



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Pseudo-Zernike Moments Based Sparse Representations for SAR Image Classification

Citation for published version:

Gishkori, S & Mulgrew, B 2018, 'Pseudo-Zernike Moments Based Sparse Representations for SAR Image Classification', *IEEE Transactions on Aerospace and Electronic Systems*.
<https://doi.org/10.1109/TAES.2018.2856321>

Digital Object Identifier (DOI):

[10.1109/TAES.2018.2856321](https://doi.org/10.1109/TAES.2018.2856321)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

IEEE Transactions on Aerospace and Electronic Systems

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Pseudo-Zernike Moments Based Sparse Representations for SAR Image Classification

Shahzad Gishkori, *Member, IEEE* and Bernard Mulgrew, *Fellow, IEEE*

Abstract—We propose radar image classification via pseudo-Zernike moments based sparse representations. We exploit invariance properties of pseudo-Zernike moments to augment redundancy in the sparsity representative dictionary by introducing auxiliary atoms. We employ complex radar signatures. We prove the validity of our proposed methods on the publicly available MSTAR dataset.

Index Terms—Sparse representations, pseudo-Zernike moments, SAR image classification, complex signatures

I. INTRODUCTION

Synthetic aperture radar (SAR) can provide all-weather imagery with a very high resolution [1]. This has naturally led to using SAR for the purpose of automatic target recognition or classification. Initial usage was military related. However, SAR imaging with the aim of classification is making very quick strides for the automotive usage as well [2]. Traditionally, a number of techniques are used for SAR image classification. A state-of-the-art review on different approaches used for automatic target recognition in SAR imagery can be found in [3]. Here, we briefly mention a couple of them. Template based classification [4] requires generation of a large number of templates for each target and then matching the test image with those templates in an exhaustive search manner. It is an effective linear approach. However, it is computationally quite expensive. Among the nonlinear approaches, support vector machine classifier (SVC) has been quite popular [5]. It is a large margin classifier and it can outperform the template based classifier. However, this approach is dependent upon accurate estimation of the pose angle which involves an extra preprocessing stage.

Recent trends in classification are based on sparse representations, also known as sparse coding [6], [7]. Initially, efforts were made to find or use a unified dictionary for all the classes, see, e.g., [8], [9] and references therein. Instead of using a single dictionary for all the classes, [10] proposed to use unit normalised measurements of the objects as the columns of an overcomplete dictionary. Coding is done through an ℓ_1 -norm minimisation problem and the classification is based on a least-squares metric w.r.t. the group of columns specific to a particular class object. This is known as sparse representation based classifier (SRC). The ease of formulating a dictionary by using the measurements of the class objects directly, made

SRC a favourable choice for classification in a wide range of fields. In SAR image classification, SRC was used in [11] from the perspective of class manifolds [12]. A class manifold is defined over the set of measurements for a particular class object, and the SAR image is claimed to lie in that manifold by using the fact that linear representation can be provided to a nonlinear manifold if a local region of the manifold is considered [13]. This local region of the manifold gives the basis for sparse representation of a test image over that manifold. Therefore, it obviates the need for a rigorous preprocessing as well as pose angle estimation. Dimensionality reduction can be achieved via random projections. However, it can result in a performance loss.

It was shown in [10], [11] that SRC can outperform linear SVC (LSVC), i.e., when a linear kernel is used. However, SRC is primarily based upon sparse reconstruction or coding, and it does not involve the classification aspect during the coding process. A number of papers have been written to incorporate this aspect in sparse representations. Some discriminative dictionary learning techniques have been proposed in [14], [15]. Similarly, joint dictionary learning and encoding has been proposed in [16] (see e.g., [17] and [18] for such recent attempts). Although, these methods provide good performance but dictionary learning, whether discriminative or not, is a computationally intensive process.

Moments based image representations have been successfully used over many decades [19], [20]. The basic idea is to derive image features which are scale-, shift- and rotation-invariant by using nonlinear combinations of the regular moments (also called geometric moments). However, the gains have been limited, primarily due to the non-orthogonality of regular moments. Orthogonal moments, e.g., Legendre moments, Zernike moments and pseudo-Zernike (PZ) moments [21], [22] have been a popular substitute in pattern recognition. Among these, PZ-moments stand apart both in terms of generating the maximum number of invariant moments as well as in terms of performance regarding noise rejection. PZ-moments have been used for radar automatic target recognition in [23] with a nearest neighbour classifier. Similarly, in [24], PZ-moments have been used for radar classification based on its micro-Doppler signatures, with an SVC. However, in both these cases, the emphasis has been on feature extraction w.r.t. PZ-moments and not on the choice of an optimal classifier.

Contributions. In this paper, we propose using PZ-moments in combination with the SRC framework (PZ-SRC), in order to gain from both optimal feature extraction as well as optimal classification. By using a finite number of PZ-moments, we reduce the dimensionality of the problem. This

S. Gishkori and B. Mulgrew are with Institute for Digital Communications (IDCOM), The School of Engineering, The University of Edinburgh, UK. Emails: {s.gishkori, bernie.mulgrew}@ed.ac.uk

This work was supported by Jaguar Land Rover and the UK-EPSC grant EP/N012240/1 as part of the jointly funded Towards Autonomy: Smart and Connected Control (TASCC) Programme.

has a direct impact on reducing the computational complexity of the proposed method, as shown in subsequent sections. Due to invariance properties of the PZ-moments, we obtain good performance, albeit in the low dimensional setting. We also introduce auxiliary atoms in the dictionary to increase the redundancy of information, which further exploits the invariance properties of the PZ-moments. Thus, information is better localised in individual class manifolds. This results in a further improvement in classification performance. Note, this also forms a unique contribution to the SRC framework, in general, as well. In order to utilise both the magnitude as well as the phase information of the complex radar signatures, we fuse the two parameters by a simple averaging mechanism (see [25] for details on the fusion mechanisms). This results in an even more informative radar signatures with direct positive impact over the classification performance. Note, a similar approach has been used in [26]. However, the feature extraction there is based on regular moments and the authors do not use auxiliary atoms. Now, in order to encode the test image in our proposed framework of PZ-SRC, we use state-of-the-art technique of iterative hard thresholding (IHT) algorithm [27]. IHT provides very fast convergence as well accuracy (in comparison to the approach of [10]), both of which are crucial in real-time radar image classification applications. We test our proposed methods on the publicly available moving and stationary target acquisition and recognition (MSTAR) dataset [28].

