AFRL-AFOSR-VA-TR-2021-0004

Geometric Invariants for Radar Motion Estimation

Matthew Ferrara
MATRIX RESEARCH, INC
230 W 41ST STREET FL 7
NEW YORK, NY, 45432-2818
US

01/04/2021
Final Technical Report

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
04-01-2021	Final	29 Aug 2019 - 31 Jan 2021

4. TITLE AND SUBTITLE
Geometric Invariants for Radar Motion Estimation

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-19-1-0382

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Matthew Ferrara

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MATRIX RESEARCH, INC
230 W 41ST STREET FL 7
NEW YORK, NY 45432-2818
US

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AF Office of Scientific Research
875 N. Randolph St. Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR RTB1

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-VA-TR-2021-0004

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We describe an algorithm for reconstructing the motion of a radar platform relative to a scene, given video phase history and an estimate of the sensor's average speed and altitude. This algorithm is a more robust alternative to previous geometric-invariant-based approaches, capable of producing motion estimates accurate enough for the formation of back projection imagery. The algorithm is modular; in the first phase, the algorithm estimates the range to fixed locations in the scene as a function of time. In the second phase, these range tracks are used by a specialized solver to recover the relative positions of the tracked locations in the scene, along with the relative position of the sensor platform as a function of time. We demonstrate the effectiveness of the algorithm by forming an image from the large-scene Gotcha dataset, without the use of GPS or inertial measurement unit data. The results are compared to those from another invariant-based algorithm.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			MICHAEL YAKES
U	U	U	UU	28	**19b. TELEPHONE NUMBER** *(Include area code)* 00000000

Table of Contents

List of Figures

1.0 SUMMARY

This report describes an algorithm for reconstructing the motion of a radar platform relative to a scene, given video phase history (VPH) and an estimate of the sensor's average speed and altitude. This algorithm is a more robust alternative to previous geometric-invariant-based approaches, capable of producing motion estimates accurate enough for the formation of backprojection imagery. The algorithm is modular; in the first phase, the algorithm estimates the range to fixed locations in the scene as a function of time. In the second phase, these range tracks are used by a specialized solver to recover the relative positions of the tracked locations in the scene, along with the relative position of the sensor platform as a function of time. We demonstrate the effectiveness of the algorithm by forming an image from the large-scene Gotcha dataset, without the use of Global Positioning System (GPS) or inertial measurement unit data. The results are compared to those from another invariant-based algorithm.

2.0 INTRODUCTION

Imaging techniques in synthetic aperture radar (SAR) applications require accurate knowledge of the sensor locations to mitigate blurring and distortion of the produced image. Because the operators used to generate SAR images depend on the relative position of the transmit and receive platforms to the scene, any error in sensor position measurements creates a mismatch between the imaging operator and the data. The mismatch manifests itself as blurring in the image, and the choice of imaging operator is limited by the availability of accurate motion estimates. Often, position estimates produced by GPS or inertial measurement devices are not sufficent for generating focused imagery.

Any motion not accounted for in the image formation processor (IFP) causes an unanticipated phase adjustment in the complex pixel values of the finished image, and so many common approaches for focusing SAR images rely on phase adjustments to compensate for unanticipated motion [2, 3, 4, 5, 6].

These methods generate a 2-D relative motion correction (range and cross-range). Phase errors, however, are the result of any unknown 3-D motion. In order to reliably produce focused images, it is necessary to estimate the motion of the target relative to the sensor in the full 3-D space.

Several 3-D motion-estimation algorithms have been presented in the literature. For example, Cantalloube and Nahum [7] developed a SAR focusing method which refined the 3-D path in a two-step process. First, a collection of classical one-dimensional one-dimensional (1-D), translation-only, subscene autofocus solutions were generated for small regions within the radar beam footprint. Then, the collection of autofocus-based range corrections was inserted into a linear least squares (LLS) model for the trajectory errors. It is important to note that in order to arrive at an LLS model, a Taylor series expansion of the motion parameters was required. Ran, et al. [8] followed a multilateration-based approach similar to Cantalloube and Nahum, in which small-scene, translation-only, motion-error estimates were combined to arrive at a 3-D trajectory-error model. While the least-squares problem formulated by Ran, et al. again utilized a low-order, truncated-Taylor-series model for the motion errors, the resulting linear system was solved in a weighted-total-least-squares manner, taking into account information regarding the signal strength, and thus reliability, of each of the local autofocus solutions.

In a pair of papers from the Environmental Research Institute of Michigan (ERIM), Werness et al. [9, 10] describe a motion compensation method based on tracking prominent points in the image. Stuff further developed this idea in a series of papers [11, 12, 13, 14, 15, 16]. Theoretically, this signal processing approach makes it possible to track relative target-sensor motion in two or three dimensions, using only the radar data collected from a standard SAR system and without restricting the unknown motion model. After extracting a set of range measurements for at least four scattering centers, the approach describes a means to reconstruct the relative positions of the scatterers by exploiting invariants in the range data. The ranges from a radar to scattering centers on any rigid body must be constrained to a submanifold of the space of possible range observations, and this manifold determines the arrangement of the scattering centers up to rotation or reflection. With this arrangement,

the relative motion of target to sensor can then be determined.

The original approach was reduced to the computation of two singular value decomposition (SVD) and a linear solve by Ferrara et al. [1]. Up to this point, the literature was constrained to monostatic radar systems and targets so far from the sensor that the curvature of electromagnetic wave fronts could be ignored. This model does not apply for many realistic imaging scenarios, so Ferrara and Arnold extended the results for monostatic, near-field radar data [17] and for bistatic, far-field data [18]. With these improvements, the geometric invariant recovery was computationally expedient, but not numerically stable. Given perfect range measurements, the method is capable of exactly recovering the relative motion of the sensor to the scene, as well as the geometric configuration of scatterers in the scene. The output shape and motion estimates quickly diverge from this ideal reconstruction as range measurement quality degrades. Because tracking objects in range is a difficult problem in itself, this method was not widely applicable.

The aim of this work is to revisit the monostatic, near-field case [17], and describe a more numerically stable technique for recovering the relative motion of a radar sensor to a scene of interest. To demonstrate the feasibility of the updated technique, we will apply it to a large-scene radar dataset [19], recover the motion of the sensor, and form a high-resolution backprojection image with the estimated sensor path.

3.0 METHODS

3.1 Radar Data Model

In radar applications, it is convenient to think of the total electric field, $\mathcal{E}^{\text{tot}}$ as the superposition of an incident field, $\mathcal{E}^{\text{in}}$, and the scattered field, $\mathcal{E}^{\text{sc}}$. With this distinction, a simplified model for the propagation of a transmitted radar signal is

$$\left(\nabla^2 - c_0^{-2}\partial_t^2\right)\mathcal{E}^{\text{in}}(t,\mathbf{x}) = j(t,\mathbf{x}), \tag{1a}$$

$$\left(\nabla^2 - c^{-2}(\mathbf{x})\partial_t^2\right)\mathcal{E}^{\text{tot}}(t,\mathbf{x}) = j(t,\mathbf{x}), \tag{1b}$$

where c_0 is the speed of light in free space, and $c(\mathbf{x})$ is a spatially-dependent wave speed [20]. We assume that the only power input comes from our antenna, so that $j(t,\mathbf{x})$ models the current density on the antenna. In this description, (1a) models the uninterrupted propagation of the incident field through free space, and (1b) includes the interaction with any scatterers.

