Micro-Doppler Effect Removal in ISAR Imaging by Promoting Joint Sparsity in Time-Frequency Domain

Lin Sun and Weidong Chen *

Key Laboratory of Electromagnetic Space Information of the Chinese Academy of Sciences, University of Science and Technology of China, Hefei 230027, China; sunlin@mail.ustc.edu.cn
* Correspondence: wdchen@ustc.edu.cn; Tel.: +86-0551-6360-7705; Fax: +86-0551-6360-1326

Received: 1 February 2018; Accepted: 19 March 2018; Published: 23 March 2018

Abstract: For micromotion scatterers with small rotating radii, the micro-Doppler (m-D) effect interferes with cross-range compression in inverse synthetic aperture radar (ISAR) imaging and leads to a blurred main body image. In this paper, a novel method is proposed to remove the m-D effect by promoting the joint sparsity in the time-frequency domain. Firstly, to obtain the time-frequency representations of the limited measurements, the short-time Fourier transform (STFT) was modelled by an underdetermined equation. Then, a new objective function was used to measure the joint sparsity of the STFT entries so that the joint sparse recovery problem could be formulated as a constrained minimization problem. Similar to the smoothed \(l_0\) (SL0) algorithm, a steepest descend approach was used to minimize the new objective function, where the projection step was tailored to make it suitable for m-D effect removal. Finally, we utilized the recovered STFT entries to obtain the main body echoes, based on which cross-range compression could be realized without m-D interference. After all contaminated range cells were processed by the proposed method, a clear main body image could be achieved. Experiments using both the point-scattering model and electromagnetic (EM) computation validated the performance of the proposed method.

Keywords: inverse synthetic aperture radar imaging; micro-Doppler effect; joint sparsity; short time Fourier transform; time-frequency domain

1. Introduction

In inverse synthetic aperture radar (ISAR) imaging, mechanical rotations or vibrations of structures on a target may introduce additional frequency modulations on the returned signal, known as the micro-Doppler (m-D) effect [1–3]. The main body image is usually blurred because of the interference of the rotating or vibrating scatterers, which are also called micromotion scatterers [3]. Therefore, the m-D effect should be removed in order to obtain a clear image of the main body.

Micromotion targets with rotating structures can often be found in real world scenarios, e.g., a ship with scanning antennas or an aircraft in flight equipped with turbofans. A micromotion scatterer with a large rotating radius will generate a m-D signal which can be depicted in the form of sinusoidal modulations in the spectrogram after range compression [4,5]. In contrast, the main body signal in the spectrogram has the shape of straight lines [5]. Based on the different shapes in the spectrogram, many methods [4–7] have been proposed to eliminate the m-D signal. Li and Ling [4] put forward an adaptive chirplet decomposition algorithm in which the returned signal is decomposed into a family of chirplet functions. The main body signal and the micromotion signal subsequently are separated according to their distinct chirp rates. However, this algorithm has high computation cost because of the large chirplet dictionary. To remove the sinusoidal m-D interference, the method in Q. Zhang et al. [6] extracts the straight lines in the spectrogram using the Hough transform. Similarly, the methods in H.C. Liu et al. and L. Sun et al. [5,7] only recover the straight lines via
the sparse representation (SR)-based algorithms where the sinusoids are eliminated. Nevertheless, the aforementioned methods [5-7] cannot remove the m-D effect generated by a micromotion scatterer with a small rotating radius, since the m-D signal has the same straight line shape as the main body signal in the spectrogram when the rotating radius is smaller than half of the range resolution.

Recently, some methods based on the time-frequency analysis have been developed to address the problem of the removal of the m-D effect under a small rotating radius. L. Stanković et al.’s method [8], based on L-statistics, performs the short-time Fourier transform (STFT) to the echoes in the contaminated range cell. It is assumed in L. Stanković et al. [8] that the m-D interference occupies a small portion of time instants at each frequency bin. Using the L-statistics, a fixed fraction of the most significant STFT entries were eliminated to achieve m-D effect removal. Subsequently, the remaining entries were utilized to recover the main body signal. However, a large amount of STFT entries corresponding to the main body might be removed together with the m-D interference, which leads to a high sidelobe level in the imaging result [9]. A method based on histogram analysis is proposed in R. Zhang et al. [9], where the STFT entries with a high frequency of occurrence were regarded as the main body components and were preserved to reconstruct the ISAR image. On the other hand, the STFT entries with a low frequency of occurrence were considered to correspond to the m-D interference and suppressed. When the m-D effect is severe, the method in R. Zhang et al. [9] cannot remove the m-D signal in the time-frequency domain completely because the m-D components might be mistaken for the main body signal. Consequently, there are some spurious points in the imaging result due to the residual m-D signal. In real world situations, the data samples of the echoes might be randomly missing at given time instants [10] where strong electromagnetic interference or sensor failure prevents effective observation. As a result, the effective pulses are limited. The L-statistics-based method is combined with the short-time compressed sensing (STCS) [11] approach in Q.K. Hou et al. [12] to reduce the m-D effect with limited pulses. The STCS was employed to recover the STFT entries where the number of the frequency bins was equal to the number of full pulses. Nevertheless, the window width in the STCS could have been much shorter than the signal duration to obtain the local frequency characteristics. Thus, the measurements in the window were much fewer than the frequency bins in the time-frequency domain. The frequency resolution of the recovered STFT entries was reduced due to the extremely limited measurements. As a result, the L-statistics-based method in Q.K. Hou et al. [12] cannot obtain the accurate support of the main body signal in the frequency dimension and some spurious points exist near the main body scatterers in the imaging result.