Organisation. Section II gives the basics of PZ-moments, Section III briefly describes the SRC method, Section IV details our proposed method of PZ-SRC, Section V provides experimental results and conclusions are given in Section VI. **Notations.** Matrices are in upper case bold while column vectors are in lower case bold, $(\cdot)^T$ denotes transpose, $[\mathbf{x}]_i$ is the i th element of \mathbf{x} , $\hat{\mathbf{x}}$ is the estimate of \mathbf{x} , \triangleq defines an entity, and the ℓ_p -norm is denoted as $\|\mathbf{x}\|_p = (\sum_{i=0}^{N-1} |[\mathbf{x}]_i|^p)^{1/p}$.

II. PSEUDO-ZERNIKE MOMENTS

Let a piecewise continuous function $s(x, y)$ (with bounded support) be the intensity function of a 2-D real image in Cartesian coordinates. The regular moments of $s(x, y)$ can be defined as

$$\mu_{p,q} = \int_x \int_y x^p y^q s(x, y) dx dy \quad (1)$$

where $\{p, q\} \in \mathbb{Z}_+$ and $p + q$ is the degree of the moments. Note, (1) represents the projection of $s(x, y)$ on monomial $x^p y^q$. Since $\{x^p y^q\}$ is not an orthogonal set, $\mu_{p,q}$ are not independent moments. In contrast, the PZ-moments are generated from a set of orthogonal polynomials. We refer to these polynomials as PZ-polynomials. The PZ-polynomials are a set of complex polynomials described as

$$z_n^m(r, \theta) = \rho_n^m(r) \exp(jm\theta) \quad (2)$$

where $r \triangleq \sqrt{x^2 + y^2}$ and $\theta \triangleq \tan^{-1}(y/x)$ are the length and angle of the position vector of a point (x, y) w.r.t. the centre of the image, respectively, $n \in \mathbb{Z}_+$ is the degree of the

polynomial with frequency m , i.e., $m \in [-n, +n]$, and

$$\rho_n^m(r) \triangleq \sum_{\kappa=0}^{n-|m|} \frac{(-1)^\kappa (2n+1-\kappa)! r^{n-\kappa}}{\kappa! (n+|m|+1-\kappa)! (n-|m|-\kappa)!} \quad (3)$$

is the radial polynomial. When defined over a unite circle, i.e., $r \leq 1$, the PZ-polynomials exhibit orthogonality, i.e.,

$$\int_0^{2\pi} \int_0^1 [z_n^m(r, \theta)]^* z_{n'}^{m'}(r, \theta) r dr d\theta = \frac{\pi}{n+1} \delta_{nn'} \delta_{mm'} \quad (4)$$

where $\delta_{ii'}$ is the Kronecker delta function. Note, it can be seen via simple enumeration that cardinality of the set of PZ-polynomials with degree $\leq n$ is, $P = (n+1)^2$. Now, the PZ-moments can be obtained by projecting the image onto the PZ-polynomials as¹

$$a_n^m = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 [z_n^m(r, \theta)]^* s(r, \theta) r dr d\theta \quad (5)$$

where $s(r, \theta) = s(x, y)|_{x=r \cos \theta, y=r \sin \theta}$. Due to (4), it can be shown that (5) generates a set of independent moments.

The invariance properties of the PZ-moments can be established via mathematical manipulations. For scale and translation invariance, one way is to use the regular moments of the image. The transformed image can be written as

$$g(x, y) = s\left(\frac{x}{v} + m_x, \frac{y}{v} + m_y\right) \quad (6)$$

where $m_x \triangleq \mu_{1,0}/\mu_{0,0}$ and $m_y \triangleq \mu_{0,1}/\mu_{0,0}$ are the centroid adjustment parameters of the image $s(x, y)$, and $v \triangleq \sqrt{\xi/\mu_{0,0}}$ is the scale adjustment parameter of the image $s(x, y)$ with a predetermined value ξ . Now, the scale- and translation-invariant PZ-moments can be generated by replacing $s(r, \theta)$ with $g(r, \theta)$ in (5), where $g(r, \theta) = g(x, y)|_{x=r \cos \theta, y=r \sin \theta}$. Since PZ-polynomials are a set of complex polynomials, the PZ-moments generated via (5) are also complex. The rotation invariance of the PZ-moments refers to the magnitude part only, i.e., $|a_n^m|$ and not the phase.

