Journal of Systems Engineering and Electronics
Vol. 19, No. 3, 2008, pp.493–498

Available online at www.sciencedirect.com
ScienceDirect
www.sciencedirect.com/science/journal/10044132

# Information compression and speckle reduction for multifrequency polarimetric SAR images based on kernel PCA*

*Li Ying, Lei Xiaogang, Bai Bendu & Zhang Yanning*
Dept. of Computer Science and Engineering, Northwest Polytechnical Univ., Xi'an 710072, P. R. China
(Received November 1, 2006)

**Abstract:** Multifrequency polarimetric SAR imagery provides a very convenient approach for signal processing and acquisition of radar image. However, the amount of information is scattered in several images, and redundancies exist between different bands and polarizations. Similar to signal-polarimetric SAR image, multifrequency polarimetric SAR image is corrupted with speckle noise at the same time. A method of information compression and speckle reduction for multifrequency polarimetric SAR imagery is presented based on kernel principal component analysis (KPCA). KPCA is a nonlinear generalization of the linear principal component analysis using the kernel trick. The NASA/JPL polarimetric SAR imagery of P, L, and C bands quadpolarizations is used for illustration. The experimental results show that KPCA has better capability in information compression and speckle reduction as compared with linear PCA.

**Keywords:** kernel PCA, multifrequency polarimetric SAR imagery, information compression, despeckling.

## 1. Introduction

The principle components analysis (PCA) has been a standard tool for image information compression and enhancement of multispectral data, such as LANDSAT, SPOT, and aircraft multispectral scanners[1−3]. The existence of corrections between spectral bands permits the PCA to condense the image information from a large number of bands into a small number of components with the additional advantages of noise reduction. Recent advances in the polarimetric synthetic aperture radar (polarimetric SAR) technology with multiple frequencies provide a rich set of multifrequency and multipolarization data for a single scene. For example, the NASA/JPL Polarimetric SAR has (P, L, C) bands, and each band has (HH, HV, VV) polarization (VH=HV, under the assumption of reciprocity), for a total of nine intensity images plus six phase difference images. The amount of information is scattered in several images, and redundancies exist between the images as indicated by the high correlations between the different bands and the polarization. The ability of the PCA to pack information and to reduce noise enables efficient automated image segmentation and /or better human interpretation.

The commonly used method for the principle components transformation of multispectral imagery proposed by Ready and Wintz [1] utilizes the covariance matrix of the spectral bands to compute the eigenvalues and eigenvectors. The first component image, which contains the maximum information is computed by the inner product between the spectral bands and the eigenvector, which is associated with the largest eigenvalue. Lee, et al.[4] proposed a speckle adjusted PCA for multifrequency polarimetric SAR, which maximizes the signal-to-speckle-noise ratio. However, the above methods are all based on the linear PCA methods. Clearly, linear PCA will not always detect all structures in a given data set. By the use of suitable nonlinear features, one can extract more information. Kernel principle component analysis (KPCA)[5−6] is very suitable to extract interesting nonlinear structures in the data. In contrast to PCA,

* This project was supported by the Specialized Research Found for the Doctoral Program of Higher Education (20070699013); the Natural Science Foundation of Shaanxi Province (2006F05); and the Aeronautical Science Foundation (05I53076).

KPCA is capable of capturing part of the higher-order statistics, which are particularly important for encoding the image structure [7]. This article presents a method of information compression and speckle reduction for multifrequency polarimetric SAR imagery based on KPCA, the experimental results of which are compared with that of linear PCA.