We can introduce the reflectivity function, $V(\mathbf{x}) = c_0^{-2} - c^{-2}(\mathbf{x})$, into the system composed of (1a) and (1b) by taking the difference of the two equations,

$$\left(\nabla^2 - c_0^{-2}\partial_t^2\right)\mathcal{E}^{\text{sc}}(t,\mathbf{x}) = -V(\mathbf{x})\partial_t^2\mathcal{E}^{\text{tot}}. \tag{2}$$

The partial differential equation in (2) is solved for $\mathcal{E}^{\text{sc}}$ by convolving the right hand side with the outgoing Green's function,

$$g(t,\mathbf{x}) = \frac{\delta(t - c_0^{-1}\left|\mathbf{x}\right|)}{4\pi\left|\mathbf{x}\right|}. \tag{3}$$

The result is the Lippmann-Schwinger integral equation,

$$\mathcal{E}^{\text{sc}}(t,\mathbf{x}) = \int\int g(t-\tau,\mathbf{x}-\mathbf{z})V(\mathbf{z})\partial_\tau^2\mathcal{E}^{\text{tot}}(\tau,\mathbf{z})\,\mathrm{d}\tau\,\mathrm{d}\mathbf{z}. \tag{4}$$

The Lippmann-Schwinger equation has better aesthetic in the frequency domain, so define the Fourier-transform of $\mathcal{E}$ to be

$$E(\omega) = \int e^{-i\omega t}\mathcal{E}(t)\,\mathrm{d}t. \tag{5}$$

Here, we use ν to denote frequency, $\omega = 2\pi\nu$ for angular frequency, and the exponential kernel sign convention of the FFTW library [21]. Additionally, we will let $k = c_0^{-1}\omega$ be the angular wavenumber. Then (4) takes the form

$$E^{\text{sc}}(\omega,\mathbf{x}) = -\int G(\omega,\mathbf{x}-\mathbf{z})V(\mathbf{z})\omega^2 E^{\text{tot}}(\omega,\mathbf{z})\,\mathrm{d}\mathbf{z}, \tag{6}$$

where

$$G(\omega,\mathbf{x}-\mathbf{z}) = \frac{e^{-ik|\mathbf{x}|}}{4\pi\left|\mathbf{x}\right|} \tag{7}$$

is the Fourier transform of g. Both V and E^{tot} are unknowns, so reconstructing V given sampled values of E^{sc} is a nonlinear inverse problem. It is common to linearize this problem by making the Born approximation, which amounts to replacing the total electric field with the incident electric field. While the Born approximation enables us to proceed with our analysis, it does not take into account the effects of multiple scattering and is not appropriate in some situations.

Assuming we can estimate the current density on our antenna, we can solve equation (1a) for the incident electric field by convolving with the outgoing Green's function. In the frequency domain, the partial differential equation (PDE) takes the form

$$(\nabla^2 + k^2)E^{\text{in}}(\omega, \mathbf{z}) = J(\omega, \mathbf{x}_t), \tag{8}$$

where we have approximated the transmit antenna with a point source at $\mathbf{x}_t$. Then we can solve for E^{in} at any point in the scene, $\mathbf{z}$,

$$E^{\text{in}}(\omega, \mathbf{z}) = \int G(\omega, \mathbf{z} - \mathbf{x}_t)J(\omega, \mathbf{x}_t)\,\mathrm{d}\mathbf{z}. \tag{9}$$

Now, the Born-approximated scattered field at the receiver position, $\mathbf{x}_r$, is

$$\begin{aligned} E_B^{\text{sc}}(\omega, \mathbf{x}_r, \mathbf{x}_t) &= -\int G(\omega, \mathbf{x}_r - \mathbf{z})\omega^2 E^{\text{in}}(\omega, \mathbf{z})V(\mathbf{z})\,\mathrm{d}\mathbf{z} \\ &= -\int G(\omega, \mathbf{x}_r - \mathbf{z})G(\omega, \mathbf{z} - \mathbf{x}_t)\omega^2 J(\omega, \mathbf{x}_t)V(\mathbf{z})\,\mathrm{d}\mathbf{z} \\ &= -\int \frac{e^{-ik(|\mathbf{x}_r - \mathbf{z}| + |\mathbf{z} - \mathbf{x}_t|)}}{16\pi^2\,|\mathbf{x}_r - \mathbf{z}|\,|\mathbf{z} - \mathbf{x}_t|}\omega^2 J(\omega, \mathbf{x}_t)V(\mathbf{z})\,\mathrm{d}\mathbf{z} \end{aligned} \tag{10}$$

If we group the product of slowly-varying terms into an amplitude variable $A(\omega, \mathbf{x}_r, \mathbf{x}_t, \mathbf{z})$, the Born-approximated field is

$$E_B^{\text{sc}}(\omega, \mathbf{x}_r, \mathbf{x}_t) = \int e^{-ik(|\mathbf{x}_r - \mathbf{z}| + |\mathbf{z} - \mathbf{x}_t|)}A(\omega, \mathbf{x}_r, \mathbf{x}_t, \mathbf{z})V(\mathbf{z})\,\mathrm{d}\mathbf{z}. \tag{11}$$

If our sensors are moving, it is more appropriate to parameterize the antenna positions with $\boldsymbol{\gamma}_r(s)$ and $\boldsymbol{\gamma}_t(s)$. Further, we note that $\boldsymbol{\rho}(\mathbf{z}, s) = |\boldsymbol{\gamma}_r(s) - \mathbf{z}| + |\mathbf{z} - \boldsymbol{\gamma}_t(s)|$ is actually the two-way range to the point $\mathbf{z}$ in the scene at time s, so that we have

$$E_B^{\text{sc}}(\omega, s) = \int e^{-ik\boldsymbol{\rho}(\mathbf{z}, s)}A(\omega, s, \mathbf{z})V(\mathbf{z})\,\mathrm{d}\mathbf{z}. \tag{12}$$

Notice that this data model also encompasses monostatic sensor configurations if we set $\boldsymbol{\gamma}_r(s) = \boldsymbol{\gamma}_t(s)$.

3.2 Imaging Operators

The data model we describe in (12) can be written as an operator on the reflectivity function,

$$\eta(\omega, s) = F[V](\omega, s). \tag{13}$$

Here we use η instead of E_B^{sc}, to note that the actual data collected is slightly different than E_B^{sc} due to matched filtering. We'll commonly refer to η as VPH. To isolate V, we want to be able to invert the operator F. One strategy is to find the adjoint operator to F, and modify that map to approximate the inverse, in the spirit of the inverse radon transform [20]. We can define the Hermitian inner products

$$\langle g, H \rangle_{\omega,s} = \int g(\omega, s) H^*(\omega, s) \, \mathrm{d}\omega \, \mathrm{d}s, \tag{14}$$

$$\langle G, h \rangle_{\mathbf{z}} = \int G(\mathbf{z}) h^*(\mathbf{z}) \, \mathrm{d}\mathbf{z}, \tag{15}$$

so that the adjoint we seek is the operator $F^\dagger$ defined by

$$\langle g, Fh \rangle_{\omega,s} = \langle F^\dagger g, h \rangle_{\mathbf{z}} \tag{16}$$

for all $g(\omega, s)$ and $h(\mathbf{z})$. Using our definition of F from (13), we want $F^\dagger$ such that

$$\int f(\omega, s) \eta^*(\omega, s) \, \mathrm{d}\omega \, \mathrm{d}s = \int (F^\dagger f)(\mathbf{x}) V^*(\mathbf{z}) \, \mathrm{d}\mathbf{z}. \tag{17}$$

It follows that

$$\int (F^\dagger f)(\mathbf{x}) V^*(\mathbf{z}) \, \mathrm{d}\mathbf{z} = \int \int f(\omega, s) e^{ik\boldsymbol{\rho}(\mathbf{z},s)} A^*(\omega, s, \mathbf{z}) \, \mathrm{d}\omega \, \mathrm{d}s V^*(\mathbf{z}) \, \mathrm{d}\mathbf{z} \tag{18}$$

so that our adjoint is

$$F^\dagger g(\mathbf{z}) = \int e^{ik\boldsymbol{\rho}(\mathbf{z},s)} A^*(\omega, s, \mathbf{z}) g(\omega, s) \, \mathrm{d}\omega \, \mathrm{d}s. \tag{19}$$

Then we can define an approximate inverse operator of the form

$$B[\eta](z) = \int e^{ik\boldsymbol{\rho}(\mathbf{z},s)} Q(\omega, s, \mathbf{z}) \eta(\omega, s) \, \mathrm{d}\omega \, \mathrm{d}s. \tag{20}$$

This operator is referred to as a filtered backprojection operator, and has a similar physical interpretation to the backprojection operator defined for the Radon transform. The exact specification of the filter Q is not important for this exposition, as we are most interested in computing appropriate estimates for $\boldsymbol{\rho}(\mathbf{z}, s)$. The formation of a quality backprojection image requires detailed knowledge of the range function, as it allows for complicated sensor motions, and it also accounts for the effects of wavefront curvature.

In some scenarios, it is appropriate to approximate the range function for image formation. When the sensor is so far from the scene of interest that wavefront curvature is not a concern, the far-field (or planar wavefront) approximation,

$$\boldsymbol{\rho}(\mathbf{z}, s) \approx |\boldsymbol{\gamma}_r(s)| + |\boldsymbol{\gamma}_t(s)| + (\hat{\boldsymbol{\gamma}}_r(s) + \hat{\boldsymbol{\gamma}}_t(s)) \cdot \mathbf{z}, \tag{21}$$

is often used. The resulting imaging operator can be applied in a series of processing steps known as the PFA [22]. This method is generally less computationally intensive than

backprojection, and the range function approximation requires less information to construct. The far-field assumption, of course, does not account for the curvature of electromagnetic wave fronts. This results in geometric distortions of the finished image, and a spatially-variant deblurring. The blurring and distortion are more severe at the edges of the image than at the center, and so PFA images are commonly constructed only for limited scene sizes [23, 24].