III. SPARSE REPRESENTATION BASED CLASSIFIER

Let a generic $\sqrt{N} \times \sqrt{N}$ image with intensity function $g(x, y)$ (or $g(r, \theta)$) is represented as an $N \times 1$ vector \mathbf{g} via a lexicographic ordering (column or row ordered). Let \mathbf{g}_j^k be the j th image measurement of the k th object class, for $j = 1, \dots, J_k$ and $k = 1, \dots, K$. Now, given a set of training image measurements $\{\mathbf{g}_j^k\}$, with $\|\mathbf{g}_j^k\|_2^2 = 1$, the SRC method defines the dictionary as

$$\mathbf{G} \triangleq [\mathbf{G}^1, \mathbf{G}^2, \dots, \mathbf{G}^K] \quad (7)$$

where \mathbf{G} is an $N \times J$ matrix with $J = \sum_{k=1}^K J_k$ and $\mathbf{G}^k \triangleq [\mathbf{g}_1^k, \mathbf{g}_2^k, \dots, \mathbf{g}_{J_k}^k]$ is an $N \times J_k$ matrix acting as a sub-dictionary for class k , for $k = 1, \dots, K$. Any test image measurement, represented as an $N \times 1$ vector $\tilde{\mathbf{y}}$ can then be decomposed or encoded according to the linear model

$$\tilde{\mathbf{y}} = \mathbf{G}\tilde{\mathbf{x}} + \tilde{\mathbf{n}} \quad (8)$$

¹Note, in case of a digital image, the integrals in the projection operations are replaced by summations.

where $\tilde{\mathbf{x}}$ is a $J \times 1$ vector of coefficients defined as, $\tilde{\mathbf{x}} \triangleq [\tilde{\mathbf{x}}^1 T, \tilde{\mathbf{x}}^2 T, \dots, \tilde{\mathbf{x}}^K T]^T$, where $\tilde{\mathbf{x}}^k$ are the coefficients w.r.t. the sub-matrix \mathbf{G}^k , and the $N \times 1$ vector $\tilde{\mathbf{n}}$ accounts for model errors with a bounded energy, i.e., $\|\tilde{\mathbf{n}}\|_2 < \tilde{\epsilon}$. It is clear from (8) that given $\tilde{\mathbf{y}}$ belongs to the k th class, $\tilde{\mathbf{x}}$ would be a sparse vector. Now, an estimate of $\tilde{\mathbf{x}}$ can be obtained by solving the following ℓ_1 -norm optimisation problem (OP).

$$\hat{\tilde{\mathbf{x}}} = \arg \min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{y}} - \mathbf{G}\tilde{\mathbf{x}}\|_2^2 + \lambda \|\tilde{\mathbf{x}}\|_1 \quad (9)$$

where $\lambda > 0$. The classification result is then obtained by finding the k for which $\|\tilde{\mathbf{y}} - \mathbf{G}^k \hat{\tilde{\mathbf{x}}}^k\|_2^2$ is minimum, for $k = 1, \dots, K$. In case of feature based representation, the SRC model takes the form,

$$\mathbf{R}\tilde{\mathbf{y}} = \mathbf{R}\mathbf{G}\tilde{\mathbf{x}} + \mathbf{R}\tilde{\mathbf{n}} \quad (10)$$

where \mathbf{R} is an $R \times N$ linear transformation matrix. When $R < N$, it can reduce the computational complexity of the SRC approach. However, it may reduce the classification performance as well. If the elements of \mathbf{R} are drawn from a normal/Gaussian distribution, i.e., $\mathcal{N}(0, 1)$, (10) is known as SRC with random projections (\mathbf{R}_ω -SRC) [11], where $R = N/\omega$ and $\omega \geq 1$.

IV. PZ-MOMENTS BASED SPARSE REPRESENTATIONS

In this paper we consider feature based sparse representations. In our case, PZ-moments form the feature set of the radar image. Since PZ-moments are generated by projecting the image onto PZ-polynomials, we can easily generate the features by converting the PZ-polynomials into a basis matrix and then projecting the image vector onto this basis matrix. Let a $P \times 1$ vector \mathbf{z}_i of PZ-polynomials, with degree $\leq n$, w.r.t. image point (r_i, θ_i) , where (r_i, θ_i) are the Polar coordinates equivalent of the image point (x_i, y_i) in Cartesian coordinates, for $i = 1, \dots, N$, be defined as

$$\mathbf{z}_i \triangleq [\gamma_0 z_0^0(r_i, \theta_i), \gamma_1 z_1^{-1}(r_i, \theta_i), \gamma_1 z_1^0(r_i, \theta_i), \gamma_1 z_1^{+1}(r_i, \theta_i), \dots, \gamma_n z_n^{-n}(r_i, \theta_i), \gamma_n z_n^{-n+1}(r_i, \theta_i), \dots, \gamma_n z_n^{+n}(r_i, \theta_i)]^T \quad (11)$$

where $\gamma_n \triangleq (n+1)/(\pi N)$ accounts for subsequent constants as well as integration to summation approximations in (5). Note, we assume that $r_i \leq 1, \forall i \in [1, N]$, which ensures that all image points are within the unit circle. The PZ-polynomials based basis matrix can then be defined as a $P \times N$ matrix \mathbf{Z} , i.e.,

$$\mathbf{Z} \triangleq [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N] \quad (12)$$

which is still a matrix of complex polynomials. Now, given a set of training image measurements $\{\mathbf{g}_j^k\}$, for $j = 1, \dots, J_k$ and $k = 1, \dots, K$, the dictionary based on PZ-moments features, with the property of rotational invariance, can be defined as a column normalised (i.e., normalised to unity) $P \times J$ matrix \mathbf{A} , i.e.,

$$\mathbf{A} \triangleq \text{abs}(\mathbf{Z}\mathbf{G}) = [\mathbf{A}^1, \mathbf{A}^2, \dots, \mathbf{A}^K] \quad (13)$$

