Minimum Entropy Autofocus Correction of Residual Range Cell Migration

Joshua M. Kantor

Abstract—In this article we present a SAR autofocus algorithm which can correct for motion errors exceeding a range resolution cell. Most traditional autofocus algorithms operate by applying a 1D phase correction in the cross-range dimension of the image spatial frequency domain [1]–[3]. When the motion errors exceed a range resolution cell, 1D phase only compensation is not sufficient. We will describe an algorithm which can deal with such errors by using a 2D phase correction. Our algorithm is based on optimizing image focus by minimizing the image entropy.

Index Terms—Synthetic Aperture Radar, SAR, autofocus.

I. INTRODUCTION

In this article we present a SAR autofocus algorithm which can correct for motion errors exceeding a range resolution cell. Most traditional autofocus algorithms operate by applying a one-dimensional phase correction in the cross-range dimension of the image spatial frequency domain [1]–[3]. These algorithms are 1D in the sense that the phase correction is constant along the range-frequency axis and only varies along the cross-range spatial frequency dimension. When the motion errors exceed a range resolution cell, 1D phase compensation is not sufficient. Motion errors of this level may arise in a variety of situations including higher resolution SAR systems, less stable platforms such as UAVs, and inverse synthetic aperture radar (ISAR) imaging where the motion is unknown and must be estimated. We will describe an autofocus algorithm which can deal with such errors by using a 2D phase correction. Rather than an arbitrary 2D phase correction we will present an algorithm that estimates a phase and phase slope at each cross-range spatial frequency. Thus for each cross-range spatial frequency, the algorithm estimates a phase adjustment and a linear phase ramp (across range spatial frequency) which effectively applies a range shift. Our algorithm is based on minimizing the image entropy which is a commonly used measure of image focus.

There have been some previously published autofocus approaches which attempt to compensate for residual range migration, such as the two approaches described in [4]. Existing approaches generally rely on either relating the phase error estimated by conventional 1D autofocus methods to a residual range shift, or operate by cross correlating range profiles to estimate residual range migration. These approaches are quite effective in many cases, however, we have found the entropy based approach described here to be more robust in some circumstances. In particular, accurately estimating range migration by cross-correlating range profiles can be difficult when the single pulse SNR is low and the scene does not have prominent point-like scatters. On the other hand, the approaches which estimate range migration using the phase derived from conventional autofocus solutions are not always valid since in some cases, some amount of range migration may be decoupled from pulse to pulse phase error. This can happen for example when there are timing errors or phase errors introduced by the radar hardware at baseband. In such cases approaches which estimate range walk solely from the phase of a conventional 1D autofocus will fail. We also note that there have been 2D variants of PGA proposed which could be applicable [5].

II. AUTOFOCUS ALGORITHM

Now let us describe our algorithm. Let \( z \in \mathbb{C}^{MN} \) denote a complex SAR image (\( M \) rows and \( N \) columns) which can be viewed as an \( MN \) dimensional vector by stacking the columns of the image. Note that we will switch between viewing the images (and their transforms) as elements of \( \mathbb{C}^{MN} \) or as \( M \times N \) arrays depending on context. Let \( x \in \mathbb{C}^{MN} \) denote the 2D FFT of the image (interpreted as an \( M \times N \) array during the FFT and then as an \( MN \) dimensional vector afterwards). We assume \( z \) has been normalized so that \( |z_k| = 1 \). Throughout the paper \( z_k \) refers to the \( k^{th} \) entry of \( z \) viewed as an \( MN \) dimensional vector.

If we let

\[
g(z_k) = -|z_k|^2 \log(|z_k|^2)
\]

then the image entropy viewed as a function of the image can be written

\[
f(z) = \sum_k g(z_k).
\]

Let \( z_0 \) denote the initial image and let \( x_0 \) denote the transformation of that image to the spatial frequency domain via a 2D FFT. We will apply a 2D phase compensation parameterized by a vector \( \phi \in \mathbb{R}^P \) to \( x_0 \). The parameter vector \( \phi \) will be transformed to a phase vector the same size as \( x_0 \) by the linear operator

\[
A : \mathbb{R}^P \rightarrow \mathbb{R}^{MN},
\]

Thus

\[
x = \text{diag}(x_0) \exp(iA\phi), \tag{1}
\]
where \( \text{diag}(x_0) \) denotes the diagonal \( MN \times MN \) matrix with \( x_0 \) along its diagonal. The new image \( z \) can be written
\[
z = F \text{diag}(x_0) \exp (iA\phi),
\]
where \( F : \mathbb{C}^{MN} \rightarrow \mathbb{C}^{MN} \) denotes the unitary linear operator corresponding to a 2D IFFT on an \( M \times N \) image. We have written (1) in the slightly odd way with the \( \exp(i) \) on \( x_0 \) to simplify the subsequent derivative with respect to \( \phi \).

