ISPRS Journal of Photogrammetry and Remote Sensing 107 (2015) 50–63

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis

Michele Volpi [a,*], Gustau Camps-Valls [b], Devis Tuia [c]

[a] CALVIN, Institute of Perception, Action and Behaviour, The University of Edinburgh, EH8 9AB Edinburgh, United Kingdom
[b] Image Processing Laboratory, Universitat de València, C/ Catedrático José Beltran, 2., E-46980 Paterna, València, Spain
[c] Multimodal Remote Sensing Group, Department of Geography, University of Zurich, 8057 Zurich-Irchel, Switzerland

ARTICLE INFO

Article history:
Received 18 June 2014
Received in revised form 11 November 2014
Accepted 11 February 2015
Available online 18 March 2015

Keywords:
Change detection
Cross-sensor
Relative spectral alignment
Kernel methods
Feature extraction
Canonical correlation analysis

ABSTRACT

In this paper we present an approach to perform relative spectral alignment between optical cross-sensor acquisitions. The proposed method aims at projecting the images from two different and possibly disjoint input spaces into a common latent space, in which standard change detection algorithms can be applied. The system relies on the regularized kernel canonical correlation analysis transformation (kCCA), which can accommodate nonlinear dependencies between pixels by means of kernel functions. To learn the projections, the method employs a subset of samples belonging to the unchanged areas or to uninteresting radiometric differences. Since the availability of ground truth information to perform model selection is limited, we propose a completely automatic strategy to select the hyperparameters of the system as well as the dimensionality of the transformed (latent) space. The proposed scheme is fully automatic and allows the use of any change detection algorithm in the transformed latent space. A synthetic problem built from real images and a case study involving a real cross-sensor change detection problem illustrate the capabilities of the proposed method. Results show that the proposed system outperforms the linear baseline and provides accuracies close the ones obtained with a fully supervised strategy. We provide a MATLAB implementation of the proposed method as well as the real cross-sensor data we prepared and employed at https://sites.google.com/site/michelevolpiresearch/codes/cross-sensor.
© 2015 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

The amount of Earth observation data available to study the Planet's surface is destined to grow exponentially, as dictated by the "big data" trends. Recent and forthcoming missions from both national and private agencies will further increase the amount of available images from remote sensing instruments. Consequently, the evolution in technologies and increased accessibility to data makes remote sensing a viable tool for many environmental studies. Moreover, image databases and derived products offer a variety of information sources to study the temporal evolution of the dynamical process occurring at the Earth's surface. However, versatile and generic methods able to extract relevant information automatically are needed. The ability of processing and integrating multi-temporal acquisitions from multiple sensors into standard remote sensing systems is of paramount importance for three main reasons. Firstly, integrating such series accommodates data complementarity: fused products extracted from complementary data possess an increased information content (Pohl and Van Genderen, 1998). Secondly, the ability of exploiting images from heterogeneous sources in multi-temporal scenarios drastically reduces constraints such as equal spectral resolution on same sensed wavelengths, possibly increasing the temporal resolution of such studies. Thirdly, response time of image processing systems, such as those required for disaster management and post-catastrophe assessment may be further reduced by allowing for integration of arbitrary multi-sensor and cross-sensor images (Roemer et al., 2010; Chatelain et al., 2008).

In this paper we distinguish *multi-sensor* from *cross-sensor*. The former corresponds to two time series of acquisitions from different sensors with single pairs of images acquired within a short time interval. The latter denotes a single time series composed of acquisitions from multiple sensors. In this paper, we consider the *cross-sensor* setting. When dealing with time series composed of

* Corresponding author. Tel.: +44 (0) 131 650 2733; fax: +44 (0)131 651 3435.
E-mail addresses: michele.volpi@ed.ac.uk (M. Volpi), gcamps@uv.es (G. Camps-Valls), devis.tuia@epfl.ch (D. Tuia).
URLs: https://sites.google.com/site/michelevolpiresearch/ (M. Volpi), http://isp.uv.es (G. Camps-Valls).

http://dx.doi.org/10.1016/j.isprsjprs.2015.02.005
0924-2716/© 2015 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

single-sensor images, it is common to observe gaps in the spatial or in the temporal sampling domains (for example due to clouds or sensor failures) (Gómez-Chova et al., 2013; Longbotham et al., 2012; Villa et al., 2013). The same issue affects also multi-sensor time series, but in this case images from the complementary sensor may be used to fill in the gap observed in the other time series (Röder et al., 2005; Desclée et al., 2013; Amorós-López et al., 2011a). However, to model the time series one has to manually combine multiple sensors acquisitions or develop specific strategies of data fusion (Schmid et al., 2005). The systematic exploitation of optical cross-sensor data in real world systems is still limited by the complexity and lack of universality of current data fusion methods. Methods such as the STARFM (Gao et al., 2006) or the one by Amorós-López et al. (2011a) allow to fuse images from different sensors and to retrieve both spectrally and spatially enhanced images. However, these methods have been developed for precise pairs of sensors with specific characteristics. The change detection tasks may be greatly simplified by generic and possibly automatic cross-sensor image integration systems.

A question of interest when dealing with cross-sensor change detection problems is how to deal with differences in the spectral domain. If spatial resolutions may be easily adjusted by gridding and resampling, detecting spectral changes in inhomogeneous domains is more delicate. Most of the times, the common spectral channels across acquisitions are manually selected so that similar wavelength intervals are used for bandwise comparison (Wulder et al., 2008). However, if sensors provide *disjoint* representations of the spectrum, such a manual approach may not be applicable. A general framework to cross-sensor spectral alignment for change detection is still missing.