where $\text{abs}(\cdot)$ is a function which generates element-wise absolute values, and $\mathbf{A}^k \triangleq \text{abs}(\mathbf{Z}\mathbf{G}^k)$ is a $P \times J_k$ matrix of PZ-moments w.r.t. \mathbf{G}^k , for $k = 1, \dots, K$. In order to

capitalise on the invariance structure provided by PZ-moments, we introduce auxiliary atoms in the dictionary (see Section IV-B for details). Thus, the dictionary can be defined as

$$\begin{aligned} \Phi &\triangleq [[\mathbf{A}^1, f(\mathbf{A}^1)], [\mathbf{A}^2, f(\mathbf{A}^2)], \dots, [\mathbf{A}^K, f(\mathbf{A}^K)]] \\ &= [\Phi^1, \Phi^2, \dots, \Phi^K] \end{aligned} \quad (14)$$

where $f(\mathbf{A}^k)$ is a $P \times L_k$ auxiliary matrix (with columns normalised to unity) and it is a function of the columns of \mathbf{A}^k , $\Phi^k \triangleq [\mathbf{A}^k, f(\mathbf{A}^k)]$ is a $P \times Q_k$ matrix with $Q_k = J_k + L_k$, for $k = 1, \dots, K$, and the over-complete dictionary Φ is a $P \times Q$ matrix with $Q = \sum_{k=1}^K Q_k$. Now, the test image $\tilde{\mathbf{y}}$ can be encoded according to the following linear model.

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{n} \quad (15)$$

where $\mathbf{y} \triangleq \text{abs}(\mathbf{Z}\tilde{\mathbf{y}})$ is the $P \times 1$ vector of PZ-moments of the test image, \mathbf{x} is the $Q \times 1$ encoded vector defined as, $\mathbf{x} \triangleq [\mathbf{x}^1 T, \mathbf{x}^2 T, \dots, \mathbf{x}^K T]^T$, where \mathbf{x}^k is the $Q_k \times 1$ encoding vector w.r.t. Φ^k , for $k = 1, \dots, K$, and \mathbf{n} is the $P \times 1$ model error vector with bounded energy, i.e., $\|\mathbf{n}\|_2 < \epsilon$. It is clear from (15), given that the test image belongs to a particular class, \mathbf{x} would be a sparse vector with nonzero elements ideally corresponding to the sub-dictionary of only a particular class.

A. Sparse Reconstruction and Classification

Since $P \ll Q$, (15) is an under-determined system of linear equations. In order to recover \mathbf{x} in (15), we use IHT as the sparse recovery algorithm. An estimate of \mathbf{x} can be obtained by processing the following iterations.

$$\hat{\mathbf{x}}^{[t+1]} = \mathcal{H}_\Gamma \left(\hat{\mathbf{x}}^{[t]} + \Phi^T (\mathbf{y} - \Phi \hat{\mathbf{x}}^{[t]}) \right) \quad (16)$$

where t is the iteration index (starting with $t = 0$) and \mathcal{H}_Γ is the hard thresholding operator defined as

$$\mathcal{H}_\Gamma(\mathbf{q}) \triangleq \mathbf{q} \mathbb{I}_{\{i | |\mathbf{q}_i| \geq [\text{Ascend}(\mathbf{q})]_\Gamma, \forall i\}} \quad (17)$$

where $\mathbb{I}_{\{\cdot\}}$ is an indicator operator which discards those elements of vector \mathbf{q} that are not in the indicator set (the set given in its subscript), and $\text{Ascend}(\mathbf{q})$ is a sorting function which sorts the elements of \mathbf{q} in an ascending order. Essentially, $\mathcal{H}_\Gamma(\mathbf{q})$ preserves only the Γ largest element magnitudes of \mathbf{q} in each iteration t . Thus, (16) approximates the ℓ_0 -norm estimate of \mathbf{x} , i.e.,

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 \quad \text{subject to } \|\mathbf{x}\|_0 \leq \Gamma \quad (18)$$

where Γ is the order of sparsity. Note, the stopping criterion of iterations in (16) can either be the maximum number of allowable iterations or the minimum residual error, i.e., $\|\mathbf{y} - \Phi \hat{\mathbf{x}}^{[t]}\|_2 / \|\mathbf{y}\|_2$. After sparse encoding of \mathbf{y} , the classification of the target image can be done by solving the following OP.

$$\hat{k} = \arg \min_k \|\mathbf{y} - \Phi^k \hat{\mathbf{x}}^k\|_2^2, \quad \text{for } k = 1, \dots, K \quad (19)$$

where $\hat{\mathbf{x}}^k$ is the estimate obtained in the PZ-SRC framework, when the stopping criterion for (18) has been achieved.

Note, the computational complexity of IHT for PZ-SRC is $\mathcal{O}(PQ)$ per iteration, where $\mathcal{O}(\cdot)$ denotes the order of complexity. Since $P \ll N$, the computational complexity of PZ-SRC is much lower than SRC.

B. Auxiliary Atoms

The auxiliary atoms can have a substantial impact on the performance of the classification. Ideally, variations in the image measurements w.r.t. different aspect angles, should not produce any variations in their respective PZ-moments. However, radar reflectivities at different aspect angles might not be uniform. Therefore, image at one aspect angle might be absolutely different from the image obtained at another aspect angle. Also, noise in the form of clutter or other artefacts can play a disruptive role. Auxiliary atoms try to recover the information lost due to these irregularities. In this section, we present a number of techniques to generate the auxiliary atoms. Note, here our focus is primarily on rotational invariance of the moments.