We would like to choose \( \phi \) so that we decrease the entropy of the image as low image entropy is a good measure of image focus or sharpness. To minimize the entropy of the image it will be useful to have an explicit formula for the gradient of the image entropy with respect to the parameter vector \( \phi \). Using the chain rule in the Wirtinger calculus [6], [7] we find that the gradient can be written \(^1\)
\[
\frac{\partial f}{\partial \phi_n} = \sum_k \frac{\partial g(z_k)}{\partial z_k} \frac{\partial z_k}{\partial \phi_n} + \sum_k \frac{\partial g(z_k)}{\partial z_k^*} \frac{\partial z_k^*}{\partial \phi_n},
\]
where
\[
\frac{\partial g(z_k)}{\partial z_k} = -(z_k^* \log(|z_k|^2) + z_k^*), \quad \frac{\partial g(z_k)}{\partial z_k^*} = -z_k^*.
\]
and the conjugate derivative \( \frac{\partial g(z_k)}{\partial z_k^*} \) is given by the conjugate of this expression.

We assume our image \( z_0 \) is either formed using the polar format algorithm, or backprojected to a grid in which the image axis are aligned with range and cross-range (at the aperture center). We refer to a column of \( x_0 \) (now viewed as an \( M \times N \) matrix) as a pseudo-pulse because it approximately coincides with the data acquired during a single pulse (or small number of pulses) of the synthetic aperture. See figure 14 for an illustration of the relationship between pseudo-pulses and true pulses. In the case of range-walk correction we can let \( \phi \) consist of the phase correction applied to the first sample of each pseudo-pulse as well as a phase slope of the correction for each pseudo-pulse. Thus \( \phi \) will be a \( 2N \) dimensional parameter vector with the first \( N \) components corresponding to the phase and the last \( N \) components corresponding to the phase slope. Explicitly the phase correction applied to the \( k^{\text{th}} \) sample of the \( l^{\text{th}} \) pseudo-pulse will be \( \phi_l + (k-1)\phi_{l+N} \). Since \( A \) is an \( MN \times 2N \) dimensional matrix we will write the matrix as \( A_{(mn)k} \) where \( 1 \leq m \leq M, 1 \leq n \leq N, \) and \( 1 \leq k \leq 2N \). For the range-walk correcting minimum entropy algorithm, the matrix \( A_{(mn)k} = \delta_{nk} \) for \( k \leq N \) and \( A_{(mn)k} = (m-1)\delta_{n(k-N)} \) for \( N < k \leq 2N \) and is 0 in all other entries. The \( \delta \) is the Kronecker delta.

Let
\[
\frac{\partial g}{\partial z^*}
\]
denote the vector whose \( i^{\text{th}} \) component is \( \frac{\partial g(z_i)}{\partial z_i^*} \). One can compute (essentially from (2) and (3)) that
\[
\left. \frac{\partial f}{\partial \phi} \right|_{\phi=0} = -2 \text{Imag} \left( A^T \text{diag}(x_0) \left( F^H \frac{\partial g}{\partial z^*} \right)^* \right). \quad \text{(5)}
\]

Working from right to left in (5), note that this can be computed explicitly by applying the conjugate of (4) to the image, taking a 2D FFT of the result (viewed as an \( M \times N \) array), conjugating the output of the FFT and then multiplying this pointwise by \( x_0 \). Finally, to compute the result of applying \( A^T \), we view it as operating on an \( M \times N \) array, and the result is obtained by summing over each column to get the first \( N \) components, and then summing over each column after weighting the \( n^{\text{th}} \) row by \((n-1)\) to get the second \( N \) components. Finally we take twice the negative imaginary component of the result.