## 2. Kernel PCA

The KPCA algorithm first maps the data set $x_1, \ldots, x_n \in R^N$ into a high dimensional feature space$F$via a function $\Phi(x)$ and computes the covariance matrix

$$C = \frac{1}{n}\sum_{j=1}^{n} \Phi\left(x_j\right) \Phi\left(x_j\right)^{\mathrm{T}} \tag{1}$$

The principal components are then computed by solving the Eigenvalue problem: find $\lambda > 0$, $V \neq 0$ with

$$\lambda V = CV = \frac{1}{n}\sum_{j=1}^{n} \left(\Phi\left(x_j\right) \cdot V\right) \Phi\left(x_j\right) \tag{2}$$

Furthermore, as seen from Eq. (2), all eigenvectors with nonzero eigenvalue must be in the span of the mapped data, i.e., $V \in span\left\{\Phi\left(x_1\right), \ldots, \Phi\left(x_n\right)\right\}$. This can be written as

$$V = \sum_{i=1}^{n} \alpha_i \Phi\left(x_i\right) \tag{3}$$

by multiplying with $\Phi\left(x_k\right)$ from the left, Eq. (2) reads

$$\lambda\left(\Phi\left(x_k\right) \cdot V\right) = \left(\Phi\left(x_k\right) \cdot CV\right), \text{ for all } k = 1, \ldots, n \tag{4}$$

As the feature space $F$ may be very high dimensional (e.g. when mapping into the space of all possible $d$-th order monomials of input space), KPCA employs Mercer kernels instead of carrying out the mapping $\Phi$ explicitly. A Mercer kernel is a function $k(x, y)$, which for all data sets $\{x_i\}$ gives rise to a positive matrix $K_{ij} = k(x_i, x_j)$. One can see that using $k$instead of a dot product in input space corresponds to mapping the data with some $\Phi$ to a feature space$F$, i.e., $k\left(x, y\right) = \left(\Phi\left(x\right) \cdot \Phi\left(y\right)\right)$. Therefore, defining an $n \times n$-matrix

$$K_{ij} := \left(\Phi\left(x_i\right) \cdot \Phi\left(x_j\right)\right) = k\left(x_i, x_j\right) \tag{5}$$

one can compute an eigenvalue problem for the expansion coefficients $\alpha_i$, which is now solely dependent on the kernel function

$$n\lambda\alpha = K\alpha\left(\alpha = \left(\alpha_1, \ldots, \alpha_n\right)^{\mathrm{T}}\right) \tag{6}$$

The solutions $\left(\lambda_k, \alpha^k\right)$ must be further normalized by imposing $\lambda_k\left(\alpha^k \cdot \alpha^k\right) = 1$ in $F$. Also, as in every PCA algorithm, the data must be centered in $F$. This can be done by simply substituting the kernel-matrix $K$ with

$$\hat{K} = K - 1_n K - K 1_n + 1_n K 1_n \tag{7}$$

where $\left(1_n\right)_{ij} = 1/n$.

For extracting the features of a new pattern $x$ with KPCA, one can simply project the mapped pattern $\Phi\left(x\right)$ onto $V^k$

$$\left(V^k \cdot \Phi\left(x\right)\right) = \sum_{i=1}^{n} \alpha_i^k \left(\Phi\left(x_i\right) \cdot \Phi\left(x\right)\right) = \sum_{i=1}^{n} \alpha_i^k k\left(x_i, x\right) \tag{8}$$

KPCA has the same mathematic and statistic features as linear PCA in the $F$ space, such as, the principal components are not correlative and they can denote the maximum deviation of the sample data. Using the principal components to reconstruct the sample data, we can obtain the minimum square error, the reduction of the samples' denotation entropy, and so on. Besides, KPCA can extract more information of the data than linear PCA, and it need not face the nonlinear optimizing problem, which the other nonlinear PCA methods have to solve; KPCA only needs to cope with the calculation of the matrix's eigenvalues.

## 3. Information compression and speckle reduction based on KPCA

In each band, the multifrequency polarimetric SAR data can be denoted as a complex scattering matrix $S$

$$S = \begin{bmatrix} S_{\mathrm{HH}}, S_{\mathrm{HV}} \\ S_{\mathrm{VH}}, S_{\mathrm{VV}} \end{bmatrix} \tag{9}$$

For a reciprocal and calibrated polarimetric radar, the relation $S_{\mathrm{HV}} = S_{\mathrm{VH}}$ is satisfied, and therefore, a complex vector $U$ can also denote the multifrequency polarimetric SAR data