The main body scatterer has a constant Doppler frequency [13]. Thus, the support of the main body signal is a slow time invariant in the time-frequency domain which indicates the joint sparsity [14,15] of the main body signal. In contrast, the Doppler frequency of the micromotion scatterer varies with the slow time [13]. Based on the distinct patterns in the time-frequency domain, this paper proposes a joint sparsity-based ISAR imaging method to remove the m-D effect generated by the micromotion scatterers with small rotating radii. Firstly, the echoes in the contaminated range cell were modelled by an underdetermined equation where the relation of the echoes and the corresponding STFT entries is built on the STFT matrix. Generally, the $l_2/l_0$-norm [16] of the STFT entry vector could be used to measure the joint sparsity in the time-frequency domain because it is equal to the number of frequency bins having nonzero STFT entries. Nonetheless, minimizing the $l_2/l_0$-norm leads to a combinatorial optimization problem [17] that is difficult to solve. In M. Bevacqua et al. [18], the $l_1$-norm of an auxiliary variable, defined as the common upper bound to the amplitudes of the electric currents, is utilized to replace the $l_2/l_0$-norm so that the new problem is tractable. However, the difference between the $l_1$-norm of the auxiliary variable and the original $l_2/l_0$-norm is significant [19] because larger coefficients are penalized more heavily than smaller coefficients in the $l_1$-norm. A reweighted $l_1$-norm method is proposed in Y. Liu et al. [19] to overcome the drawback of the $l_1$-norm, but this method is lack of convergence guarantee [20]. Recently, the $l_2/l_1$-norm was utilized to promote the joint sparsity of highly conductive scatterer in microwave imaging [21]. Nonetheless, the $l_2/l_1$-norm is a loose approximation to the original $l_2/l_0$-norm and often leads to suboptimal solutions [22]. Different from
where \( \hat{\mathbf{v}} \) with velocity \( \mathbf{v} \) is defined as
\[
\parallel \mathbf{v} \parallel \text{with complex elements.}
\]

The transmitted chirp signal is
\[
s(t, t_m) = \text{rect} \left( \frac{t}{T} \right) \exp \left( j2\pi \left( f_c t + \frac{1}{2} \gamma t^2 \right) \right)
\]
where \( t \) and \( t_m \) denote the fast time and the slow time, respectively. \( T \) is the pulse duration, \( f_c \) is the carrier frequency, and \( \gamma \) is the chirp rate. \( \text{rect}(\cdot) \) represents the rectangular function, which is defined as
\[
\text{rect} \left( \frac{t}{T} \right) = \begin{cases} 1, & |t| \leq T/2 \\ 0, & |t| > T/2. \end{cases}
\]
where $s_{\text{ref}}(\hat{t}, t_m) = \text{rect}\left(\frac{\hat{t} - 2 R_u(t_m)}{T_{\text{ref}}}\right) \exp\left(j 2\pi \left(f_c \left(\hat{t} - 2 R_u(t_m) c^{-1}\right) + t_m\right) + \frac{1}{2} \gamma \left(\hat{t} - 2 R_u(t_m) c^{-1}\right)^2\right)$ \hspace{1cm} (3)

where $T_{\text{ref}}$ is the duration of the reference signal with $T_{\text{ref}} > T$, and $c$ denotes the speed of light. After dechirping and removing the residual video phase, the echo can be expressed as

$$s_e(\hat{t}, t_m) = \sum_{p=1}^{p} \sigma_p \text{rect}\left(\frac{\hat{t} - 2 R_p(t_m)}{T_{\text{ref}}}\right) \exp\left(-j \frac{4\pi}{c} \left(f_c + \gamma \left(\hat{t} - 2 R_u(t_m) c^{-1}\right)\right) A_P(t_m)\right)$$

$$+ \sum_{q=1}^{q} \sigma_q \text{rect}\left(\frac{\hat{t} - 2 R_q(t_m)}{T_{\text{ref}}}\right) \exp\left(-j \frac{4\pi}{c} \left(f_c + \gamma \left(\hat{t} - 2 R_u(t_m) c^{-1}\right)\right) A_Q(t_m)\right) \hspace{1cm} (4)$$

where $\sigma_p$ and $\sigma_q$ denote the backscattering coefficients of the $p$th main body scatterer and the $q$th micromotion scatterer, respectively. After the Fourier transform to Equation (4) along the $\hat{t}$ dimension, the range compression is achieved and the spectrogram can be written as

$$s_r(\hat{f}, t_m) = \sum_{p=1}^{p} A_P \sin c\left(T(\hat{f} + 2 \frac{\gamma}{c} A_P(t_m))\right) \exp\left(-j \frac{4\pi}{c} \Delta R_P(t_m)\right)$$

$$+ \sum_{q=1}^{q} A_Q \sin c\left(T(\hat{f} + 2 \frac{\gamma}{c} A_Q(t_m))\right) \exp\left(-j \frac{4\pi}{c} \Delta R_Q(t_m)\right) \hspace{1cm} (5)$$

where $\hat{f}$ is the fast frequency and $\lambda$ is the wavelength. $A_P = \sigma_p T$ and $A_Q = \sigma_q T$ denote the complex coefficients of scatterers $p$ and $q$, respectively.

