathbf{y}_j = \mathbf{A}_j \mathbf{x}_j, \quad (2)$$

where $\mathbf{A}_j \in \mathbb{R}^{M_j \times N}$ is the measurement matrix, $\mathbf{y}_j \in \mathbb{R}^{M_j}$ is the measurement vector, and M_j is the number of measurements.

2.2. The Joint Recovery Strategy for the JSM-1 Model

JRS is the first proposed algorithm for the JSM-1 model. The main idea is to stack the common component \mathbf{z}_c and all innovation components $\{\mathbf{z}_j\}_{j=1}^J$ together to form one sparse vector $\tilde{\mathbf{z}}$ with multiple dimension $(J+1)N$, and the similar process for measurement vectors $\{\mathbf{y}_j\}_{j=1}^J$ and measurement matrices $\{\mathbf{A}_j\}_{j=1}^J$, as

$$\tilde{\mathbf{z}} := \begin{bmatrix} \mathbf{z}_c \\ \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_J \end{bmatrix}, \tilde{\mathbf{y}} := \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_J \end{bmatrix}, \tilde{\mathbf{A}} := \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_J & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_J \end{bmatrix}. \quad (3)$$

The J measurement Equations (2) can be formulated in one compact form as

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}} \tilde{\mathbf{z}}, \quad (4)$$

resulting in the formulation of a standard CS problem. Then the sparse vector $\tilde{\mathbf{z}}$ is reconstructed based on Equation (4) via well developed algorithms of CS, such as orthogonal matching pursuit (OMP) and BP. Finally, the jointly sparse vectors $\{\mathbf{x}_j\}_{j=1}^J$ are reconstructed by adding the reconstructed common component to each reconstructed innovation component.

However, the computational costs of JRS for the JSM-1 model increase dramatically when the dimension of sparse vector increases. So it is necessary to develop another more efficient recovery strategy for the JSM-1 model.

3. Signal Model of the Dual-Channel SAR System

This section introduces the signal model of a dual-channel SAR system, which is the same as that in [34]. The readers can refer to [34] for more details about the signal model. The geometry relationship between the flying platform and moving target is shown in Figure 1, where v and H denote velocity and height of the platform respectively. Channel one periodically transmits and receives pulses, while channel two receives pulses. The distance between the two channels is d . t_m is the slow time in azimuth. PRF represents the pulse repetition frequency. The target is denoted by P . R_B denotes the nearest slant range between the platform and target, whereas $R_1(t_m)$ is the instantaneous slant range between channel one and the moving target, and the similar definition for $R_2(t_m)$.

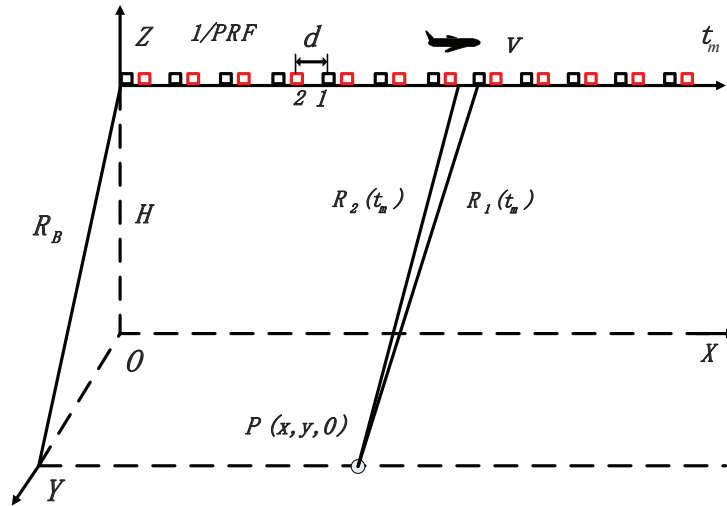


Figure 1. Geometry of the dual-channel SAR system.

In order to facilitate the analysis, the slant range history geometry of the moving target is shown in Figure 2. v_a and v_r denote the along-track and cross-track velocities of the moving target on the slant range plane, respectively. The instantaneous slant ranges $R_1(t_m)$ and $R_2(t_m)$ are respectively defined as

$$R_1(t_m; R_B) = \sqrt{(vt_m - v_a t_m)^2 + (R_B - v_r t_m)^2} \approx R_B - v_r t_m + \frac{(v - v_a)^2}{2R_B} t_m^2, \quad (5)$$

and

$$R_2(t_m; R_B) = \sqrt{(vt_m - v_a t_m - d)^2 + (R_B - v_r t_m)^2} \approx R_B - v_r t_m + \frac{[(v - v_a)t_m - d]^2}{2R_B}, \quad (6)$$

where

$$t_m \in \{t_1, t_2, \dots, t_i, \dots, t_M\} - \frac{t_1 + t_M}{2},$$

and M denotes the number of pulses transmitted by antennas.

For channel one, the received signal of moving target after demodulation and range compression can be expressed as

$$s_1(t, t_m) = \sigma G \text{sinc} \left(\Delta B \left(t - \frac{2R_1(t_m; R_B)}{c} \right) \right) \omega_a(t_m) \exp \left(-j \frac{4\pi}{\lambda} R_1(t_m; R_B) \right). \quad (7)$$

The received signal for channel two is similarly defined as

$$s_2(t, t_m) = \sigma G \text{sinc} \left(\Delta B \left(t - \frac{R_1(t_m; R_B) + R_2(t_m; R_B)}{c} \right) \right) \omega_a(t_m) \exp \left(-j \frac{2\pi}{\lambda} (R_1(t_m; R_B) + R_2(t_m; R_B)) \right). \quad (8)$$

In the above expressions, t is the fast time in range, σ denotes the complex reflectivity of the target, and G is the range compression gain. ΔB is the bandwidth of transmitted signal. $\text{sinc}(\cdot)$ is the envelope after range compression. $\omega_a(\cdot)$ is the azimuth windowing function. c and λ represent the speed of light and the wavelength of transmitted signal, respectively.

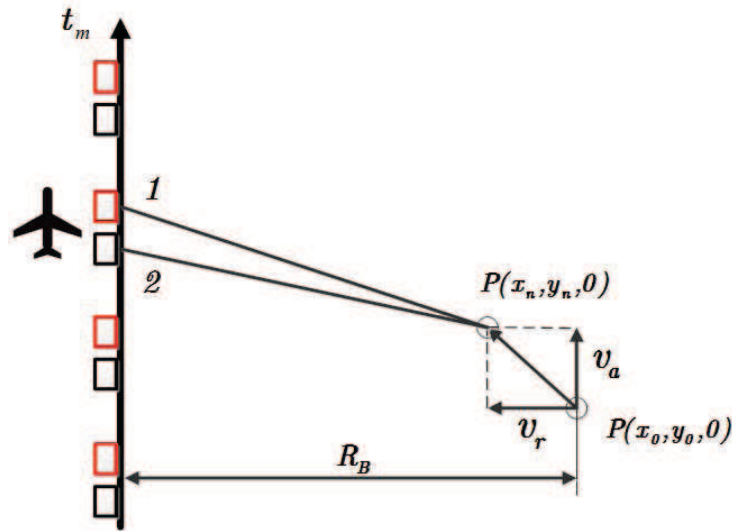


Figure 2. Slant range history geometry of the moving target.

4. Distributed Compressed Sensing Based Dual-Channel SAR-GMTI System

4.1. Sparse Representation for Individual Channel

In this section, the random sampling mode in the azimuth direction [34] is adopted for the dual-channel SAR system, which is shown in Figure 3. In this mode, the pulses in azimuth are randomly transmitted and received, which results in time gaps within a coherent processing interval (CPI) where no echoes are recorded. Moreover, we assume that the sensors remain still between transmission and reception of a pulse.

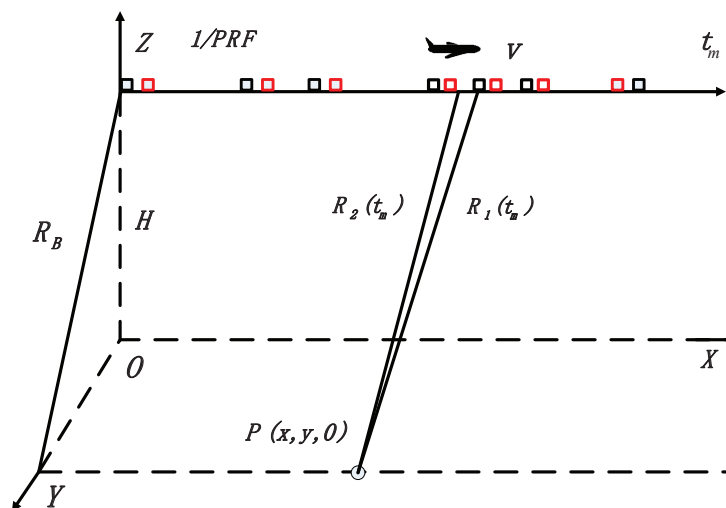


Figure 3. Random sampling mode in the azimuth direction.

For channel one, we can obtain Equation (9) by substituting Equation (5) into Equation (7).

$$s_1(t, t_m) = \sigma G \text{sinc} \left(\Delta B \left(t - \frac{2R_B}{c} \right) \right) \omega_a(t_m) \exp \left(-j \frac{4\pi}{\lambda} R_B \right) \exp \left(j 2\pi f_{dc} t_m + j \pi \gamma_m(R_B) t_m^2 \right), \quad (9)$$

where $f_{dc} = 2v_r/\lambda$ is the Doppler centroid frequency and $\gamma_m(R_B) = -2(v - v_a)^2/(\lambda R_B)$ is the corresponding Doppler chirp rate. Thus Equation (9) can be expanded as

$$s_1(t, t_m) = \sigma G \text{sinc} \left(\Delta B \left(t - \frac{2R_B}{c} \right) \right) \omega_a(t_m) \exp \left(-j \frac{4\pi}{\lambda} R_B \right) \exp \left(j 2\pi \frac{v_r^2 R_B}{\lambda (v - v_a)^2} \right) \times \exp \left(-j 2\pi \frac{(v - v_a)^2}{\lambda R_B} \left(t_m - \frac{v_r R_B}{(v - v_a)^2} \right)^2 \right). \quad (10)$$

Assuming that the cross-track velocity of the moving point target is relatively slow, the range migration through resolution cells does not take place.

Remark: If the cross-track velocity of the moving target is high enough, the Doppler centroid frequency f_{dc} may exceed the limit of PRF, which induces the Doppler ambiguity problem and an additional range walk [35]. For the condition of high cross-track velocity, an additional range walk correction step is required after range compression. As the cross-track velocity is unknown, it is required to search the Doppler ambiguity number and estimate the cross-track velocity. In recent years, many algorithms have been proposed to estimate the Doppler ambiguity number and the cross-track velocity, e.g., the road slope-aided algorithm [36] and the range walk correcting-based algorithm [37]. In the future work, we will combine the proposed HVB-DCS algorithm and the above-mentioned estimation algorithms to tackle the high cross-track velocity problem.

The received signal at a given range bin data, after demodulation and range compression, can be rewritten as

$$s_1(t_m) = \rho^{(1)} \omega_a(t_m) \exp \left(-j 2\pi \frac{(v - v_a)^2}{\lambda R_B} \left(t_m - \frac{v_r R_B}{(v - v_a)^2} \right)^2 \right), \quad (11)$$

where

$$\rho^{(1)} = \sigma G \text{sinc} \left(\Delta B \left(t - \frac{2R_B}{c} \right) \right) \exp \left(-j \frac{4\pi}{\lambda} R_B \right) \exp \left(j 2\pi \frac{v_r^2 R_B}{\lambda (v - v_a)^2} \right). \quad (12)$$

We assume that there are no more than N_0 scattering centers which can be distinguished in the synthetic aperture time T_0 . The coordinates of these scatterers in the azimuth direction are $x_n (n = 1, 2, \dots, N_0)$. The received signal at a given range bin data, after demodulation and range compression, can then be represented as a combination of the echo signals from N_0 scattering centres as,

$$s_1(t_m) = \sum_{n=1}^{N_0} \rho_n^{(1)} \omega_a(t_m - x_n/v) \exp \left(-j 2\pi \frac{(v - v_a)^2}{\lambda R_B} \left(t_m - x_n/v - \frac{v_r R_B}{(v - v_a)^2} \right)^2 \right), \quad (13)$$

where

$$\rho_n^{(1)} = \sigma_n G \text{sinc} \left(\Delta B \left(t - \frac{2R_B}{c} \right) \right) \exp \left(-j \frac{4\pi}{\lambda} R_B \right). \quad (14)$$

The along-track velocity of the target defocuses the image, thus in order to facilitate the analysis, we only consider the cross-track velocity and let $v_a = 0$.

