# MONOTONIC ITERATIVE ALGORITHM FOR MINIMUM-ENTROPY AUTOFOCUS

*Thomas J. Kragh*

Lincoln Laboratory
Massachusetts Institute of Technology
tkragh@ll.mit.edu

## ABSTRACT

Radar imaging modalities such as SAR, ISAR, and GMTI suffer image focus degradation in the presence of phase errors in the received signal due to unknown platform motion or other signal propagation delays. Data-adaptive methods such as minimum-entropy autofocus have shown promising results for estimating phase errors under conditions of low SNR, low image contrast, and substantial phase errors. However, since they involve iteratively minimizing a non-quadratic objective function, it remains an open question under what conditions these algorithms converge and if the resulting solution is optimal. In this paper we present an iterative algorithm for minimum-entropy autofocus developed using techniques collectively known as Optimization Transfer or MM-Algorithms that is guaranteed to converge. We present results using representative SAR imagery collected by the Lincoln Multi-mission ISR Testbed (LiMIT) sensor platform.

## 1. INTRODUCTION

Radar imaging modalities such as SAR or GMTI suffer image focus degradation in the presence of phase errors in the received signal due to unknown platform motion or other propagation delays. Compared to traditional autofocus algorithms such as Phase Gradient Autofocus (PGA) [1,2], image-entropy based autofocus algorithms [3–6] are known to be more robust under stressing conditions such as low SNR, low image contrast, and substantial phase errors but are typically more computationally intensive. In this paper we will present an iterative algorithms for minimum-entropy autofocus using a coordinate descent optimization scheme, where we minimize the image entropy by sequentially updating the phase parameters one at a time. We will derive this algorithms using a technique from the optimization literature collectively known as Optimization Transfer or MM-algorithms [7–9], and the resulting iterative algorithm is guaranteed to have monotone convergence properties. In addition we will also present a

fast, computationally efficient version that uses a Fast Fourier Transform (FFT) to update the phase parameters simultaneously, but does not have guaranteed monotonic convergence.

## 2. IMAGE MODEL

Let $\tilde{y}_{mk} \in \mathbb{C}$ be one of $N$ radar range profile measurements, each consisting of $M$ discrete range bins, arranged into an $M \times N$ matrix which we will refer to as the pulse history. In our notation, the row index corresponds to range bins while the column index corresponds to each range profile. Under various approximations and assumption [10], the inverse discrete Fourier transform along the column index results in a complex-valued SAR, ISAR, or GMTI image

$$y_{mn} = 1/N \sum_{k=0}^{N-1} e^{j2\pi kn/N} \tilde{y}_{mk},$$

where $y_{mn} \in \mathbb{C}$ is the value of the $(m,n)$th image pixel and the column index $n$ corresponds to either the cross-range or Doppler coordinate.

We will assume that the pulse history has been corrupted by a constant phase error on a pulse-by-pulse basis, which we will compensate for by applying a phase correction to each range profile. Let $\phi = \{\phi_1, \ldots, \phi_N\}$ be a vector of phase corrections, let

$$\tilde{z}_{mk} = e^{j\phi_k} \tilde{y}_{mk}, \; k = 1, \cdots, N \qquad (1)$$

be the phase-corrected pulse history, and let

$$z_{mn}(\phi) = 1/N \sum_{k=0}^{N-1} e^{j2\pi kn/N} e^{j\phi_k} \tilde{y}_{mk} \qquad (2)$$

be the corresponding phase-corrected image. For notational simplicity, we will implicitly assume for the remainder of this paper that the phase-corrected image is a function of the phase parameter $\phi$, i.e. $z_{mn} = z_{mn}(\phi)$. In the next section we will derive an iterative technique for determining the optimal phase parameter vector such that the phase-corrected image $z_{mn}$ is "well-focused" as measured by the image entropy.

