



ELSEVIER

Contents lists available at ScienceDirect

## Remote Sensing of Environment

journal homepage: [www.elsevier.com/locate/rse](http://www.elsevier.com/locate/rse)

## Randomized kernels for large scale Earth observation applications

Adrián Pérez-Suay\*, Julia Amorós-López, Luis Gómez-Chova, Valero Laparra, Jordi Muñoz-Marí, Gustau Camps-Valls

Image Processing Laboratory (IPL), Parc Científic, Universitat de València, C/ Catedrático José Beltrán, Paterna 2. 46980, València, Spain

## ARTICLE INFO

## Article history:

Received 1 July 2016

Received in revised form 13 January 2017

Accepted 11 February 2017

Available online xxx

## Keywords:

Machine learning

Image classification

Biophysical parameter estimation

Kernel machines

Random features

Cloud screening

IASI/MetOp

SEVIRI/MSG

Copernicus Sentinel-2

Emulation

Model inversion

## ABSTRACT

Current remote sensing applications of bio-geophysical parameter estimation and image classification have to deal with an unprecedented big amount of heterogeneous and complex data sources. New satellite sensors involving a high number of improved time, space and wavelength resolutions give rise to challenging computational problems. Standard physical inversion techniques cannot cope efficiently with this new scenario. Dealing with land cover classification of the new image sources has also turned to be a complex problem requiring large amount of memory and processing time. In order to cope with these problems, statistical learning has greatly helped in the last years to develop statistical retrieval and classification models that can ingest large amounts of Earth observation data. Kernel methods constitute a family of powerful machine learning algorithms, which have found wide use in remote sensing and geosciences. However, kernel methods are still not widely adopted because of the high computational cost when dealing with large scale problems, such as the inversion of radiative transfer models or the classification of high spatial-spectral-temporal resolution data. This paper introduces to the remote sensing community an efficient kernel method for fast statistical retrieval of atmospheric and biophysical parameters and image classification problems. We rely on a recently presented approximation to shift-invariant kernels using projections on random Fourier features. The method proposes an explicit mapping function defined through a set of projections randomly sampled from the Fourier domain. It is proved to approximate the implicit mapping of a kernel function. This allows to deal with large-scale data but taking advantage of kernel methods. The method is simple, computationally very efficient in both memory and processing costs, and easily parallelizable. We show that kernel regression and classification is now possible for datasets with millions of samples. Examples on atmospheric parameter retrieval from hyperspectral infrared sounders like IASI/Metop; large scale emulation and inversion of the familiar PROSAIL radiative transfer model on Sentinel-2 data; and the identification of clouds over landmarks in time series of MSG/Seviri images show the efficiency and effectiveness of the proposed technique.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

## 1.1. Remote sensing and the big data challenge

Earth-observation (EO) satellites provide a unique source of information to address some of the challenges of the Earth system science (Berger et al., 2012). EO deals with the important objective of monitoring and modelling the processes on the Earth surface and their interaction with the atmosphere. To accomplish this ambitious goal, EO deploys both data acquired by remote sensing airborne and

satellite sensors, as well as quantitative *in situ* measurements of biophysical parameters (Camps-Valls et al., 2011). Predictive models of biophysical parameters and classification of remotely sensed images are thus relevant outputs to monitor our Planet.

In this context, current EO applications of image classification and biophysical parameter estimation, have to deal with an unprecedented big amount of heterogeneous and complex data sources. Spatio-temporally explicit quantitative methods are a requirement in a variety of Earth system data processing applications. Optical Earth observing satellites for example, endowed with a high temporal resolution, enable the retrieval and hence monitoring of climate and bio-geophysical variables (Dorigo et al., 2007; Schaepman et al., 2009). The super-spectral Copernicus Sentinel-2 (S2) (Drusch et al., 2012) and the forthcoming Sentinel-3 mission (Donlon et al., 2012), as well as the planned EnMAP (Stuffer et al., 2007),

\* Corresponding author.

E-mail address: [adrian.perez@uv.es](mailto:adrian.perez@uv.es) (A. Pérez-Suay).

HyspIRI (Roberts et al., 2012), PRISMA (Labate et al., 2009) and FLEX (Kraft et al., 2013), will soon provide unprecedented data streams. Very high resolution (VHR) sensors like Quickbird, Worldview-2 and the recent Worldview-3 (Longbotham et al., 2014) also pose big challenges to data processing. The challenge is not only attached to optical sensors. Infrared sounders, like the Infrared Atmospheric Sounding Interferometer (IASI) (Tournier et al., 2002) sensor on board the Metop satellite series, impose even larger constraints: the orbit time of Metop satellites (101 min), the large spectral resolution (8461 spectral channels between  $645\text{ cm}^{-1}$  and  $2760\text{ cm}^{-1}$ ), and the spatial resolution ( $60 \times 1530$  samples) of the IASI instrument yield several hundreds of gigabytes of data to be processed daily. EO radar images also increased in resolution, and the current platforms, such as ERS-1/2, ENVISAT, RadarSAT-1/2, TerraSAR-X, and Cosmo-SkyMED give rise to extremely fine resolution data that call for advanced scalable processing methods. Besides, we should not forget the availability of the extremely large remote sensing data archives already collected by several past missions, such as ENVISAT, Seviri/MSG, Cosmo-SkyMED, Landsat, or SPOT.

These large-scale data problems require enhanced processing techniques that should be accurate, robust and fast. Standard physical inversion techniques and parametric classification algorithms cannot cope (nor adapt to) this new scenario efficiently. Over the last few decades a wide diversity of methods have been developed to tackle particular EO processing tasks, but only a few of them made it into operational processing chains, and many of them are only in its infancy.