1) *Fixed Auxiliary Atoms (AuxFix)*: In case the measurements are obtained at random aspect angles, we propose to constitute the auxiliary atoms as an overall average of the PZ-moments based measurements of each class, i.e.,

$$f(\mathbf{A}^k) = \sum_{j=1}^{J_k} \mathbf{a}_j^k \quad (20)$$

where $\mathbf{a}_j^k \triangleq \text{abs}(\mathbf{Z}\mathbf{g}_j^k)$, for $k = 1, \dots, K$. AuxFix causes the effect of irregular reflectivities to be averaged out. Here, $L_k = 1$, for $k = 1, \dots, K$.

2) *Moving Average Based Auxiliary Atoms (AuxMov)*: In case the measurements are arranged in the order of increasing aspect angles around the object, a moving average of atoms over each class can constitute the auxiliary atoms, i.e.,

$$f_j(\mathbf{A}^k) = \sum_{w=-W_k/2}^{+W_k/2} \mathbf{a}_{w+j}^k \quad (21)$$

where W_k is the window size for the k th class, for $k = 1, \dots, K$, and $j = 1, \dots, J_k$. We can see that the window is centred over the j th column of \mathbf{A}^k . Note, in case $(w+j) < 1$ or $(w+j) > J_k$, \mathbf{a}_{w+j}^k can be considered as zero vectors. Here, $L_k = J_k$. The auxiliary matrix can be formed as

$$f(\mathbf{A}^k) = [f_1(\mathbf{A}^k), f_2(\mathbf{A}^k), \dots, f_{J_k}(\mathbf{A}^k)]. \quad (22)$$

3) *Correlation Based Auxiliary Atoms (AuxCorr)*: An optimal method is to find correlated atoms w.r.t. every training measurement for each class, i.e., the columns of \mathbf{A}^k . The auxiliary atoms can then be generated based on a minimum correlation value, i.e.,

$$f_j(\mathbf{A}^k) = \sum_{\substack{l=1 \\ \mathbf{a}_j^k T \mathbf{a}_l^k > \Upsilon}}^{J_k} \mathbf{a}_l^k \quad (23)$$

where $\mathbf{a}_j^k T \mathbf{a}_l^k$ performs the inner product, Υ is the correlation threshold, and $j = 1, \dots, J_k$. Here, $L_k = J_k$. The auxiliary matrix can be formed according to (22). This procedure ensures that all informative measurements, i.e., measurements with high mutual correlation, are accounted for.

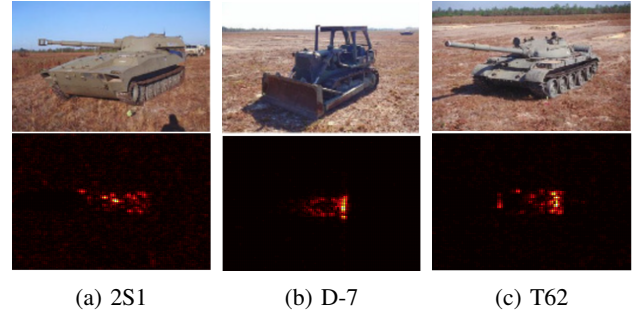


Fig. 1: MSTAR Targets.

C. Complex Signatures

We can see from the previous sections that most of the classification strategies use only the intensities or magnitudes of the images. However, a radar signature contains information both in the magnitude as well as in the phase. One way to benefit from both sources of information is to create an image of the target scene from each of these sources and then use an image fusion technique to optimally combine them into a single fused image. This resulting image will have more information of the target scene than the magnitude-only and the phase-only images. To this end, we combine the magnitude and the phase of the radar signatures via an averaging fusion mechanism, and use the fused image to create the PZ-moments. Thus, the fused image has the generalized form, $\alpha[\log(\text{abs}(\{\mathbf{g}_j^k\}))] + \beta[\text{phase}(\{\mathbf{g}_j^k\})]$, where $\text{phase}(\cdot)$ is an element-wise phase generating function, and, $0 < \alpha \leq 1$ and $0 < \beta \leq 1$ denote the magnitude and phase fractions, respectively.

Note, we use an averaging fusion mechanism due to the ease of its implementation and the generality of its applicability across all classes of the target objects. However, other image fusion techniques, optimal for each target class (based on different performance metrics), can also be used to further enhance the classification performance. See [25] for more details on different image fusion techniques.

V. EXPERIMENTAL RESULTS

In this section we present experimental results of our proposed methods. We use the publicly available MSTAR dataset. The MSTAR data was collected by Sandia National Laboratory by using an X-band SAR (in spotlight mode) with 1 foot resolution. We consider three targets from this dataset, i.e., 2S1 tank, D-7 land clearing vehicle and T62 tank (so $K = 3$). Figure 1 shows the optical and SAR (magnitude only) images, for one aspect angle of these targets. For the purpose of training, a total of $J_k = 299$ measurements are considered, for each target, at a radar elevation angle of 17° . The measurements have been taken at sequentially increasing aspect angles of approximately 1.2° , i.e, covering the complete angular range of 360° . Note, the measurements are in the form of 96×96 SAR images. These images are vectorised for the sake of processing. Thus, $N = 9216$. For the purpose of testing, a total of 273 image measurements (for each class) are considered, which have been taken at