## 2.1. Image Entropy

Let $z_{mn} \in \mathbb{C}$ be the $(m, n)$th pixel of the phase-corrected image, let $|z_{mn}|^2 = z_{mn}z_{mn}^*$ be the pixel intensity, let $E_{\boldsymbol{z}} = \sum_{m,n} |z_{mn}|^2$ be the total energy, and let $p_{mn} = |z_{mn}|^2/E_{\boldsymbol{z}}$ the normalized pixel intensity. The image entropy[1,2] is defined as

$$\Phi = -\sum_{m,n} p_{mn} \ln p_{mn}.$$

Since the phase-corrected image and intensity is a function of the phase parameters, i.e. $z_{mn} = z_{mn}(\boldsymbol{\phi})$, we will express the image entropy as a function of phase parameters $\boldsymbol{\phi}$,

$$\Phi(\boldsymbol{\phi}) = -1/E_{\boldsymbol{z}} \sum_{m,n} |z_{mn}|^2 \ln |z_{mn}|^2 + \ln E_{\boldsymbol{z}} \qquad (3)$$

where the minimum-entropy phase estimate is defined as

$$\hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} \Phi(\boldsymbol{\phi}). \qquad (4)$$

Note that the image entropy is invariant to both changes in scale (i.e. $\Phi(\alpha\boldsymbol{\phi}) = \Phi(\boldsymbol{\phi})$) as well as shifting or permutation of the image pixels [3]. Thus the image entropy is invariant to both constant and linear-phase terms and can only estimate higher-order components of the phase error function. This is not necessarily a problem since it is only the higher-order phase error components that defocus the image, while the linear phase terms merely corresponds to a simple circular-shift of the image.

Since there is no closed-form solution for (4), one needs to use an iterative numerical minimization procedure to solve for the optimal phase parameters. As an alternative to the general-purpose gradient based techniques of [3, 5] or Stage-by-Stage Approaching (SSA) technique of [4, 6], we will use Optimization Transfer based techniques to derive an iterative algorithm that has guaranteed monotone convergence.

## 3. OPTIMIZATION TRANSFER

The concept behind Optimization Transfer or MM-Algorithms [7–9] is to transform the hard problem of directly minimizing the objective function $\Phi(\boldsymbol{\phi})$ into a sequence of more easily solvable problems. This is done by constructing a sequence of surrogate functions $\Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}) \geq \Phi(\boldsymbol{\phi})$ that are an upper-bound to the objective at the $l$-th iteration, where $\boldsymbol{\phi}^{(l)}$ is the $l$-th estimate of the phase parameters. If the surrogate functions are constructed correctly, then the resulting sequence of

---

[1]Without loss of generality we could normalize $\boldsymbol{z}$ to unit energy or redefine the image entropy as $-\sum_{m,n} |z_{mn}|^2 \ln |z_{mn}|^2$. Since the total energy $E_{\boldsymbol{z}}$ is invariant to phase, both cases would result in the same phase estimate.

[2]Note that there is no information-theoretic connection between the image entropy and the Shannon entropy of $\boldsymbol{z}$. We are simply using the image entropy as a mathematically convenient measure of image focus. Other measures include the image contrast [11] as well as the Rényi entropy, of which the Shannon entropy [3,4,6] and image sharpness [5] are special cases.

minimizers $\{\boldsymbol{\phi}^{(l)}\}_{l=1,2,\dots}$ will converge to a minimizer of the original objective function, i.e.

$$\lim_{l\to\infty} \arg\min_{\boldsymbol{\phi}} \Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}) = \arg\min_{\boldsymbol{\phi}} \Phi(\boldsymbol{\phi}).$$

In order to guarantee that the objective function decreases (or at least does not increase) at each iteration, we will use the following *monotonicity condition* in constructing our surrogate function

$$\Phi(\boldsymbol{\phi}) - \Phi(\boldsymbol{\phi}^{(l)}) \leq \Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}) - \Theta(\boldsymbol{\phi}^{(l)}; \boldsymbol{\phi}^{(l)}) \leq 0 \quad \forall \boldsymbol{\phi},$$

where $\Phi(\boldsymbol{\phi}) - \Phi(\boldsymbol{\phi}^{(l)})$, the change in our objective function at the $l$-th iteration, is bounded above by zero. For a differentiable surrogate function the following three conditions are sufficient to ensure monotonicity [7].

$$
\begin{aligned}
\Theta(\boldsymbol{\phi}^{(l)}; \boldsymbol{\phi}^{(l)}) &= \Phi(\boldsymbol{\phi}^{(l)}) \\
\frac{\partial}{\partial \phi_k} \Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^{(l)}} &= \frac{\partial}{\partial \phi_k} \Phi(\boldsymbol{\phi})|_{\boldsymbol{\phi}=\boldsymbol{\phi}^{(l)}}, \; k = 1, \cdots, N \\
\Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}) &\geq \Phi(\boldsymbol{\phi}) \quad \forall \boldsymbol{\phi}
\end{aligned}
\qquad (5)
$$

The first two conditions are simply that the surrogate is tangent to the objective function at the current iterate $\boldsymbol{\phi}^{(l)}$, while the third conditions requires the surrogate to lie above the objective function for all values of the parameter vector $\boldsymbol{\phi}$, or least all values within an interval of interest. Since phase is $2\pi$-periodic, we will only require the last condition over the interval $[-\pi, \pi)$.

