SHIP WAKE DETECTION FOR SAR IMAGES WITH COMPLEX BACKGROUNDS BASED ON MORPHOLOGICAL DICTIONARY LEARNING

Guozheng Yang[1, 2], Jing Yu[3], Chuangbai Xiao[3], Weidong Sun[1]

[1] State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, China
[2] Beijing Institute of Remote Sensing Information, China
[3] College of Computer Science and Technology, Beijing University of Technology, China
gzyangbj@yeah.net, jing.yu@bjut.edu.cn, cbxiao@bjut.edu.cn, wdsun@tsinghua.edu.cn

ABSTRACT

The ship wake detection of SAR images is useful not only in estimating the speed and the direction of moving ships, but also in finding small ships which are hard to be detected. The traditional ship wake detection methods of SAR images can achieve satisfactory results in simple backgrounds, but hardly work in complex backgrounds. In this paper, we propose a novel method based on the morphological component analysis and the dictionary learning to detect ship wakes in complex backgrounds. In our method, the SAR image is decomposed into a cartoon component containing ship wakes and a sea-background texture component by adaptively learning the ship wake dictionary and the sea-background texture dictionary; and then the shearlet transform is used to enhance ship wakes in the cartoon component. Experimental results show our method outperforms the traditional methods for SAR images in complex backgrounds.

Index Terms—SAR image, ship wake detection, dictionary learning, morphological component analysis (MCA), shearlet transform

1. INTRODUCTION

Sailing ships cause wakes. The ship wake detection is greatly useful not only in estimating the speed and the direction of moving ships, but in finding small ships [1]. Nowadays, the synthetic aperture radar (SAR) is widely used, but serious speckle noises and roughly imaging conditions make the ship wake detection of SAR images difficult. Related work dates from late 1980s, much of which uses the Radon transform or the Hough transform after speckle noises are reduced. Ref. [1] detects ship wakes with the Radon transform. Ref. [2] uses the normalized Hough transform and the constant false alarm rate (CFAR) method for the detection. Ref. [3] uses the normalized Radon transform for the detection.

All of these traditional methods can work well under simple background conditions, but fail under complex background conditions. To solve this problem, we propose a novel ship wake detection method based on the morphological component analysis (MCA) and the dictionary learning for complex backgrounds. In the proposed method, the MCA is used to decompose the SAR image into a cartoon component containing ship wakes and a sea-background texture component, as well as, the ship wake and the background texture dictionaries are respectively obtained by learning the decomposed components. Then, the cartoon component is transformed into the frequency domain by the shearlet transform and some high-frequency coefficients are reconstructed. Finally, the reconstructed cartoon component is binarized and ship wakes are detected with the parallel coordinate (PC) transform. Compared with those traditional methods which pursue a high detection rate by despeckling, our method can eliminate the impacts of complex backgrounds on ship wakes. Thus, the proposed method outperforms those traditional ones.

The remainder is organized as follows. Sec. 2 reviews the principles of the sparse representation and the MCA. Sec. 3 elaborates the proposed method based on the MCA and the dictionary learning. Sec. 4 compares the proposed method with two other methods. Sec. 5 draws some conclusions.

2. SPARSE REPRESENTATION AND MORPHOLOGICAL COMPONENT ANALYSIS

A signal $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$ is usually expressed as,

$$\boldsymbol{x} = \sum_{i=1}^{m} \boldsymbol{\varphi}_i \alpha_i = \boldsymbol{\Phi\alpha} \tag{1}$$

where $\boldsymbol{\alpha}$ is a representation coefficient, $\boldsymbol{\Phi}$ is a dictionary. When nonzero entries of $\boldsymbol{\alpha}$ are few, it is a sparse representation coefficient (SRC). Solving the SRC is called sparse coding and is expressed as follows,

