

文章编号: 2095-4980(2015)06-0937-05

一种正则化图像特征增强中正则参数选择方法

陈 杭, 叶春茂, 张伯彦

(中国航天科工集团第二研究院 北京无线电测量研究所, 北京 100854)

摘 要: 正则化图像特征增强算法能够在保持图像较高分辨力的情况下抑制散斑和旁瓣。选择合适的正则化参数对于正则化算法所形成图像的质量至关重要。本文应用推广到非二次型正则化情形的广义交叉验证(GCV)、稳健广义交叉验证(RGCV)和 Stein 无偏风险估计(SURE)方法研究正则参数选择策略, 推导了特征增强正则化方法中 GCV, RGCV 和 SURE 函数的直接计算公式, 并提出了修正特征项后快速求取正则解的算法以及一般的不动点迭代算法, 从而实现了正则参数的自适应选择。数值仿真和实测数据处理结果均说明所提方法的有效性。

关键词: 正则参数; 广义交叉验证; 鲁棒广义交叉验证; Stein 无偏风险估计; 特征增强

中图分类号: TN957

文献标识码: A

doi: 10.11805/TKYDA201506.0937

Hyper-parameter selection in feature-enhanced regularization method

CHEN Hang, YE Chunmao, ZHANG Boyan

(Beijing Institute of Radio Measurement, China Aerospace Science and Technology Corporation, Beijing 100854, China)

Abstract : The feature-enhanced regularization-based radar image formation technique can effectively obtain high resolution image with speckle and sidelobe artifacts suppressed. Hyper-parameter selection is vital for the quality of the regularizing image. The Generalized Cross Validation(GCV), Robust Generalized Cross Validation(RGCV) and Stein's Unbias Risk Estimator(SURE) methods are applied in the non-quadratic regularization and the close form expressions of GCV, RGCV and SURE function are deduced. A fast algorithm and a generalized fix-point iteration algorithm are proposed to solve the regularization problem when the regularizing item is amended. The algorithm can be used for adaptive selection of the hyper-parameter. Numeric simulation proves the effectiveness of the proposed method.

Key words: regularization parameter; Generalized Cross Validation; Robust Generalized Cross Validation; Stein's Unbias Risk Estimator; feature enhancement

作为一种高动态成像传感器设备, 合成孔径雷达(Synthetic Aperture Radar, SAR)所提供的图像会遇到低分辨、散斑和旁瓣干扰等问题, 从而严重影响了基于图像的目标检测与识别。Çetin 等^[1]提出了一种基于特征增强的 SAR 正则化图像形成方法, 这种方法是广泛运用的实值图像恢复与重构正则化方法的推广, 它通过选取合适正则项表征图像某种特征信息并联合测量信息, 获取较高分辨、散斑和旁瓣抑制的图像。Önhon 等^[2]在特征增强正则方法的基础上提出了联合 SAR 成像和相位误差校正的正则算法, 与传统成像算法相比能够显著提高分辨力。Samadi 等^[3]给出了多特征增强 SAR 成像的稀疏表示框架, 通过不同感兴趣特征的特征字典矩阵表示不同特征, 从而将多种特征统一于同一框架下。这些正则化方法都通过正则参数来平衡数据保真项与正则项, 从而获得稳定解, 所以正则参数的选择对于图像质量有着至关重要的影响。文献[1-3]中均采取人为选择正则参数的策略而非自适应选择参数, 因此研究非 Tikhonov 正则下正则参数的选取方法很有必要。

目前对于正则化参数选择方法主要有: a) 传统 Tikhonov 正则化中 Stein 无偏风险估计(SURE)和广义交叉验证(GCV)方法推广到非二次型、非线性情形; b) L 曲线准则方法; c) 图像质量度量准则方法; d) 其他准则方法, 如乘积准则。与 GCV 和 SURE 方法相比, 其他几项准则理论依据薄弱, 是启发式方法, 应用效果有限。GCV 和 SURE 方法在 Tikhonov 正则中可以得到预计风险的封闭形式解析式, 但在非二次型正则化中, 则难以获得预计风险表示式。Ramani 等^[4]通过迭代方法计算量测对估计的雅可比矩阵, 从而计算预估风险, 但这

收稿日期: 2014-11-26; 修回日期: 2014-12-30

基金项目: 国家自然科学基金资助项目(61271417)