## 3.1. Surrogate function

We can write the image entropy (3) in the following form

$$\Phi(\boldsymbol{\phi}) = -1/E_{\boldsymbol{z}} \sum_{m,n} f(|z_{mn}|^2) + \ln E_{\boldsymbol{z}},$$

where $f(p) = -p \ln p$ is a real-valued concave function for $p > 0$. Let $g(p; q) = f(q) - (1 + \ln q)(p - q)$ be the 1st-order Taylor series of $f(p)$ about $q$. Since $f(p)$ is concave it is trivial to show that $g(p; q)$ satisfies the monotonicity conditions in (5), as illustrated in Figure 1. With this result one can show that the image entropy is upper-bounded by the following expression

$$
\begin{aligned}
\Phi(\boldsymbol{\phi}) &\leq -1/E_{\boldsymbol{z}} \sum_{m,n} g(|z_{mn}|^2; |z_{mn}^{(l)}|^2) + \ln E_{\boldsymbol{z}} \\
&= -1/E_{\boldsymbol{z}} \sum_{m,n} |z_{mn}|^2 \ln |z_{mn}^{(l)}|^2 + \ln E_{\boldsymbol{z}} \\
&= \Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}),
\end{aligned}
$$

where the terms $\boldsymbol{\phi}^{(l)}$ and $z_{mn}^{(l)}$ are the phase parameter estimate and phase-corrected image at the $l$-th iteration, respectively, and

$$\Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}) = -1/E_{\boldsymbol{z}} \sum_{m,n} |z_{mn}|^2 \ln |z_{mn}^{(l)}|^2 + \ln E_{\boldsymbol{z}} \qquad (6)$$

is a surrogate for the image-entropy objective function (3) that satisfies the monotonicity conditions of (5). Figure 2 illustrates the variation in the image entropy and surrogate function with respect to a single phase parameter for a representative SAR image.

Although the surrogate function $\Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)})$ is linear in the intensity $|z_{mn}|^2$, jointly minimizing it for the next iterate

$$\boldsymbol{\phi}^{(l+1)} = \arg\min_{\boldsymbol{\phi}} \Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}), \qquad (7)$$

is still a hard problem due to the non-linear coupling of the phase parameters within the intensity term $|z_{mn}|^2$. As an alternative to joint minimization, we will derive a coordinate descent algorithm where we minimize the surrogate function one parameter at a time while holding all others fixed, as well as show that this minimizer has a closed-form solution.

## 3.2. Coordinate Descent

A coordinate descent algorithm sequentially minimizes the objective function with respect to a single parameter while holding the remaining parameters constant. Thus by definition a coordinate decent algorithm has monotone convergence in the objective function[3]. For our application we will apply coordinate descent to minimize the surrogate function, and since the surrogate function satisfies the monotonicity conditions a decrease in the surrogate implies a decrease in the objective function at each iteration.

Let $\boldsymbol{\phi}^{(l,k)}$ be the estimate of the phase parameters at the $l$-th iteration where the first $k-1$ parameters have already been updated

$$\boldsymbol{\phi}^{(l,k)} = \{\phi_1^{(l+1)}, \ldots, \phi_{k-1}^{(l+1)}, \phi_k^{(l)}, \phi_{k+1}^{(l)}, \ldots, \phi_N^{(l)}\},$$

where we define an iteration as a complete cycle through all $N$ phase parameters. Similarly, define our free parameter $\boldsymbol{\phi}$ as

$$\boldsymbol{\phi} = \{\phi_1^{(l+1)}, \ldots, \phi_{k-1}^{(l+1)}, \phi_k, \phi_{k+1}^{(l)}, \ldots, \phi_N^{(l)}\}.$$