It is assumed that the target has a constant velocity $v_x$ in a short coherent processing interval (CPI). Thus, the aspect angle variation of the target is calculated to be $v_x t_m / R_u$. Because the target is far from the radar, the instantaneous distances from $p$ and $q$ to the reference point can be approximately expressed as

$$\Delta R_P(t_m) = x_p \frac{v_xt_m}{R_u} + y_p - y_u \hspace{1cm} (6)$$

$$\Delta R_Q(t_m) = x_q \frac{v_xt_m}{R_u} + y_q - y_u + r_q \sin(\omega_q t_m + \theta_q) \hspace{1cm} (7)$$

Substituting Equations (6) and (7) into Equation (5) yields

$$s_r(\hat{f}, t_m) \approx \sum_{p=1}^{p} A_P \sin c\left(T\left(\hat{f} + 2 \frac{\gamma}{c} (y_p - y_u)\right)\right) \exp\left(-j \frac{4\pi}{c} \left(x_p \frac{v_xt_m}{R_u} + y_p - y_u\right)\right)$$

$$+ \sum_{q=1}^{q} A_Q \sin c\left(T\left(\hat{f} + 2 \frac{\gamma}{c} (y_q - y_u + r_q \sin(\omega_q t_m + \theta_q))\right)\right) \exp\left(-j \frac{4\pi}{c} \left(x_q \frac{v_xt_m}{R_u} + y_q - y_u + r_q \sin(\omega_q t_m + \theta_q)\right)\right) \hspace{1cm} (8)$$

It can be seen from Equation (8) that the peaks of the spectrogram are located at

$$\hat{f}_p = -\frac{2\gamma}{c} (y_p - y_u) \hspace{1cm} (9)$$

$$\hat{f}_q = -\frac{2\gamma}{c} (y_q - y_u + r_q \sin(\omega_q t_m + \theta_q)) \hspace{1cm} (10)$$

and the range resolution is $r_r = c/(2\gamma T)$. For the targets with large rotors, the rotating radius $r_q$ might be much larger than the range resolution. Equation (8) indicates that the signal of the micromotion scatterer migrates through the range cells in the spectrogram and has the shape of a sinusoid. In contrast, the position of the main body signal is slow time invariant, which has the shape of a straight line. The sinusoid in the spectrogram will seriously affect the cross-range compression, i.e., the integration in the slow time domain, and lead to a smeared ISAR image. To remove the
sinusoidal m-D interference, the method in Q. Zhang et al. [6] extracts the straight lines in the spectrogram using the Hough transform. Similarly, the methods in H.C. Liu et al. and L. Sun et al. [5,7] only recover the straight lines where the sinusoids are eliminated. However, when the target has small rotors, \( r_q \) might be smaller than half of the range resolution. In this condition, the signal of the micromotion scatterer is located in a constant range cell and also has the shape of a straight line. The spectrogram can be approximately expressed as

\[
s_r(f, t_m) \approx \sum_{p=1}^{P} A_p \sin c \left( T \left( f + \frac{2\pi}{c} (y_p - y_u) \right) \right) \exp \left( -j \frac{4\pi}{\lambda} (x_p \frac{\nu_x}{R_u} + y_p - y_u) \right) \\
+ \sum_{q=1}^{Q} A_q \sin c \left( T \left( f + \frac{2\pi}{c} (y_{q'} - y_u) \right) \right) \exp \left( -j \frac{4\pi}{\lambda} (x_{q'} \frac{\nu_x}{R_u} + y_{q'} - y_u + r_q \sin (\omega_q t_m + \theta_q)) \right). \tag{11}
\]

Because the main body signal and the signal of the micromotion scatterer have the same straight line shapes in the spectrogram, the methods in H.C. Liu et al., Q. Zhang et al., and L. Sun et al. [5–7] are invalid to remove the m-D effect. According to Equation (11), the Doppler frequencies of the main body scatterer and the micromotion scatterer are

\[
f_p = -\frac{2\nu_x}{\lambda R_u} \tag{12}
\]

\[
f_q = -\frac{2}{\lambda} \left( x_{q'} \frac{\nu_x}{R_u} + r_q \omega_q \cos (\omega_q t_m + \theta_q) \right). \tag{13}
\]

It is obvious from Equations (12) and (13) that the main body scatterer has a constant Doppler frequency while the Doppler frequency of the micromotion scatterer is slow time variant. Therefore, it is possible to separate the main body signal from the micromotion signal in the time-frequency domain.

![Figure 1](image_url). Inverse synthetic aperture radar (ISAR) imaging geometry with rotating scatterers.

3. Joint Sparsity of Main Body Signal in Time-Frequency Domain

Typical time-frequency transforms include the Wigner–Ville distribution (WVD) [27] and the STFT. Although WVD has better time-frequency resolution than STFT, it suffers from cross-terms for multiple signal components. Thus, we chose the STFT to achieve the time-frequency representations of the echoes in the contaminated range cell. In order to obtain the STFT entries, the echoes in the contaminated range cell were first segmented into narrow time intervals so that the signal in each segment could be considered stationary. Then, the Fourier transform was carried out for the
windowed signal in each segment to obtain the local spectral representations. In this paper, we chose the rectangular window function and the segmentation of the echoes could be realized by a sliding window with proper time step.

