KERNEL METHODS IN ORTHOGONALIZATION OF HYPERSPECTRAL DATA

Allan A. Nielsen

Technical University of Denmark
DTU Space – National Space Institute
Presently located at DTU Informatics – Informatics and Mathematical Modelling
Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark
Tel +45 4525 3425, Fax +45 4588 1397
E-mail aa@space.dtu.dk, http://www.imm.dtu.dk/∼aa

ABSTRACT

Principal component analysis (PCA) is the mother of all linear orthogonal transformations for data compression and dimensionality reduction of correlated multivariate data. This contribution describes a kernel version of PCA and it also sketches kernel versions of maximum autocorrelation factor (MAF) analysis and minimum noise fraction (MNF) analysis. In kernel methods the actual observations are replaced by high (in principal infinite) dimensional nonlinear mappings of the observations. In the dual formulation or Q-mode analysis applied in the kernel versions, these mappings are implicitly defined by a kernel function. This type of analysis allows for nonlinearities in the original data, and especially the kernel MAF/MNF transformations tend to focus on outlying observations on a slower changing background. Data examples include analysis of simple differences of multi- and hyperspectral Earth observation image data, hyperspectral NIR images for food quality control, and irregularly spaced geochemical data for geological mapping. The example shown here illustrates the successful application of kernel PCA and especially kernel MAF analysis to change detection in HyMap data covering a small agricultural area near Lake Waging-Taching, Bavaria, in Southern Germany.

Figure 1 shows HyMap bands 27 (828 nm), 81 (1,648 nm) and 16 (662 nm) as RGB, 30 June 2003 8:43 UTC (left) and 4 August 2003 10:23 UTC (right). The images consist of 400 by 270 (5 m by 5 m) pixels.

In the change detection analysis all 126 spectral bands of the HyMap are included. Figure 2 shows kernel principal components 1-3 (left) and kernel maximum autocorrelation factors 1-3 (right) of simple band-by-band difference images as RGB. All bands are stretched linearly between mean minus and plus three standard deviations. In this representation no-change areas will appear as grayish and change regions will appear in saturated colours. Change on the ground is due to growth of the main crop types such as maize, barley and wheat. On pastures, which are constantly being grazed, in forest stands and in the lake to the south, no change is observed. Furthermore, solar effects cause edge effects where height differences occur (both solar elevation and azimuth have changed). We see that both types of kernel analysis emphasize change and that unlike kernel PCA, kernel MAF analysis seems to focus on the most conspicuous changes and that it gives a strong discrimination between change and no-change regions.

Ordinary linear PCA or MAF analysis (not shown) does not give the beautiful discrimination between change and no-change regions offered by kernel MAF analysis.

The HyMap data are kindly provided by Andreas Müller and co-workers, DLR German Aerospace Center, Oberpfaffenhofen, Germany. Thanks to both Andreas Müller and Mort Canty, Research Center Jülich, Germany, for many years of interesting cooperation on the analysis of multi- and hyperspectral image data.

A few good references in this kernel context are listed below.[1–6]

REFERENCES

1. A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, 1988.
2. B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.
3. J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, 2004.
4. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
5. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, third edition, Cambridge University Press, 2007.
6. A. A. Nielsen and M. J. Canty, "Kernel principal component analysis for change detection," SPIE Europe Remote Sensing Conference, Cardiff, Great Britain, 15-18 September 2008, Internet http://www.imm.dtu.dk/pubdb/p.php?5667.

Figure 1. HyMap bands 27 (828 nm), 81 (1,648 nm) and 16 (662 nm) as RGB, 30 June 2003 8:43 UTC (left) and 4 August 2003 10:23 UTC (right).

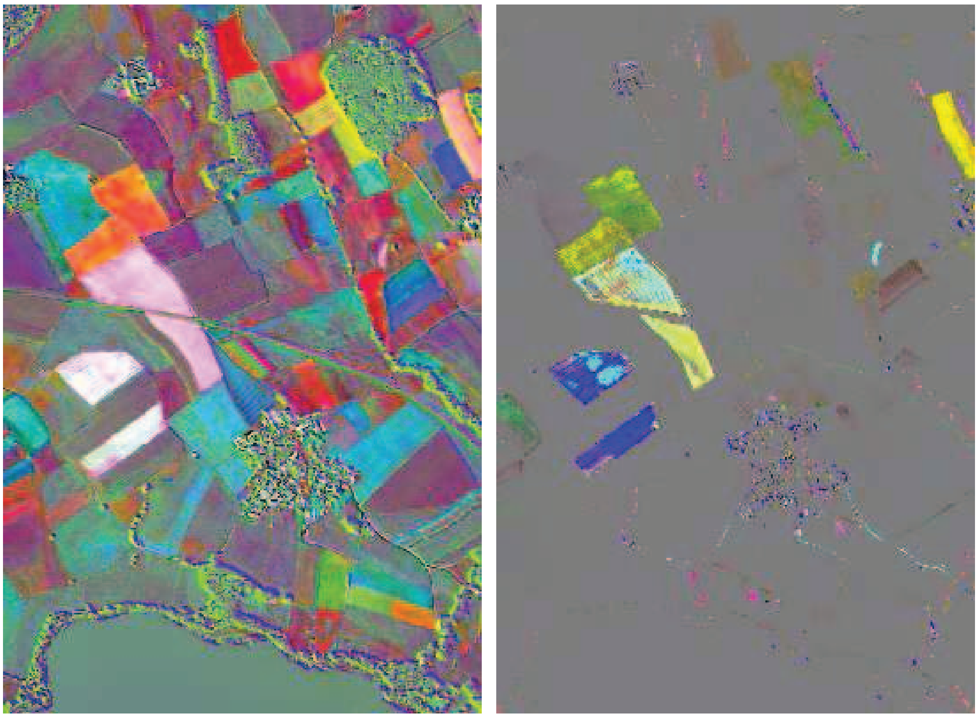

Figure 2. Kernel principal components 1-3 (left) and kernel maximum autocorrelation factors 1-3 (right) of 126 simple difference images as RGB. All bands are stretched linearly between mean minus and plus three standard deviations.