With this choice of free vs. fixed parameters, our surrogate reduces to a scalar function $\psi$ of the phase parameter $\phi_k$,

$$\psi(\phi_k) = \Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l,k)})$$

and the joint minimization problem in (7) reduces to the following scalar minimization problem

$$\phi_k^{(l+1)} = \arg\min_{\phi_k} \psi(\phi_k). \qquad (8)$$

In Appendix A we show that $\psi$ has the following functional form

$$\psi(\phi_k) = A_k \cos(\phi_k - \phi_k^{(l)}) + B_k \sin(\phi_k - \phi_k^{(l)}) + C_k,$$

---

[3]Monotone convergence in the objective function $\{\Phi(\boldsymbol{\phi}^{(l)})\}_{l=1,2,\ldots}$ does not necessarily imply that the iterates $\{\boldsymbol{\phi}^{(l)}\}_{l=1,2,\ldots}$ converge to a fixed-point. This requires some additional conditions which are beyond the scope of this paper.

where the constants $A_k$, $B_k$, and $C_k$ are functions of the surrogate and its derivatives evaluated at the current iterate $\boldsymbol{\phi}^{(l,k)}$. By solving for $\psi'(\phi_k) = 0$ we arrive at a closed-form solution for (8), which results in the following two update equations:

$$\phi_k^{(l+1)} = \phi_k^{(l)} + \tan^{-1}(B_k/A_k) \qquad (9)$$

and

$$z_{mn}^{(l,k+1)} = z_{mn}^{(l,k)} + \frac{1}{N}e^{j2\pi kn/N}\left(e^{j\phi_k^{(l+1)}} - e^{j\phi_k^{(l)}}\right)\tilde{y}_{mk}, \qquad (10)$$

where within each iteration the phase estimates are updated sequentially, then applied as a rank-1 update to the focused image $z_{mn}$. In order to pick the correct branch of the tangent function, we can ensure that the solution is a minimizer (and not a maximizer) by testing that the second derivative $\psi''(\phi_k) > 0$, or equivalently

$$A_k \cos(\phi_k - \phi_k^{(l)}) + B_k \sin(\phi_k - \phi_k^{(l)}) < 0.$$

Note that we could of used coordinate descent to iteratively minimize our objective function in (3) directly, however since there is no closed-form solution for the minimizer we would had to resort to a computationally expensive numerical line-search.

## 3.3. Simultaneous Update

The computational complexity of the coordinate-descent based algorithm is dominated by evaluating the gradient terms $A_k$, $B_k$ in (9) as well as the the rank-1 update in (10). Evaluating these terms for an $M \times N$ pixel image involves $\mathcal{O}(MN)$ operations per estimate per iteration, or $\mathcal{O}(MN^2)$ per iteration.

From the definition of the gradient terms $A_k$ in (A5), (A8) and $B_k$ in (A4), (A9) we note that they can be expressed in terms of a Discrete Fourier Transform. A significant reduction in algorithm complexity can be achieved by simultaneously evaluating all $N$ gradient terms with a Fast Fourier Transform algorithm and then jointly applying the phase estimates, resulting in $\mathcal{O}(MN \ln N)$ operations per iteration. However, we now lose any guarantee of strict monotonicity.

## 4. RESULTS

In Figure 3(a) we show a SAR image collected by the Lincoln Multi-mission ISR Testbed (LiMIT) sensor platform, a B707 aircraft with a nose-mounted phased-array antenna. The radar was operating at a center frequency of $f_c = 9.7$GHz with a bandwidth of 180MHz. We simulated the effect of platform motion by applying the slowly varying phase function shown in black in Figure 5, resulting in the defocused image shown in Figure 3(b).

We then applied both the Coordinate Descent and Simultaneous Update algorithms outlined in §3.2-3.3 to the defocused SAR image shown in Figure 3(b) and iterated until the

relative change in the image entropy was less then $10^{-4}$. The resulting minimum-entropy focused SAR image is shown in Figure 3(c). The Coordinate Descent (CD) based algorithm converged within 14 iterations as shown in Figure 4, taking a little over 40 seconds per iteration in Matlab on a 2.4GHz AMD Opteron CPU for a $500^2$ pixel image, for a total computation time of 9.3 minutes. In comparison, the FFT-based Simultaneous Update (SU) algorithm took a total of 21 iterations at only 0.18 seconds per iteration, for a total computation time of 3.8 seconds. For comparison purposes, we also included the performance of a eigenvector-based PGA algorithm [2], which took a total of 3.0 seconds for 5 iterations.