To demonstrate the joint sparsity of the main body signal in the time-frequency domain, we conducted the following simulation. The target model was composed of nine main body scatterers and two micromotion scatterers, as described in Figure 2. The micromotion scatterers \(q_1\) and \(q_2\) rotated around the origin counterclockwise with radii of 0.1 m and 0.2 m, and with rotating frequencies of 10 Hz and 3 Hz, respectively. The initial rotating phases were \(\pi/2\) rad and 0 rad, respectively. The target moved along the cross-range direction with a velocity of 300 m/s. The radar carrier frequency was 10 GHz, the bandwidth was 600 MHz, the pulse repetition interval (PRI) was 1 ms, and the CPI was 1 s.

According to the radar system parameters, the range resolution was calculated to be 0.25 m, which was larger than twice the rotating radii of both \(q_1\) and \(q_2\). Thus, the micromotion signal did not migrate through the range cells in the spectrogram. As shown in Figure 3a, the micromotion signal was located in the 40th range cell and had the same straight line shape as the main body signal. In this situation, the methods in H.C. Liu et al., Q. Zhang et al., and L. Sun et al. [5–7] were invalid to eliminate the micromotion components. As a result from the m-D effect, the application of a fast Fourier transform (FFT) to the spectrogram along the slow time dimension would lead to a blurred main body image, as depicted in Figure 3b. We selected 50 cross-range cells where the target was located in the imaging result to display.

![Diagram](image)

**Figure 2.** Target model.

![Diagram](image)

**Figure 3.** (a) Spectrogram; (b) Imaging result by the FFT-based method.

To obtain the time-frequency representations, we applied the STFT to the echoes in the 40th range cell of the spectrogram. The time step was 10 PRI and the window width was 30 PRI. Thus, there were...
30 frequency bins in the time-frequency domain. As shown in Figure 4, the main body signal was located at the 15th and 16th frequency bins. Consequently, the support of the main body signal in the time-frequency domain was slow time invariant which indicated the joint sparsity of the main body signal. In contrast, the micromotion scatterers’ Doppler frequencies were slow time variant and the micromotion signal had the shape of sinusoids, as illustrated in Figure 4. Because of the distinct patterns, the micromotion signal could be removed by promoting the joint sparsity in the time-frequency domain, which is detailed in the next section.

4. M-D Effect Removal by Promoting Joint Sparsity

4.1. Problem Formulation for Joint Sparse Recovery

The echoes in the range cell contaminated by the m-D effect are denoted by \( s \), and we assume that the number of the received pulses is \( V \). The window width and the time step of the STFT are denoted by \( N \) and \( G \), respectively. Thus, the number of the segments in the STFT is \( M = \lceil (V - N) / G \rceil + 1 \), where \( \lceil \cdot \rceil \) denotes rounding to the nearest integer towards infinity.

The discrete Fourier matrix for each segment is

\[
Q = \begin{bmatrix}
1 & 1 & \cdots & 1 \\
1 & W^{-1} & \cdots & W^{-(N-1)} \\
\vdots & \vdots & \ddots & \vdots \\
1 & W^{-(N-1)} & \cdots & W^{-(N-1)(N-1)}
\end{bmatrix}_{N \times N}
\]  

(14)

where \( W = e^{j2\pi / N} \). Denoting the STFT entries in Figure 4 by \( X \), the corresponding STFT entry vector is \( x \), which is generated by stacking the columns of \( X \) into a single vector. Then, the STFT of \( s \) can be written in a compact matrix form

\[
x = Hs
\]

(15)

where \( H \in \mathbb{C}^{K \times V} \) denotes the STFT matrix

\[
H = \begin{bmatrix}
Q & 0_{N \times (V-N)} \\
0_{N \times G} & Q & 0_{N \times (V-G-N)} \\
0_{N \times 2G} & Q & 0_{N \times (V-2G-N)} \\
\vdots & \vdots & \vdots \\
0_{N \times (V-N)} & Q & 0_{N \times (V-N)}
\end{bmatrix}_{K \times V}
\]  

(16)

and \( K = MN \) is the number of the STFT entries in \( x \). \( 0_{N \times G} \) denotes a matrix of size \( N \times G \), whose elements are equal to 0.
The echoes in the contaminated range cell can be modeled as follows:

\[ s = Dx + \phi \]  

(17)

where \( \phi \in \mathbb{C}^{V \times 1} \) denotes the noise, and \( D = (H^H H)^{-1} H^H \) is the Moore–Penrose inverse of the STFT matrix \( H \). However, the data samples of the echoes might be randomly missing in real world scenarios where strong electromagnetic interference or sensor failure prevents effective observation. The data samples under strong electromagnetic interference must be discarded and the remaining \( L \) samples constitute the measurement vector \( y \in \mathbb{C}^{L \times 1} \), which can be expressed as

\[ y = Ax + n \]  

(18)

where \( n \in \mathbb{C}^{L \times 1} \) is the noise. \( A \in \mathbb{C}^{L \times K} \) is formed by omitting the rows of \( D \) which correspond to the missing samples. Because the remaining samples are fewer than the STFT entries to recover, i.e., \( L < K \), Equation (18) is underdetermined.