The imaging process may be further simplified by first motion-compensating the data as well as possible with an estimate for

$$\boldsymbol{\rho}(\mathbf{z}, s) \approx |\boldsymbol{\gamma}_r(s)| + |\boldsymbol{\gamma}_t(s)|\,, \tag{22}$$

and taking the 2-D inverse Fourier transform of the data. The resulting image,

$$\mathcal{I}(\mathbf{z}) = \mathcal{F}^{-1}[e^{ik(|\boldsymbol{\gamma}_r(s)|+|\boldsymbol{\gamma}_t(s)|)}\eta(\omega, s)], \tag{23}$$

we refer to as a range-doppler image. An improvement on this approach performs a partial reformatting of the data after motion-compensation, which eliminates the linear component of any range migration remaining in the data [25]. The transformation amounts to a frequency-dependent resampling of the VPH in the time dimension, resulting in a keystone-shaped signal support in the reformatted domain. This tranformation also alters the axes of the output image, so that we refer to images of this type as keystone-formatted or range-velocity images. Both range-doppler and range-velocity images are less expensive to form than PFA images, and can be of reasonable quality even when the rough stand-off of the sensors is the only available positional information. The downside of these images is that the increase in uncompensated sensor motion means that the blurring and distortion encountered in large PFA images is more severe in range-doppler images.

3.3 Geometric Invariants

Given a VPH, suppose there are several coordinates, $\mathbf{x}_n \in \mathbb{R}^3$, within the scene which we have tracked in range. When these landmarks do not move relative to each other, we say that the target is rigid. Assuming a rigid target, we group the N ideal point scatterers in a shape matrix,

$$X = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_N^T \end{bmatrix}. \tag{24}$$

We assume that the coordinates will be non-coplanar, so that X is full rank. Similarly, define the coordinates of our moving sensor at time t_m as $\mathbf{y}_m \in \mathbb{R}^3$, and the matrix of these coordinates as $Y \in \mathbb{R}^{M\times 3}$. Then the range from scatterer n at time t_m is

$$R_{mn} = \|\mathbf{y}_m - \mathbf{x}_n\| \tag{25}$$

Concatenating these range measurements gives the $M \times N$ range data matrix, R. Each column of R now contains the range to a single coordinate as a function of time. We sometimes refer to such a range history as a range track, and call R the track matrix. For

now, assume we are given the matrix R; we will discuss how to obtain such a matrix in Section 3.6. Our goal is to reconstruct the matrices X and Y, corresponding to the shape of our target and the motion of the sensor platform. In the following section, we will use tilde notation to distinguish the measured from noise-free versions of various quantities, i.e., $\tilde{R} = R + \text{noise}$, will represent measured ranges to scatterers while, R will represent the noise-free quantity which exactly satisfies invariant equations.

3.4 Multidimensional Unfolding

Multidimensional unfolding originated as a technique for data visualization in psychometrics [26]. Given two distinct classes and a set of distance measurements between members of each class, the goal is to create a two- or three-dimensional representation of the objects that preserves the relative distances. Our problem is a member of a subclass of multidimensional unfolding (MDU) problems, known as metric MDU. Given the matrix of measured ranges $\tilde{R}$, we define the matrix of measured squared distances as $\tilde{D} = \tilde{R} \odot \tilde{R}$, where $\odot$ is the Hadamard (or element-wise) product of matrices. If X and Y were known, then a noise-free D would be equivalent to

$$D(X,Y) = \text{diag}(YY^T)\mathbf{1}^T - 2YX^T + \mathbf{1}\text{diag}(XX^T)^T. \tag{26}$$

We use $\text{diag}(A)$ to denote the vector of entries on the diagonal of A, and $\mathbf{1}$ is a vector with every entry equal to one. The standard MDU approach to this problem is nearly identical to the approach used for multidimensional scaling (MDS). Many students encounter MDS as a solution to the problem of estimating the geographical coordinates of a set of cities, given the pairwise distances between the cities. The approach to the classical MDU / MDS problems begins by double-centering and scaling the data matrix,

$$C \coloneqq -\frac{1}{2} J_M D J_N = (J_M Y)(J_N X)^T, \tag{27}$$

where $J_p = I_p - \frac{1}{p}\mathbf{1}\mathbf{1}^T$, and I_p is the $p \times p$ identity matrix. Then, compute a matrix factorization like the SVD, $C = U\Sigma V^T$, to see that

$$\begin{aligned} X &= V\Sigma_x + \mathbf{1}\boldsymbol{x}_c^T, \\ Y &= U\Sigma_y + \mathbf{1}\boldsymbol{y}_c^T, \\ \Sigma &= \Sigma_y \Sigma_x^T. \end{aligned} \tag{28}$$

where $\mathbf{1}^T V = \mathbf{0}$ and $\mathbf{1}^T U = \mathbf{0}$. To fully specify an estimate of the coordinates X and Y, we must determine the scaling parameters in the matrices $\Sigma_x, \Sigma_y \in \mathbb{R}^{3\times 3}$, as well as the geometric centroids of the two point configurations, $\boldsymbol{x}_c, \boldsymbol{y}_c \in \mathbb{R}^3$. Without loss of generality, we choose $\boldsymbol{x}_c$ to be the origin of the coordinate system, so that $\boldsymbol{x}_c = 0$. For MDS problems, $\Sigma_y = \Sigma_x$ and $\boldsymbol{x}_c = \boldsymbol{y}_c = 0$, so determination of the scaling matrices is trivial. For MDU problems, current approaches substitute the representations in (28) back into (26) to obtain a system of equations. The solution of this system determines the scaling matrices.

We note that the matrix V corresponds to the affine invariants VV^T described by Ferrara et al. [1, 17], and that the system of equations considered in those works is a subsystem of the

one used in MDU applications. The estimation of affine invariants is relatively stable in the presence of noise, thanks to the numerical properties of the SVD. Geometric invariants are quantities that depend only on V and Σ_x. Since estimates of Σ_x are the solution of a system of (polynomial) equations, any perturbation of the input range data greatly alters the output shape estimate. This instability is especially pronounced when the shape matrix is nearly rank deficient, or when the sensor path is constrained to certain quadric surfaces [1].

3.5 Stabilized Estimation

To identify appropriate estimates for X and Y, we will describe an approach for minimizing a loss function inspired by the strain objective from MDS,

$$S(X,Y) = \frac{1}{2}\left\| D(X,Y) - \tilde{D} \right\|_F^2, \tag{29}$$

where $\|A\|_F^2$ is the squared Frobenius norm of A. The MDU process suggests a natural segmentation of the strain objective. Let

$$\begin{aligned}
\boldsymbol{\eta}_x &\coloneqq \frac{1}{N} D^T \mathbf{1} - \frac{1}{NM}(\mathbf{1}^T D \mathbf{1})\mathbf{1} \\
&= \operatorname{diag}(XX^T) - \frac{1}{N}\operatorname{tr}(XX^T)\mathbf{1} - 2X\mathbf{y}_c \\
&= J_N \operatorname{diag}(XX^T) - 2X\mathbf{y}_c, \\
\boldsymbol{\eta}_y &\coloneqq \frac{1}{M} D \mathbf{1} - \frac{1}{NM}(\mathbf{1}^T D \mathbf{1})\mathbf{1} \\
&= \operatorname{diag}(YY^T) - \frac{1}{M}\operatorname{tr}(YY^T)\mathbf{1} \\
&= J_M \operatorname{diag}(YY^T), \\
\eta_{xy} &\coloneqq \frac{1}{NM}\mathbf{1}^T D \mathbf{1} \\
&= \frac{1}{M}\operatorname{tr}(YY^T) + \frac{1}{N}\operatorname{tr}(XX^T), \\
C &\coloneqq -\frac{1}{2} J_N D J_M \\
&= J_N Y X^T,
\end{aligned} \tag{30}$$

where we use $\operatorname{tr}(A)$ to denote the trace of A and the coordinate-system choice $\mathbf{x}_c = \mathbf{0}$ to simplify the equations. Substituting the expressions in (28), we have

$$\begin{aligned}
\boldsymbol{\eta}_x &= J_N \operatorname{diag}(V\Sigma_x\Sigma_x^T V^T) - 2V\Sigma_x \mathbf{y}_c, \\
\boldsymbol{\eta}_y &= J_M \operatorname{diag}(U\Sigma_y\Sigma_y^T U^T) + 2U\Sigma_y \mathbf{y}_c, \\
\eta_{xy} &= \frac{1}{M}\operatorname{tr}(U\Sigma_y\Sigma_y^T U^T) + \|\mathbf{y}_c\|^2 + \frac{1}{N}\operatorname{tr}(V\Sigma_x\Sigma_x^T V^T), \\
C &= U\Sigma_y\Sigma_x^T V^T.
\end{aligned} \tag{31}$$