In terms of estimation performance, both the Coordinate Descent and Simultaneous Update based minimum-entropy algorithms achieved the ideal image entropy of 8.57 (the entropy of the original image shown in Figure 3(a)) with an overall phase estimate error of 2.4 degrees RMS. Although the PGA algorithm quickly decreased the image entropy in the first few iterations as shown in Figure 4, after 5 iterations it achieved a steady-state image entropy of 8.58 with an overall phase estimate error of 5.6 degrees RMS.

## 5. CONCLUSIONS

We have presented two monotonic iterative algorithms for minimum-entropy based autofocus of SAR imagery. The coordinate descent based algorithm has guaranteed monotonic convergence, but is decidedly slower computationally compared to the "gold-standard" of PGA. As for the simultaneous update algorithm, although it does not decrease the entropy as fast per iteration as the coordinate-descent algorithm, it can be implemented very efficiently with a FFT and is computationally competitive with PGA.

## 6. REFERENCES

[1] D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and C. V. Jakowatz, Jr., "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, Jul 1994.

[2] C. V. Jakowatz, Jr. and D. E. Wahl, "Eigenvector method for maximum-likelihood estimation of phase errors in synthetic-aperture radar imagery," *Journal of the Optical Society of America A*, vol. 10, no. 12, pp. 2539–2546, Dec 1993.

[3] A. F. Yegulalp, "Minimum entropy SAR autofocus," in *Proceedings, Adaptive Sensor Array Processing (ASAP) Workshop*. Lexington, MA: MIT Lincoln Laboratory, March 1999, pp. 25–36. [Online]. Available: http://www.ll.mit.edu/asap/

[4] L. Xi and J. Ni, "Autofocusing of ISAR images based on entropy minimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 4, pp. 1240–1252, Oct 1999.

[5] J. R. Fienup, "Synthetic-aperture radar autofocus by maximizing sharpness," *Optics Letters*, vol. 25, no. 4, pp. 221–223, February 2000.

[6] R. L. Morrison, Jr. and D. C. Munson, "An experimental study of a new entropy-based SAR autofocus technique," in *Proceedings, International Conference on Image Processing (ICIP)*, vol. 2. Rochester, NY: IEEE, Sep 2002, pp. II–441–II–444.

[7] H. Erdoğan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Transactions on Medical Imaging*, vol. 18, no. 9, pp. 801–814, Sept. 1999.

[8] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *Journal Comp. and Graph. Stat.*, vol. 9, no. 1, pp. 1–20, Mar. 2000.

[9] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.

[10] C. V. Jakowatz, Jr., D. E. Wahl, *et al.*, *Spotlight-mode Synthetic Aperture Radar: A Signal Processing Approach*. Boston: Kluwer Academic Publishers, 1996.

[11] F. Berizzi and G. Corsini, "Autofocusing of inverse synthetic aperture radar images using contrast optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 3, pp. 1185–1191, Jul 1996.

## A. DERIVATIVES OF THE SURROGATE FUNCTION

The surrogate function defined previously in (6) is given by

$$\Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)}) = -1/E_{\boldsymbol{z}} \sum_{m,n} |z_{mn}|^2 \ln |z_{mn}^{(l)}|^2 + \ln E_{\boldsymbol{z}}$$

where the terms $\boldsymbol{\phi}^{(l)}$ and $z_{mn}^{(l)}$ are the phase parameter estimate and phase-corrected image at the $l$-th iteration, respectively. Let $\psi_k = \psi(\phi_k)$ be the surrogate function evaluated over $\phi_k$ while holding all other parameters fixed, i.e.

$$\psi_k \triangleq \Theta(\boldsymbol{\phi}; \boldsymbol{\phi}^{(l)})\Big|_{\boldsymbol{\phi} = \{\phi_1^{(l)}, \dots, \phi_k, \dots, \phi_N^{(l)}\}} \quad \text{(A1)}$$