To facilitate the notation, we rearrange the STFT entries in the \( i \)th frequency bin into a vector

\[ b_i = [x_i, x_{N+i}, \ldots, x_{MN+i}], 1 \leq i \leq N. \]  

(19)

According to the analysis in Section 3, the micromotion components in the time-frequency domain can be eliminated by promoting the joint sparsity of \( x \). The joint sparse recovery problem is usually formulated by the following \( l_2/l_0 \) minimization problem

\[ \hat{x} = \arg \min_x \|x\|_{2,0} \text{ s.t. } \|y - Ax\|_2 < \epsilon \]  

(20)

where \( \epsilon \) bounds the \( l_2 \)-norm of the noise in the measurements. \( \|x\|_{2,0} \) stands for the \( l_2/l_0 \)-norm of \( x \) defined as [16]

\[ \|x\|_{2,0} = \|\beta\|_0 \]  

(21)

where \( \beta = [\|b_1\|_2, \|b_2\|_2, \ldots, \|b_N\|_2]^T \). It is obvious that \( \|x\|_{2,0} \) is equivalent to the number of frequency bins which have nonzero STFT entries. Therefore, by solving the problem in Equation (20), we can find the solution \( \hat{x} \) which occupies the fewest frequency bins and the joint sparsity in the time-frequency domain can be promoted.

However, the problem in Equation (20) is essentially a combinatorial optimization problem [17], which is difficult to solve. Inspired by the smoothed \( l_0 \) (SL0) algorithm, we propose a novel objective function to replace \( \|x\|_{2,0} \), so that the new problem is tractable. In the following subsections, the original SL0 algorithm is briefly reviewed at first. Then, we elaborate the proposed ISAR imaging method.

4.2. Brief Review of SL0 Algorithm

The basic idea of the SL0 algorithm is to approximate the \( l_0 \)-norm of \( x \) with the following function [23]

\[ F_\sigma(x) = \sum_{i=1}^{MN} f_\sigma(x_i) \]  

(22)

where

\[ f_\sigma(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right) \]  

(23)

and

\[ \lim_{\sigma \to 0} f_\sigma(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x \neq 0 \end{cases} \]  

(24)
It can be seen from Equations (22) and (24) that \( \|x\|_0 \approx MN - F_\sigma(x) \) for small values of \( \sigma \). Furthermore, this approximation becomes equality as \( \sigma \to 0 \). The SL0 algorithm obtains the sparse solution of Equation (18) by solving the following problem

\[
\tilde{x} = \arg \max_x F_\sigma(x) \text{ s.t. } \|y - Ax\|_2 \leq \epsilon. \tag{25}
\]

To avoid trapping into local maxima, a decreasing sequence of \( \sigma \) is utilized, i.e., \([\sigma_1, \ldots, \sigma_\eta]\), where \( \sigma_i > \sigma_j \) for \( i > j \). \( \eta \) is the iteration number and the current maximizer of \( F_\sigma(x) \) is used as the starting point of the next iteration. To maximize \( F_\sigma(x) \) for a fixed \( \sigma \), there are two steps in each iteration: the unconstrained maximization step \( x' \leftarrow x + \mu \nabla F_\sigma(x) \) and the projection to the feasible set

\[
x = \arg \min_x \|x - x'\|_2 \text{ s.t. } y = Ax
= x' - \hat{A}^\dagger (Ax' - y) \tag{26}
\]

where \( \hat{A}^\dagger = A^H (AA^H)^{-1} \), and \( \nabla F_\sigma(x) \) is the gradient of \( F_\sigma(x) \).

### 4.3. Proposed Method

Although it has been shown in H. Mohimani et al. [23] that the SL0 algorithm is effective to obtain the sparse solution of Equation (18), the joint sparsity of \( x \) cannot be promoted. It can be seen from Equation (22) that the SL0 algorithm only promotes the sparsity of the individual entries in \( x \), and \((MN - F_\sigma(x))\) is merely able to approximate \( \|x\|_0 \) rather than \( \|x\|_{2,0} \), which measures the joint sparsity of \( x \). As a result, the SL0 algorithm is inappropriate to remove the m-D effect in ISAR imaging for micromotion targets. In this paper, we propose a smoothed \( l_2/l_0 \) (SL2L0) algorithm, where \( \|x\|_{2,0} \) in Equation (20) is approximated by a new objective function

\[
U_\sigma(x) = N - \sum_{i=1}^{N} f_\sigma \left( \|b_i\|_2^2 \right)
= N - \sum_{i=1}^{N} \exp \left( -\frac{\sum_{j=1}^{M} |x_{(j-1)N+i}|^2}{2\sigma^2} \right). \tag{27}
\]

Since \( \|x\|_{2,0} \) is equivalent to the number of frequency bins which have nonzero STFT entries, it is obvious from Equations (19), (24) and (27) that \( U_\sigma(x) = \|x\|_{2,0} \) when \( \sigma \to 0 \). Therefore, the joint sparsity can be promoted by minimizing the new objective function \( U_\sigma(x) \) with small values of \( \sigma \). The joint sparse recovery problem can be formulated as

\[
\overline{x} = \arg \min_x U_\sigma(x) \text{ s.t. } \|y - Ax\|_2 \leq \epsilon. \tag{28}
\]