种方法相当复杂, 计算量非常大。Lukas^[5]指出当量测维数较小时, GCV 方法所选择的正则参数偏小, 文中提出 RGCV 方法, 在某些情况下比 GCV 方法更可靠。Batu 等^[6-7]对于点特征增强情形推导了预估风险解析式, 解决了点特征增强情形下正则参数选择问题, 但对于区域特征情形下正则参数选择没有给出对应方法。本文推导了区域特征情形下 GCV 和 SURE 的直接计算式, 这一直接计算式适用于区域特征和点特征。考虑到数值计算稳定性, 对于区域特征项作了近似修正, 并给出了特征项修正后的正则化快速求解算法和一般的不动点迭代方法。

1 问题描述

SAR^[1]和逆合成孔径雷达^[8](Inverse Synthetic Aperture Radar, ISAR)成像在内的许多估计问题, 可建模为:

$$\mathbf{g} = \mathbf{T}\mathbf{f} + \mathbf{w} \quad (1)$$

式中: $\mathbf{T}$ 为复成像投影矩阵; $\mathbf{g}$ 为量测复向量; $\mathbf{f}$ 为真实图像复矢量; $\mathbf{w}$ 为复量测噪声。

Çetin^[1]所提出的特征增强正则化方法框架如下:

$$J(\mathbf{f}) = \|\mathbf{T}\mathbf{f} - \mathbf{g}\|_2^2 + \lambda_1 \|\mathbf{P}\mathbf{f}\|_k^k + \lambda_2 \|\mathbf{D}|\mathbf{f}|\|_p^p \quad (2)$$

式中: $\|\mathbf{T}\mathbf{f} - \mathbf{g}\|_2^2$ 为数据保真项, 其中 $\mathbf{f}$ 为图像复矢量; $\lambda_1 \|\mathbf{P}\mathbf{f}\|_k^k$ 为点特征项, $\mathbf{P}$ 为加权矩阵, 一般取为单位阵; $\lambda_2 \|\mathbf{D}|\mathbf{f}|\|_p^p$ 为区域特征项, $\mathbf{D}$ 为差分算子矩阵, $|\mathbf{f}|$ 为向量 $\mathbf{f}$ 对应的模矢量; $\|\cdot\|_p$ 为向量 $\mathbf{p}$ 范数。

特征增强正则化求解就是求取使目标函数 $J(\mathbf{f})$ 取最小值的图像复矢量 $\mathbf{f}_\lambda$ 。

1.1 区域特征增强项修正的正则化问题

考虑到需要保持在零值附近的可导性和数值计算稳定性, 将正则化目标函数修正为:

$$J_{\varepsilon, \beta}(\mathbf{f}) = \|\mathbf{T}\mathbf{f} - \mathbf{g}\|_2^2 + \lambda_1 \sum_{i=1}^n [|\mathbf{f}_i|_\beta^2 + \varepsilon]^2 + \lambda_2 \sum_{i=1}^m [(|\mathbf{D}|\mathbf{f}|_\beta)_i^2 + \varepsilon]^2 \quad (3)$$

式中: $|\mathbf{f}|_\beta$ 为近似模值矢量, $\mathbf{f}$ 第 i 个元素, 记为 $\mathbf{f}_i$, 则 $|\mathbf{f}|_\beta$ 第 i 个元素为 $|\mathbf{f}_i|_\beta = \sqrt{|\mathbf{f}_i|^2 + \beta^2}$; m 为矩阵 $\mathbf{D}$ 的行数; n 为向量 $\mathbf{f}$ 的维数, 为近似参数; ε, β 为近似参数, 当 ε, β 均趋于 0 时, 式(3)和式(2)有相同的特征项。

1.2 无偏风险估计、广义交叉验证及改进方法

SURE 和 GCV 方法^[7]是对预估风险 $\|\mathbf{T}\mathbf{f}_{\text{true}} - \mathbf{T}\mathbf{f}_\lambda\|_2^2$ 产生某种不依赖于真实图像矢量 $\mathbf{f}_{\text{true}}$ 的估计量, 这 2 种估计量分别是 SURE 函数和 GCV 函数。在本文中若同时存在点特征和区域特征项则 $\lambda = (\lambda_1, \lambda_2)$ 。

定义 SURE 函数:

$$U(\lambda) = -N\sigma^2 + \|\mathbf{H}\mathbf{f}_\lambda - \mathbf{g}\|_2^2 + 2\sigma^2 \text{tr}(\text{Re}\{\mathbf{T}_\lambda\}) \quad (4)$$