Since the surrogate function is linear in the intensity $|z_{mn}|^2$, the derivatives of $\psi_k$ will be proportional to the derivatives of the intensity,

$$\frac{\partial}{\partial \phi_k} |z_{mn}|^2 = \frac{\partial z_{mn}}{\partial \phi_k} z_{mn}^* + z_{mn} \frac{\partial z_{mn}^*}{\partial \phi_k} \quad \text{(A2)}$$

where

$$\frac{\partial z_{mn}}{\partial \phi_k} = \frac{j}{N} e^{j2\pi kn/N} \tilde{z}_{mk} \quad \text{(A3)}$$

follows immediately from (2), and similarly for the conjugate terms. By combining terms in (A2) and (A3) we arrive at the following expression for the first derivative of the surrogate function

$$\psi_k' = \frac{2}{NE_{\boldsymbol{z}}} \operatorname{Im} \left\{ \sum_m \tilde{z}_{mk} \sum_n e^{j2\pi kn/N} z_{mn}^* \ln |z_{mn}^{(l)}|^2 \right\} \quad \text{(A4)}$$

and similarly for the next two derivatives

$$\psi_k'' = \frac{2}{NE_{\boldsymbol{z}}} \operatorname{Re} \left\{ \sum_m \tilde{z}_{mk} \sum_n e^{j2\pi kn/N} z_{mn}^* \ln |z_{mn}^{(l)}|^2 \right\}$$
$$- \frac{2}{N^2 E_{\boldsymbol{z}}} \sum_m \left\{ |\tilde{y}_{mk}|^2 \sum_n \ln |z_{mn}^{(l)}|^2 \right\} \quad \text{(A5)}$$

and

$$\psi_k''' = -\frac{2}{NE_{\boldsymbol{z}}} \operatorname{Im} \left\{ \sum_m \tilde{z}_{mk} \sum_n e^{j2\pi kn/N} z_{mn}^* \ln |z_{mn}^{(l)}|^2 \right\}. \quad \text{(A6)}$$

From inspection of (A4) and (A6) we note that $\psi_k$ satisfies the linear differential equation $\psi_k' = -\psi_k'''$ and thus has a solution of the following form,

$$\psi(\phi_k) = A_k \cos(\phi_k - \phi_k^{(l)}) + B_k \sin(\phi_k - \phi_k^{(l)}) + C_k. \quad \text{(A7)}$$

The coefficients $A_k$, $B_k$, $C_k$ can be solved for by evaluating $\psi_k$ and its derivatives at the current iterate $\phi_k^{(l)}$, resulting in

$$A_k = -\psi''(\phi_k^{(l)}) \quad \text{(A8)}$$
$$B_k = \psi'(\phi_k^{(l)}) \quad \text{(A9)}$$
$$C_k = \psi(\phi_k^{(l)}) + \psi''(\phi_k^{(l)}). \quad \text{(A10)}$$
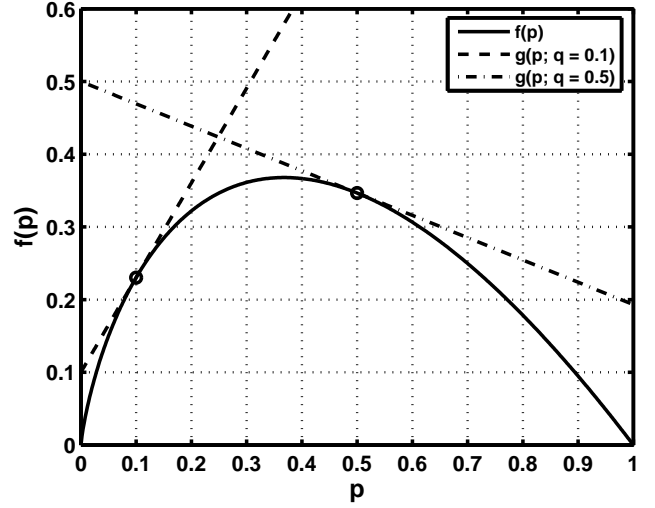