Similar to the SL0 algorithm, a decreasing sequence of \( \sigma \) is used to avoid trapping into local minimum and a steepest descend approach is employed to minimize \( U_\sigma(x) \) in each iteration, i.e.,

\[
x' = x - \left( 2\mu \sigma^2 \right) \nabla U_\sigma(x) \tag{29}
\]

where \( \mu \) is the parameter controlling the step-size \( (2\mu \sigma^2) \). The gradient \( \nabla U_\sigma(x) \) in Equation (29) is calculated to be

\[
\nabla U_\sigma(x) = 1_N \otimes \alpha \otimes x \tag{30}
\]

where \( 1_N \) is a vector of length \( N \) with its entries equal to 1, \( \otimes \) is the Kronecker product, \( \odot \) is the Hadamard product, and \( \alpha \in \mathbb{C}^{N \times 1} \) is
\[
\alpha = \left[ \frac{1}{2\sigma^2} \exp \left( -\frac{\sum_{j=1}^{M} x_{(j-1)N+1}^2}{2\sigma^2} \right) \right] T . \tag{31}
\]

Since the micromotion components are eliminated by promoting the joint sparsity of \( x \), we modify the projection step in Equation (26) as follows:

\[
x = \arg \min_{x} \| x - x' \|_1 \text{ s.t. } \| y - Ax \|_2 \leq \lambda \tag{32}
\]

where \( \lambda \) bounds the \( l_2 \)-norm of the micromotion components in the echoes. Equation (32) can be rewritten into an unconstrained optimization problem

\[
z = \arg \min_{z} \frac{1}{2} \| y' - Az \|_2 + \tau \| z \|_1 \tag{33}
\]

where \( z = x - x' \) and \( y' = y - Ax' \). \( \tau \) is the regularization parameter related to \( \lambda \). \( g(z) = \| z \|_1 \) is called the regularizer which indicates that the optimization problem in Equation (33) is convex. We use the fast iterative shrinkage-thresholding algorithm (FISTA) \([28,29]\) to solve the problem in Equation (33) efficiently. The key concept in FISTA is known as Moreau’s proximal operator, or proximal operator for short, which is expressed as

\[
\text{prox}_{\theta g}(z_k) = \frac{z_k}{|z_k|} \max(|z_k| - \theta \tau, 0), 1 \leq k \leq K,
\]

and \( \theta \) in Equation (34) is a positive constant with \( \theta \geq 1/\| A \|_F^2 \).

In summary, the successive steps of the proposed SL2L0 algorithm are given in Algorithm 1.

**Algorithm 1 Proposed smoothed \( l_2/l_0 \) (SL2L0) algorithm**

**Input:**
- The echo sampling vector \( y \), the overcompleted dictionary \( A \), positive constants \( \{\mu, \theta, \epsilon\} \), the maximum number of outer loop iterations \( \eta \), the maximum number of inner loop iterations \( \phi \), a suitable decreasing sequence \( \{\sigma_1, \ldots, \sigma_{\eta}\} \), the tolerance values \( \epsilon z \) and \( \epsilon x \); initialization \( x^{(1)} = A^t y \) and \( t^{(1)} = 1 \).

**Iteration:**

For \( i = 1, \ldots, \eta \)

- \( \sigma = \sigma_i \);
- \( x' = x^{(i)} - (2\mu\sigma^2) \nabla U_e(x^{(i)}) \);
- \( y' = y - Ax' \);
- \( z^{(0)} = x^{(i)} - x' \);
- \( q^{(1)} = z^{(0)} \);
- For \( j = 1, \ldots, \phi \)
  - \( z^{(j)} = \text{prox}_{\theta g}(q^{(j)} - \theta A^H(Aq^{(j)} - y')) \);
  - \( t^{(j+1)} = \frac{1 + \sqrt{1 + 4t^{(j)}}}{2} \);
  - \( q^{(j+1)} = q^{(j)} + \frac{t^{(j+1)}}{t^{(j)}} (z^{(j)} - z^{(j-1)}) \);
  - if \( \| z^{(j)} - z^{(j-1)} \|_2 / \| z^{(j-1)} \|_2 < \epsilon z \), then break;
- \( x^{(i+1)} = x' + z^{(j)} \);
- if \( \| y - Ax^{(i+1)} \|_2 < \epsilon x \) or \( \| x^{(i+1)} - x^{(i)} \|_2 / \| x^{(i)} \|_2 < \epsilon x \), then break;

End

**Output:** \( \bar{x} = x^{(i+1)} \)
By promoting the joint sparsity in the time-frequency domain, the output $\tilde{x}$ of the SL2L0 algorithm only contains the main body components. Based on the recovered STFT entries, we can obtain the main body echoes free of m-D effect as $\tilde{y} = A\tilde{x}$. Assuming that the residual noise in $\tilde{y}$ is $e \in \mathbb{C}^{L \times 1}$, $\tilde{y}$ can be rewritten as

$$\tilde{y} = \Omega p + e$$

where $p \in \mathbb{C}^{V \times 1}$ is the cross-range compression result of $\tilde{y}$, and $\Omega \in \mathbb{C}^{L \times V}$ is a partial Fourier matrix with its $(l,v)$th element given by $\exp(-j2\pi f_v t_l)$. $f_v$ is the $v$th Doppler frequency bin, and $t_l$ is the $l$th sampling point of the slow time. Subsequently, we used the sparse Bayesian learning (SBL) [25] algorithm to achieve the cross-range compression of $\tilde{y}$ by recovering $p$ in Equation (35).