式中: $\mathbf{H}$ 为复值离散 SAR 算子(Complex-Valued Discrete SAR Operator); $\mathbf{T}_\lambda = \mathbf{T}\mathbf{J}_\Pi^{-1}\mathbf{T}^H$, $\mathbf{J}_\Pi = \frac{\partial^2 J}{\partial \mathbf{f}^T \partial \mathbf{f}^*}$; N 为量测矢量 $\mathbf{g}$ 的维数; $\mathbf{f}_\lambda$ 为正则参数为 λ 时增强图像矢量; σ 为式(1)中量测噪声 $\mathbf{w}$ 分量的标准差。

定义 GCV 函数如下:

$$GCV(\lambda) = \frac{\|\mathbf{T}\mathbf{f}_\lambda - \mathbf{g}\|_2^2}{[\frac{1}{N} \text{tr}(\mathbf{I}_N - \mathbf{T}_\lambda)]^2} \quad (5)$$

式(5)中各个变量含义与式(4)中相同。

GCV 方法与 SURE 方法相比, 不需要量测噪声的先验信息, 适用范围更广。但 GCV 方法在量测维数较小时会出现正则参数过小的问题, 文献[5]中提出一种 RGCV 方法, 在很多情况下 RGCV 方法比 GCV 方法更为可靠和精准。将 RGCV 从 Tikhonov 正则化情形推广到本文所处理的特征增强正则化情形, 定义 RGCV 函数:

$$RGCV(\lambda) = [\gamma + (1 - \gamma) \text{tr}(\mathbf{T}_\lambda^H \mathbf{T}_\lambda) / N] \cdot GCV(\lambda) \quad (6)$$

式中 γ 为鲁棒参数, 当 $\gamma = 1$ 时, RGCV 方法就完全退化成 GCV 方法。

对于正则参数 λ , λ 取值很小时, 数据保真项起主要作用, 反映到正则求解中即问题的病态性没有很大改善, 所求图像质量没有提升; 而 λ 取值很大时, 特征正则项处于主导地位, 所获取的正则解与量测偏差较远, 因此正则参数的选取实质上是数据保真项和特征正则项作用的折中处理。GCV 方法以使 $GCV(\lambda)$ 取最小值的正

则化参数值为 GCV 最优参数, 同样的, RGCV, SURE 方法选择使对应评价函数取最小值的正则化参数值。

计算准则函数关键在于 Hessian 矩阵 J_{ff} , 在 Tikhonov 正则化方法中能够推导该矩阵的表达式, 因而耗费很小的运算量就可以计算准则函数, 而在非二次型正则化方法中, 只有某些简单情形能够直接计算 Hessian 矩阵, 例如 Batu 所提算法针对点特征增强情形可直接计算 J_{ff} 。本文将给出 J_{ff} 的直接计算式, 从而可以在特征增强情形中实现正则参数的快速选择。

2 算法提出

2.1 Hessian 矩阵计算公式

利用 Hjørungnes^[9]给出的复值矩阵导数理论和复变量拆分方法, 得到 J_{ff} 的计算公式:

$$J_{ff} = T^H T + \lambda_1 \text{diag}\left\{\frac{k}{2}(|f_i|^2 + \varepsilon)^{\frac{k}{2}-2} \left[\frac{k}{2}|f_i|^2 + \varepsilon\right]\right\} + \frac{\lambda_2 P}{4} G \otimes [\text{vec}\left(\frac{f_i}{|f_i|_\beta}\right) \text{vec}\left(\frac{f_i}{|f_i|_\beta}\right)^H] + \frac{\lambda_2 P}{4} \text{diag}(\eta_i \frac{|f_i|_\beta^2 + \beta^2}{|f_i|_\beta^3}) \quad (7)$$

在本文中 $\text{vec}(x_i)$ 为以 x_i 为第 i 个元素的列向量, $\text{diag}(x_i)$ 为主对角线上第 i 个元素为 x_i 的对角矩阵。符号 $\otimes$ 定义的运算为 Hardmard 乘积。式(7)中,

$$\begin{aligned} G &= D^T \text{diag}\{\tau_i\} D & \tau_i &= [(D|f|_\beta)_i^2 + \varepsilon]^{\frac{p}{2}-2} [(p-1)(D|f|_\beta)_i^2 + \varepsilon] \\ \eta &= D^T \omega & \omega &= \text{vec}((D|f|_\beta)_i [(D|f|_\beta)_i^2 + \varepsilon]^{\frac{p}{2}-1}) \end{aligned} \quad (8)$$