Next, we represent the signal received by channel one in a standard CS framework. We first build the measurement matrix according to Equation (13), as

$$\Phi_1 = \left[\mathbf{s}_0\left(-\frac{N}{2} + 1\right), \mathbf{s}_0\left(-\frac{N}{2}\right), \dots, \mathbf{s}_0(0), \dots, \mathbf{s}_0\left(\frac{N}{2} - 1\right), \mathbf{s}_0\left(\frac{N}{2}\right) \right]_{M \times N}, \quad (15)$$

where,

$$\mathbf{s}_0(i) = [s_0(t_1 - i\Delta\tau), s_0(t_2 - i\Delta\tau), \dots, s_0(t_M - i\Delta\tau)]_M^T, \quad (16)$$

$$s_0(t_m - i\Delta\tau) = \begin{cases} \exp\left(j\pi\gamma_m(R_B)(t_m - i\Delta\tau)^2\right) & , \quad |t_m - i\Delta\tau| \leq T/2, \\ 0 & , \quad |t_m - i\Delta\tau| > T/2, \end{cases} \quad (17)$$

$$i \in \left\{ -\frac{N}{2} + 1, \dots, 0, \dots, \frac{N}{2} \right\},$$

$\gamma_m(R_B) = -2(v - v_a)^2 / (\lambda R_B)$, T is the full synthetic aperture time, $\Delta\tau = 1/\text{PRF}$, and $N \geq (T_0 + T)/\Delta\tau$. N is the number of samples in the azimuth direction. M is far less than N . The measurement matrix captures the contribution to the received signal of a point target.

We then define the measurement vector $\mathbf{s}_1 \in \mathcal{C}^M$ as the received signal for a given range bin after range compression by channel one, $\mathbf{s}_1 = [s_1(t_1), s_1(t_2), \dots, s_1(t_M)]^T$, and the sparse vector $\boldsymbol{\rho}_1 \in \mathcal{C}^N$ as the complex image for a given range bin, $\boldsymbol{\rho}_1 = [\rho_1^{(1)}, \rho_2^{(1)}, \dots, \rho_N^{(1)}]^T$. Thus we can obtain the standard equation in CS for channel one as

$$\mathbf{s}_1 = \Phi_1 \boldsymbol{\rho}_1. \quad (18)$$

Similarly, for channel two, Equation (6) is substituted into Equation (8), resulting in

$$s_2(t, t_m) = \sigma G \text{sinc}\left(\Delta B \left(t - \frac{2R_B}{c}\right)\right) \omega_a(t_m) \exp\left(-j\frac{4\pi}{\lambda} R_B\right) \exp\left(j2\pi \frac{v_r^2 R_B}{\lambda(v - v_a)^2}\right) \exp\left(-j\pi \frac{d^2}{2\lambda R_B}\right) \\ \times \exp\left(-j2\pi \frac{(v - v_a)^2}{\lambda R_B} \left(t_m - \frac{v_r R_B}{(v - v_a)^2} - \frac{d}{2(v - v_a)}\right)^2\right) \exp\left(j\frac{2\pi}{\lambda} \frac{v_r d}{(v - v_a)}\right). \quad (19)$$

Comparing Equation (19) (for channel two) with Equation (10) (for channel one), we can find that there is a phase difference $\exp(-j\pi d^2 / (2\lambda R_B))$ and a time delay $d / (2(v - v_a))$ between the two channels. Thus the measurement matrix for channel two can be represented as

$$\Phi_2 = \left[\mathbf{s}_0' \left(-\frac{N}{2} + 1\right), \mathbf{s}_0' \left(-\frac{N}{2}\right), \dots, \mathbf{s}_0'(0), \dots, \mathbf{s}_0' \left(\frac{N}{2} - 1\right), \mathbf{s}_0' \left(\frac{N}{2}\right) \right]_{M \times N}, \quad (20)$$

where,

$$\mathbf{s}_0'(i) = [s_0'(t_1 - i\Delta\tau), s_0'(t_2 - i\Delta\tau), \dots, s_0'(t_M - i\Delta\tau)]_M^T, \quad (21)$$

$$s_0'(t_m - i\Delta\tau) = \begin{cases} \exp\left(j\pi\gamma_m(R_B) \left(t_m - i\Delta\tau - \frac{d}{2(v - v_a)}\right)^2\right) & , \quad \left|t_m - i\Delta\tau - \frac{d}{2(v - v_a)}\right| \leq T/2, \\ 0 & , \quad \left|t_m - i\Delta\tau - \frac{d}{2(v - v_a)}\right| > T/2, \end{cases} \quad (22)$$

$$i \in \left\{ -\frac{N}{2} + 1, \dots, 0, \dots, \frac{N}{2} \right\}.$$

We define the measurement vector $\mathbf{s}_2 \in \mathcal{C}^M$ as the received signal for a given range bin after range compression by channel two, $\mathbf{s}_2 = [s_2(t_1), s_2(t_2), \dots, s_2(t_M)]^T$ and the sparse vector $\boldsymbol{\rho}_2 \in \mathcal{C}^N$ as $\boldsymbol{\rho}_2 = [\rho_1^{(2)}, \rho_2^{(2)}, \dots, \rho_N^{(2)}]^T$. Thus, we obtain the standard equation in CS for channel two as

$$\mathbf{s}_2 = \Phi_2 \boldsymbol{\rho}_2. \quad (23)$$

As we can see, the measurement matrices are built using the echo signal models of moving targets as in Equations (13) and (19). In practice, it is difficult to construct the measurement matrices since the along-track velocity v_a and cross-track velocity v_r of the moving target are unknown. Thus, in order to facilitate the analysis and construction of the measurement matrices, we set the values of v_a and v_r to zero. The effects of the nonzero cross-track velocity and nonzero along-track velocity on SAR imaging are analyzed in detail through simulations performed in Section 6.1.

4.2. JSM-1 Model

For the i^{th} scattering center, its corresponding reflection coefficients in channel one and channel two can be respectively represented as

$$\rho_i^{(1)} = \sigma_i \text{Gsinc} \left(\Delta B \left(t - \frac{2R_B}{c} \right) \right) \exp \left(-j \frac{4\pi}{\lambda} R_B \right) \exp \left(j \frac{2\pi v_r^2 R_B}{\lambda v^2} \right), \quad (24)$$

and

$$\begin{aligned} \rho_i^{(2)} &= \sigma_i \text{Gsinc} \left(\Delta B \left(t - \frac{2R_B}{c} \right) \right) \exp \left(-j \frac{4\pi}{\lambda} R_B \right) \exp \left(j \frac{2\pi v_r^2 R_B}{\lambda v^2} \right) \\ &\times \exp \left(-j \pi \frac{d^2}{2R_B \lambda} \right) \exp \left(j \frac{2\pi}{\lambda} \left(v_r \frac{d}{v} \right) \right). \end{aligned} \quad (25)$$

Equations (24) and (25) demonstrate the close connection between the two channels, i.e., except the items of a fixed phase difference $\exp(-j\pi d^2 / (2\lambda R_B))$, and a varying phase difference $\exp(j \frac{2\pi}{\lambda} (v_r \frac{d}{v}))$ changing with cross-track velocity in Equation (25), the other items on the reflection coefficients of the two channels are the same.

First, we consider compensating for the fixed phase difference on the i^{th} reflection coefficient of channel two, $\rho_i^{(2)}$.

$$\begin{aligned} \rho_i^{(2)'} &= \rho_i^{(2)} \exp \left(j \pi d^2 / 2\lambda R_B \right) = \sigma_i \text{Gsinc} \left(\Delta B \left(t - \frac{2R_B}{c} \right) \right) \\ &\cdot \exp \left(-j \frac{4\pi}{\lambda} R_B \right) \exp \left(j \frac{2\pi v_r^2 R_B}{\lambda v^2} \right) \exp \left(j \frac{2\pi}{\lambda} \left(v_r \frac{d}{v} \right) \right), i = 1, \dots, N. \end{aligned} \quad (26)$$

The compensating procedure in Equation (26) is repeated for N reflection coefficients of channel two, resulting in a new sparse vector, i.e., $\rho_2' = [\rho_1^{(2)'}, \rho_2^{(2)'}, \dots, \rho_N^{(2)'}]^T$, namely the compensated reflection coefficient vector. Furthermore, we can obtain the relationships between the reflection coefficient vectors of channel two, respectively after and before the compensating procedure, together with the reflection coefficient vector of channel one, as

$$\begin{bmatrix} \rho_1 \\ \rho_2' \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \exp \left(j \pi \frac{d^2}{2\lambda R_B} \right) \mathbf{I} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}, \quad (27)$$

where \mathbf{I} denotes an identity matrix of appropriate size.

Secondly, we build the JSM-1 model based on the original reflection coefficient vector of channel one, ρ_1 , and the compensated reflection coefficient vector of channel two, ρ_2' . The nonzero elements of ρ_1 consist of the reflection coefficients from both the stationary clutter and moving targets, which are denoted as $\rho_s^{(1)}$ and $\rho_m^{(1)}$, respectively. Similarly, the nonzero elements of ρ_2' consist of the reflection coefficients from both the stationary clutter and moving targets, which are denoted as $\rho_s^{(2)'}$ and $\rho_m^{(2)'}$, respectively. On the one hand, the reflection coefficients of stationary clutter are common to both ρ_1 and ρ_2' , since $v_r = 0$ and we have $\rho_s^{(1)} = \rho_s^{(2)'}$. On the other hand, the reflection coefficients of moving targets are different for channel one and channel two, since $v_r \neq 0$, thus $\rho_m^{(1)} \neq \rho_m^{(2)'}$. Let \mathbf{z}_c denote

the vector consisting of the reflection coefficients of stationary clutter, \mathbf{z}_1 and \mathbf{z}_2 denote the vectors consisting of the reflection coefficients of moving targets observed in channel one and channel two, respectively. Then we have

$$\begin{bmatrix} \rho_1 \\ \rho_2' \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{z}_c \\ \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}. \quad (28)$$

We define $\mathbf{y} = [\mathbf{s}_1^T \quad \mathbf{s}_2^T]^T$, and can obtain

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} \Phi_1 & \mathbf{0} \\ \mathbf{0} & \Phi_2 \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2' \end{bmatrix} = \begin{bmatrix} \Phi_1 & \mathbf{0} \\ \mathbf{0} & \Phi_2 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \exp\left(-j\pi\frac{d^2}{2\lambda R_B}\right)\mathbf{I} \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2' \end{bmatrix} \\ &= \begin{bmatrix} \Phi_1 & \mathbf{0} \\ \mathbf{0} & \Phi_2 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \exp\left(-j\pi\frac{d^2}{2\lambda R_B}\right)\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{z}_c \\ \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} \\ &= \begin{bmatrix} \Phi_1 & \mathbf{0} \\ \Phi_2 \exp\left(-j\pi\frac{d^2}{2\lambda R_B}\right) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{z}_c \\ \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}. \end{aligned} \quad (29)$$

In Equation (29), the first equality is based on Equations (18) and (23), and the second equality holds considering the inverse of Equation (27).

Equation (29) is a standard JSM-1 model. By jointly reconstructing the common and innovation components of the above, the moving targets \mathbf{z}_1 and \mathbf{z}_2 can be separated. However, the computational costs of some existing algorithms for solving the above model increase dramatically when the dimension of sparse vector increases. So we develop a more efficient recovery strategy for this JSM-1 model.

5. The Hierarchical Variational Bayesian Based DCS Algorithm

5.1. Proposed Algorithm

In Bayesian modeling, all unknowns are treated as stochastic quantities with assigned probability distributions. Considering the effects of noise, the standard JSM-1 model can be presented as

$$\mathbf{y}_k = \mathbf{A}_k(\mathbf{z}_c + \mathbf{z}_k) + \mathbf{v}_k, \quad (30)$$

where $\mathbf{y}_k \in \mathbb{R}^M$, $\mathbf{A}_k \in \mathbb{R}^{M \times N}$ and $\mathbf{v}_k \in \mathbb{R}^M$ denote the measurement vector, the measurement matrix and noise vector, of channel k , respectively; $\mathbf{z}_c, \mathbf{z}_k \in \mathbb{R}^N$ denote the common component and the innovation component respectively.