The original shape and motion are not exactly recoverable, as $D(X,Y) = D(X\mathcal{O}, Y\mathcal{O})$ for any $\mathcal{O} \in SO(3)$. In other words, any rotated or reflected version of X and Y will produce

the same set of pairwise distances. Since this is the case, we choose to fix $\mathbf{y}_c = r\mathbf{e}_1$, where $r = \|\mathbf{y}_c\|$ and $\mathbf{e}_1$ is the unit vector in the first dimension (x direction). Define the residuals

$$\begin{aligned} f(\Sigma_x, \Sigma_y) &= U\Sigma_y\Sigma_x^T V^T - C, \\ h(\Sigma_y, r) &= J_M \mathrm{diag}(U\Sigma_y\Sigma_y^T U^T) + 2rU\Sigma_y\mathbf{e}_1 - \boldsymbol{\eta}_y, \\ g(\Sigma_x, r) &= J_N \mathrm{diag}(V\Sigma_x\Sigma_x^T V^T) - 2rV\Sigma_x\mathbf{e}_1 - \boldsymbol{\eta}_x, \\ k(\Sigma_x, \Sigma_y, r) &= \frac{1}{M}\mathrm{tr}(U\Sigma_y\Sigma_y^T U^T) + \frac{1}{N}\mathrm{tr}(V\Sigma_x\Sigma_x^T V^T) + r^2 - \eta_{xy}, \end{aligned} \tag{32}$$

and let

$$\begin{aligned} F(\Sigma_x, \Sigma_y) &= \frac{1}{2}\left\|f(\Sigma_x, \Sigma_y)\right\|_F^2, \\ H(\Sigma_y, r) &= \frac{1}{2}\left\|h(\Sigma_y, r)\right\|^2, \\ G(\Sigma_x, r) &= \frac{1}{2}\left\|g(\Sigma_x, r)\right\|^2, \\ K(\Sigma_x, \Sigma_y, r) &= \frac{1}{2}\left|k(\Sigma_x, \Sigma_y, r)\right|^2. \end{aligned} \tag{33}$$

Finally, construct the new objective

$$L(\Sigma_x, \Sigma_y, r) = 2F(\Sigma_x, \Sigma_y) + \mu_1 H(\Sigma_y, r) + \mu_2 G(\Sigma_x, r) + \mu_3 K(\Sigma_x, \Sigma_y, r). \tag{34}$$

For $\mu_1 = \frac{N}{2}, \mu_2 = \frac{M}{2}$, and $\mu_3 = \frac{MN}{2}$, this objective is equivalent to the strain objective (29). To see this, decompose D and $\tilde{D}$ into orthogonal components,

$$\begin{aligned} D &= J_M \mathrm{diag}(YY^T)\mathbf{1}^T + \mathbf{1}(J_N \mathrm{diag}(XX^T) - 2X\mathbf{y}_c)^T \\ &\quad - 2J_N YX^T + \left(\frac{1}{M}\mathrm{tr}(YY^T) + \frac{1}{N}\mathrm{tr}(XX^T)\right)\mathbf{1}\mathbf{1}^T, \\ \tilde{D} &= \boldsymbol{\eta}_y\mathbf{1}^T + \mathbf{1}\boldsymbol{\eta}_x^T - 2C + \eta_{xy}\mathbf{1}\mathbf{1}^T, \end{aligned} \tag{35}$$

so that the Frobenius norm decomposes as

$$\begin{aligned} \left\|D - \tilde{D}\right\|_F^2 &= \left\|J_M \mathrm{diag}(YY^T)\mathbf{1}^T - \boldsymbol{\eta}_y)\mathbf{1}\right\|_F^2 + \left\|\mathbf{1}(J_N \mathrm{diag}(XX^T) - 2X\mathbf{y}_c - \boldsymbol{\eta}_x)^T\right\|_F^2 \\ &\quad + 4\left\|J_N YX^T - C\right\|_F^2 + \left\|\left(\frac{1}{M}\mathrm{tr}(YY^T) + \frac{1}{N}\mathrm{tr}(XX^T) - \eta_{xy}\right)\mathbf{1}\mathbf{1}^T\right\|_F^2. \end{aligned} \tag{36}$$

Substitute the relations in (28) to find

$$\begin{aligned} \left\|D - \tilde{D}\right\|_F^2 &= \left\|h(\Sigma_y, r)\mathbf{1}\right\|_F^2 + \left\|\mathbf{1}g(\Sigma_x, r)^T\right\|_F^2 + 4\left\|f(\Sigma_x, \Sigma_y)\right\|^2 + \left\|k(\Sigma_x, \Sigma_y)\mathbf{1}\mathbf{1}^T\right\|^2 \\ &= N\left\|h(\Sigma_y, r)\right\|_F^2 + M\left\|g(\Sigma_x, r)\right\|_F^2 + 4\left\|f(\Sigma_x, \Sigma_y)\right\|_F^2 + MN\left|k(\Sigma_x, \Sigma_y)\right|^2. \end{aligned} \tag{37}$$

Decomposing the objective in this manner allows us to penalize the sub-objectives individually. By manipulating these penalty terms, we describe an alternating minimization

Input: R, U, τ_H, τ_G
Output: X, Y

$\tilde{D} \leftarrow \tilde{R} \odot \tilde{R}$
Compute $C, \boldsymbol{\eta}_x, \boldsymbol{\eta}_y$, and η_{xy} according to (31).
Compute V as right singular vectors of C.
$r^{(0)} \leftarrow \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} R_{mn}$
$\Sigma_y^{(0)} \leftarrow \left[\frac{1}{2r} U^T \boldsymbol{\eta}_y \quad \mathbf{0} \quad \mathbf{0}\right]$
$\Sigma_x^{(0)} \leftarrow \left[\frac{1}{2r} V^T \boldsymbol{\eta}_x \quad \mathbf{0} \quad \mathbf{0}\right]$
$\mu_1 \leftarrow N, \mu_2 \leftarrow M, \mu_3 \leftarrow MN$
$L^{(0)} \leftarrow \infty$
$L^{(1)} \leftarrow L(\Sigma_x^{(0)}, \Sigma_y^{(0)}, r^{(0)})$
$i \rightarrow 1$
while $L^{(i-1)} > L^{(i)}$ do
    $r^{(i)}, \Sigma_x^{(i)}, \Sigma_y^{(i)} \leftarrow BFGS(r^{(i-1)}, \Sigma_x^{(i-1)}, \Sigma_y^{(i-1)})$
    $L^{(i+1)} \leftarrow L(\Sigma_x^{(i)}, \Sigma_y^{(i)}, r^{(i)})$
    if $H(\Sigma_y^{(i)}, r^{(i)}) < M\tau_H$ then
        $\mu_1 \leftarrow 0$
    else
        $\mu_1 \leftarrow N$
    end if
    if $G(\Sigma_x^{(i)}, r^{(i)}) < N\tau_G$ then
        $\mu_2 \leftarrow 0$
    else
        $\mu_2 \leftarrow M$
    end if
    $i \leftarrow i + 1$
end while
return $X \leftarrow V\Sigma_x^{(i-1)}, Y \leftarrow U\Sigma_y^{(i-1)} + r\mathbf{e}_1$

Figure 1. Algorithm 1: Stabilized Shape and Motion Estimation. U is the optional user-defined basis for the sensor motion, and the scalars τ_H, τ_G, are per-entry error thresholds for the sub-objectives H and G.

procedure in Fig. 1. Between adjustments of the penalty scalars, we compute a local minima of L using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [27].

While U and V may be estimated using the SVD of C as in the usual MDU approach, we may also manually select bases for X and Y by manually specifying V and U. In our particular case, we do not have enough information about the target shape to manually specify V. However, if we assume that $\mathbf{y}_m = y(t_m)$, where $y(t) \in \mathbb{R} \rightarrow \mathbb{R}^3$ is smooth ($y(t) \in C^\infty$), then it makes sense to choose U to reflect this property. One reasonable choice for U is the basis of Legendre polynomials. We form the orthonormal columns of U by sampling the first p Legendre polynomials at equispaced values in the interval $[-1, 1]$, where p is the desired dimension of the motion basis.