2.2 状态扩张法

Çetin^[11]对于 $\beta=0$ 的情形提出了一种拟牛顿算法用于求解正则解, 可以利用该算法提供存在 β 修正时正则解求取算法的初值。很多图像增强过程都是在传统方法所成图像基础上作增强处理的, 此时成像投影矩阵为单位阵。针对这种情形下面将给出一种快速求解算法, 这种算法通过扩张优化状态变量将非二次型优化问题转换为二次型优化问题, 因此称这种方法为状态扩张法。求取 $|f_\lambda|$ 快速迭代算法, 即状态拓张算法为:

Step 1: 初始化。给定 $|f^0|$ 。

Step 2: 迭代求解。对第 $n+1$ 次迭代, 有:

$$\begin{aligned} b_i &= \frac{p}{2} [(D|f^n|_\beta)_i^2 + \varepsilon]^{\frac{p}{2}-1} \quad (1 \leq i \leq M) & c_i &= \frac{k}{2} [|f_i|^2 + \varepsilon]^{\frac{k}{2}-1} \quad (1 \leq i \leq N) \\ A &= D^T \text{diag}(b_i) D & h^n &= \text{vec}\left(\frac{\beta^2}{|f_i^n|_\beta + |f_i^n|}\right) \\ |f^{n+1}| &= (I_N + \lambda_1 \text{diag}(c_i) + \lambda_2 A)^{-1} (|g| - \lambda_2 A h^n) \end{aligned} \quad (9)$$

Step 3: 终止条件。

$$\frac{\| |f^{n+1}| - |f^n| \|_2}{\| |f^n| \|_2} < \delta_{ES} \quad (10)$$

如果满足式(10), 则迭代终止, 否则返回 Step 2。

Step 4: 取 $|f_\lambda| = |f^{n+1}|$, 而 f_λ 的辐角与 g 一致。

在式(9)中采用共轭梯度方法求解, 不必做高阶矩阵求逆 $(I_N + \lambda_1 \text{diag}(c_i) + \lambda_2 A)^{-1}$ 。式(10)中 δ_{ES} 和下文式(13)中 δ_{EP} 均为预设的算法终止参数, 一般取为 10^{-5} 。

2.3 不动点迭代法

对于一般的成像投影矩阵 T , 本节提出不动点迭代法实现对式(3)的优化求解。

运用文献[9]中的复值导数理论, 得到正则解应该满足:

$$\begin{aligned} H(f) f &= v(f) \\ H(f) &= T^H T + \lambda_1 \text{diag}\left\{\frac{k}{2}(|f_i|^2 + \varepsilon)^{\frac{k}{2}-2} \left[\frac{k}{2}|f_i|^2 + \varepsilon\right]\right\} + \frac{\lambda_2 P}{2} \text{diag}\left(\frac{f_i}{|f_i|_\beta}\right) D^T \text{diag}([(D|f|_\beta)_i^2 + \varepsilon]^{\frac{p}{2}-1}) D \text{diag}\left(\frac{f_i^*}{|f_i|_\beta}\right) \\ v(f) &= T^H g - \frac{\lambda_2 P}{2} \text{diag}\left(\frac{f_i}{|f_i|_\beta}\right) D^T \text{diag}([(D|f|_\beta)_i^2 + \varepsilon]^{\frac{p}{2}-1}) D \text{diag}\left(\frac{\beta^2}{|f_i|_\beta}\right) \end{aligned} \quad (11)$$

式(11)可以导出不动点迭代格式:

$$\mathbf{H}(\mathbf{f}^n) \mathbf{f}^{n+1} = \mathbf{v}(\mathbf{f}^n) \quad (12)$$

$\mathbf{f}^n$ 为正则解求取过程中第 n 个迭代周期下的优化向量。

根据 $\mathbf{H}(\mathbf{f})$ 的定义可知, $\mathbf{H}(\mathbf{f})$ 为正定 Hermite 矩阵, 每个迭代周期可以唯一确定下一个优化向量 $\mathbf{f}^{n+1}$ 。

不动点迭代算法的终止条件为:

$$\frac{\|\mathbf{f}^{n+1} - \mathbf{f}^n\|_2}{\|\mathbf{f}^n\|_2} < \delta_{\text{FP}} \quad (13)$$

3 数值仿真

图 1 显示了第 1 个仿真的数据生成过程: 图 1(a)中给出了一个具备区域特征的复信号 $x[n]$, 将其作离散傅里叶变换得到图 1(b)和图 1(c)中的 $X[k]$, 在 $X[k]$ 上叠加图 1(d)中的复噪声得到量测信号, 见图 1(b)和图 1(c)。由于离散傅里叶变换对应的投影矩阵 $\mathbf{T}$ 是列满秩的, 因此可以利用最小二乘方法得到 $x[n]$ 的最小二乘估计 (LSE), 在图 1(a)中将其与 $x[n]$ 作对比。