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \ \text{ s.t. } \ \|\boldsymbol{x} - \boldsymbol{\Phi\alpha}\|_2 \le \epsilon \tag{2}$$

where $\|\cdot\|_0$ is a ℓ_0-norm, ϵ is the difference between $\boldsymbol{x}$ and its recovery. The orthogonal matching pursing (OMP) method [4] and the iterative thresholding (IT) method [5] are both widely used to solve (2). Furthermore, there are two ways to construct dictionary $\mathbf{\Phi}$: analytic dictionary and dictionary learning. The former is simple and fast, but it can only sparsely represent specific signals. Therefore, the latter is pursued and the popular K-SVD is presented [6]. But two shortages make the K-SVD impractical. On the one hand, the SVD makes the speed of the K-SVD very slow, on the other hand, samples must be fixed. Therefore, Ref. [7] proposes a block coordinate descent (BCD) dictionary learning method, recording the information of the signal $\boldsymbol{x}_i$ and its SRC $\boldsymbol{\alpha}_i$ in the matrix $\mathbf{U}$ and $\mathbf{V}$ as follows,

$$\begin{aligned} \mathbf{U}^{(t)} &= \mathbf{U}^{(t-1)} + \boldsymbol{\alpha}_i^{(t)} \left(\boldsymbol{\alpha}_i^{(t)}\right)^T \\ \mathbf{V}^{(t)} &= \mathbf{V}^{(t-1)} + \boldsymbol{x}_i^{(t)} \left(\boldsymbol{\alpha}_i^{(t)}\right)^T \end{aligned} \tag{3}$$

Then, dictionary $\mathbf{\Phi}$ can be updated column-by-column,

$$\begin{aligned} \boldsymbol{\psi} &= \frac{1}{\mathbf{U}(i,i)} \left(\boldsymbol{v}_i - \mathbf{\Phi}\boldsymbol{u}_i\right) + \boldsymbol{\phi}_i \\ \boldsymbol{\phi}_i &= \frac{1}{\max\left(\|\boldsymbol{\psi}\|_2, 1\right)} \boldsymbol{\psi} \end{aligned} \tag{4}$$

where $\boldsymbol{u}_i$ is the ith column of $\mathbf{U}$, also called an atom, $\boldsymbol{v}_i$ is the ith column of $\mathbf{V}$, $\boldsymbol{\phi}_i$ is the ith column of $\mathbf{\Phi}$.

In MCA, $\boldsymbol{x}$ can be seen as a mixture of K components,

$$\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \ldots + \boldsymbol{x}_s + \ldots + \boldsymbol{x}_K \tag{5}$$

If each component is sparse only in its corresponding dictionary, $\boldsymbol{x}$ will be uniquely decomposed, and this is the principle of the MCA. Let $\boldsymbol{x}_s$ represent the outline of the signal $\boldsymbol{x}$, and to make $\boldsymbol{x}_s$ smoother, a regularization term $\|\boldsymbol{x}_s\|_{\mathrm{TV}}$ may be introduced. Thus, the MCA can be seen as a ℓ_0-norm minimization (L0NM) problem as follows [5],

$$\min \left(\sum_{i=1}^{K} \|\boldsymbol{\alpha}_i\|_0 + \gamma \|\boldsymbol{x}_s\|_{\mathrm{TV}} \right) \text{s.t.} \left\| \boldsymbol{x} - \sum_{i=1}^{K} \boldsymbol{x}_i \right\|_2 \leq \epsilon \tag{6}$$

where $\boldsymbol{\alpha}_i$ is the SRC of $\boldsymbol{x}_i$, γ is a regularization coefficient, and ϵ is the residual. Ref. [5] introduces the IT method and analytic dictionaries to solve (6). But Ref. [8] observes that when SAR images are decomposed by the MCA and the complex textures exists, the results might be unsatisfactory. Therefore, an adaptive MCA (AMCA) method is proposed [8]. For an image with a texture component $\boldsymbol{x}_1$ and any other component $\boldsymbol{x}_2$, the AMCA is expressed as follows,