TABLE I: Performance Comparison of Different Classifiers

	$\Omega_1(\%)$	$\Omega_2(\%)$	$\Omega_3(\%)$	$\Omega(\%)$
LSVC	91.94	98.90	92.67	94.50
PZ-LSVC	93.40	99.26	96.33	96.33
SRC	95.97	99.26	93.04	96.09
R ₀₅ -SRC	95.97	99.26	92.31	95.24
R ₁₅ -SRC	95.24	98.53	90.84	94.87
R ₃₅ -SRC	92.67	98.53	88.64	93.28
PZ-SRC	96.70	99.26	96.33	97.43
PZ-SRC (ℓ_1)	97.07	98.53	95.97	97.19
PZ-C (ℓ_2)	95.60	97.06	95.60	96.09

different aspect angles over the complete angular range of 360° , with a radar elevation angle of 15° . Aspect angles of the testing measurements are different from the training measurements. Thus, pose angle estimation is a valid issue. We define the classification/recognition accuracy/performance for the k th class as

$$\Omega_k \triangleq 100 \left(\frac{TP_k}{273} \right) \quad (24)$$

where TP_k are the true positives of the target class k , for $k = 1, \dots, K$, and the overall performance is defined as

$$\Omega \triangleq \frac{1}{K} \sum_{k=1}^K \Omega_k. \quad (25)$$

Note, both Ω and Ω_k quantify performance in percentages. For PZ-moments, we consider $n = 10$ as the degree of the polynomials which generates $P = 121$ PZ-moments. In comparison to N , this is a dimensionality reduction by a factor of over 70. Note, the value of n can impact the performance of classification. Generally, higher values of n can represent an image better. However, very large values can cause numerical instabilities. Therefore, we select a moderate value of n . Few tests on the training data can also give a good idea over the choice of n . For sparse reconstruction, we use IHT for PZ-SRC (as well as for SRC and R $_{\omega}$ -SRC, for a fair comparison) and consider the order of sparsity $\Gamma = 5$. Note, the parameter Γ is a tuning parameter and can be selected based on different cross validation approaches.

In terms of experimental results, we first consider the magnitude only radar signatures. Table I shows the classification performance results of different classifiers. Note, we also consider PZ-moments based LSVC (PZ-LSVC) for the sake of comparison. We can see that SRC outperforms LSVC for all target classes. PZ-LSVC is slightly better than SRC in the overall performance, which shows that the usage of PZ-moments improves the performance of a classifier, in general. We also provide the performance results of random projections based SRC, i.e., R $_{\omega}$ -SRC for $\omega = 05, 15, 35$. As mentioned in Section III, increasing value of ω causes a reduction in the computational complexity of the SRC. However, we can see from Table I that it also causes a proportional reduction in the classification performance. In comparison, PZ-SRC shows better classification performance than all the classifiers, in every category. Note, despite showing

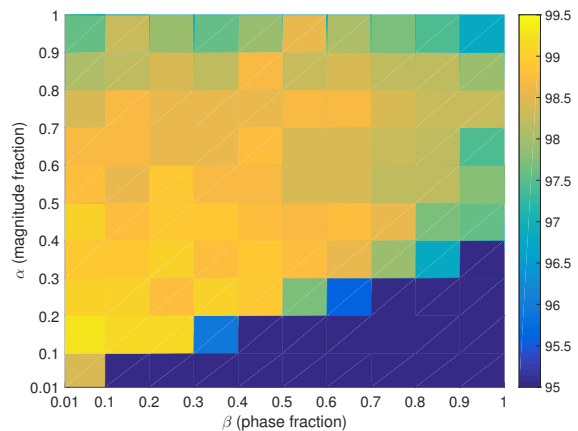


Fig. 2: $\Omega(\%)$ for PZ-SRC (Complex)

improved performance, the computational complexity of PZ-SRC is much less than other methods. In the current setting, the computational complexity of PZ-SRC is less than R₇₀-SRC, whereas, its performance is significantly higher than R₃₅-SRC. Thus, we can say that in comparison to other methods, PZ-SRC provides significantly higher classification performance at a significantly lower computational complexity. We also compare the performance of PZ-SRC with reconstruction approaches different from (18). In (18), we use an ℓ_0 -norm penalty to obtain a sparse solution. Since solving an ℓ_0 -norm problem is NP hard, we have used the IHT algorithm, which is a greedy algorithm, to approximate its solution. Another popular way to solve an ℓ_0 -norm problem is to relax it by an ℓ_1 -norm which is also known as basis pursuit (BP) [7] or least absolute shrinkage and selection operator (LASSO) [29]. We denote PZ-SRC with ℓ_1 -norm based solver as PZ-SRC (ℓ_1). From Table I, we can see that the difference in performance between PZ-SRC and PZ-SRC (ℓ_1) is not significant. However, in our experiments we observed that the former converged by an order of magnitude faster than the latter. Apart from an ℓ_1 -norm based reconstruction, we also compare the performance of PZ-SRC with ℓ_2 -norm based reconstruction approach. This is known as Tikhonov regularisation. We denote it as PZ-C (ℓ_2). As explained in Section IV, our OP is an under-determined problem. In principle an ℓ_2 -norm based approach can solve such an OP. However, our OP also requires a sparse solution which cannot be achieved by an ℓ_2 -norm based approach. Therefore, we can see from Table I that PZ-SRC outperforms PZ-C (ℓ_2).

Now, in the light of Section IV-C, we evaluate the performance of complex radar signatures, for varying values of α (magnitude fraction) and β (phase fraction). Figure 2 presents these performance results. We can see, in comparison to 97.43% of the magnitude only signatures in Table I, using complex signatures can enhance the performance of PZ-SRC to 99.14%, for $\alpha, \beta = 0.1$. For the sake of simplicity, we keep $\alpha, \beta = 0.5$ for which $\Omega = 98.41\%$, in this paper. For the rest of the experiments, we use fused complex signatures. We first obtain classification results by considering AuxFix of Section IV-B1 as auxiliary atoms. Table II shows the

TABLE II: Confusion Matrix for PZ-SRC (AuxFix)

	2S1	D-7	T62	$\Omega_k(\%)$	$\Omega(\%)$
2S1	264	0	9	96.70	-
D-7	0	271	2	99.26	-
T62	0	1	272	99.63	-
	-	-	-	-	98.53