As our optimization will employ a quasi-Newton update, it is also useful to have a formula for the Hessian. Observe that
\[
\frac{\partial^2 f}{\partial \phi_n \partial \phi_m} = \sum_k \frac{\partial^2 g(z_k)}{\partial z_k \partial z_k^*} \frac{\partial z_k}{\partial \phi_n} \frac{\partial z_k^*}{\partial \phi_m}, \quad \text{(6)}
\]
\[
+ \sum_k \frac{\partial^2 g(z_k)}{\partial z_k^* \partial z_k} \frac{\partial z_k^*}{\partial \phi_n} \frac{\partial z_k}{\partial \phi_m},
\]
\[
+ \sum_k \frac{\partial^2 g(z_k)}{\partial z_k \partial z_k^*} \frac{\partial z_k}{\partial \phi_n} \frac{\partial z_k^*}{\partial \phi_m},
\]
\[
+ \sum_k \frac{\partial^2 g(z_k)}{\partial z_k^* \partial z_k} \frac{\partial z_k^*}{\partial \phi_n} \frac{\partial z_k}{\partial \phi_m}.
\]
It turns out that the first two terms involving mixed derivatives of \( g \) with respect to \( z \) and \( z^* \) are the most significant so we will focus on them. Keeping these terms only we can write an expression for the (approximate) Hessian as
\[
\frac{\partial^2 f}{\partial \phi \partial \phi^T} \approx \left( A^T \text{diag}(x^*) F^H \text{diag} \left(-\log(|z_k|^2) - 2\right) A + A^T \text{diag}(x) F^T \text{diag} \left(-\log(|z_k|^2) - 2\right) F^* \text{diag}(x^*) A \right),
\]
where \( \text{diag} \left(-\log(|z_k|^2) - 2\right) \) denotes the matrix whose diagonal is the vector of size \( MN \) formed by applying \( -\log(|z|^2) - 2 \) to \( z \) component wise. Unfortunately, computing the full Hessian matrix is rather computationally expensive. Luckily this can be avoided. First, note that it is easy to numerically compute the result of applying the above expression to a vector in terms of FFTs and IFFT. To minimize the entropy using Newton’s method we need to update the parameter vector on each iteration by subtracting \( \left( \frac{\partial^2 f}{\partial \phi \partial \phi^T} \right)^{-1} \frac{\partial f}{\partial \phi} \) from the current parameter estimate. One approach to performing the optimization is to compute the parameter update by using a solver which can compute \( \left( \frac{\partial^2 f}{\partial \phi \partial \phi^T} \right)^{-1} \frac{\partial f}{\partial \phi} \) using only matrix vector products with the Hessian such as for example MATLAB’s cgs solver.

\(^1\)Throughout the paper \( \cdot^* \) denotes conjugation.
For large images we found using the full matrix to be rather slow even when it was used only indirectly in the manner just described. In figure 1 we show the Hessian matrix for one example image. Note that it has a block structure corresponding to the decomposition of the parameter vector into the phase components and phase-slope components. Note that it has a significant main diagonal and a significant diagonal in the off-diagonal block (which we will call a “cross-diagonal”). We have found that approximating the Hessian simply by the diagonal and a few off diagonals as well as the cross-diagonal (diagonal of one of the off-diagonal blocks) and a few off-cross-diagonal terms yields good results. In this case the approximate Hessian is sparse which substantially speeds up the Newton update if a sparse matrix solver is used.

To approximate the Hessian by the significant diagonals we need an easy way to compute the diagonal components. Let $\tilde{x}_{mn}$ denote the frequency domain data arranged as a matrix and let $\tilde{x}_{mn}$ denote the result of applying an inverse Fourier transform along the columns. Also let $\tilde{\tilde{x}}_{mn}$ be the inverse Fourier transform of $m\tilde{x}_{mn}$ along the columns i.e. inverse Fourier transform of the frequency domain data linearly weighted in the range frequency dimension. Then for $1 \leq l < N$ and $1 \leq l + n < N$

$$\frac{\partial^2 f}{\partial \phi_l \partial \phi_{l+n}} \approx \frac{2}{N} \Re \left( \sum_k \tilde{x}_{kl} \tilde{x}_{k,l+n}^* \right)$$

(8)

$$\sum_p \exp \left( -2\pi i \frac{(p-1)n}{N} \right) \left( -\log(|z_{kp}|^2) - 2 \right)$$

and again for $1 \leq l < N$ and $1 \leq l + n < N$,

$$\frac{\partial^2 f}{\partial \phi_{l+N+n} \partial \phi_l} \approx \frac{2}{N} \Re \left( \sum_k \tilde{x}_{kl} \tilde{x}_{k,l+n}^* \right)$$

(9)

$$\sum_p \exp \left( -2\pi i \frac{(p-1)n}{N} \right) \left( -\log(|z_{kp}|^2) - 2 \right)$$

The scale factors of $1/N$ are valid if the FFT and IFFT are normalized to be unitary (which note is not the case for e.g. the default MATLAB normalization).

III. EXAMPLE RESULTS

In figure 2 we show a region of a focused image. The region shown is a 256x256 chip of a SAR image. The actual autofocus algorithms are run on a larger 512x512 chip containing the region shown as a subchip and the images are 90% filled in the spatial frequency domain. We show the smaller subchip in the figures for visual interpretability. In figure 3, we have added range walk corresponding to a residual linear range migration across the aperture and then tried to focus the image using conventional 1D autofocus. Note how the 1D autofocus is unable to compensate for the range walk. In figure 4 we show the results of applying the 2D minimum entropy autofocus algorithm described here that also estimates the range walk. In figure 5 we show the estimated range walk together with the true added range walk; note the good agreement between truth and the estimate.

To show the power of this technique we now look at an example where the residual range walk is random from pulse to pulse. The range walk on each pulse is drawn from a Gaussian distribution with a standard deviation of 6 cm (the
Fig. 3: Image with added range walk after 1D minimum entropy autofocus. Note that 1D autofocus does not restore image focus.