By applying a suitable prior distribution to \mathbf{z}_c and \mathbf{z}_k , the sparsity can be guaranteed. However, it is important to allow the flexibility to model local characteristics of the signal, as the simple stationary sparse prior distribution is unable to meet the demand. For this reason, we propose an HVB-DCS algorithm for the JSM-1 model.

We adopt zero-mean Gaussian prior distributions for the common component and innovation components, respectively, which are given as

$$p(\mathbf{z}_c | \alpha_c) = \prod_{n=1}^N \mathcal{N}(z_c^n | 0, (\alpha_c^n)^{-1}), \quad (31)$$

and

$$p(\mathbf{z}_k | \alpha_k) = \prod_{n=1}^N \mathcal{N}(z_k^n | 0, (\alpha_k^n)^{-1}), \quad (32)$$

where z_c^n, z_k^n, α_c^n , and α_k^n are the n^{th} element of $\mathbf{z}_c, \mathbf{z}_k, \alpha_c$, and α_k , respectively. The precision parameters α_c and α_k are constrained by treating them as random variables, with Gamma prior distributions as

$$p(\alpha_c; a, b) = \prod_{n=1}^N \text{Gamma}(\alpha_c^n | a, b), \quad (33)$$

and

$$p(\alpha_k; c_k, d_k) = \prod_{n=1}^N \text{Gamma}(\alpha_k^n | c_k, d_k). \quad (34)$$

Furthermore, we assume that the m^{th} element of noise vector, $v_k^m (m = 1, \dots, M)$, obeys an independent and identically distributed (i.i.d) zero-mean Gaussian distribution with inverse variance β , i.e., $v_k^m \sim \mathcal{N}(v_k^m | 0, \beta^{-1})$. Thus we can obtain the posterior distribution of noise vector \mathbf{v}_k as the product of that of the individual element as,

$$p(\mathbf{v}_k | \beta) = \prod_{m=1}^M \mathcal{N}(v_k^m | 0, \beta^{-1}). \quad (35)$$

We further assume a Gamma distribution as prior for the noise inverse variance β

$$p(\beta; e, f) = \text{Gamma}(\beta | e, f), \quad (36)$$

where e and f represent the shape and scale parameters of the Gamma distribution, respectively.

We define $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$, $\mathbf{Z} = \{\mathbf{z}_c, \mathbf{z}_1, \dots, \mathbf{z}_k, \alpha_c, \dots, \alpha_k, \beta\}$, and $\boldsymbol{\theta} = \{a, b, c_1, \dots, c_k, d_1, \dots, d_k, e, f\}$ as the sets of measurement vectors, the hidden variables, and the hyperparameters of the imposed prior, respectively. Figure 4 describes the relationships among the measurements (indicated as a doubly circled node), the hidden variables (indicated as single circled nodes) and the hyperparameters (indicated as square nodes). The directed edges of the graphical model represent the dependencies among the variables. For instance, the measurements \mathbf{Y} relies on the hidden variables $\mathbf{z}_c, \mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k$ and β , while the hidden variable \mathbf{z}_c stochastically depends on the hidden variable α_c , and α_c further relies on the model parameters a, b .

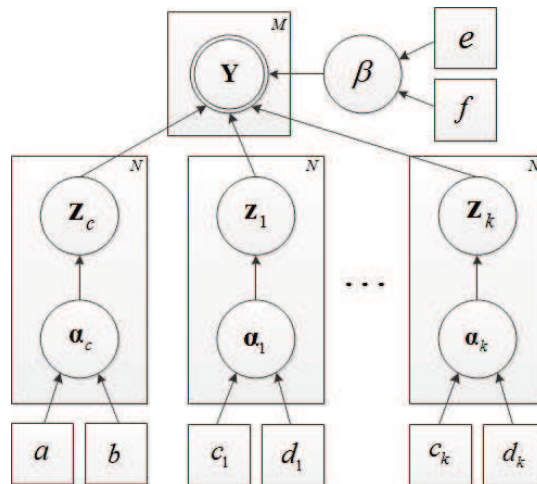


Figure 4. Graphical model for the JSM-1 model. Doubly circled node represents measurements, while single circled nodes represent hidden variables. Nodes denoted with squares correspond to hyperparameters. The directed edges represent the dependencies among the variables.

Next, we are to pursuit the posterior distribution of the hidden variable, $p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta})$. In [24], VB introduces a variational distribution $q(\mathbf{Z})$ to approximate the true posterior distribution $p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta})$. First, the log marginal likelihood $\log p(\mathbf{Y};\boldsymbol{\theta})$ can be represented as

$$\log p(\mathbf{Y};\boldsymbol{\theta}) = F(q(\mathbf{Z}),\boldsymbol{\theta}) + \text{KL}(q(\mathbf{Z})||p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta})), \quad (37)$$

where $F(q(\mathbf{Z}),\boldsymbol{\theta})$ is the free energy,

$$F(q(\mathbf{Z}),\boldsymbol{\theta}) = \int q(\mathbf{Z}) \log \left(\frac{p(\mathbf{Z},\mathbf{Y};\boldsymbol{\theta})}{q(\mathbf{Z})} \right) d\mathbf{Z}, \quad (38)$$

and $\text{KL}(q(\mathbf{Z})||p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta}))$ is the Kullback-Leibler (KL) divergence between the true posterior $p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta})$ and the variational distribution $q(\mathbf{Z})$,

$$\text{KL}(q(\mathbf{Z})||p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta})) = \int q(\mathbf{Z}) \log \left(\frac{q(\mathbf{Z})}{p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta})} \right) d\mathbf{Z}. \quad (39)$$

The goal is to approximate the true posterior distribution by minimizing the KL. Due to the fact that $\text{KL}(q(\mathbf{Z})||p(\mathbf{Z}|\mathbf{Y};\boldsymbol{\theta})) \geq 0$, that objective can be achieved by maximizing $F(q(\mathbf{Z}),\boldsymbol{\theta})$. For the JSM-1 model, we should not only find separable functions that approximate the posterior distribution of \mathbf{z}_c and \mathbf{z}_k , but also make the integral $F(q(\mathbf{Z}),\boldsymbol{\theta})$ tractable.

In order to meet the requirements, we can assume that $q(\mathbf{Z})$ has a factorized form Equation (40). This factorized form stems from theoretical physics where it is called mean field theory [38].

$$q(\mathbf{Z}) = q(\mathbf{z}_c)q(\mathbf{z}_1) \cdots q(\mathbf{z}_k)q(\boldsymbol{\alpha}_c)q(\boldsymbol{\alpha}_1) \cdots q(\boldsymbol{\alpha}_k)q(\boldsymbol{\beta}). \quad (40)$$

Optimizing the free energy in Equation (38) is realized by taking functional derivatives with respect to each of $q(\cdot)$ distributions while fixing the other distributions and setting $\partial F(q)/\partial q(\cdot) = 0$. Furthermore, the computation of $\partial F(q)/\partial q(\cdot) = 0$ (Assuming that $q(\mathbf{Z}) = \prod_{i=1}^K q(\mathbf{z}_i)$) can be expressed as

$$q(\mathbf{z}_j) \propto \exp \left(\langle \ln p(\mathbf{Y},\mathbf{Z};\boldsymbol{\theta}) \rangle_{i \neq j} \right), \quad (41)$$

where $\langle \cdot \rangle_{i \neq j}$ denotes an expectation with respect to all factors except $q(\mathbf{z}_j)$.

Next, the approximate posterior distribution of each part in Equation (40) is calculated according to Equation (41). We first calculate the variational distribution for common component \mathbf{z}_c , via Equation (42).

$$\begin{aligned} q(\mathbf{z}_c) &\propto \exp \left(\langle \ln p(\mathbf{Y},\mathbf{Z};\boldsymbol{\theta}) \rangle_{q(\mathbf{Z})/q(\mathbf{z}_c)} \right) \propto \exp \left(\sum_{k=1}^K \langle \ln p(\mathbf{y}_k|\mathbf{z}_c,\mathbf{z}_k,\boldsymbol{\beta}) \rangle_{q(\mathbf{z}_k)q(\boldsymbol{\beta})} + \langle \ln p(\mathbf{z}_c|\boldsymbol{\alpha}_c) \rangle_{q(\boldsymbol{\alpha}_c)} \right) \\ &\propto \mathcal{N}(\mathbf{z}_c; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c). \end{aligned} \quad (42)$$

Thus \mathbf{z}_c is confirmed to be Gaussian distributed, with covariance matrix $\boldsymbol{\Sigma}_c$ and mean vector $\boldsymbol{\mu}_c$, where

$$\boldsymbol{\Sigma}_c = \left(\sum_{k=1}^K \langle \boldsymbol{\beta} \rangle \mathbf{A}_k^T \mathbf{A}_k + \boldsymbol{\Lambda}_{\langle \boldsymbol{\alpha}_c \rangle} \right)^{-1}, \quad (43)$$

$$\boldsymbol{\mu}_c = \langle \boldsymbol{\beta} \rangle \boldsymbol{\Sigma}_c \sum_{k=1}^K \mathbf{A}_k^T (\mathbf{y}_k - \mathbf{A}_k \boldsymbol{\mu}_k). \quad (44)$$

In the above equations, $\boldsymbol{\Lambda}_{\langle \boldsymbol{\alpha}_c \rangle} \in \mathbb{R}^{N \times N}$ is a diagonal matrix with hyperparameters α_c^n ($n = 1, \dots, N$), and K is the number of channels.

Secondly, we calculate the variational distributions for innovation components \mathbf{z}_k , $k = 1, \dots, K$ via Equation (45).

$$\begin{aligned}
q(\mathbf{z}_k) &\propto \exp \left(\langle \ln p(\mathbf{Y}, \mathbf{Z}; \boldsymbol{\theta}) \rangle_{q(\mathbf{Z})/q(\mathbf{z}_k)} \right) \propto \exp \left(\langle \ln p(\mathbf{y}_k | \mathbf{z}_c, \mathbf{z}_k, \beta) \rangle_{q(\mathbf{z}_c)q(\beta)} + \langle \ln p(\mathbf{z}_k | \boldsymbol{\alpha}_k) \rangle_{q(\boldsymbol{\alpha}_k)} \right) \\
&\propto \exp \left(-\frac{\langle \beta \rangle}{2} \left\langle \|\mathbf{y}_k - \mathbf{A}_k(\mathbf{z}_c + \mathbf{z}_k)\|_2^2 \right\rangle_{q(\mathbf{z}_c)} - \frac{1}{2} \mathbf{z}_k^T \boldsymbol{\Lambda}_{\langle \boldsymbol{\alpha}_k \rangle} \mathbf{z}_k \right) \propto \mathcal{N}(\mathbf{z}_k; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k).
\end{aligned} \quad (45)$$

Thus, \mathbf{z}_k is also Gaussian distributed, with covariance matrix $\boldsymbol{\Sigma}_k$ and mean vector $\boldsymbol{\mu}_k$, where

$$\boldsymbol{\Sigma}_k = \left(\langle \beta \rangle \mathbf{A}_k^T \mathbf{A}_k + \boldsymbol{\Lambda}_{\langle \boldsymbol{\alpha}_k \rangle} \right)^{-1}, \quad (46)$$

$$\boldsymbol{\mu}_k = \langle \beta \rangle \boldsymbol{\Sigma}_k \mathbf{A}_k^T (\mathbf{y}_k - \mathbf{A}_k \boldsymbol{\mu}_c). \quad (47)$$

In the above equations, $\boldsymbol{\Lambda}_{\langle \boldsymbol{\alpha}_k \rangle} \in \mathbb{R}^{N \times N}$ is a diagonal matrix with hyperparameters α_k^n ($k = 1, \dots, K$; $n = 1, \dots, N$).