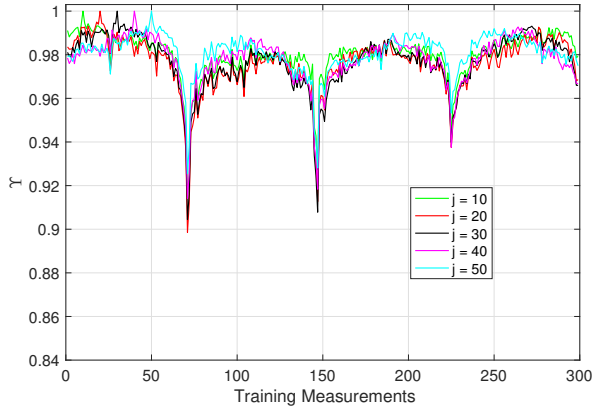


Fig. 3: Correlations among PZ-moments based Measurements.

confusion matrix in this regard. The performance improvement has been encouraging, with $\Omega = 98.53\%$. Next, we simulate the classification problem by considering AuxMov of Section IV-B2 as auxiliary atoms. Table III shows the performance of PZ-SRC for varying sizes of W_k (same $\forall k$). We can see that the classification performance is affected by changing size of W_k . The best performance is achieved when W_k/J_k is a multiple of 0.5. If all the test measurements are divided into four quadrants, with each quadrant corresponding to a range of aspect angles of approximately 90° , then $W_k/J_k = 0.5$ essentially corresponds to the numerical size of one quadrant, when the best performance is achieved. This can be explained as follows. Due to the rotational invariance properties of the PZ-moments, measurements at consecutive aspect angles are correlated to each other, in general, with some variations mostly because of radar reflectivity irregularities. However, measurements at the boundary of two quadrants correspond to the fine corners of the considered rectangular-shaped targets, and these measurements are highly uncorrelated with all the measurements in the preceding or the succeeding quadrant. This phenomenon can be seen in Figure 3, which shows the correlations of a few training measurements (PZ-moments) with the rest of the measurements in a k th target class. We can see that mutual correlations are minimum at quadrants, i.e., when $W_k/2 = 75, 150, 225$. Thus, it is better to exploit only the correlated measurements for generating auxiliary atoms and that happens when the size of W_k is such that it contains most of the correlated measurements of a quadrant or its multiple. Since, correlation is an important parameter for generating auxiliary atoms, we next assess the classification performance by considering AuxCorr of Section IV-B3. Table IV shows the classification performance with varying Υ . We can see that best performance is achieved for $\Upsilon = 0.94$. This

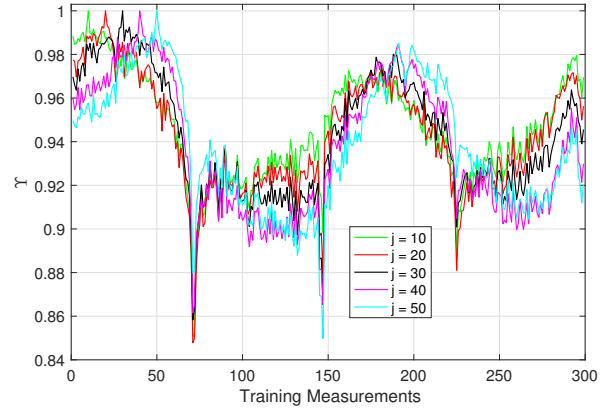


Fig. 4: Correlations among Test Measurements.

is quite understandable. A higher value of Υ does not collect enough number of informative measurements and a lower value of Υ involves noisy or non-informative measurements. This can be seen from Figure 3 as well. We also provide a confusion matrix regarding the performance of PZ-SRC with fused complex signatures and using $W_k/J_k = 0.5$, in Table V. An overall performance of 98.90% is achieved. Note, the overall performance can reach over 99.50% for $\alpha, \beta = 0.1$ (see Section IV-C). This enhanced classification performance has been achieved at a very low computational cost. In our case, the computational complexity of PZ-SRC with a maximum number of J auxiliary atoms is less than R_{35} -SRC, whereas, its classification performance is over 5% higher than R_{35} -SRC (see Table I). A comparison of computation time, using a Macbook Pro (2.2 GHz Intel Core i7, 16 GB 1600 MHz DDR3), is given in Table VI. We also compare the performance of PZ-C (ℓ_2) with our proposed methods of PZ-SRC. Table VII shows this comparison for magnitude only measurements, complex measurements and when using auxiliary atoms from AuxCorr of Section IV-B3. We can see that the holistic approach of PZ-C (ℓ_2) is not able to cope with the proposed modifications. However, the sparsity preserving approach of PZ-SRC exhibits substantial gains and can perform more than 4% higher than PZ-C (ℓ_2), for a similar setting. Note, in order to better appreciate the invariance properties of the PZ-moments, we also plot the correlations among original test measurements, i.e., without PZ-moments, in Figure 4. We can see that the correlation structure is quite inconsistent in comparison to the PZ-moments as in Figure 3.

VI. CONCLUSIONS

In this paper, we presented sparse representations for radar image classification by using pseudo-Zernike moments. We obtained a reduction in the dimensionality of the problem without compromising the performance. We exploited the invariance properties of the pseudo-Zernike moments to generate auxiliary atoms to complement the dictionary, which resulted in an enhanced classification performance. We used a fusion strategy to gain both from the magnitude as well as the phase of the radar signatures. We proved the validity of our proposed methods via numerical experiments on the MSTAR dataset.