Fig. 4: Image with added range walk after 2D minimum entropy autofocus that estimates range walk.

Fig. 5: Estimated vs added range walk. Note that the image chip was 512x512 but since it is 90% filled in the spatial frequency domain, there are 461 effective pseudo-pulses.

Fig. 6: Image with added noisy range walk after 1D minimum entropy autofocus. Note that while the 1D autofocus does maintain the image sharpness in this case, it cannot fully control the IPR sidelobes leading to a substantial increase in the image noise floor.

Fig. 7: Image with added noisy range walk after 2D minimum entropy autofocus that estimates range walk.

image resolution is 0.25 m). In figure 6 we show the result of applying a conventional 1D autofocus algorithm to the degraded image. Note that while this amount of range walk is less than a resolution cell, because it varies rapidly the 1D phase compensation does not fully mitigate it. Although the 1D autofocus does improve the focus of the image in this case, there is a substantial increase in the image noise floor that the 1D autofocus cannot fully eliminate. In figure 7 we show the result of applying the 2D minimum entropy algorithm which restores the image quality to that of the original undegraded image. In figure 8 we show the added range walk and the residual after applying the correction from the 2D minimum entropy algorithm. Note that the 2D minimum entropy algorithm has reduced the residual to effectively a slowly varying bias on the order of a wavelength (~ 3 cm) which has negligible impact on the image focus.
IV. PRACTICAL CONSIDERATIONS AND LIMITATIONS

The concepts of the basic range walk autofocus algorithm described in the previous section can be applied in a variety of ways. Due to the fact that multiple 2D FFTs must be performed at each iteration of the autofocus algorithm, the algorithm is much faster when applied to smaller images or image chips. We have found that a useful and efficient way to apply the algorithm in practical cases of larger images is to run it on small image subchips. From the phase and phase slope estimate produced by applying the autofocus algorithm to the subchip, one can estimate the residual range migration as a true time delay and phase correction at the pulse level which can then be applied to form a new focused image.

In order to successfully apply the correction to a small chip, it is necessary that the full blurred extend of most scatterers in the chip be confined to the chip extent. A reasonable approach we have found is to first apply conventional phase only autofocus to a coarse resolution image (coarse enough that the range walk does not exceed the resolution at which the image is formed). Note that this strategy of using coarse resolution images to get a phase only autofocus solution is similar to the approach taken in one of the approaches described in [4]. This phase only solution can be used to partially correct the full resolution image which will still be defocused due to the range walk, but the energy from most scatterers will be relatively compressed in cross-range. The range migration autofocus can then be run on a chip of the partially focused image. The solution from that chip can then be extrapolated and applied to the full image.

As an example of this, figure 9 shows a region of a clean image with a resolution of 1 ft. Figure 10 shows a case where we have added additional range migration to the raw data. Figure 11 shows the result of applying the phase only autofocus solution derived from a coarse resolution image. Note the blurring due to residual range migration. Figure 12 shows the image after applying the range migration correction estimated from running the proposed autofocus algorithm on the subchip and then extrapolating that solution back to the full image. In this case the full image was 3001x3001 pixels.
but the size of the region used for autofocus was 512x1024 pixels (note that the image chips shown are part of the image not used for autofocus). Figure 13 shows the estimated range migration correction. The estimate tracks the true range walk well except at the edges of the aperture where no estimate was made (this is simply a consequence of the fact that more pulses were collected than necessary to obtain the output resolution so the image does not contain data from the pulses at the edge of the aperture).

Note that the primary limitation of the applying the autofocus in this manner (to a subchip first and then extrapolated to the full image) is that using a chip which is small in cross-range (relative to the original image) will limit the ability to correct range migration errors that vary quickly (are high frequency). If higher frequency range migration errors must be estimated than the chip extent in cross-range must be chosen large enough to resolve those errors.

Let us emphasize that in this example the range migration was introduced as a true range shift at the phase history level. A basic fact which the polar format algorithm takes advantage of is that the radar pulses actually correspond to radial lines in the spatial frequency domain [8]. Thus even though the range migration introduced corresponds to a phase slope along these radial lines, approximating the range shift as a phase slope along the pseudo-pulses is shown to be valid in this case. The relationship of pseudo-pulses to true pulses is illustrated in figure 14. Although the pseudopulse approximation is valid in many cases we should also make clear that this approximation will break down for very high resolution images where the SAR integration angle is large. Developing an efficient way of overcoming this limitation is the subject of future work.

V. SUMMARY

We have developed an autofocus algorithm capable of focusing SAR imagery in the presence of range errors exceeding a range resolution cell. We have developed a fast quasi-Newton update scheme for the phase and phase slope correction and shown its applicability to real imagery.

REFERENCES