Thirdly, we calculate the variational distribution for the prior of common component.

$$\begin{aligned}
q(\alpha_c^n) &\propto \exp \left(\langle \ln p(z_c^n | \alpha_c^n) + \ln p(\alpha_c^n; a, b) \rangle_{q(z_c)q(\beta)} \right) \\
&\propto \exp \left(\frac{1}{2} \ln \alpha_c^n - \frac{1}{2} \langle (z_c^n)^2 \rangle \alpha_c^n + (a-1) \ln \alpha_c^n - b \alpha_c^n \right) \\
&\propto \exp \left(\left(\frac{1}{2} + a - 1 \right) \ln \alpha_c^n - \left(b + \frac{1}{2} \langle (z_c^n)^2 \rangle \right) \alpha_c^n \right) \propto \text{Gamma}(\alpha_c^n; \tilde{a}, \tilde{b}_n).
\end{aligned} \quad (48)$$

Thus α_c^n is distributed as $\text{Gamma}(\alpha_c^n; \tilde{a}, \tilde{b}_n)$, where

$$\tilde{a} = a + \frac{1}{2}, \tilde{b}_n = b + \frac{1}{2} \langle (z_c^n)^2 \rangle. \quad (49)$$

Similarly, we can obtain the variational distributions for the prior of innovation components, as

$$q(\alpha_k^n) \propto \text{Gamma}(\alpha_k^n; \tilde{c}_k, \tilde{d}_k^n), \quad (50)$$

where

$$\tilde{c}_k = c_k + \frac{1}{2}, \tilde{d}_k^n = d_k + \frac{1}{2} \langle (z_k^n)^2 \rangle. \quad (51)$$

Finally, we calculate the approximate posterior distribution for the parameter of inverse variance β .

$$\begin{aligned}
q(\beta) &\propto \exp \left(\langle \ln p(\mathbf{Y}, \mathbf{Z}; \boldsymbol{\theta}) \rangle_{q(\mathbf{Z})/q(\beta)} \right) \propto \exp \left(\sum_{k=1}^K \langle \ln p(\mathbf{y}_k | \mathbf{z}_c, \mathbf{z}_k, \beta) \rangle_{q(\mathbf{z}_c)q(\mathbf{z}_k)} + \ln p(\beta | e, f) \right) \\
&\propto \exp \left(\frac{KM}{2} \ln \beta - \frac{\beta}{2} \sum_{k=1}^K \left\langle \|\mathbf{y}_k - \mathbf{A}_k(\mathbf{z}_c + \mathbf{z}_k)\|_2^2 \right\rangle_{q(\mathbf{z}_c)q(\mathbf{z}_k)} + (e-1) \ln \beta - f \beta \right) \\
&\propto \text{Gamma}(\beta; \tilde{e}, \tilde{f}).
\end{aligned} \quad (52)$$

Thus β is distributed as $\text{Gamma}(\beta; \tilde{e}, \tilde{f})$, where

$$\begin{aligned}
\tilde{e} &= e + \frac{KM}{2}, \\
\tilde{f} &= f + \frac{1}{2} \left(\sum_{k=1}^K \left\langle \|\mathbf{y}_k - \mathbf{A}_k(\mathbf{z}_c + \mathbf{z}_k)\|_2^2 \right\rangle_{q(\mathbf{z}_c)q(\mathbf{z}_k)} \right).
\end{aligned} \quad (53)$$

The variational optimization proceeds by iteratively updating Equations (42), (45), (48), (50), and (52) until convergence occurs to hyperparameters $\boldsymbol{\theta}$. Finally, we can obtain the reconstructed signal by applying the maximum a posteriori estimation.

$$\hat{\rho}_k = \arg \max_{\mathbf{z}_c + \mathbf{z}_k} p(\mathbf{Z} | \mathbf{Y}; \boldsymbol{\theta}) = \arg \max_{\mathbf{z}_c} q(\mathbf{z}_c) + \arg \max_{\mathbf{z}_k} q(\mathbf{z}_k) = \boldsymbol{\mu}_c + \boldsymbol{\mu}_k, \quad k = 1, 2, \dots, K.$$

The proposed HVB-DCS algorithm for solving the JSM-1 reconstruction problem is summarized in Algorithm 1.

Algorithm 1: HVB-DCS Algorithm

Input: A set of measurement vectors $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K\}$ and corresponding measurement matrices $\{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_K\}, k = 1, \dots, K$.

Output: The reconstructed signal $\hat{\rho}_k = \mu_c + \mu_k, k = 1, \dots, K$.

Initialize the hyperparameters. Set the initial values of the variables $(a, b, \{c_k, k = 1, \dots, K\}, \{d_k, k = 1, \dots, K\}, e, f)$ as 10^{-6} .

Compute the variational distribution for the common component.

Compute $\Sigma_c = \left(\sum_{k=1}^K \langle \beta \rangle \mathbf{A}_k^T \mathbf{A}_k + \Lambda_{\langle \alpha_c \rangle} \right)^{-1}$, and $\mu_c = \langle \beta \rangle \Sigma_c \sum_{k=1}^K \mathbf{A}_k^T (\mathbf{y}_k - \mathbf{A}_k \mu_k)$.

Compute the variational distribution for the prior of the common component.

Update $q(\alpha_c^n)$, compute

$$\tilde{a} = a + \frac{1}{2}, \tilde{b}_n = b + \frac{1}{2} \langle (z_c^n)^2 \rangle.$$

Compute the variational distributions for the innovation components.

Compute $\Sigma_k = \left(\langle \beta \rangle \mathbf{A}_k^T \mathbf{A}_k + \Lambda_{\langle \alpha_k \rangle} \right)^{-1}$, and $\mu_k = \langle \beta \rangle \Sigma_k \mathbf{A}_k^T (\mathbf{y}_k - \mathbf{A}_k \mu_c)$.

Compute the variational distributions for the prior of the innovation components.

Update $q(\alpha_k^n)$, compute

$$\tilde{c}_k = c_k + \frac{1}{2}, \tilde{d}_k^n = d_k + \frac{1}{2} \langle (z_k^n)^2 \rangle.$$

Compute the variational distribution for the prior of noise vector.

Update $q(\beta)$, compute

$$\tilde{e} = e + \frac{KM}{2},$$

$$\tilde{f} = f + \frac{1}{2} \left(\sum_{k=1}^K \langle \|\mathbf{y}_k - \mathbf{A}_k(\mathbf{z}_c + \mathbf{z}_k)\|_2^2 \rangle_{q(\mathbf{z}_c)q(\mathbf{z}_k)} \right).$$

Iterate steps 2, 3, 4, 5 and 6 until convergence occurs to hyperparameters.

Output $\hat{\rho}_k = \mu_c + \mu_k$ for $k = 1, \dots, K$.

5.2. Complexity Analysis

We present the computational complexity of the proposed HVB-DCS algorithm, as well as the comparison to the CS-based joint sparsity recovery algorithm and the RD algorithm. The computational complexity of the HVB-DCS algorithm at each iteration is dominated by the inversion operations on an $N \times N$ matrix in Equations (43) and (46), and the multiplications of a matrix and a vector in Equations (44) and (47). The computational complexities of the matrix inversion and matrix-vector multiplication are $\mathcal{O}(N^3)$ and $\mathcal{O}(N^2)$, respectively. By using matrix inversion lemma [39], the complexity of the matrix inversion can be reduced to $\mathcal{O}(M^3)$. Thus the overall computational complexity of the HVB-DCS algorithm is $\mathcal{O}(N_t(K+1)M^3)$, where N_t is the total number of iterations. The small values of N_t and M will result in acceptable computational complexity. In contrast, a large N will lead to high computational complexity, which makes the HVB-DCS algorithm impractical. To circumvent the high computational complexity problem, we can resort to a number of accelerating algorithms [40,41] to develop a more computationally efficient algorithm in the future work.

For the CS-based joint sparsity recovery algorithm, the dimensions of the sparse vector and the measurement vectors are $3N$ and $2M$, respectively. Thus the computational complexity of the CS-based algorithm is $\mathcal{O}(12NM^2)$. For the RD algorithm, the computational complexity is dominated by the fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT), which have computational complexities of $\mathcal{O}(M \log_2 M)$ and $\mathcal{O}(M/2 \log_2 M)$, respectively. Thus, the overall computational complexity of the RD algorithm is $\mathcal{O}(K(3N/2 + 1)M \log_2 M)$.

6. Simulations and Experiments

In this section, the performance of the proposed HVB-DCS algorithm is verified through both simulations and experiments. First, the effects of along-track and cross-track velocities on the SAR images reconstructed using the HVB-DCS algorithm are analyzed in Section 6.1. Secondly, the proposed HVB-DCS algorithm is compared with some classical GMTI algorithms, e.g., the RD [42] and CS [43]

algorithms, based on point target simulation, in Section 6.2. Thirdly, the effects of problem size and data rate on complexity of the HVB-DCS algorithm are analyzed in Section 6.3. Finally, in Section 6.4, the proposed HVB-DCS algorithm is applied to the real background cluttered environment, where the real SAR data are collected by RADARSAT-1 satellite in the Vancouver region [42].

6.1. The Effects of Along-Track and Cross-Track Velocities on SAR Imaging

6.1.1. The Effect of Cross-Track Velocity

First, we consider the influence of cross-track velocity on SAR imaging and take channel one as an example. The parameters of the simulated SAR radar system are listed in Table 1. Suppose that a moving target is at (R_B, x_0) , with a given cross-track velocity v_r and along-track velocity $v_a = 0$ m/s. The received signal of the moving target by channel one, after demodulation and range compression, which is presented in Equation (13), is rewritten as

$$s_1(t_m) = \rho_0^{(1)} \omega_a(t_m - k\Delta\tau) \exp\left(-j2\pi \frac{v^2}{\lambda R_B} \left(t_m - k\Delta\tau - \frac{v_r R_B}{v^2}\right)^2\right), \quad (54)$$

where $k\Delta\tau = x_0/v$. Since the CS-based SAR imaging algorithm applied in the azimuth direction is essentially searching the most matched atom from a redundant dictionary Equation (15) to $s_1(t_m)$, the degree of match is mainly affected by the exponential term in Equation (54), whereas the minor shift of the azimuth windowing function has little influence. Assuming that the most matched atom in the redundant dictionary Φ_1 is

$$s_0(t_m - k'\Delta\tau) = \omega_a(t_m - k'\Delta\tau) \exp\left(-j2\pi \frac{v^2}{\lambda R_B} (t_m - k'\Delta\tau)^2\right). \quad (55)$$

When $k'\Delta\tau = k\Delta\tau + \frac{v_r R_B}{v^2}$, $s_1(t_m)$ matches the atom most, resulting in $k' = k + \frac{v_r R_B}{v^2 \Delta\tau}$. Obviously, for a stationary clutter with $v_r = 0$ m/s, the reflection coefficient of the stationary clutter locates at the $(N/2 + k)^{th}$ pixel in the azimuth direction. For a moving target with $v_r \neq 0$, the cross-track velocity of the moving target leads to an azimuth deviation of $\frac{v_r R_B}{v^2 \Delta\tau}$ pixels, compared with that of the stationary clutter, on the SAR image reconstructed using the HVB-DCS algorithm.

The influence of cross-track velocity is demonstrated in Figure 5. In Figure 5, a static scattering center and a moving scattering center with the cross-track velocity of $v_r = 0.5$ m/s and the along-track velocity of $v_a = 0$ m/s are both at (7071, 0) m. The amplitude of reflection coefficient of the moving target is half that of the stationary clutter. The moving target has a deviation of 47 pixels in the azimuth direction compared with that of the stationary clutter on SAR image, which completely conforms to $\frac{v_r R_B}{v^2 \Delta\tau}$.

Table 1. SAR radar system parameters for simulation.

Parameter	Value
Wavelength (m)	0.03
Range bandwidth (MHz)	150
Platform height (m)	5000
Platform velocity (m/s)	150
Incidence angle (°)	45
PRF (Hz)	300
Aperture size (m)	2
Channel distance (m)	1

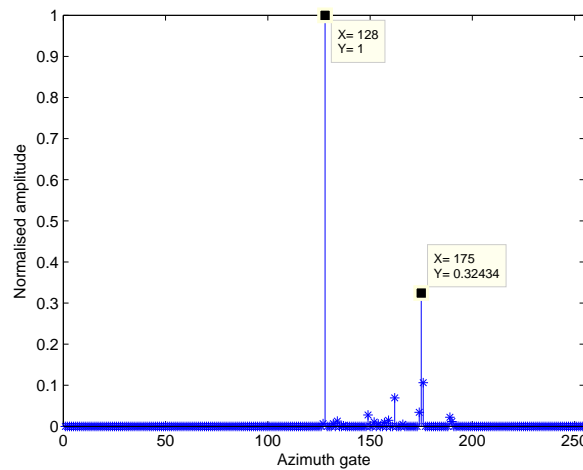


Figure 5. The influence of cross-track velocity on SAR image.