Although the discussions in this section are focused on the echoes in a single range cell, the proposed method is also suitable for other contaminated range cells. After all of the contaminated range cells are processed by the proposed method, a clear main body image can be achieved.

5. Experimental Results and Performance Comparisons

In this section, we discuss the experiments conducted using the echoes generated by the point-scattering model and EM computation, respectively. To validate the effectiveness, the imaging results of the proposed method and other existing methods are compared. The parameters of the proposed method used in the experiments are set as follows: $\mu = 1, \theta = 1/\|A\|_F^2, \eta = 40, \phi = 20, \sigma_i = 0.8^{-1}$ for $i = 1, \ldots, \eta$, $e_z = 10^{-4}$, and $e_x = 10^{-4}$. Since $\epsilon$ is related to the noise level, this parameter is tuned manually in each experiment.

5.1. Experiments Using the Point-Scattering Model

The target model is depicted in Figure 2, and the radar system parameters are the same as in Section 3. We obtained the limited measurements by choosing 500 pulses randomly in the slow time domain, and the signal-to-noise ratio (SNR) was 20 dB. The number of cross-range cells to be recovered was 1000. The window width of the STFT was 30 PRI and the time step was 10 PRI. According to the above parameters, the size of $A$ was calculated to be $500 \times 2940$. Similar to Figure 3b, we selected 50 cross-range cells to display in the imaging results.

To eliminate the micromotion components of the echoes in the contaminated range cell, both the method in R. Zhang et al. [9] and the L-statistics-based method utilize the time-frequency representations obtained by the STCS [11]. However, part of the main body signal in the time-frequency domain is missing due to the discarded pulses, as illustrated in Figure 5. Additionally, there is a strong m-D signal which has the shape of sinusoids. As a result, the histogram analysis in R. Zhang et al. [9] could not be used to remove the m-D signal, which might have been mistaken for the main body signal. Figure 6a shows the imaging result produced by the method in R. Zhang et al. [9]. Many spurious points exist in the 40th range cell due to the m-D interference and the main body scatterers are defocused. Because the measurements within the sliding window are much fewer than the frequency bins in STCS, the frequency resolution in the time-frequency domain is lower than the reciprocal of the CPI. Therefore, the L-statistics-based method cannot obtain the accurate support of the main body scatterers in the cross-range domain. As depicted in Figure 6b, some spurious points exist near the main body scatterers. In addition, one of the main body scatterers in the 40th range cell cannot be recovered because part of the main body signal is missing in Figure 5. In contrast to STCS, the proposed SL2L0 algorithm can remove the m-D signal by promoting the joint sparsity in the time-frequency domain, where only the main body signal with constant Doppler frequencies are preserved, as shown in Figure 7a. Based on the recovered STFT entries, the proposed method achieves a clear main body image in Figure 7b.

In addition to the visual results, we introduce the normalized mean square error (NMSE) and the image contrast (IC) as two metrics to compare the performance of different methods. A lower NMSE
means the image has less recovery error, and a higher IC indicates the image is better focused. The NMSE is computed by [30]

\[ \text{NMSE} = \frac{1}{U} \sum_{k=1}^{U} \left\| \frac{1}{\|I\|_F} I - \frac{1}{\|\hat{I}_k\|_F} \hat{I}_k \right\|_F \]  

(36)

where \( U = 100 \) independent trials are conducted, and \( \hat{I}_k \in \mathbb{C}^{R \times V} \) denotes the recovered result of the true main body image \( I \) in the \( k \)th trial. The IC is computed by X. Z. Gao et al. [24]

\[ \text{IC} = \frac{\sqrt{\text{Ave} \left( \hat{I}_k^2 - \text{Ave} \left( \hat{I}_k^2 \right) \right)^2}}{\text{Ave} \left( \hat{I}_k^2 \right)} \]  

(37)

where \( \hat{I}_k^2 \) represents the intensity matrix of the recovered main body image with its \((r, v)\)th element \( \hat{I}_k^2(r, v) \) equal to \( |\hat{I}_k(r, v)|^2 \). \( \text{Ave}(\cdot) \) denotes the mean operator, which is defined as

\[ \text{Ave}(I) = \frac{1}{RV} \sum_{r=1}^{R} \sum_{v=1}^{V} I(r, v) \]  

(38)

Figure 8 gives the NMSEs and ICs obtained by different methods against the undersampling factor [31], which is calculated as \( \psi = L/V \), i.e., the number of measurements in slow time dimension divided by the number of cross-range cells to recover. For each \( \psi \), the IC is averaged over 100 independent trials. It can be observed from Figure 8 that the proposed method achieves the lowest NMSEs and the highest ICs among all of the methods, which indicate the least recovery error and the best focused quality, respectively.

**Figure 5.** Time-frequency representations obtained by the short-time compressed sensing (STCS).

**Figure 6.** (a) Imaging result by the method in [9]; (b) Imaging result by the L-statistics-based method.
Within each pulse, there were 128 complex range samples. The window width and the time step of the STFT were 20 PRI and 2 PRI, respectively. According to the aforementioned parameters, the size of \( A \) was calculated to be \( 128 \times 2380 \).