$$\min \left(\frac{1}{2} \|\boldsymbol{x} - \boldsymbol{x}_1 - \boldsymbol{x}_2\|_2 + \mu \sum_{k=1}^{K} \|\boldsymbol{\alpha}_{1,k}\|_1 + \mu \|\boldsymbol{\alpha}_2\|_1 \right) \tag{7}$$

which is a ℓ_1-norm minimization problem, and is equivalent to the L0NM problem, and $\boldsymbol{\alpha}_{1,k}$ is the kth block-SRC of $\boldsymbol{x}_1$, μ is a regularization coefficient. Three steps are used to solve (7): sparse coding, components updating, and texture dictionary learning from the decomposed $\boldsymbol{x}_1$.

3. THE PROPOSED METHOD

Our method introduces the MCA to separate the cartoon component. After the cartoon component is enhanced, ship wakes in it are detected based on the PC transform.

3.1. Cartoon component separation based on MCA

The SAR image $\mathbf{X}$ can be regarded as a superposition of a cartoon component $\mathbf{S}$ containing ship wakes, a sea background texture component $\mathbf{T}$, and a residual $\mathbf{R}$ with speckle noises. That is $\mathbf{X} = \mathbf{S} + \mathbf{T} + \mathbf{R}$.

According to (6) and (7), the separation of $\mathbf{S}$ and $\mathbf{T}$ is expressed as follows,

$$\begin{gathered} \min_{\mathbf{S},\mathbf{T}} \left(\sum_{i=1}^{M} \|\boldsymbol{\alpha}_{\mathbf{S},i}\|_1 + \sum_{j=1}^{N} \|\boldsymbol{\alpha}_{\mathbf{T},j}\|_1 + \gamma \|\mathbf{S}\|_{\mathrm{TV}} \right) \\ \text{s.t. } \|\mathbf{X} - \mathbf{S} - \mathbf{T}\|_2 \leq \epsilon \end{gathered} \tag{8}$$

The solution of (8) is outlined in **Alg. 1**.

Algorithm 1- Image decomposition and dictionaries learning algorithm based on the MCA

Input: SAR image $\mathbf{X}$, background texture dictionary $\mathbf{\Phi_T}$ and ship wake dictionary $\mathbf{\Phi_S}$, the number of dictionary updating iterations N_{learn}, the number of MCA maximum allowed iterations N_{iter}, regularization coef. γ, stopping condition ϵ, ratio coef. η.

Initialize: Normalize $\mathbf{X}$: $\mathbf{X} = \text{mat2gray}(\mathbf{X})$, set $\mathbf{S}^{(0)} = \mathbf{0}$, $\mathbf{T}^{(0)} = \mathbf{0}$, $\mathbf{R}^{(0)} = \mathbf{X}$, and obtain $\lambda^{(0)}: \lambda^{(0)} = \eta \min(\lambda_1, \lambda_2)$. where, $\lambda_1 = \max \left\| \mathbf{\Phi_S^+} \mathbf{R_S^{(0)}} \right\|_\infty$, $\lambda_2 = \max \left\| \mathbf{\Phi_T^T} \mathbf{R_T^{(0)}} \right\|_\infty$, $\mathbf{\Phi_S^+}$ is the pseudo-inverse of $\mathbf{\Phi_S}$, $\mathbf{\Phi_T^T}$ is the transpose of $\mathbf{\Phi_T}$; matrix $\mathbf{R_S^{(0)}}$ comes from cutting $\mathbf{R}^{(0)}$ by the atom size of $\mathbf{\Phi_S}$ and matrix $\mathbf{R_T^{(0)}}$ comes from cutting $\mathbf{R}^{(0)}$ by the atom size of $\mathbf{\Phi_T}$.