Standard approaches to perform cross-sensor change detection require pixel-level labels representing either the spectral classes or the type of transitions occurring across images. Post-classification comparison is usually employed in the former case (Singh, 1989; Mubea and Menz, 2012), while for the latter direct multi-date classification is often preferred (Singh, 1989; Turker and Asik, 2005; Qin et al., 2013; Volpi et al., 2013b). In the first case, thematic classification maps obtained independently for each acquisition are compared and the segmented multi-temporal information is summarized in a single map layer. However, this system does not account for temporal dependencies and the final results strongly depend on the quality of single *independent* classifications. In the second case, image channels are simply stacked and a single thematic classification is performed on the entire multi-temporal set, implicitly accounting for temporal dependencies. Compared to post-classification approaches it tends to provide more precise change detection maps, since not affected by the propagation of the errors of single maps. However, this procedure requires labeled multi-temporal signals describing all the transitions of interest, which are rarely available. Also, dealing with a high dimensional problem with few training samples will lead to the curse of dimensionality and robust supervised learners must be employed. This may occur easily when dealing with hyperspectral data or adding spatial-contextual descriptors in direct multi-date scenarios (Volpi et al., 2013b).

There are few techniques able to perform general cross-sensor change detection which not rely on the limiting assumption of having an exhaustive training set. Parametric models such as the one developed by Fernàndez-Prieto and Marconcini (2011) perform selective change detection on a subset of changes described by labels. Since relying on joint probabilities modeled independently on the images, the method is independent of the spectral characteristics of the image. Alberga (2009) presented a method which exploits spectrally invariant measures to depict differences in the spatial arrangement of pixels between the multi-temporal acquisitions. The measures proposed consider only the relative local illumination differences and consequently the dimensionality of the spectral domain does not constraint the problem. This may be seen as comparing texture from local patches working on general single channel images. The last family of cross-sensor methods rely on joint transformations of the spectral domain of the images. These techniques perform a relative radiometric alignment of images by training a multi-output regression model to predict all the pixels of the post-event image on the basis of a subset of pairs of unchanged pixels across times (Yang and Lo, 2000; Healey et al., 2006). For change detection purposes, the residual image (difference between actual and estimated pixel values) are compared as if employing standard difference image analysis, e.g. change vector analysis (CVA) (Bovolo et al., 2011). A similar reasoning is behind the multivariate alteration detection (MAD), which relies on the canonical correlation analysis (CCA) transformation (Nielsen et al., 1998; Nielsen, 2002, 2007; Canty, 2010; de Carvalho Júnior et al., 2013). CCA is a linear rotation-based method that aims at finding an optimal linear combination of the two disjoint groups of features maximizing correlation (Hotelling, 1936). In this case, instead of fitting a least-squares regression on unchanged samples, MAD-based approaches transforms the pixels so that their projection maximizes the correlation. The multivariate difference of the projected data is used as a change indicator in standard change detection routines. Note that the multiple output regression schemes and CCA are intimately related, as illustrated in de la Torre (2012).

This latter family of methods estimates a rotation from the spectral channels directly (also known as the 'primal formulation'). Despite the simplicity and generating light computational efforts, this assumption of linearity between spectral channels may be not sufficient to properly model cross-sensor data. The pair of acquisitions may show differences due to complex light interactions, seasonality, local atmospheric conditions, illumination changes and other possibly nonlinear spectral transformations affecting unevenly the image in both spatial and spectral domains but not representing ground cover changes (Tuia et al., 2014; Gómez-Chova et al., 2013). Nonlinear learning methods have been deeply studied for remote sensing applications, in particular for pixel classification tasks. Classical examples of employed classification models are support vector machines and kernel methods (Gómez-Chova et al., 2008; Maulik and Chakraborty, 2013) and neural networks (Li et al., 2014; Shao and Lunetta, 2012; Amorós-López et al., 2011b). In this paper, we focus on kernel methods, since offering sound theoretical guarantees and flexible solutions (Camps-Valls & Bruzzone, 2009). Moreover, kernel methods are well suited for dealing with heterogeneous sources of information (Camps-Valls et al., 2008; Tuia et al., 2010), feature extraction and dimensionality reduction (Arenas-García et al., 2013), regression and function approximation (Camps-Valls & Bruzzone, 2009) and change detection (Bovolo et al., 2008; Volpi et al., 2013b). The main intuition behind kernel methods is that a nonlinear problem in the original input space may be transformed into a linear one by recasting it into a higher dimensional space. To achieve this, one usually has to find a mapping function of the original data samples to a space in which a selected linear method works. However, the optimal mapping function is not known in advance and the proper estimation directly from the data may be difficult and computationally unfeasible. To alleviate this issue kernel methods theory shows that a dot product between samples mapped into a higher dimensional space may be replaced by a valid kernel function, only taking as argument the samples in their input space. For more details on SVM and kernel methods, we refer to (Shawe-Taylor and Cristianini, 2004; Schölkopf and Smola, 2002).

In this paper, we propose an approach relying on the kernel extension of the canonical correlation analysis, the kernel CCA (kCCA), to perform relative spectral alignment of cross-sensor acquisitions. Kernel CCA is a nonlinear variant of the CCA, which aims at computing a projection of the samples maximizing the

Download English Version:

https://daneshyari.com/en/article/557179

Download Persian Version:

https://daneshyari.com/article/557179

Daneshyari.com