Figure 7. (a) STFT entries recovered by the SL2L0 algorithm; (b) Imaging result by the proposed method.

Figure 8. Performance of different methods: (a) Normalized mean square error (NMSE) against undersampling factor; (b) Image contrast (IC) against undersampling factor.

5.2. Experiments Using EM Computation

To further testify the performance of the proposed method, we used the echoes of a ship target generated by the graphic electromagnetic computing (GRECO) technique [32], which combines physical optical (PO) theory and physical theory of diffraction (PTD). The target model is described in Figure 9a, where the dashed circle indicates the antenna rotating with a frequency of 1 Hz. Figure 9b depicts the imaging geometry. The ship was 6.9 km from the airborne radar and located at the origin. The radar operated at 5 GHz, the bandwidth was 1 GHz, and the radar moved along the opposite direction of the X-axis with a constant velocity of 300 m/s. 256 pulses were collected in a CPI of 2.3 s and the SNR was 20 dB. We obtained the limited measurements by choosing 128 pulses randomly. Within each pulse, there were 128 complex range samples. The window width and the time step of the STFT were 20 PRI and 2 PRI, respectively. According to the aforementioned parameters, the size of \( A \) was calculated to be \( 128 \times 2380 \).

The imaging results by different methods are shown in Figure 10. It can be seen from Figure 10a that the m-D effect led to a blurred main body image by the FFT-based method. Though the method in R. Zhang et al. [9] and the L-statistics-based method suppressed the m-D effect to some degree, there were residuals and some spurious points exist in Figure 10b,c. In contrast, the proposed method could remove the m-D effect through the SL2L0 algorithm and a clear image with the least spurious points is achieved in Figure 10d.
To quantitatively evaluate the performance of different methods, we used the image entropy (IE). Generally, a smaller IE indicates a better focused quality of the image. The IE is defined by W. Qiu et al. [33]

\[
IE = - \sum_{r=1}^{R} \sum_{v=1}^{V} I(r, v) \ln[I(r, v)] 
\]

(39)

where \(I\) is the power normalized image computed by

\[
I = \frac{I^2}{\sum_{r=1}^{R} \sum_{v=1}^{V} |I(r, v)|^2} 
\]

(40)

Figure 9. (a) Ship model; (b) Imaging geometry.

Figure 10. (a) Imaging result by the FFT-based method; (b) Imaging result by the method in R. Zhang et al. [9]; (c) Imaging result by the L-statistics-based method; (d) Imaging result by the proposed method.
The IEs of the imaging results are given in Table 1. The proposed method achieved the minimum IE which verifies the superiority of the proposed method to other methods.

The computational time of different methods was recorded based on the MATLAB code and summarized in Table 2, where the results are given by the averages of 50 independent trials conducted on an Intel Core i7 machine at 3.50 GHz. The computational time of the FFT-based method was not recorded since it could not suppress the m-D effect, as shown in Figure 10a. It can be seen from Table 2 that the proposed method is more computationally efficient than the method in R. Zhang et al. [9] and the L-statistics-based method.

### Table 1. Performance Evaluation by Image Entropy.

<table>
<thead>
<tr>
<th>Method</th>
<th>Image Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>The FFT-based method</td>
<td>5.977</td>
</tr>
<tr>
<td>The method in [9]</td>
<td>4.865</td>
</tr>
<tr>
<td>The L-statistics-based method</td>
<td>4.803</td>
</tr>
<tr>
<td>The proposed method</td>
<td>4.204</td>
</tr>
</tbody>
</table>

### Table 2. Computational Time.

<table>
<thead>
<tr>
<th>Method</th>
<th>Computational Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The method in [9]</td>
<td>5.529</td>
</tr>
<tr>
<td>The L-statistics-based method</td>
<td>3.406</td>
</tr>
<tr>
<td>The proposed method</td>
<td>1.811</td>
</tr>
</tbody>
</table>

6. Conclusions

In ISAR imaging, the m-D effect generated by the micromotion scatterers with small rotating radii leads to a blurred main body image. It is claimed that the Doppler frequency of the micromotion scatterer is slow time variant while the main body scatterer has a constant Doppler frequency, indicating the joint sparsity property in the time-frequency domain. To remove the m-D effect, this paper proposes a novel algorithm named SL2L0, where a new objective function is used to measure the joint sparsity and is minimized in a way similar to the SL0 algorithm. We have tailored the projection step in the SL0 algorithm to make it suitable for m-D effect removal. Experimental results showed that the proposed SL2L0 algorithm can effectively remove the m-D effect by promoting the joint sparsity in the time-frequency domain and a clear main body image can be achieved after cross-range compression. Diverse metrics, including the NMSE, IC, and IE, were employed to validate the superiority of the proposed method to other existing methods.

Acknowledgments: This work was supported by the National Natural Science Foundation of China under Grant No. 61431016.

Author Contributions: L.S. conceived the experiments, conducted the theoretical analyses, and wrote the paper; W.C. contributed analysis tools, revised the manuscript, and provided valuable suggestions.

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

References

11. Liu, Z.; Wei, X.Z.; Li, X. Aliasing-free micro-Doppler analysis based on short-time compressed sensing. *IET Signal Process.* 2013, 8, 176–187. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).