6.1.2. The Effect of Along-Track Velocity

For the analysis on along-track velocity, it is assumed that the moving target is at $(R_B, 0)$, with a given along-track velocity v_a and cross-track velocity of $v_r = 0$ m/s. Thus the received signal of the moving target, after demodulation and range compression, can be written as

$$s_1(t_m) = \rho_0^{(1)} \omega_a(t_m) \exp \left(-j2\pi \frac{(v - v_a)^2}{\lambda R_B} t_m^2 \right). \quad (56)$$

Similarly, for channel one, the received signal after range compression can be represented linearly by the atoms in the redundant dictionary Φ_1 , as

$$s_1(t_m) = \sum_{i=-N/2+1}^{N/2} C_i^{(1)} \omega_a(t_m - i\Delta\tau) \exp \left(-j2\pi \frac{v^2}{\lambda R_B} (t_m - i\Delta\tau)^2 \right), \quad (57)$$

where $C_i^{(1)}$ denotes the reconstructed reflection coefficient in channel one. Substituting Equation (56) into Equation (57), and dividing both sides by the item $\exp \left(-j2\pi \frac{v^2}{\lambda R_B} t_m^2 \right)$, results in

$$\rho_0^{(1)} \omega_a(t_m) \exp \left(-j2\pi \frac{(v_a^2 - 2vv_a)}{\lambda R_B} t_m^2 \right) = \sum_{i=-N/2+1}^{N/2} D_i^{(1)} \omega_a(t_m - i\Delta\tau) \exp(-ji\omega t_m), \quad (58)$$

where $\omega = \frac{-4\pi v^2}{\lambda R_B} \Delta\tau$, and $D_i^{(1)} = C_i^{(1)} \exp \left(-j \frac{2\pi v^2}{\lambda R_B} (i\Delta\tau)^2 \right)$. When the along-track velocity v_a of the moving target equals zero, from Equation (58), we can see that the left side of the equation equals a constant. Thus on the right side of the equation, the most sparse solution is $D_i^{(1)} = 0$ ($i \neq 0$) and $D_0^{(1)} = \rho_0^{(1)}$. In general, the along-track velocity of the target is much less than that of the flying platform, resulting in a slow variation of the terms on the left side of Equation (58). Thus the left side of Equation (58) contains only a few low frequency components. With the increase of left side of Equation (58), the high frequency component increases slowly. $D_i^{(1)}$ can be approximated to the coefficients of expansion of the slowly varying function under different frequency components. Therefore, with the increase of the left side of the Equation (58), the nonzero value $D_i^{(1)}$ will spread to both sides, centering at $D_0^{(1)}$, which leads to the defocus of the moving target on the SAR image. The influence of along-track velocity can be shown in Figure 6, where v_a is set as different values of

0 m/s, 1 m/s and 2 m/s. It can be seen from Figure 6 that as the along-track velocity increases, the azimuth defocus of the moving target on the SAR image becomes more serious.

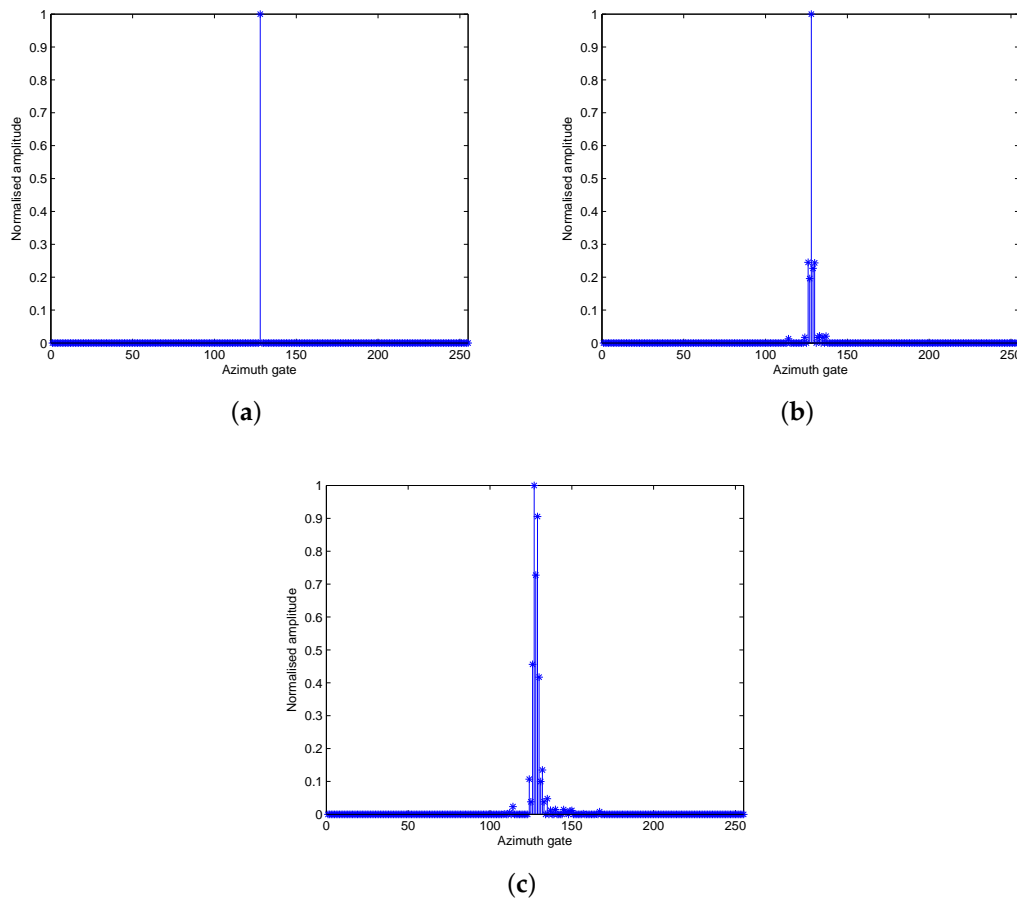


Figure 6. The influence of along-track velocity on SAR image: (a) $v_a = 0$ m/s; (b) $v_a = 1$ m/s; (c) $v_a = 2$ m/s.

The defocus of the moving target on the SAR image can be treated as a basis mismatch problem [44]. In this work, the measurement matrices are built based on the echo signal models of two channels (Equations (13) and (19) in Section 4.1), assuming that the along-track velocity v_a equals zero. However, in practice, this assumption does not always hold true. In this case, the constructed measurement matrix does not exactly match the underlying true one, leading to the basis mismatch problem. A parameter β_d is defined in [44] to quantitatively evaluate the difference between the assumed basis and the true one. The larger the absolute value of v_a is, the higher the mismatch level of the constructed measurement matrix will be, resulting in a higher β_d . Moreover, the defocus effect can be evaluated by the reconstruction error of the sparse vector recovered with mismatched basis. Based on [44], an upper bound of the ℓ_1 -norm reconstruction error is defined as:

$$\|\rho_k - \hat{\rho}_k\|_1 \leq N\beta_d \|\rho_k\|_{q'}, \quad (59)$$

where ρ_k and $\hat{\rho}_k$ denote the true and reconstructed sparse vector respectively, N is the dimension of the sparse vector, $\|\cdot\|_q$ represents the ℓ_q -norm, and $1 \leq q \leq \infty$. Thus, the azimuth defocus on the SAR image is bounded given v_a . In addition, please refer to [45,46] for more analyses of basis mismatch problem in CS-based SAR imaging and the methods to deal with it.

Furthermore, the effect of v_a on the azimuth defocus is analyzed through simulation. The degree of azimuth defocus is represented by the relative reconstruction error, which is defined as

$$err \triangleq \|\rho_k - \hat{\rho}_k\|_1 / \|\rho_k\|_1. \quad (60)$$

Figure 7 shows the relative reconstruction error err versus v_a . The result illustrates that the reconstruction error err is low when the absolute value of v_a is relatively small.

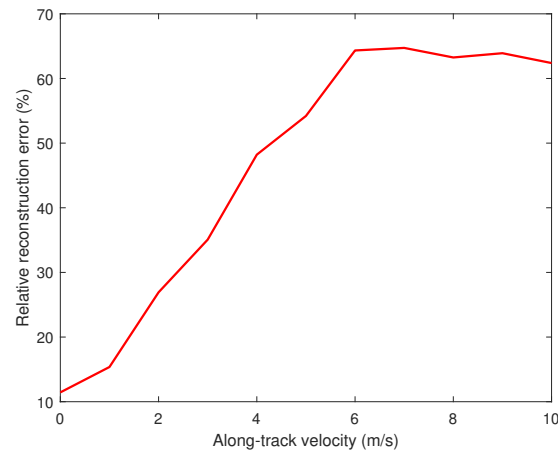


Figure 7. Relative reconstruction error versus along-track velocity.

6.2. The Simulation on Point Targets

In this section, the performance of the HVB-DCS algorithm is verified by simulations on point targets. The nearest slant range R_B is set as 7071 m. Four scattering centers are located at the range bin $R_B = 7071$ m, including three static scattering centers, and a moving scattering center with cross-track velocity of $v_r = 0.5$ m/s and along-track velocity of $v_a = 0$ m/s. The three stationary targets are located at -5 m, 0 m, $+5$ m in the azimuth direction, and the moving target is located at 0 m in the azimuth direction. The amplitudes of reflection coefficient of three stationary targets are the same, which are twice that of the moving target. The parameters of the simulated SAR radar system are the same as those listed in Table 1.

In the field of SAR-GMTI, the proposed HVB-DCS algorithm is compared to the RD [42] and CS [43] algorithms. The flow charts of the RD based GMTI system, the CS-based GMTI system, and the HVB-DCS based GMTI system are shown in Figures 8–10, respectively. For the RD based GMTI system and CS-based GMTI system, an additional DPCA [47] procedure is needed for moving-target indication, which is shown in both Figures 8 and 9. DPCA can effectively suppress clutter and detect slow moving targets. In contrast, in the HVB-DCS based GMTI system, the moving targets (the innovation components) and the stationary targets (the common component) can be directly distinguished using the DCS framework. Thus the DPCA is no longer required in the HVB-DCS based GMTI system and removed (Figure 10). This significantly simplifies the architecture of the GMTI system, and saves the computing time.

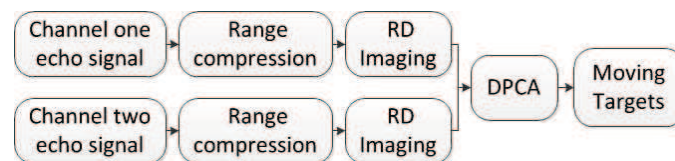


Figure 8. Flow chart of the RD based GMTI system.

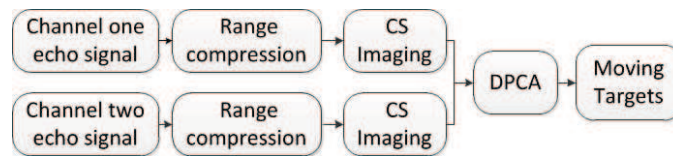


Figure 9. Flow chart of the CS-based GMTI system.

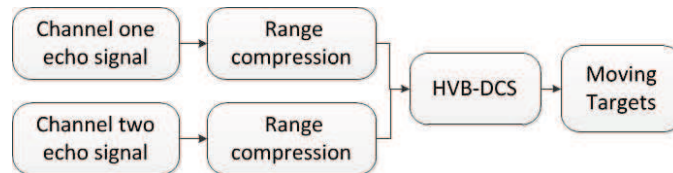


Figure 10. Flow chart of the HVB-DCS based GMTI system.