for $j = 1$ **to** N_{learn} **do**

  for $i = 0$ **to** $(N_{\text{iter}} - 1)$ **do**

  - Obtain $\mathbf{S_R^{(i)}} = \mathbf{S}^{(i)} + \mathbf{R}^{(i)}$, cut it into grids to form column matrix $\mathbf{S_R^{B(i)}}$ by the atom size of $\mathbf{\Phi_S}$, threshold columns of matrix $\mathbf{\Phi_S^+ S_R^{B(i)}}$ by $\lambda^{(i)}$ to $\mathbf{A_S^{(i)}}$; Compute $\mathbf{S_R^{B(i+1)}} = \mathbf{\Phi_S A_S^{(i)}}$, rearrange to $\mathbf{S}^{(i+1)}$ with compensation $\left\|\mathbf{S}^{(i+1)}\right\|_{\mathrm{TV}}$.
  - Obtain $\mathbf{T_R^{(i)}} = \mathbf{T}^{(i)} + \mathbf{R}^{(i)}$, cut it into grids to form column matrix $\mathbf{T_R^{B(i)}}$ by the atom size of $\mathbf{\Phi_T}$, threshold columns of matrix $\mathbf{\Phi_T^T T_R^{B(i)}}$ by $\lambda^{(i)}$ to $\mathbf{A_T^{(i)}}$; Compute $\mathbf{T_R^{B(i+1)}} = \mathbf{\Phi_T A_T^{(i)}}$, rearrange to $\mathbf{T}^{(i+1)}$.
  - Obtain $\lambda^{(i+1)}$ similar to obtain $\lambda^{(0)}$, then stop iterations when $\lambda^{(i+1)} \leq \epsilon$.

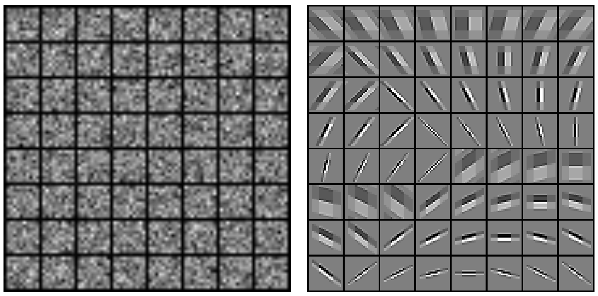
  end for

  - Update dictionaries $\mathbf{\Phi_T}$ and $\mathbf{\Phi_S}$, using the BCD-based dictionary learning method.

end for

Output: cartoon component $\mathbf{S}$ and background texture component $\mathbf{T}$, final dictionary $\mathbf{\Phi_S}$ and $\mathbf{\Phi_T}$.

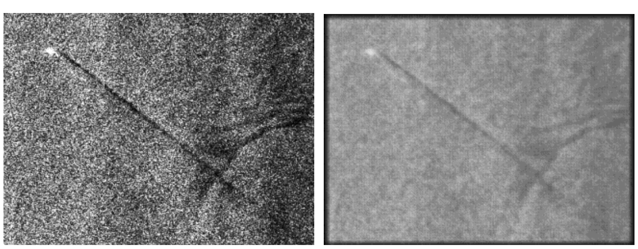
The dictionary $\mathbf{\Phi_T}$ is initially random, the atom size of which is 10×10 pixels and the total number of atoms is 64. The dictionary $\mathbf{\Phi_S}$ is initially from a 4-layer shearlet transform with 18 decomposition directions in each layer, the atom size of which is 20×20 pixels and the total number of atoms is also 64. Figs. 1(a) and 1(b) show the initial background texture dictionary $\mathbf{\Phi_T}$ and the initial ship wake dictionary $\mathbf{\Phi_S}$, respectively.