3.6 Range Tracking

We will illustrate the range tracking procedure using the large-scene Gotcha data made publicly available by the Air Force Research Laboratory (AFRL) [19]. First, we must identify features for tracking. Suppose we have a rough estimate of the sensor path. For this example, we hypothesized a straight-line path for the sensor using the aircraft's average speed, average altitude, and a rough estimate of sensor standoff. Using this sensor path, we form a defocused backprojection image with a small subaperture of the full data set, shown in Fig. 2. We then apply a median filter to the output image, and identify pixels in the original image which are brighter than their corresponding pixel in the median-filtered image by a set threshold (40 dB). The 26 peaks selected in this example are circled in Fig. 2.

For each of these selected image coordinates, we digitally spotlight that region by motion compensating the VPH to the image coordinate, and applying a smooth window function to the range-doppler image formed form the motion compensated VPH. Given our hypothesized sensor path, $\gamma(s)$, and the current image coordinate, $\mathbf{x}_n$, the hypothesized range to that tile is $\rho_n(s) = \|\gamma(s) - \mathbf{x}_n\|$. Our model for the defocused spotlight image is

$$\mathcal{I}(\mathbf{z}) = W\mathcal{F}^{-1}[e^{ik(\rho_n(s)+\epsilon(s))}\eta(\omega, s)], \tag{38}$$

where W is a smooth window function, and $\epsilon(s)$ is the residual range error remaining to the tile. The non-linear component of this phase error is responsible for the cross-range blurring in the example image tile on the left-hand side of Fig. 3.

Traditional autofocus routines are well suited to focus these small range-doppler image tiles. For this example, we have chosen to use the multiple-aperture map drift technique [28]. While other techniques like phase gradient autofocus (PGA) [29], and image-metric-based autofocus routines [30, 31] are also available, variants of the map drift technique are well suited for estimating the low-order, smooth phase errors characteristic of uncompensated sensor motion. The result of applying our implementation of the multiple-aperture map drift technique to an image tile is shown on the right-hand side of Fig. 3. The range corrections applied by an autofocus routine provide an estimate of $\epsilon(s)$, and we then update our estimate of the range to the tile center to be $\rho(s) + \epsilon(s)$. Sampling this expression at discrete slow times provides one column of the range track matrix, R.

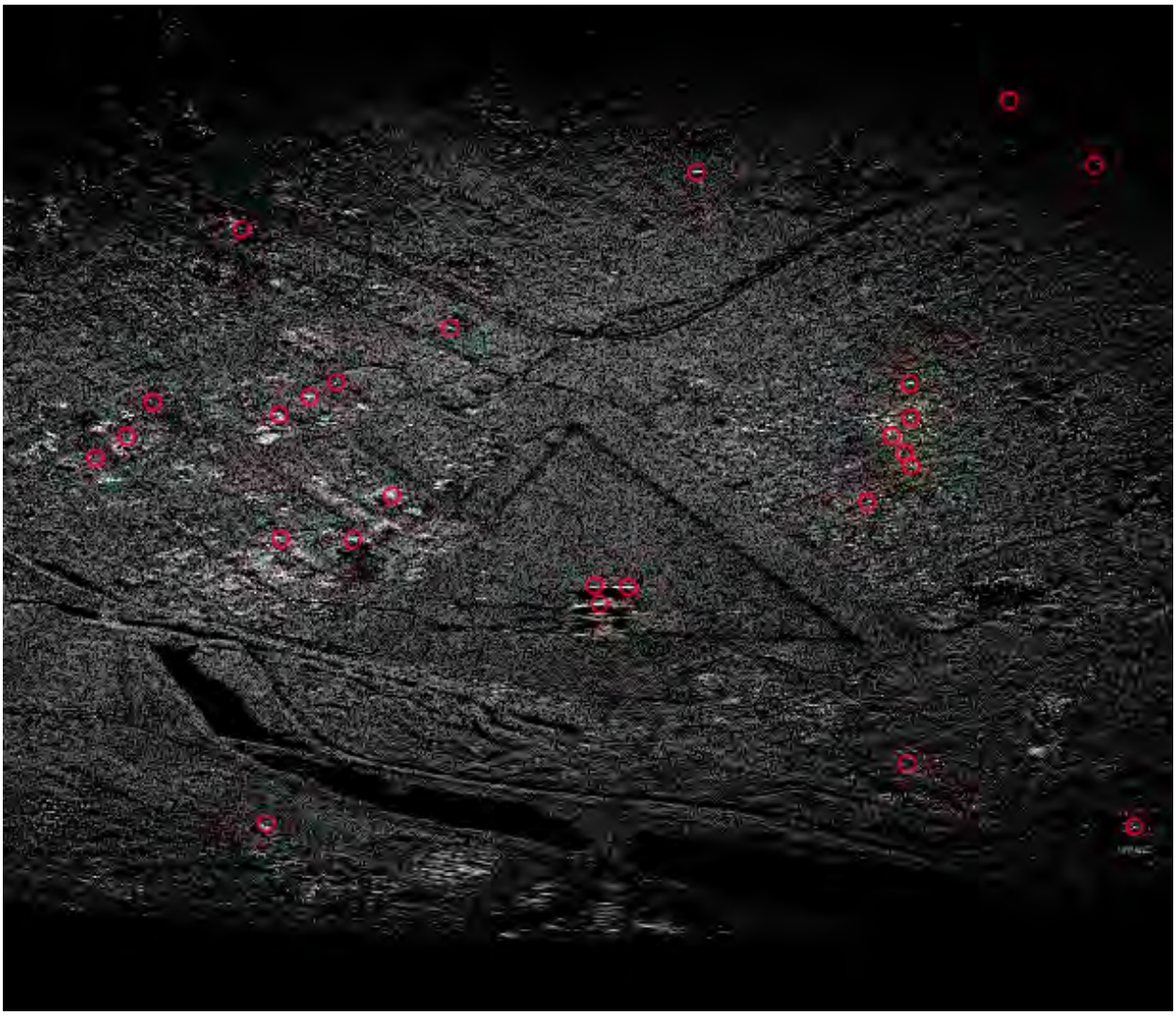

Figure 2. Ground plane backprojection image formed from a hypothesized straight-line flight path. Forming a coarsely sampled image with a small subset of the available aperture allows us to identify regions of interest for tracking.

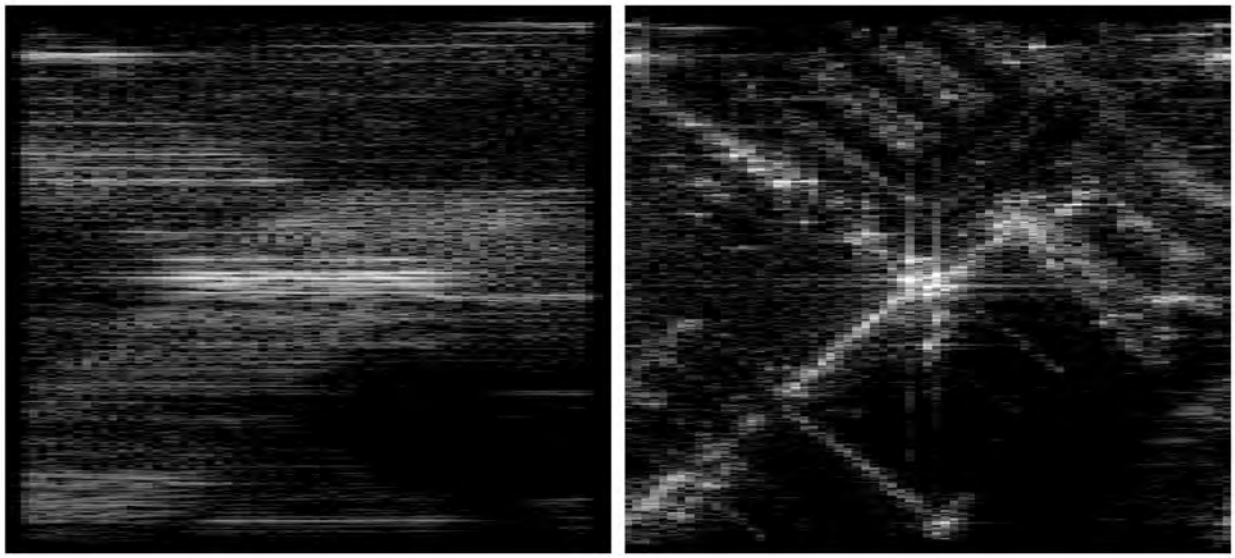

Figure 3. Left: Digital spotlight, range-doppler image of a sample image coordinate. Image is defocused due to residual range migration. Right: The same image after application of our map drift autofocus routine.