Figures 8–10 show that all the three algorithms, i.e., RD, CS and HVB-DCS algorithms, can effectively identify the moving target. For the RD algorithm, we use the oversampled raw data to image. The result of the RD algorithm is shown in Figure 11. As a widely used SAR imaging algorithm, the RD algorithm has very high running speed, but the on-board memory required by the sampling rate is large. For the CS algorithm, the simulated radar system works in the random sampling mode introduced in Section 4.1. In our experiment, we choose randomly half of transmitted pulses in azimuth and recover the complex-valued SAR image based on the CS algorithm. The result is shown in Figure 12, which shows that the image reconstructed by using the CS algorithm is more clear, and the side-lobes are effectively suppressed. However, the computational cost of the CS algorithm is relatively high. For the HVB-DCS algorithm, we further drop randomly 62.5% of the transmitted pulses in azimuth to recover the SAR image. In our experiment, we reduce to about 37.5% of the original samples in the azimuth direction and reconstruct the SAR image using the HVB-DCS algorithm. The result is shown in Figure 13, which shows that the HVB-DCS algorithm has a great suppression effect on stationary clutter and is capable of identifying the moving target accurately. The effects of both the data rate in azimuth, and noise/clutter level on the detection performance of moving targets are tested by simulations. The adopted metric is the improvement factor (IF), which is defined as $IF = SCNR_{out}/SCNR_{in}$, where $SCNR_{in}$ and $SCNR_{out}$ are the signal-to-clutter-noise ratios of input signal and output signal, respectively. We test the HVB-DCS and CS algorithms to detect moving targets by varying the data rate in the azimuth direction, at different $SCNR_{in}$ levels, except that the performance of RD algorithm is tested at different $SCNR_{in}$ levels with data rate 100%. The IFs averaged over 100 Monte Carlo trials are shown in Figure 14. For the HVB-DCS algorithm, at the same $SCNR_{in}$ level, the IF first increases dramatically with the data rate in azimuth. When the data rate in the azimuth direction exceeds 37.5%, the IF grows slowly. For the CS algorithm, at the clutter/noise levels of $SCNR_{in} = -11$ dB and $SCNR_{in} = -6$ dB, the IF is negative infinite when the data rate in azimuth is lower than 37.5%. Furthermore, at the clutter/noise level of $SCNR_{in} = -2$ dB, the IFs are negative infinite when the data rate in azimuth is lower than 25%. With the data rate in azimuth approaching 37.5%, the IF increases to a value larger than zero. When the data rate in the azimuth direction increases to more than 50%, the IF gradually increases with the data rate in azimuth at the same $SCNR_{in}$ level. The main reason is that when the data rate in azimuth decreases, the reconstruction error increases gradually, thus affecting the performance of clutter suppression. Moreover, it is noted that the IFs of HVB-DCS algorithm with the data rate 37.5% in azimuth are almost at the same level with those of CS algorithm with the data rate 50% in azimuth. The IFs of RD algorithm (indicated by asterisks) are the lowest among the three algorithms. This verifies that the proposed HVB-DCS algorithm outperforms the CS and RD algorithms in terms of clutter suppression and moving target detection performance.

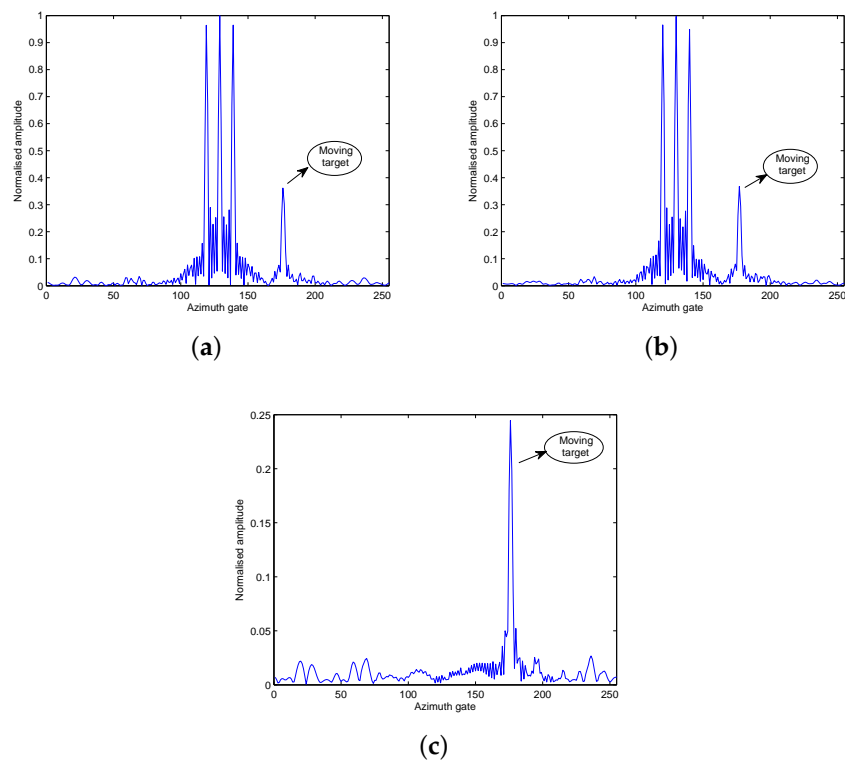


Figure 11. RD based reconstruction results: (a) channel 1 with oversampled raw data; (b) channel 2 with oversampled raw data; (c) DPCA with oversampled raw data.

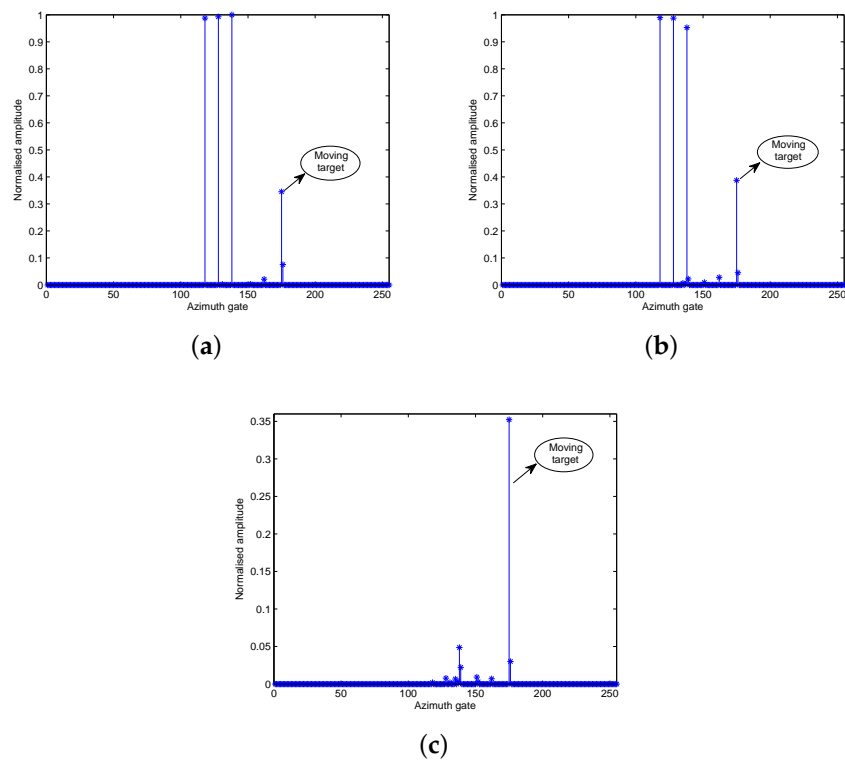


Figure 12. CS-based reconstruction results: (a) channel 1 with data rate 50% in azimuth; (b) channel 2 with data rate 50% in azimuth; (c) DPCA with data rate 50% in azimuth.

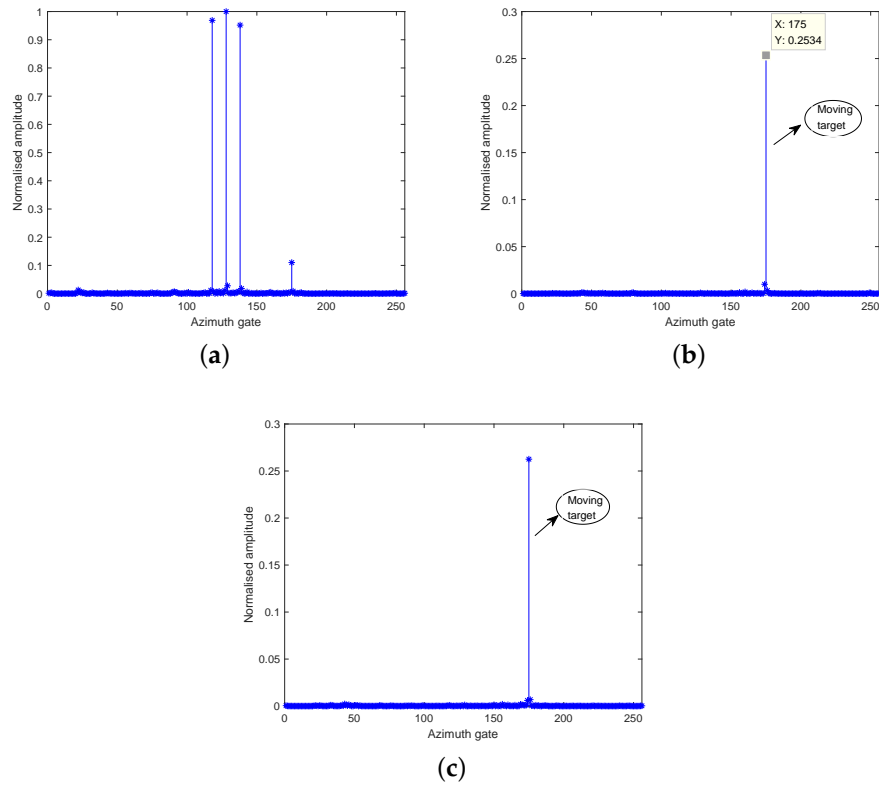


Figure 13. HVB-DCS based reconstruction results with data rate 37.5% in azimuth: (a) static scattering centers; (b) moving target in channel one; (c) moving target in channel two.

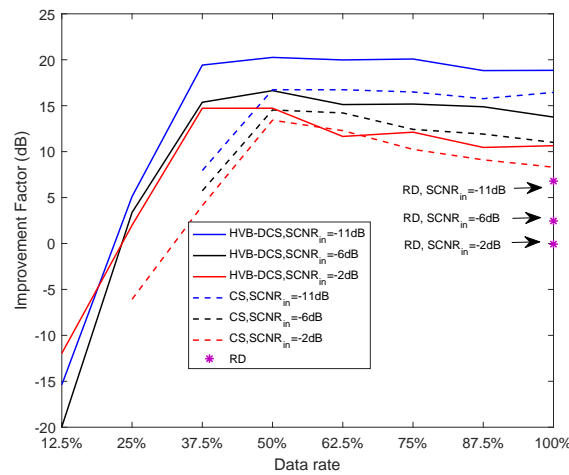


Figure 14. IFs of the RD, CS and HVB-DCS algorithms, based GMTI system for the varying data rates in azimuth and different $SCNR_{in}$ levels.

To examine the reconstruction accuracy of the proposed HVB-DCS algorithm, we introduce the reconstruction error, which is defined as

$$e_{rec} = \frac{\sum_{l=1}^L \|\hat{x}_l - x_l\|_2}{\sum_{l=1}^L \|x_l\|_2}, \quad (61)$$

where L is the number of range bins, x_l and \hat{x}_l denote the true and reconstructed sparse vector at the l^{th} range bin, respectively. The reconstruction errors of the RD, CS and HVB-DCS algorithms,

are calculated by varying data rates in azimuth, at different SCNR_{in} levels. The results are shown in Figure 15. The reconstruction error of the HVB-DCS algorithm is lower than 0.2 when the data rate is not less than 37.5%, whereas the reconstruction error of the CS algorithm is lower than 0.3 when the data rate is not less than 50%. In addition, compared to the CS algorithm, the HVB-DCS algorithm has smaller reconstruction errors when the data rate exceeds 37.5%. The main reasons are due to the following two aspects. First, the DCS exploits both intra-signal and inter-signal correlations, which ensures a good correlation between the channels in the case of lower data rate. Secondly, the hierarchical structure is more suitable for complex models. Moreover, the higher the noise/clutter level is, the larger the reconstruction errors are, which shows that the detection performance degrades with the increase of the noise/clutter level.

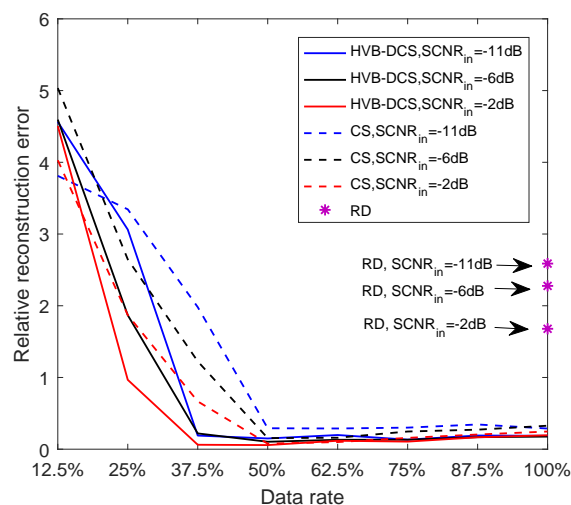


Figure 15. Reconstruction errors of three algorithms, i.e., RD, CS and HVB-DCS algorithms, based GMTI system for the varying data rates in azimuth and different SCNR_{in} levels.