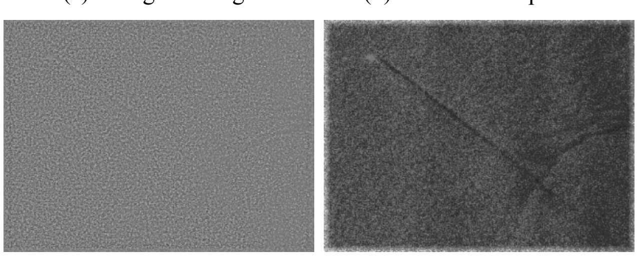
(a) Initial $\mathbf{\Phi_T}$ (b) Initial $\mathbf{\Phi_S}$

Fig. 1 Initial dictionaries

When the proposed method is applied to Fig. 2(a) and the related parameters in **Alg. 1** are set to be $N_{\text{learn}} = 40$, $N_{\text{iter}} = 10$, $\gamma = 0.01$, $\tau = 0.04$, and $\eta = 0.01$, the component decomposition results will be shown as Figs. 2(b)-2(d), and the final background texture dictionary $\mathbf{\Phi_T}$ and the final ship wake dictionary $\mathbf{\Phi_S}$ are shown as Fig. 3(a) and 3(b). From these results we can see that, 1) the cartoon component containing ship wakes is effectively separated from the sea-background texture component, and the speckles are also decomposed into the residual, and 2) the final sea background texture dictionary and the final ship wake dictionary are more suitable to represent the real sea backgrounds and the ship wakes.

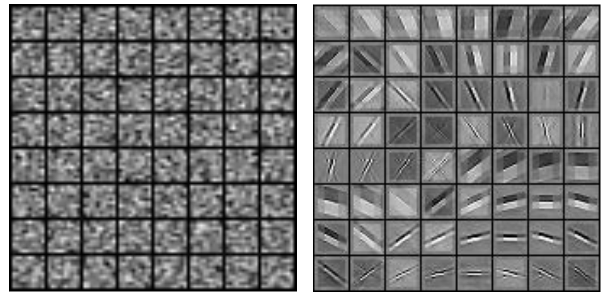


(a) Original image (b) Cartoon component



(c) Background texture component (d) The residual

Fig. 2 MCA results of a SAR image



(a) Final $\mathbf{\Phi_T}$ (b) Final $\mathbf{\Phi_S}$

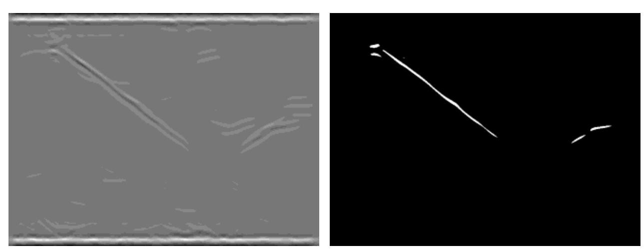
Fig. 3 Final dictionaries of the proposed method

3.2. Cartoon component enhancement

Since ship wakes in the cartoon component exhibit the directional high-frequency characteristics, and shearlet basis functions can effectively respond to them, the proposed method enhances ship wakes by extracting some shearlet high-frequency coefficients of the cartoon component, and reconstructing these coefficients.

The basis function of the shearlet transform is shown as $\psi_{j,k,\boldsymbol{m}}(\boldsymbol{p}) = |\mathbf{A}|^{\frac{j}{2}} \psi\left(\mathbf{B}^k \mathbf{A}^j \boldsymbol{p} - \boldsymbol{m}\right)$, where $\mathbf{A}$ is a dilation matrix, controlling the scale decomposition, $\mathbf{B}$ is a shearing matrix, controlling the direction decomposition; j is the number of levels of the scale decomposition and k is that of the direction decomposition; $\boldsymbol{p} = (x, y)$ and $\boldsymbol{m} = (m_x, m_y)$. Then, the shearlet transform is expressed as $\boldsymbol{\alpha} = \langle \mathbf{S}, \psi_{j,k,\boldsymbol{m}} \rangle$ [9].

The 4-layer with 34 directions decomposed in each layer shearlet transform is applied to Fig. 2(b). Some high-frequency coefficients of the 3rd layer are reconstructed as Fig. 4(a). Then, Fig. 4(a) is binarized as Fig. 4(b). It is obvious that the ship wake is highlighted.