**Fig. 1**. Plot of the function $f(p) = -p \ln p$ as well as its 1st-order Taylor series $g(p; q)$ for two different values of the tangent point $q$.
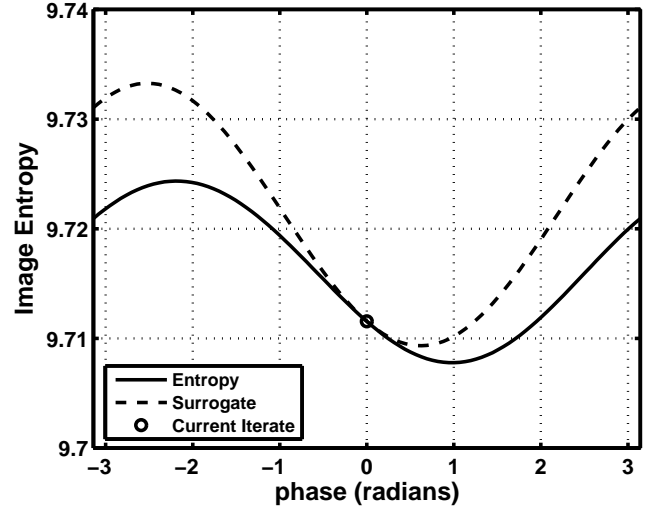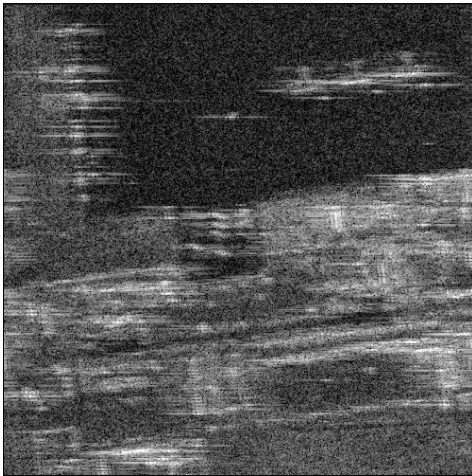


**Fig. 2**. Change in the image entropy and surrogate function with respect to a single phase parameter for a representative SAR image.

(a) Original image (entropy = 8.57)



(b) Defocused image (entropy = 9.70)



(c) Focused image (entropy = 8.57)

**Fig. 3**. Original, defocused, and focused SAR images (50dB dynamic range).
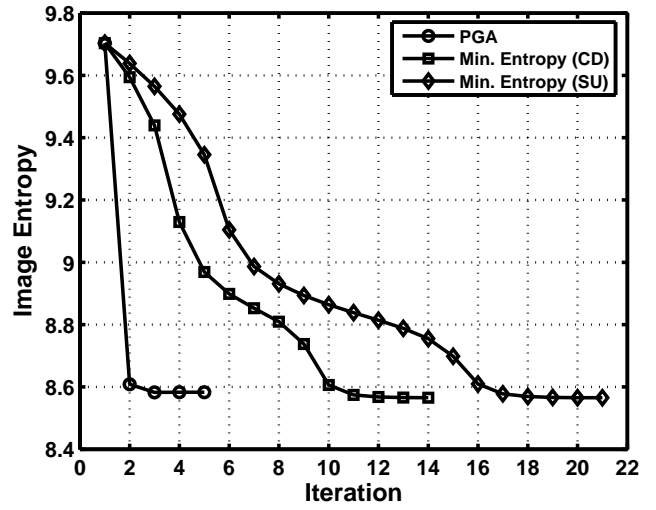


**Fig. 4**. Image entropy per iteration for both Coordinate Descent (CD) and Simultaneous Update (SU) based minimum-entropy algorithms, as compared to 5 iterations of PGA.
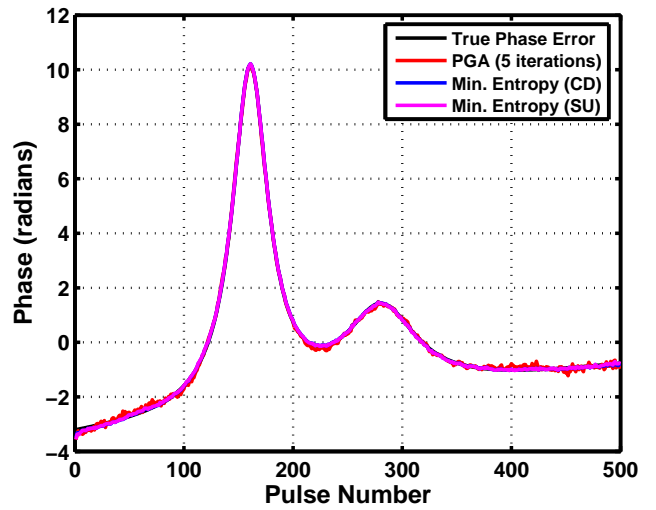


**Fig. 5**. True vs. estimated phase error for both Coordinate Descent (CD) and Simultaneous Update (SU) based minimum-entropy algorithms as well as 5 iterations of PGA.