TABLE III: Performance with Varying W_k in (21)

W_k	29	59	89	119	149	179	209	239	269	299	328	358	388	418	448
W_k/J_k	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
$\Omega(\%)$	98.16	98.41	98.65	98.77	98.90	98.65	98.65	98.77	98.90	98.90	98.90	98.77	98.65	98.77	98.90

TABLE IV: Performance with Varying Υ in (23)

Υ	0.98	0.97	0.96	0.95	0.94	0.93	0.92
$\Omega(\%)$	98.29	98.53	98.65	98.77	98.90	98.53	98.53

TABLE V: Confusion Matrix for PZ-SRC (AuxMov)

	2S1	D-7	T62	$\Omega_k(\%)$	$\Omega(\%)$
2S1	266	0	7	97.43	-
D-7	0	271	2	99.26	-
T62	0	0	273	100	-
	-	-	-	-	98.90

ACKNOWLEDGEMENTS

This work has been approved for submission by TASSC-PATHCAD Sponsor, Chris Holmes, Senior Manager Research, Research Department, Jaguar Land Rover, Coventry, UK.

REFERENCES

[1] W. Carrara, R. Goodman, and R. Majewski, *Spotlight Synthetic Aperture Radar*. Boston: Artech House, 1995.
 [2] S. Palm, R. Sommer, M. Caris, N. Pohl, A. Tessmann, and U. Stilla, "Ultra-high resolution SAR in lower terahertz domain for applications in mobile mapping," in *2016 GeMiC*, March 2016, pp. 205–208.
 [3] K. El-Darymli, E. W. Gill, P. McGuire, D. Power, and C. Moloney, "Automatic target recognition in synthetic aperture radar imagery: A state-of-the-art review," *IEEE Access*, vol. 4, pp. 6014–6058, 2016.
 [4] G. J. Owirka, S. M. Verbout, and L. M. Novak, "Template-based SAR ATR performance using different image enhancement techniques," in *of Proc. SPIE*, 1999, pp. 302–319.
 [5] Q. Zhao and J. C. Principe, "Support vector machines for SAR automatic target recognition," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 2, pp. 643–654, Apr 2001.
 [6] B. Olshausen and D. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607–609, 1996.
 [7] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, pp. 33–61, 1998.
 [8] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, Nov 2006.

TABLE VI: Comparison of Computation Time

	SRC	PZ-SRC	R ₀₅ -SRC	R ₃₅ -SRC
Time (sec)	111	4.23	20	5.5
$\Omega(\%)$	96.09	98.90	95.24	93.28

TABLE VII: PZ-SRC Vs PZ-C (ℓ_2)

	Magnitude Only $\Omega(\%)$	Complex $\Omega(\%)$	AuxCorr $\Omega(\%)$
PZ-C (ℓ_2)	96.09	95.72	94.87
PZ-SRC	97.43	98.41	98.90

[9] K. Huang and S. Aviyente, "Sparse representation for signal classification," in *Proceedings of the 19th International Conference on NIPS*. Cambridge, MA, USA: MIT Press, 2006, pp. 609–616.
 [10] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, Feb 2009.
 [11] J. J. Thiagarajan, K. N. Ramamurthy, P. Knee, A. Spanias, and V. Berisha, "Sparse representations for automatic target classification in sar images," in *2010 4th ISCCSP*, March 2010, pp. 1–4.
 [12] M. B. Wakin, D. L. Donoho, H. Choi, and R. G. Baraniuk, "The multiscale structure of non-differentiable image manifolds," in *SPIE Conference Series*, 2005, pp. 413–429.
 [13] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, 2000.
 [14] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Metaface learning for sparse representation based face recognition," in *2010 IEEE International Conference on Image Processing*, Sept 2010, pp. 1601–1604.
 [15] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2010, pp. 3501–3508.
 [16] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher discrimination dictionary learning for sparse representation," in *2011 International Conference on Computer Vision*, Nov 2011, pp. 543–550.
 [17] Y. Gao, J. Ma, and A. L. Yuille, "Semi-supervised sparse representation based classification for face recognition with insufficient labeled samples," *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2545–2560, May 2017.
 [18] J. Lu, G. Wang, and J. Zhou, "Simultaneous feature and dictionary learning for image set based face recognition," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 4042–4054, Aug 2017.
 [19] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. on Info. Theory*, vol. 8, no. 2, pp. 179–187, February 1962.
 [20] F. A. Sadjadi and E. L. Hall, "Three-dimensional moment invariants," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 2, no. 2, pp. 127–136, Feb. 1980.
 [21] C.-H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 496–513, Jul. 1988.
 [22] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 5, pp. 489–497, May 1990.
 [23] C. Clemente, L. Pallotta, I. Proudler, A. D. Maio, J. J. Soraghan, and A. Farina, "Pseudo-Zernike-based multi-pass automatic target recognition from multi-channel synthetic aperture radar," *IET Radar, Sonar Navigation*, vol. 9, no. 4, pp. 457–466, 2015.
 [24] C. Clemente, L. Pallotta, A. D. Maio, J. J. Soraghan, and A. Farina, "A novel algorithm for radar classification based on Doppler characteristics exploiting orthogonal pseudo-Zernike polynomials," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 1, pp. 417–430, Jan. 2015.
 [25] F. Sadjadi, "Comparative image fusion analysis," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, June 2005, pp. 8–8.
 [26] —, "Adaptive object classification using complex SAR signatures," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2016, pp. 299–303.
 [27] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.
 [28] E. R. Keydel, S. W. Lee, and J. T. Moore, "MSTAR extended operating conditions: a tutorial," *Proc. SPIE*, vol. 2757, pp. 228–242, 1996.
 [29] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.