### 1.2. Machine learning for Earth observation data analysis

In order to cope with these problems, statistical learning (also known as machine learning) has greatly helped in the last years to develop statistical retrieval and classification models that can ingest large amount of Earth observation data. Machine learning has become a standard paradigm for the analysis of remote sensing and geoscience data, at both local and global scales (Camps-Valls et al., 2011). Machine learning actually constitute a relevant alternative to parametric and physically-based models, which rely on established physical relations and implement complex combinations of scientific hypotheses, and give rise to too rigid solutions and eventual model discrepancies (see Berger et al., 2012 and references therein).

Alternatively, the framework of *statistical inference and machine learning* is concerned about developing *data-driven models* and they solely rely on the “unreasonable effectiveness of data” (Halevy et al., 2009). The field has proven successful in many disciplines of Science and Engineering (Hastie et al., 2009) and, in general, nonlinear and nonparametric model instantiations typically lead to more flexible and improved performance over physically-based approximations.

In the last decade, machine learning has attained outstanding results in the estimation of climate variables and related bio-geophysical parameters at local and global scales, and on the classification of remote sensing images (Camps-Valls et al., 2011). Current operational vegetation products, like leaf area index (LAI), are typically produced with neural networks (Bacour et al., 2006; Baret et al., 2013; Duveiller et al., 2011). Gross Primary Production (GPP) as the largest global  $\text{CO}_2$  flux is estimated using ensembles of random forests and neural networks (Beer et al., 2010; Jung et al., 2011). Similarly, the contribution of supervised classifiers has been improving the efficacy of the land cover/use mapping methods since the 1970s: Gaussian models such as Linear Discriminant Analysis (LDA) were replaced in the 1990s by non-parametric models able to fit the distribution observed in data of increasing dimensionality, which were later superseded by decision trees (Friedl and Brodley, 1997; Hansen et al., 1996) and then by neural networks (NN, Bischof and Leona, 1998, Bischof et al., 1992, Bruzzone and Fernández Prieto, 1999).

### 1.3. Kernel machines and random features for efficient EO data processing

The last decade kernel methods emerged as a family of powerful machine learning algorithms, and found wide use in remote sensing and geosciences (Camps-Valls and Bruzzone, 2009; Camps-Valls et al., 2011). In the last decade, a kernel method called support vector machines (SVM, Camps-Valls and Bruzzone, 2005, Camps-Valls et al., 2004, Foody and Mathur, 2004, Huang et al., 2002, Melgani and Bruzzone, 2004) was gradually introduced in the field, and quickly became a standard for image classification. Further SVM developments considered the simultaneous integration of spatial, spectral and temporal information (Benediktsson et al., 2005; Camps-Valls et al., 2008; Fauvel et al., 2008; Pacifici et al., 2009; Tuia et al., 2009), the richness of hyperspectral imagery (Camps-Valls and Bruzzone, 2005; Plaza et al., 2009), and exploiting the power of clusters of computers (Muñoz-Marí et al., 2009; Plaza et al., 2008). We observed a similar adoption of kernel machines for biophysical parameter retrieval: support vector regression showed high efficiency in modelling LAI, fCOVER and evapotranspiration (Durbha et al., 2007; Yang et al., 2006), and kernel methods like Gaussian Processes (GPs) (Rasmussen and Williams, 2006) recently provided excellent results in retrieving vegetation parameters (Camps-Valls et al., 2016; Lázaro-Gredilla et al., 2014; Pasolli et al., 2010; Verrelst et al., 2013a, 2012, 2013b).

However, kernel methods are still not widely adopted because of the high computational cost when dealing with large scale problems, such as the inversion of radiative transfer models or the classification of high spatial-spectral-temporal resolution data. Roughly speaking, given  $n$  examples available to develop the models, kernel methods typically need to store in memory *kernel matrices*  $\mathbf{K}$  of size  $n \times n$  and to process them using standard linear algebra tools (matrix inversion, factorization, eigendecomposition, etc.). This is an important constraint that hamper its applicability to large scale EO data processing.

In this paper, we introduce a kernel method for efficiently approximate kernels, that make nonlinear classification possibly with millions of examples. We will focus on the two most relevant EO data problems: statistical retrieval of bio-geo-physical parameters and image classification problems. The method proposes an explicit mapping function defined through a set of projections randomly sampled from the Fourier domain. It is proved to approximate the implicit mapping of a kernel function,  $\mathbf{K}$ . This allows to deal with large-scale data but taking advantage of kernel methods. The method is simple, computationally very efficient in both memory and processing costs, and easily parallelizable through standard divide-and-conquer strategies as the ones proposed in Zhang et al. (2013).

The contributions of this paper are as follows: (1) the introduction to the remote sensing community of this new efficient method to perform nonlinear regression and classification; (2) the extension of the method to work with other than Fourier bases, such as wavelets, stumps, and Walsh expansions that can cope with other data characteristics; and (3) to give experimental evidences in several illustrative and challenging real problems in EO data processing: atmospheric and biophysical parameter retrieval, model inversion and emulation of radiative transfer models, and remote sensing image classification using data from multispectral data. In particular, we will show that kernel regression/classification is now possible in applications involving datasets with millions of examples and high dimensionality. The efficiency, accuracy and effectiveness of the technique is illustrated in atmospheric parameter retrieval from hyperspectral infrared sounders like IASI, large scale emulation and inversion of the familiar PROSAIL radiative transfer model on Sentinel-2 data, and the identification of clouds over landmarks in time series of MSG/Seviri images. In addition, we will show that the method is very simple to implement, computationally very efficient

Download English Version:

<https://daneshyari.com/en/article/8866961>

Download Persian Version:

<https://daneshyari.com/article/8866961>

[Daneshyari.com](https://daneshyari.com)