6.3. Effects of Problem Size and Data Rate on Complexity of HVB-DCS

The goal of this section is to analyze the computational complexity of the HVB-DCS algorithm in terms of problem size and data rate using simulations. The computational complexity of the HVB-DCS algorithm is further compared with those of the RD and CS-based algorithms. First, we set the data rate as 37.5% and 50% for the HVB-DCS and CS-based algorithms respectively, and 100% for the RD algorithm. Meanwhile, we vary the dimension of sparse vector N (also, the number of samples in the azimuth direction) in the range of $[100, 900]$. We plot the runtime (averaged over 100 trials) of the HVB-DCS, RD and CS-based algorithms versus N in Figure 16a. The results show that the RD algorithm achieves the shortest runtime. When the data rate is at the level of 37.5% and $N \leq 500$, the proposed HVB-DCS algorithm takes almost the same runtime with the CS-based algorithm to perform moving target indication. When the data rate is at the level of 37.5% and $N > 500$, the proposed HVB-DCS algorithm takes longer runtime than the CS-based algorithm. When the data rate is at the level of 50%, the runtime of HVB-DCS is shorter than that of the CS-based algorithm in the range of $[100, 900]$.

In addition, we examine the runtime of the HVB-DCS and CS-based algorithms versus data rate when setting N at different levels, i.e., 100, 500 and 900 respectively. The data rate varies in the range of $[12.5\%, 100\%]$ for the two algorithms. The results are shown in Figure 16b. When $N = 100$, the runtime of HVB-DCS is shorter than that of the CS-based algorithm at any data rate in the range of $[12.5\%, 100\%]$. When $N = 500$ and the data rate is larger than 37.5%, HVB-DCS takes shorter runtime than that of the CS-based algorithm, and the similar result occurs when $N = 900$ and the data rate is larger than 50%.

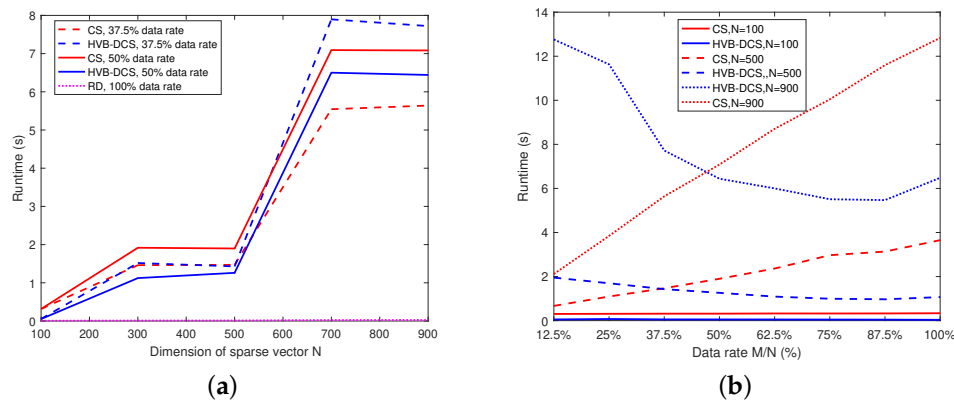


Figure 16. Complexity comparisons of the RD, CS and HVB-DCS algorithms: (a) runtime versus sparse vector dimension N ; (b) runtime versus data rate M/N .

6.4. Experiment on Real SAR Data

In order to verify the performance of the proposed HVB-DCS algorithm in practice, we carry out an experiment on data from a real background cluttered environment, which are collected by the satellite RADARSAT-1 in the Vancouver region [42]. The parameters of RADARSAT-1 satellite are shown in Table 2.

Table 2. RADARSAT-1 parameters.

Parameter	Value
Slant range of scene center (km)	988.65
Beam squint angle (rad)	0.0279
Effective radar velocity (m/s)	7062
PRF (Hz)	1256.98
Sampling rate (MHz)	32.317
Range FM rate (MHz/ μ s)	0.72135
Pulse duration (μ s)	41.75
Radar center frequency (MHz)	5300

In the experiment, the real SAR data from single-channel SAR system (RADARSAT-1) is used to simulate dual-channel raw data. The detailed procedures are as follows. According to the imaging geometry models for point targets as in Equations (13) and (19) in Section 4.1, there is a phase difference $\exp(-j\pi d^2 / (2\lambda R_B))$ and a time delay $d / (2(v - v_a))$ between the received signals from two channels. First, all the real data from the first channel is delayed Δt time, and then multiplied with $\exp(-j\pi d^2 / (2\lambda R_B))$, to produce the raw data from the second channel of the simulated dual-channel SAR system. Secondly, in order to simulate the channel imbalance, some random amplitudes and phase errors are added to the raw data of the second channel. Finally, the echo signals of nine targets (including five stationary targets and four moving targets moving) are added to the raw data of both channels. Based on the raw data from the simulated dual-channel SAR system, three algorithms, e.g., the RD, CS and HVB-DCS algorithms, are utilized to detect the moving targets.

Figure 17 shows the SAR image of the observation scene, which contains the static clutter, five stationary point targets and, four moving targets with the same cross-track velocity of $v_r = 5$ m/s, and the same along-track velocity of $v_a = 5$ m/s. The radar image covers an area of about 2.8 km in azimuth and 1.9 km in range. The corresponding parameters of nine targets are listed in Table 3. The four moving targets in full SAR RADARSAT-1 image are marked with the red rectangles in Figure 17. We use the oversampled raw data in azimuth for the RD algorithm, and 50% of the

oversampled raw data in azimuth for CS algorithm and 37.5% of the oversampled raw data in azimuth for the HVB-DCS algorithm.

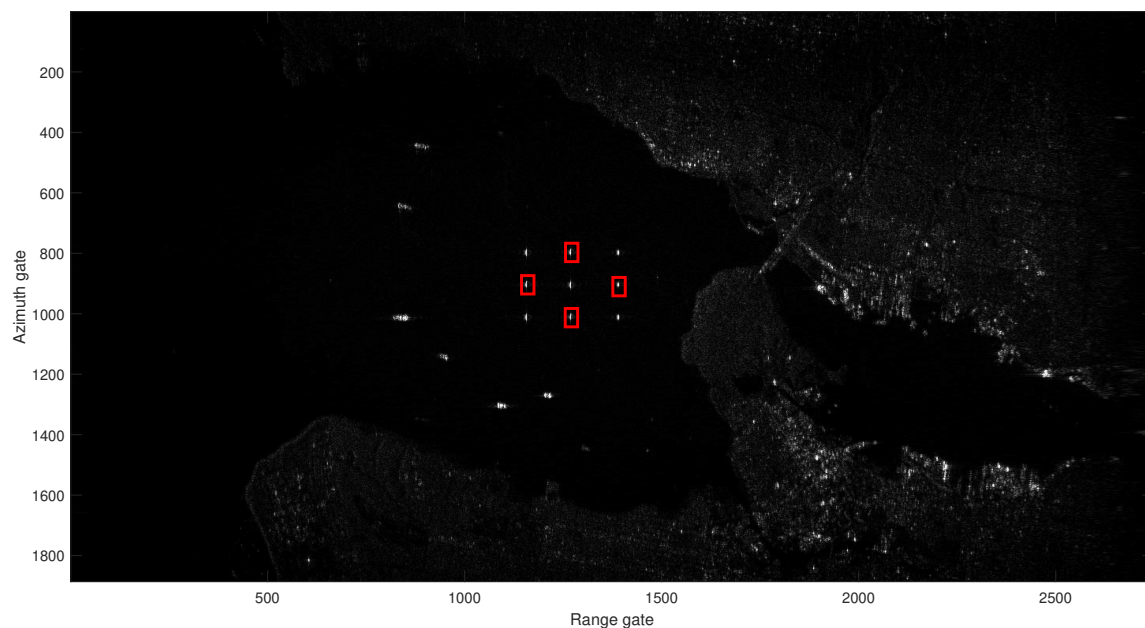


Figure 17. Ground truth scene containing the static clutter, five stationary targets and four targets with same cross-track velocity of $v_r = 5$ m/s, and same along-track velocity of $v_a = 5$ m/s. The red rectangle indicates the moving target.

Table 3. Simulation parameters of five stationary targets and four moving targets.

No.	Azimuth Coordinate (m)	Nearest Slant Range (km)	Along-Track Velocity v_a (m/s)	Across-Track Velocity v_r (m/s)
1	−500	0.8	5	5
2	−400	987.65	0	0
3	800	987.65	0	0
4	200	988.17	0	0
5	−1100	988.17	5	5
6	100	988.17	5	5
7	−500	988.73	5	5
8	−400	988.73	0	0
9	800	988.73	0	0

GMTI results (Figure 18) show that all of the three algorithms can distinguish the moving targets from the stationary clutter successfully. Moreover, the final SCNR_{out} of the three algorithms are calculated to evaluate detection performance. The initial SCNR_{in} equals -69 dB. The $\text{SCNR}_{\text{outs}}$ of the SAR image recovered based on the RD, CS and HVB-DCS algorithms are -13 dB, -12 dB and -5 dB, respectively. Note that, there is an obvious increase in the SCNR. The RD algorithm has the lowest SCNR_{out} . The SAR image of CS-based algorithm has a much lower SCNR_{out} (-12 dB) compared with the proposed HVB-DCS algorithm (-5 dB). This is due to the reason that in the CS-based GMTI system, the mismatch between channels leads to the failure in energy cancelation of static targets and clutters, whereas the proposed HVB-DCS algorithm takes advantage of the correlation between two channels and suppresses the clutter and noise efficiently, even at a much lower data rate in the azimuth direction.

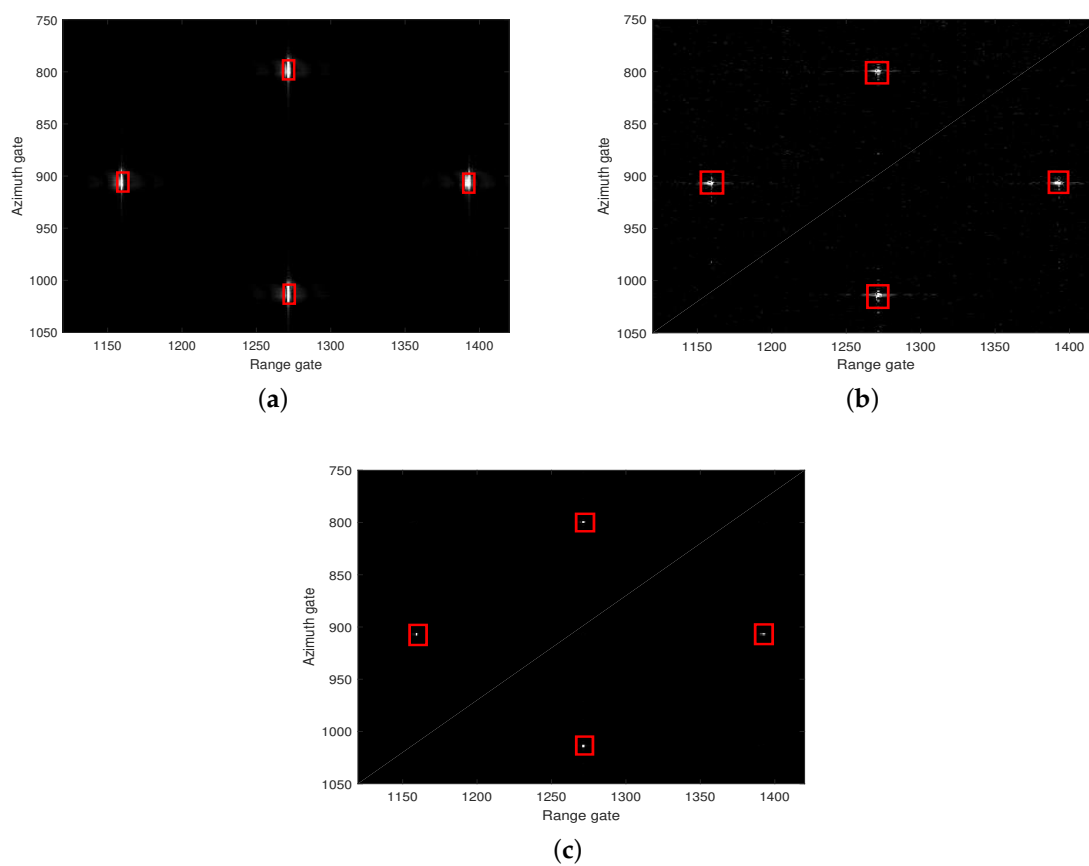


Figure 18. GMTI results: (a) RD with oversampled raw data; (b) CS with 50% of the oversampled raw data; (c) HVB-DCS with 37.5% of the oversampled raw data. The red rectangle indicates the moving target. The partial enlarged result is shown in the upper right corner of the SAR image.