4.0 RESULTS

To demonstrate the efficacy of this technique, we will use the range tracks generated in Section 3.6 for the large-scene Gotcha dataset to estimate the sensor trajectory relative to the scene. We will then compare this to the sensor coordinates provided with the dataset, and produce a ground plane backprojection image with the estimated trajectory. For comparison, we include a range-doppler image of the scene without any additional processing, as well as a PFA image generated with an alternative geometric-invariant-based approach [1, 17].

The large-scene Gotcha dataset consists of phase history collected from a monostatic, airborne platform at X-band with 600 MHz of bandwidth. The full dataset contains measurements over approximately 3.5 degrees of aperture in azimuth, and supports an image with approximately 0.3 meter resolution. The dataset also includes the relative sensor path, measured by onboard GPS. The collection was performed with a pulsed radar system at a pulse repetition frequency (PRF) of 7502 pulses per second, with the full collection consisting of 30,000 pulses. The portion of Wright-Patterson Air Force Base illuminated by the antenna footprint is approximately 6km in diameter, and so an image of the entire scene is too large to capture with a single range-doppler image. Fig. 4 is an example range-doppler image formed from only the first sixth of the collected data. We see that even over this shorter aperture, objects near the edge of scene are severely defocused. Further, the image appears distorted when compared to a photograph of the scene. This is especially apparent in the shape of the runway in the center-right of the image (it should be straight).

Estimating the sensor path with the previous geometric invariants technique is partially successful for this 5,000-pulse example. The image in Fig. 5 is a polar format image constructed with the estimated sensor motion. Image focus is significantly improved compared to the range-doppler image, however, the image is still distorted, and the edges are still out-of-focus. Further, we note that the sensor motion estimated by this technique is not sufficiently accurate to be used for the formation of a backprojection image. Generally, the shape and motion estimates will be scaled improperly. Even when this scaling is ignored and the only affine invariants of the target are used for motion estimation, the algorithm fails when the number of pulses exceeds approximately 10,000.

After tile selection and tracking, we are left with 11 tracks to process with Algorithm 1 (see Fig. 1). The estimated target shape is then aligned with the initial tile coordinates, and the aligned sensor path estimate is compared to the validation sensor path included with the dataset. We can see in Fig. 6 that Algorithm 1 recovered the approximate tile coordinates, with some small corrections. More importantly, we see in Fig. 7 that the output motion estimate is a very close to the validation sensor path. The resulting estimate is somewhat misaligned in the altitude dimension, but even with this deviation, the result is sufficient for forming a focused backprojection image. Fig. 8 is a backprojection image formed with the entire 30,000-pulse dataset. We can see that this image is free of the geometric distortion and blurring artifacts present in the previous imagery.

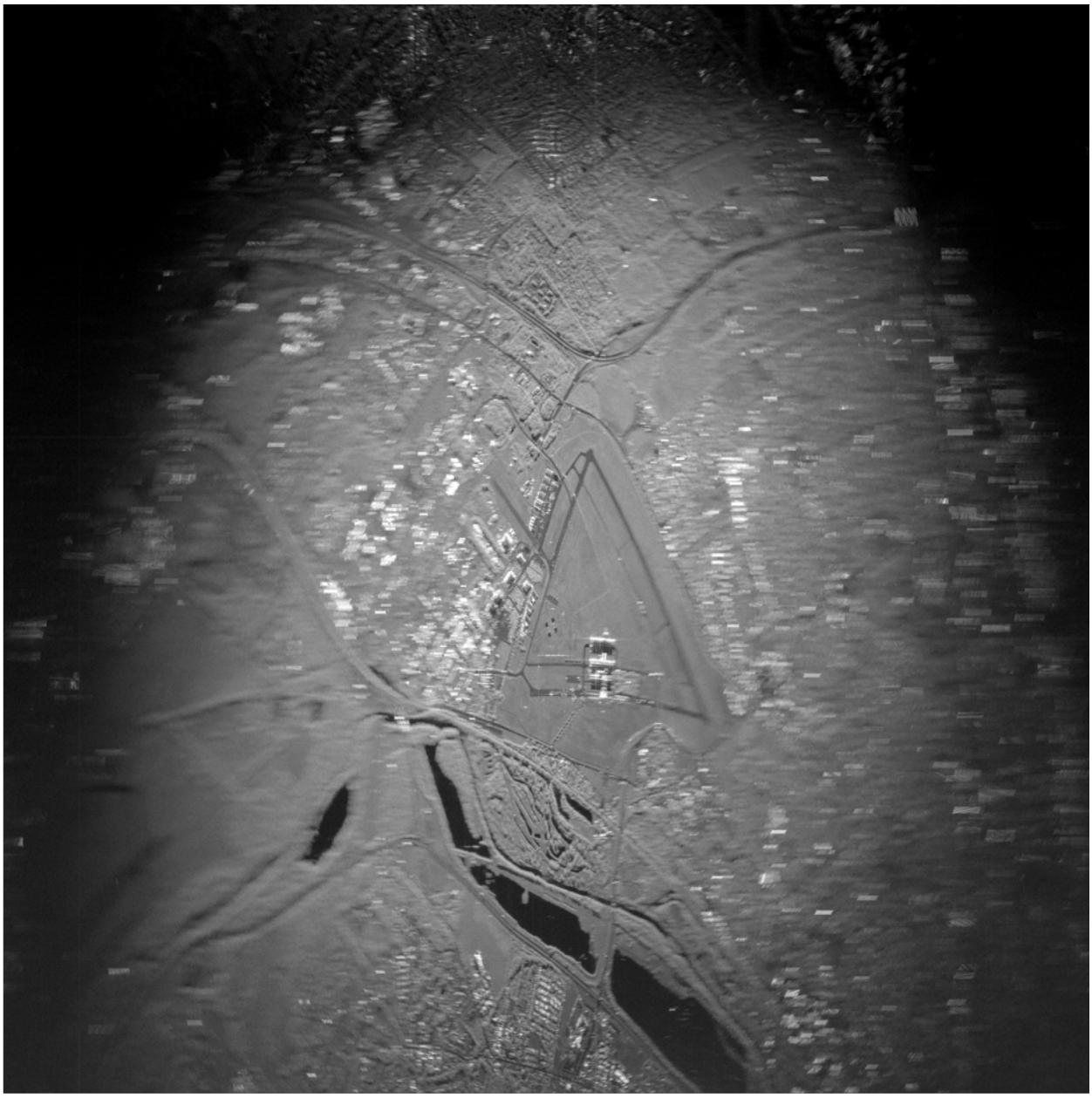

Figure 4. Range-doppler image formed from the first 5,000 pulses of the dataset. Notice the severity of the blur increases with distance from image center, and the distortion of the runway near the center-right of the image.