运用本文提出的不动点迭代方法和准则函数计算方法, 自适应地选取正则化参数, 求取正则化解, 结果见图 2 和图 3。图 1 中主要考虑区域特征, 令 λ_1 很小或者等于 0, 重点关注 λ_2 的选取问题。图 2 给出了 3 种准则函数随 λ_2 (即图 2 中 λ) 变化情况, 并根据曲线变化确定了各个准则函数对应的最优正则参数。

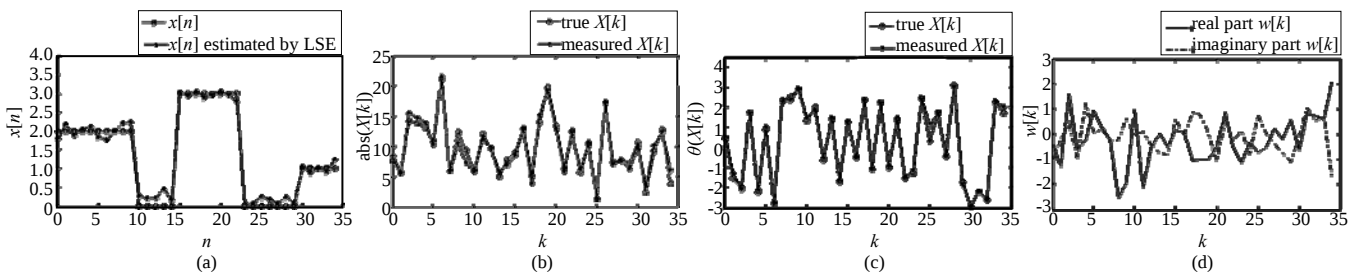


Fig.1 Formation of simulation data

图 1 仿真数据生成

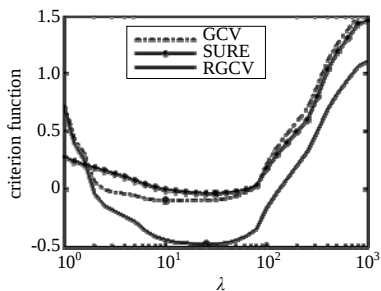


Fig.2 Criterion function curves

图 2 准则函数随正则参数变化曲线

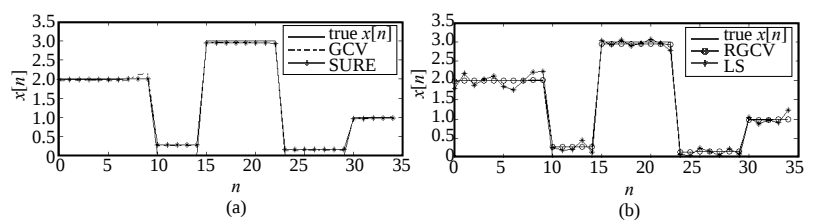


Fig.3 Simulation results of different methods

图 3 正则增强结果

图 3 是根据图 2 所选取的最优正则参数获得的正则解。根据图 3, 3 种准则下最优正则参数所对应的增强结果都较最小二乘估计结果好。GCV 最优参数明显小于 RGCV 和 SURE 正则参数。SURE 和 RGCV 下正则增强结果几乎没有差别, 而 GCV 下所增强结果要劣于另外两者, 这与 SURE 利用了噪声先验信息及 RGCV 是 GCV 的某种改良是一致的。

上述仿真情形是利用仿真生成数据, 下面处理一个 ISAR 图像真实数据, 见图 4(a)。考

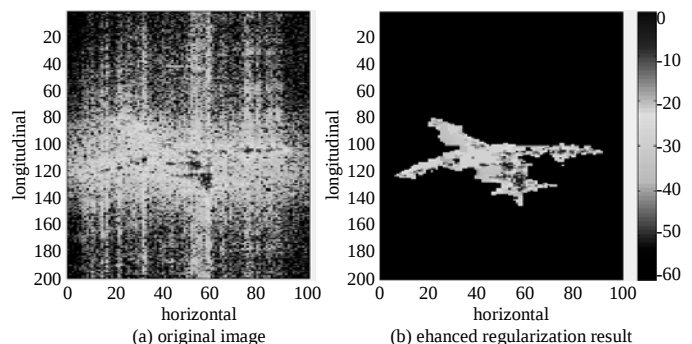


Fig.4 Enhanced regularization results of ISAR image

图 4 ISAR 图像正则增强实验

虑到计算条件, 将 ISAR 图像划分为 8 部分, 每部分都有 50×50 个像素点, 对每部分图像作正则化增强。为突出目标区域特征的保持能力, 指定点特征正则参数和区域特征正则参数的比值为 1:10, 所得特征增强图像主要体现区域特征, 故进一步作点特征增强处理, 其正则参数的选择完全按照上面所提出的 RGCV 方法, 所得结果见图 4(b)。