7. Conclusions

In this paper, we propose a novel HVB-DCS algorithm for GMTI using the dual-channel SAR system. The results from both the simulation on point targets and the experiment on data from the real background cluttered environment show that the proposed HVB-DCS algorithm can successfully detect multiple moving targets, while suppressing the clutter efficiently. The proposed algorithm achieves better detection performance for the GMTI system, in the metric of IF and reconstruction error, compared with the RD and CS algorithms.

Author Contributions: J.L., J.J. conceived and designed the experiments; X.T. and J.J. performed the experiments; X.T. analyzed the data; J.L., X.T. and J.J. wrote the paper; K.H. checked the paper.

Funding: This work was funded by the Natural Science Foundations of China (Nos. 61573276 and 61473217), the National 973 project of China through Grant 2013CB329405 and the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (61221063).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhu, S.; Liao, G.; Qu, Y.; Zhou, Z.; Liu, X. Ground Moving Targets Imaging Algorithm for Synthetic Aperture Radar. *IEEE Trans. Geosci. Remote Sens.* **2011**, *1*, 462–477. [[CrossRef](#)]
2. Wang, W.; Liao, G.; Zhu, S.; Zhang, J. Compressive sensing-based ground moving target indication for dual-channel synthetic aperture radar. *IET Radar Sonar Navig.* **2013**, *8*, 858–866. [[CrossRef](#)]

3. Zheng, M.; Yang, R. SAR moving targets detection based on dpca and interferometric processing. *J. Electron. Inf. Technol.* **2003**, *11*, 1525–1530.
4. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *4*, 1289–1306. [[CrossRef](#)]
5. Babacan, S.D.; Molina, R.; Katsaggelos, A.K. Bayesian Compressive Sensing Using Laplace Priors. *IEEE Trans. Image Process.* **2010**, *1*, 53–63. [[CrossRef](#)] [[PubMed](#)]
6. Ji, S.; Xue, Y.; Carin, L. Bayesian compressive sensing. *IEEE Trans. Signal Process.* **2008**, *6*, 2346–2356. [[CrossRef](#)]
7. Figueiredo, M.A.T. Adaptive sparseness for supervised learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *9*, 1150–1159. [[CrossRef](#)]
8. Candes, E.J. The restricted isometry property and its implications for compressed sensing. *C. R. Math.* **2008**, *9*, 589–592. [[CrossRef](#)]
9. Baraniuk, R.; Steeghs, P. Compressive Radar Imaging. In Proceedings of the 2007 IEEE Radar Conference, Boston, MA, USA, 17–20 April 2007; pp. 128–133. [[CrossRef](#)]
10. Alonso, M.T.; Lopez-Dekker, P.; Mallorqui, J.J. A Novel Strategy for Radar Imaging Based on Compressive Sensing. *IEEE Trans. Geosci. Remote Sens.* **2010**, *12*, 4285–4295. [[CrossRef](#)]
11. Sun, C.; Wang, B.; Fang, Y.; Song, Z.; Wang, S. Multichannel and Wide-Angle SAR Imaging Based on Compressed Sensing. *Sensors* **2017**, *2*, 295. [[CrossRef](#)] [[PubMed](#)]
12. Fang, J.; Xu, Z.; Zhang, B.; Hong, W.; Wu, Y. Fast Compressed Sensing SAR Imaging Based on Approximated Observation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *1*, 352–363. [[CrossRef](#)]
13. Li, B.; Liu, F.; Zhou, C.; Lv, Y.; Hu, J. Phase Error Correction for Approximated Observation-Based Compressed Sensing Radar Imaging. *Sensors* **2017**, *3*, 613. [[CrossRef](#)] [[PubMed](#)]
14. Zhu, X.X.; Bamler, R. Tomographic SAR Inversion by L_1 -Norm Regularization—The Compressive Sensing Approach. *IEEE Trans. Geosci. Remote Sens.* **2010**, *10*, 3839–3846. [[CrossRef](#)]
15. Stojanovic, I.; Karl, W.C. Imaging of Moving Targets with Multi-Static SAR Using an Overcomplete Dictionary. *IEEE J. Sel. Top. Signal Process.* **2010**, *1*, 164–176. [[CrossRef](#)]
16. Sun, K.; Zhang, H.; Li, G.; Meng, H.; Wang, X. A novel STAP algorithm using sparse recovery technique. In Proceedings of the 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; pp. 336–339. [[CrossRef](#)]
17. Parker, J.T.; Potter, L.C. A Bayesian perspective on sparse regularization for STAP post-processing. In Proceedings of the 2010 IEEE Radar Conference, Washington, DC, USA, 10–14 May 2010; pp. 1471–1475. [[CrossRef](#)]
18. Baron, D.; Wakin, M.B.; Duarte, M.F.; Sarvotham, S.; Baraniuk, R.G. *Distributed Compressed Sensing*; Technical Report ECE-0612; Electrical and Computer Engineering Department, Rice University: Houston, TX, USA, 2006.
19. Baron, D.; Duarte, M.F.; Sarvotham, S.; Wakin, M.B.; Baraniuk, R.G. An Information-Theoretic Approach to Distributed Compressed Sensing. In Proceedings of the Allerton Conference on Communication Control & Computing, Monticello, IL, USA, 26–28 September 2005.
20. Schnelle, S.R.; Laska, J.N.; Hegde, C.; Duarte, M.F.; Davenport, M.A.; Baraniuk, R.G. Texas hold ‘Em algorithms for distributed compressive sensing. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Dallas, TX, USA, 14–19 March 2010; pp. 2886–2889. [[CrossRef](#)]
21. Valsesia, D.; Coluccia, G.; Magli, E. Joint recovery algorithms using difference of innovations for distributed compressed sensing. In Proceedings of the Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 3–6 November 2013; pp. 414–417.
22. Sundman, D.; Chatterjee, S.; Skoglund, M. Parallel pursuit for distributed compressed sensing. In Proceedings of the IEEE Global Conference on Signal and Information Processing, Austin, TX, USA, 2–5 December 2013; pp. 783–786.
23. Matamoros, J.; Fosson, S.M.; Magli, E.; Anton-Haro, C. Distributed ADMM for in-network reconstruction of sparse signals with innovations. *IEEE Trans. Signal Inf. Process. Netw.* **2015**, *4*, 225–234. [[CrossRef](#)]
24. Tzikas, D.G.; Likas, A.C.; Galatsanos, N.P. The variational approximation for Bayesian inference. *IEEE Signal Process. Mag.* **2008**, *6*, 131–146. [[CrossRef](#)]
25. Chen, W.; Wassell, I.J. Variational Bayesian Algorithm for Distributed Compressive Sensing. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 4889–4894. [[CrossRef](#)]

26. Chen, W.; Wassell, I.J. A Decentralized Bayesian Algorithm for Distributed Compressive Sensing in Networked Sensing Systems. *IEEE Trans. Wirel. Commun.* **2016**, *2*, 1282–1292. [\[CrossRef\]](#)
27. Lustig, M.; Donoho, D.; Pauly, J.M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* **2007**, *6*, 1182–1195. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Herman, M.A.; Strohmer, T. High-resolution radar via compressed sensing. *IEEE Trans. Signal Process.* **2009**, *6*, 2275–2284. [\[CrossRef\]](#)
29. Mamaghanian, H.; Khaled, N.; Atienza, D.; Vanderghenst, P. Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes. *IEEE Trans. Biomed. Eng.* **2011**, *9*, 2456–2466. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Liu, J.; Zhang, G.; Zhao, K.; Jiang, X. Compressive holography algorithm for the objects composed of point sources. *Appl. Opt.* **2017**, *1*, 530–542. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Candes, E.J.; Plan, Y. Matrix completion with noise. *Proc. IEEE* **2010**, *6*, 925–936. [\[CrossRef\]](#)
32. Karakus, C.; Gurbuz, A.C.; Tavli, B. Analysis of energy efficiency of compressive sensing in wireless sensor networks. *IEEE Sens. J.* **2013**, *5*, 1999–2008. [\[CrossRef\]](#)
33. Li, S.; Qi, H. A douglas rachford splitting approach to compressed sensing image recovery using low-rank regularization. *IEEE Trans. Image Process.* **2015**, *11*, 4240–4249. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Zhu, S.; Liao, G.; Wang, W.; Zeng, C.; Yang, D. Wideswath synthetic aperture radar ground moving targets indication with low data rate based on compressed sensing. *IET Radar Sonar Navig.* **2013**, *9*, 1027–1034. [\[CrossRef\]](#)
35. Yang, J.; Liu, C.; Wang, Y. Imaging and parameter estimation of fast-moving targets with single-antenna SAR. *IEEE Geosci. Remote Sens. Lett.* **2014**, *2*, 529–533. [\[CrossRef\]](#)
36. Wang, Z.R.; Xu, J.; Huang, Z.Z.; Zhang, X.D.; Xia, X.G.; Long, T. Road-Aided Doppler Ambiguity Resolver for SAR Ground Moving Target in the Image Domain. *IEEE Geosci. Remote Sens. Lett.* **2016**, *10*, 1552–1556. [\[CrossRef\]](#)
37. Dong, Q.; Xing, M.D.; Xia, X.G.; Zhang, S.; Sun, G.C. Moving Target Refocusing Algorithm in 2-D Wavenumber Domain after BP Integral. *IEEE Geosci. Remote Sens. Lett.* **2018**, *1*, 127–131. [\[CrossRef\]](#)
38. Parisi, G.L. *Statistical Field Theory*; Addison-Wesley: Reading, MA, USA, 1988.
39. Tylavsky, D.J.; Sohie, G.R.L. Generalization of the matrix inversion lemma. *Proc. IEEE* **1986**, *7*, 1050–1052. [\[CrossRef\]](#)
40. Al-Shoukairi, M.; Schniter, P.; Rao, B.D. A GAMP-Based Low Complexity Sparse Bayesian Learning Algorithm. *IEEE Trans. Signal Process.* **2018**, *2*, 294–308. [\[CrossRef\]](#)
41. Yang, L.; Fang, J.; Duan, H.; Li, H.; Zeng, B. Fast Low-Rank Bayesian Matrix Completion with Hierarchical Gaussian Prior Models. *IEEE Trans. Signal Process.* **2018**, *11*, 2804–2817. [\[CrossRef\]](#)
42. Cumming, I.G.; Bennett, J.R. Digital Processing of SEASAT SAR Data. In Proceedings of the IEEE 1979 International Conference on Acoustics, Speech and Signal Processing, Washington, DC, USA, 2–4 April 1979; Volume 4, pp. 2–4.
43. Wang, W.; Zhu, Y.; Zhao, H.; Wu, S. Clutter Suppression and GMTI with Sparse Sampled Data for Dual-channel SAR. In Proceedings of the 2014 IEEE Radar Conference, Cincinnati, OH, USA, 19–23 May 2014; pp. 0280–0284.
44. Chi, Y.; Scharf, L.L.; Pezeshki, A.; Calderbank, A.R. Sensitivity to Basis Mismatch in Compressed Sensing. *IEEE Trans. Signal Process.* **2011**, *5*, 2182–2195. [\[CrossRef\]](#)
45. Khwaja, A.; Zhang, X.P. Compressed sensing based image formation of SAR/ISAR data in presence of basis mismatch. In Proceedings of the 2012 19th IEEE International Conference on Image Processing, Orlando, FL, USA, 30 September–3 October 2012; pp. 901–904.
46. Bernhardt, S.; Boyer, R.; Marcos, S.; Larzabal, P. Compressed Sensing with Basis Mismatch: Performance Bounds and Sparse-Based Estimator. *IEEE Trans. Signal Process.* **2016**, *13*, 3483–3494. [\[CrossRef\]](#)
47. Wang, H.S.C. Mainlobe clutter cancellation by DPCA for space-based radars. In Proceedings of the IEEE Aerospace Applications Conference Digest, Crested Butte, CO, USA, 3–8 February 1991; p. 1.