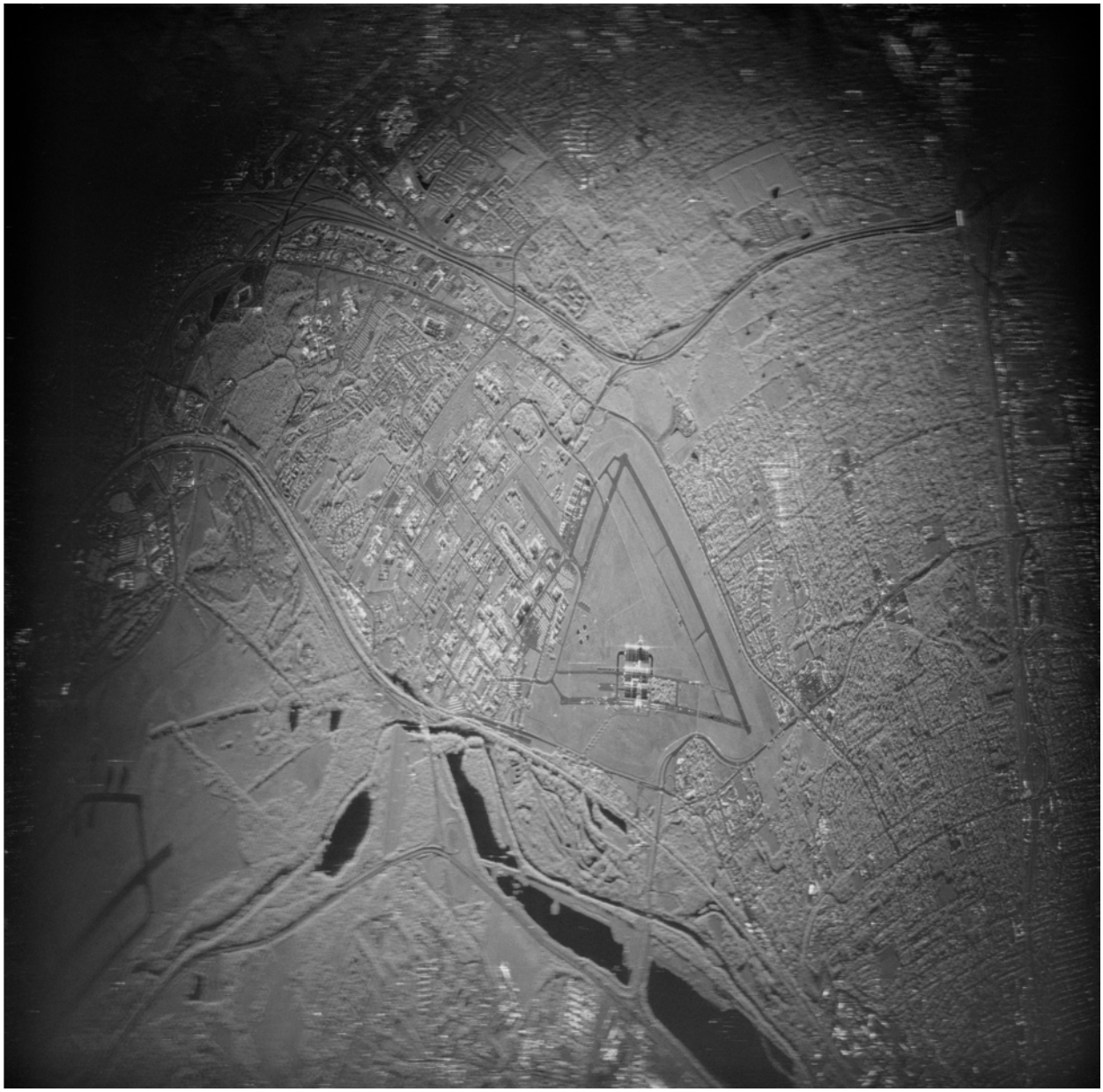

Figure 5. Image formed using the alternative geometric-invariants-based approach [1]. The estimates were accurate enough to form a PFA image from the first 5000 pulses, but were not sufficent to form a backprojection image.

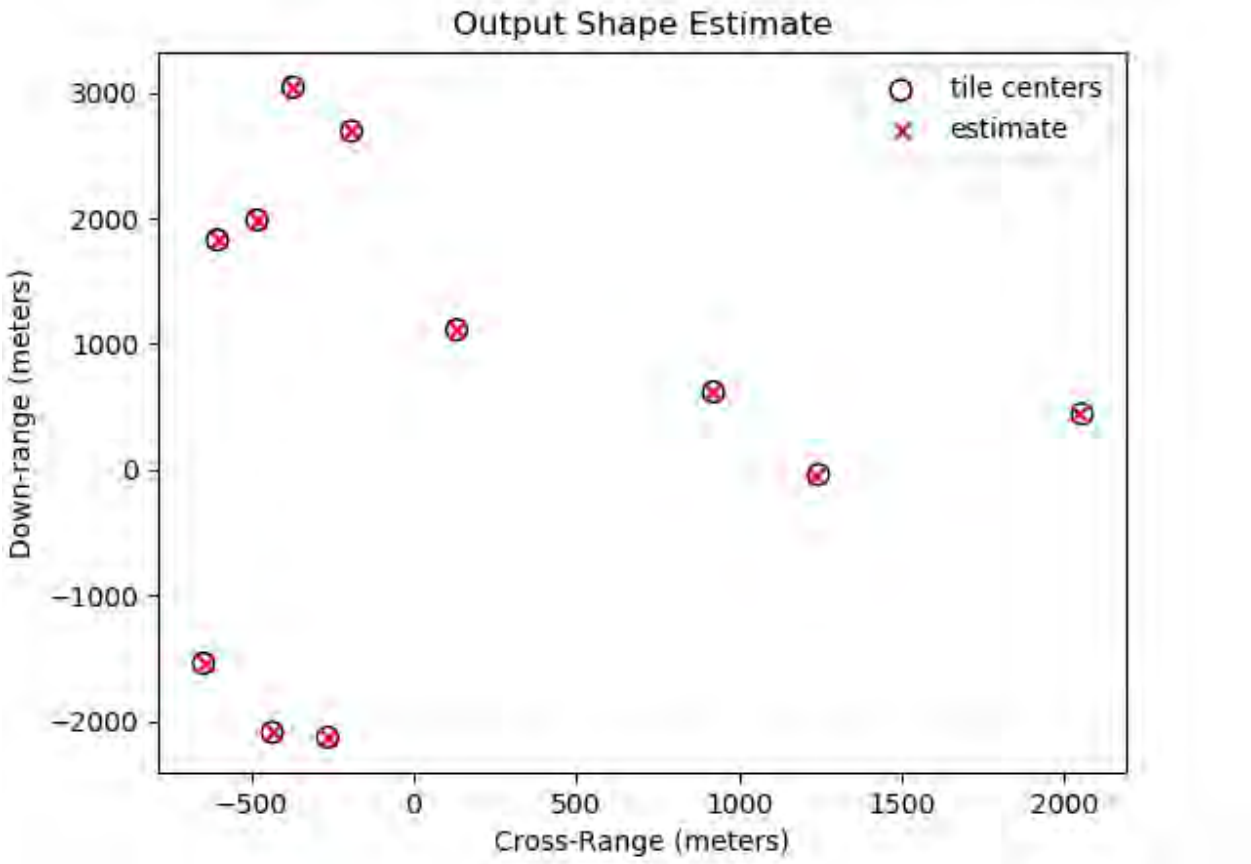

Figure 6. Shape estimate from Algorithm 1 aligned with originally selected tile center coordinates, projected into the ground plane. Algorithm 1 accurately recovers the scale of the geometric configuration of tile centers, even though the coordinates are nearly co-planar.

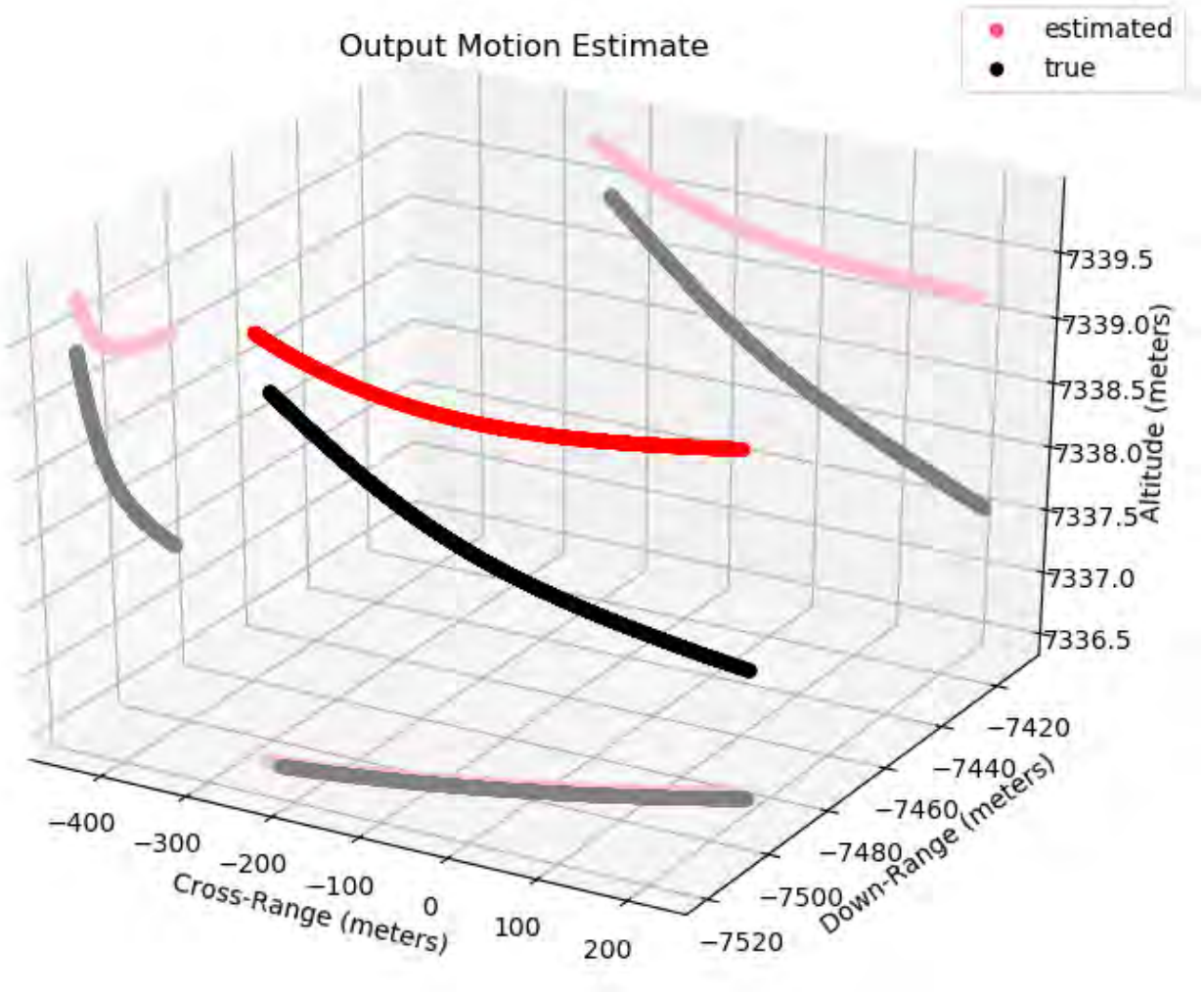

Figure 7. Motion estimate from Algorithm 1 aligned with original tile center coordinates. The estimated path very accurately matches the length and curvature of the true sensor path. The estimated path does not perfectly align with the true path, due to errors resulting from aligning the 3-D shape estimate to the 2-D tile centers.