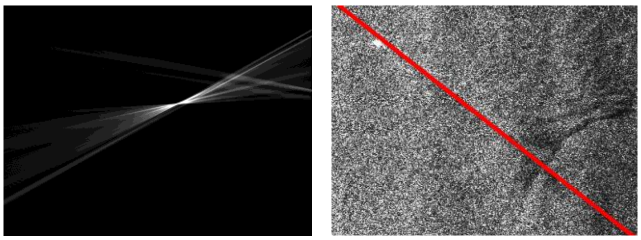
(a) Reconstructed cartoon component (b) Binary image of the cartoon component

Fig. 4 Enhancement of the cartoon component

3.3. Line detection based on PC transform

Finally, the line detection method based on the PC transform [10] is used to detect a continuous ship wake in the binary cartoon component. The basic ideas of the PC transform are that 1) a point in the Cartesian coordinate system (CCS) is converted into a line in the parallel coordinate system (PCS), 2) all points on a CCS line whose slope is negative are mapped as a group of PCS lines, which intersects at a point, and 3) a point in the PCS is remapped as a CCS line.

The above line detection method is applied to Fig. 4(b). Fig. 5(a) is a PC transform domain containing a ship wake. Fig. 5(a) is thresholded and clustered, and then one peak can be obtained. Finally, this peak is mapped back to the CCS, and the ship wake is detected. For easy observation, the detected ship wake is indicated in bold as shown in Fig. 5(b).



(a) PC transform domain (b) Line detection result

Fig. 5 Line detection based on PC transform

4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed method, 21 ERS-2 images of size 300×400 pixels are used, and experimental results are quantitatively and qualitatively compared with those of Refs. [2-3]. In the experiments, optimal parameters for each method are empirically selected, and the recall and the precision of all images for each method are calculated for the quantitative comparison, and at the same time, the detection results of each method are arranged column-by-column for the qualitative comparison.

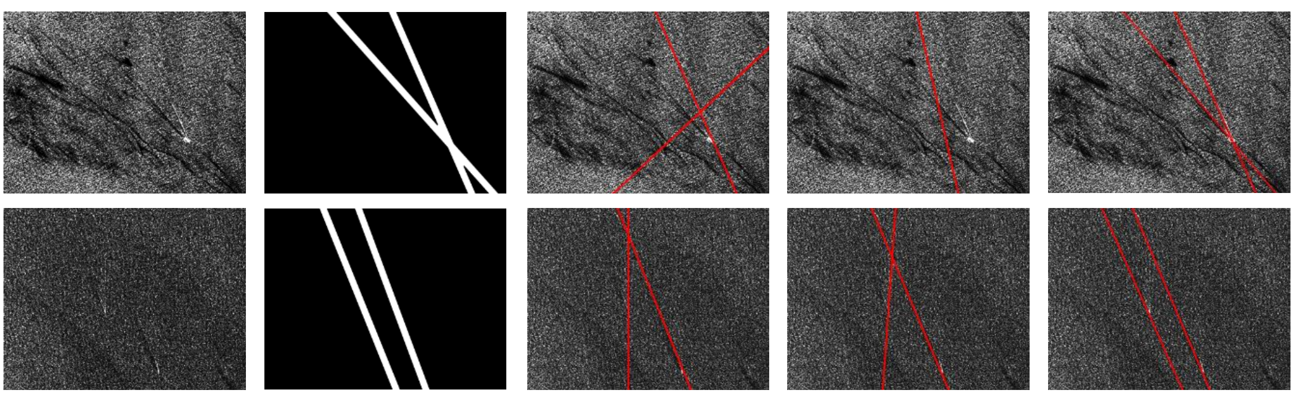
The recall and the precision are calculated as follows [11]:

$$\begin{cases} \mathrm{Recall} = \dfrac{\mathrm{TP}}{\mathrm{nP}} \\ \mathrm{Precision} = \dfrac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \end{cases} \tag{9}$$

where TP represents the number of detected ship wake lines, FP represents the number of false detection, and nP represents the total number of ship wake lines. Tab. 1 lists the recall and the precision of each method. We see that the recall and the precision of the proposed method are higher than the two other methods. Figs. 6(c)-6(e) show the detection results of all the three methods on 2 out of 21 SAR images shown as Fig. 6(a). Fig. 6(b) is the ground truth of Fig. 6(a). We see that the performance of the proposed method is also better than the two other methods. So we draw the conclusion that the proposed method outperforms other methods for complex backgrounds.