通过对原始图像作点和区域增强, 图 4(b)与图 4(a)相比, 目标飞机轮廓较清晰, 目标体中散射强度更为平滑, 图像散斑和旁瓣被抑制, 更容易识别, 图像质量显著提高。

4 结论

本文从特征增强正则化方法中参数选取问题出发, 应用 SURE, GCV 和 RGCV 方法到复值非 Tikhonov 正则中, 推导了特征增强正则化中 SURE, GCV 和 RGCV 函数的计算方法, 提出了一种修正正则项后正则解快速求取算法即状态拓张法, 针对一般的修正正则项后正则求解给出了不动点迭代法。数值仿真和实际数据处理均验证了本文方法的有效性。

实质上, 正则项表征图像区域的某种先验特征信息, 正则参数数值应当取决于这种特征的强弱。本文所给出的准则函数计算方法可以对某一组确定的正则参数进行评价, 因此有助于实现正则参数自适应选取。结合本文中给出的 2 种正则求解算法, 本文给出了 SAR 图像的自适应的特征增强正则化的全部处理过程, 这个处理方法对于增强实际 SAR 图像的质量具有一定作用。

参考文献:

- [1] Çetin M, Karl W C. Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization[J]. IEEE Transactions on Image Processing, 2001, 10(4):623–631.
- [2] Özben Önhon N, Müjdat Çetin. A sparsity-driven approach for joint SAR imaging and phase error correction[J]. IEEE Transactions on Imaging Processing, 2012, 21(4):2075–2088.
- [3] Sadeh Samadi, Müjdat Çetin, Mohammad Ali Masnadi-Shirazi. Multiple feature-enhanced synthetic aperture radar imaging[C]// Proceedings of SPIE 7337, Algorithms for Synthetic Aperture Radar Imagery XVI. Orlando, Florida, USA: SPIE, 2009.
- [4] LIU Zhihao, Rosen J, Nielsen J, et al. Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods[J]. IEEE Transactions on Image Processing, 2012, 21(8):3659–3672.
- [5] Mark A Lukas. Robust generalized cross-validation for choosing the regularization parameter[J]. Inverse Problems, Institute of Physics Publishing(IOP), 2006, 22(5):1883–1902.
- [6] Özge Batu, Müjdat Çetin. Hyper-parameter selection in non-quadratic regularization-based radar image formation[C]// Proceedings of SPIE 6970, Algorithms for Synthetic Aperture Radar Imagery XV. Orlando, Florida, USA: SPIE, 2008.
- [7] Özge Batu, Müjdat Çetin. Parameter selection in sparsity-driven SAR imaging[J]. IEEE Transactions on Aerospace and Electronic Systems, 2011, 47(4):3040–3050.
- [8] 冯灿, 肖亮, 韦志辉. 压缩感知 ISAR 成像的全变差优化最小算法[J]. 太赫兹科学与电子信息学报, 2013, 11(5):775–781. (FENG Can, XIAO Liang, WEI Zhihui. ISAR compressive imaging based on Majorization-Minimization of total variation[J]. Journal of Terahertz Science and Electronic Information Technology, 2013, 11(5):775–781.)
- [9] Are Hjørungnes, David Gesbert. Complex-valued matrix differentiation: techniques and key results[J]. IEEE Transactions on Signal Processing, 2007, 55(6):2740–2746.

作者简介:



陈 杭(1990–), 男, 湖北省黄冈市人, 在读硕士研究生, 主要研究方向为雷达数据处理及成像处理. email: chenhanghb@126.com.

叶春茂(1981–), 男, 湖北省咸宁市人, 博士, 高级工程师, 主要研究方向为雷达系统总体设计、雷达信号、成像处理与目标识别。

张伯彦(1960–), 女, 北京市人, 博士, 研究员, 主要研究方向为雷达目标探测与跟踪, 以及信息融合。