Figure 8. Image formed using the stabilized technique and the entire 30,000-pulse dataset. The new technique produces motion estimates accurate enough for the backprojection algorithm. This image is free of the geometric distortion and edge-blurring effects in Fig. 4 and 5. This image appears darker than Fig. 5 because the dynamic range has been significantly increased due to better focus quality.

5.0 CONCLUSIONS

The improved numerical performance of Algorithm 1 allows for the generation of large-scene backprojection imagery, even for scenarios in which only a rough estimate of the sensor trajectory is available. This is a significant improvement over previous geometric-invariants-based methods, as the limited quality of those motion estimates restrict the user to the formation of PFA images over relatively short apertures.

The technique demonstrated here is modular, with room for more sophisticated autofocus techniques to be used in the range tracking step (Section 3.6). Since our technique recovers the motion of the sensor relative to the target, the method is equally applicable for inverse synthetic aperture radar (ISAR) scenarios. Such an algorithm would be useful for target recognition and target motion estimation problems. At this point, further research is required to create a robust tracker capable of providing high-quality track estimates for ISAR image formation.

Finally, Algorithm 1 presents a new approach in the search for geometric invariants for bistatic collection geometries. The extension of these results to bistatic scenarios would allow for high-quality image formation in a wide range of operating conditions.

6.0 REFERENCES

[1] M. Ferrara, G. Arnold, and M. Stuff, "Shape and motion reconstruction from 3D-to-1D orthographically projected data via object-image relations," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, pp. 1906–1912, Oct 2009.

[2] J. R. Fienup and A. Kowalczyk, "Detecting moving targets in SAR imagery by using a phase-error correction algorithm," in Proceedings of SPIE, vol. 2487, pp. 17–21, 1995.

[3] J. R. Fienup, "Detecting moving targets in SAR imagery by focusing," IEEE Transactions on Aerospace and Electronic Systems, vol. 37, pp. 794–809, July 2001.

[4] S. Barbarossa and A. Scaglione, "Autofocusing of SAR images based on the product high-order ambiguity function," IEE Proceedings - Radar, Sonar and Navigation, vol. 145, pp. 269–273, Oct 1998.

[5] J. R. Moreira and W. Keydel, "A new MTI-SAR approach using the reflectivity displacement method," IEEE Transactions on Geoscience and Remote Sensing, vol. 33, pp. 1238–1244, Sep 1995.

[6] D. W. Warner, D. C. Ghiglia, A. Fitzgerrell, and J. Beaver, "Two-dimensional phase gradient autofocus," in Image Reconstruction from Incomplete Data, vol. 4123, pp. 162–174, International Society for Optics and Photonics, 2000.

[7] H. M. J. Cantalloube and C. E. Nahum, "Multiscale local map drift driven multilateration SAR autofocus using fast polar format image synthesis," in 8th European Conference on Synthetic Aperture Radar, pp. 1–4, June 2010.

[8] L. Ran, Z. Liu, L. Zhang, T. Li, and R. Xie, "An autofocus algorithm for estimating residual trajectory deviations in synthetic aperture radar," IEEE Transactions on Geoscience and Remote Sensing, vol. PP, no. 99, pp. 1–18, 2017.

[9] S. A. Werness, M. A. Stuff, and J. R. Fienup, "Two-dimensional imaging of moving targets in SAR data," in 1990 Conference Record Twenty-Fourth Asilomar Conference on Signals, Systems and Computers, 1990., vol. 1, pp. 16–, Oct 1990.

[10] S. A. S. Werness, W. G. Carrara, L. S. Joyce, and D. B. Franczak, "Moving target imaging algorithm for SAR data," IEEE Transactions on Aerospace and Electronic Systems, vol. 26, pp. 57–67, Jan 1990.

[11] M. A. Stuff, R. C. Sullivan, Jr., B. J. Thelen, and S. A. Werness, "Automated two- and three-dimensional, fine-resolution radar imaging of rigid targets with arbitrary unknown motion," 1994.

[12] M. A. Stuff, "Three-dimensional analysis of moving target radar signals: methods and implications for ATR and feature-aided tracking," 1999.

[13] M. A. Stuff, "Three-dimensional invariants of moving targets," 2000.

[14] M. A. Stuff, Derivation and Estimation of Euclidean Invariants of Far Field Range Data. PhD thesis, University of Michigan, 2002.

[15] M. A. Stuff, P. Sanchez, and M. Biancalana, "Extraction of three-dimensional motion and geometric invariants from range dependent signals," Multidimensional Syst. Signal Process., vol. 14, pp. 161–181, Jan. 2003.

[16] M. A. Stuff, M. Biancalana, G. Arnold, and J. Garbarino, "Imaging moving objects in 3D from single aperture synthetic aperture radar," in Proceedings of the 2004 IEEE Radar Conference (IEEE Cat. No.04CH37509), pp. 94–98, April 2004.

[17] M. Ferrara and G. Arnold, "Shape and motion estimation from near-field echo-based sensor data," SIAM Journal of Imaging Sciences, vol. 2, no. 2, 2009.

[18] G. Arnold, M. Ferrara, and J. T. Parker, "Multiple-object shape and motion reconstruction with missing radar data," Proc. SPIE, vol. 8746, pp. 87460G–87460G–15, 2013.

[19] Air Force Research Laboratory (AFRL), "Large Scene Gotcha Data Example." Sensor Data Management System (SDMS).

[20] M. Cheney and B. Borden, Fundamentals of Radar Imaging. Society for Industrial and Applied Mathematics, 2009.

[21] M. Frigo and S. G. Johnson, "The design and implementation of fftw3," Proceedings of the IEEE, vol. 93, no. 2, pp. 216–231, 2005.

[22] A. W. Doerry, "Basics of polar-format algorithm for processing synthetic aperture radar images," Sandia National Laboratories report SAND2012-3369, Unlimited Release, 2012.

[23] A. W. Doerry, Wavefront curvature limitations and compensation to polar format processing for synthetic aperture radar images. United States. Department of Energy, 2006.

[24] C. Musgrove, "Polar format algorithm: survey of assumptions and approximations.," tech. rep., Sandia National Laboratories, 2012.

[25] D. Kirkland, Linear and Second Order Keystone Transforms and Their Applications. Defence Research and Development Canada, 2013.

[26] P. H. Schönemann, "On metric multidimensional unfolding," Psychometrika, vol. 35, no. 3, pp. 349–366, 1970.

[27] J. Nocedal and S. Wright, Numerical optimization. Springer Science & Business Media, 2006.

[28] W. Carrara, R. Goodman, and R. Majewski, Spotlight Synthetic Aperture Radar: Signal Processing Algorithms,(ser. Artech House remote sensing library). Norwood, MA, USA: Artech House, 1995.

[29] D. E. Wahl, P. Eichel, D. Ghiglia, and C. Jakowatz, "Phase gradient autofocus-a robust tool for high resolution sar phase correction," IEEE Transactions on Aerospace and Electronic Systems, vol. 30, no. 3, pp. 827–835, 1994.

[30] J. R. Fienup, "Synthetic-aperture radar autofocus by maximizing sharpness," Opt. Lett., vol. 25, pp. 221–223, Feb 2000.

[31] F. Berizzi, G. Corsini, M. Diani, and M. Veltroni, "Autofocus of wide azimuth angle sar images by contrast optimisation," in IGARSS'96. 1996 International Geoscience and Remote Sensing Symposium, vol. 2, pp. 1230–1232, IEEE, 1996.