Table 1 Quantitative comparison of our method with Refs.[2-3]

	Ref [2]	Ref [3]	Our method
Recall	0.41	0.22	**0.72**
Precision	0.39	0.24	**0.70**



(a) Original images (b) Ground truth (c) Results of Ref. [2] (d) Results of Ref. [3] (e) Results of our method

Fig. 6 Experimental comparison results of the related methods

5. CONCLUSIONS

The detection of ship wakes in SAR images under complex background conditions is a challenging task. The traditional ship wake detection methods only work in simple backgrounds but hardly work in complex backgrounds. In this paper, we propose a novel ship wake detection method based on the MCA and the dictionary learning for complex backgrounds. In the proposed method, the cartoon component containing ship wakes is separated from the sea-background texture component by using the MCA and an online updating mechanism for the ship wake dictionary and the sea-background texture dictionary. Then, some high-frequency coefficients of the cartoon component are extracted and reconstructed by the shearlet transform to enhance ship wakes. Finally, all of the ship wakes are detected by the PC transform. The quantitative and qualitative experimental results demonstrate that, our proposed method outperforms other traditional ship wake detection methods.

6. REFERENCES

[1] M. T. Rey, J. K. Tunaley, J. T. Folinsbee, P. A. Jahans, J. A. Dixon, and M. R. Vant, “Application Of Radon Transform Techniques To Wake Detection In Seasat-A SAR Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 4, pp. 553–560, Jul. 1990.

[2] J. Ai, X. Qi, W. Yu, Y. Deng, F. Liu, L. Shi, and Y. Jia, “A Novel Ship Wake CFAR Detection Algorithm Based on SCR Enhancement and Normalized Hough Transform,” *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 681–685, Jul. 2011.

[3] X. Xing, K. Ji, H. Zou, J. Sun, S. Zhou, and Z. Li, “An enhancing normalized radon transform method for ship wake detection in SAR imagery,” in *9th European Conference on Synthetic Aperture Radar, 2012. EUSAR*, Nuremberg, Germany, 2012, pp. 559–562.

[4] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” presented at the 1993 Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, 1993, 1993, pp. 40–44 vol.1.

[5] M. J. Fadili, J.-L. Starck, J. Bobin, and Y. Moudden, “Image Decomposition and Separation Using Sparse Representations: An Overview,” *Proc. IEEE*, vol. 98, no. 6, pp. 983–994, Jun. 2010.

[6] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation,” *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[7] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, “Online Learning for Matrix Factorization and Sparse Coding,” *J Mach Learn Res*, vol. 11, pp. 19–60, 2010.

[8] G. Peyré, J. Fadili, and J. Starck, “Learning the Morphological Diversity,” *SIAM J. Imaging Sci.*, vol. 3, no. 3, pp. 646–669, 2010.

[9] G. Kutyniok and D. Labate, “Introduction to Shearlets,” in *Shearlets*, G. Kutyniok and D. Labate, Eds. Birkhäuser Boston, 2012, pp. 1–38.

[10] S. El Mejdani, R. Egli, and F. Dubeau, “Old and new straight-line detectors: Description and comparison,” *Pattern Recognit.*, vol. 41, no. 6, pp. 1845–1866, 2008.

[11] S. Agarwal, A. Awan, D. Roth. “Learning to detect objects in images via a sparse, part-based representation,” *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 26, no. 11, pp. 1475–1490, 2004.