ERCIM "Alain Bensoussan" Fellowship Scientific Report

Fellow: Mónica Benito Bonito
Visited Location : INRIA
Duration of Visit: 8 months

I - Scientific activity (1 page at maximum)

The research carried out during this training program has been focussed on:

1) Dimensionality reduction, reconstruction and classification for hyperspectral image data, in Colaboration with the Laboratoire Planétologie de Grenoble.

2) Variable selection and dimensionality reduction for dealing with complex data.

Dimensionality reduction, reconstruction and classification for hyperspectral image data.

Multispectral methods for deriving information about the Earth's resources is an important topic for its applicability in many disciplines. Since advances in sensor technology have made possible the collection of such image data with several hundred spectral bands, conventional image processing methods would not be appropriate when dealing with such high dimensional data. The motivation for this work is to take advantage of these spectral and spatial variation to derive useful information in a high dimensional setting.

The matter of how spectral variations are represented mathematically and conceptually is an important first step in defining how the extraction of the desired information should proceed. The way in which multispectral data is represented quantitatively is displaying pixels as points in a N-dimensional space (as much as spectral bands). We consider factorial models (Kroonemberg, 1983) to reduce the dimensionality of the spectral bands to a lower dimensionality space. Moreover data compression, a second goal is feature extraction for classification procedures. The goal of an optimum feature extraction method for classification applications is not only to reduce data dimensionality for reducing computational costs, but also improve the efficiency of classification by extracting features that maximize the separation between spectrum classes. Factorial models are successfully applied for hyperspectral image reconstruction and perform better than traditional methods as Principal Component Analysis (PCA).

For classification procedures, K-means algorithms may be implemented as an alternative to that supervised techniques based on the construction of binary decision trees (Bittencourt and Clarke, 2004).

Variable selection and dimensionality reduction when dealing with complex data

This work bridges the gap between variable selection methods and dimensionality reduction algorithms when considering a sample of complex data, as face images.

Dimensionality reduction algorithms do not, in general, benefit from using informative, uncorrelated features. On the other hand, while dimensionality reduction algorithms do well on sets of correlated features, variable selection methods performs poorly. They fail to pick relevant variables, because the score they assign to correlated features is too similar, and none of the variables is strongly preferred over another. Hence, variable selection and dimensionality reduction algorithms have complementary advantages and disadvantages. The proposed work combines both methodologies. The combined optimization function searches for the variable selection that maximizes the information content of the data's most informative directions (Benito and Peña, 2005).

II- Publication(s) during your fellowship

Please insert the title(s), author(s) and abstract(s) of the published paper(s). You may also mention the paper(s) which were prepared during your fellowship period and are under reviewing.

Title: "Detecting defects with image data"
Authors: Mónica Benito and Daniel Peña
Submitted to *Computational Statistics and Data Analysis.*

Abstract. Quality control using continuous monitoring from images is emerging as an active research area. These applications require of adaptive statistical techniques in order to detect and isolate process abnormalities. In this work we consider monitoring schemes for images which takes into account both the spatial dependency and the changes in local variability and show its good performance in quality process monitoring.

Key words: Bootstrap, Spatial variability, Variance ratio, Uniform gray-scales, Variogram.

(this article is continuation of my research during my doctoral thesis)

Title: "Variable selection and dimensionality reduction for image data: an unified approach"
Authors: Mónica Benito
(*In progess*)

Abstract. Variable selection plays an important role in image analysis and recognition, where the essential information distributed over a large number of pixels need to be captured by a subset of significant variables (pixels). This scheme becomes hard to do when considering a sample of similar objects and we want to use the overall information. The method we propose extract the most important variables using the most informative directions in the sample given by dimensionality reduction procedures.

Key words: Backward elimination, Unsupervised feature selection, Singular Value Decomposition, Most informative directions.

III -Attended Seminars, Workshops, and Conferences

Please identify the name(s), date(s) and place(s) of the events in which you participated during your fellowship period.

Attended Course

Intensive course in French Language - University of Grenoble, 2nd May-2nd June.

Attended Seminars

"Introduction aux ondelettes" – Paulo Goncalves, INRIA Rhone-Alpes.
28th February and 1st March.

" Real Time Human-Computer Interaction Systems for Users with Disabilities" – Benjamin Waber. INRIA Rhone-Alpes, 30th June

"Blind Vision" – Shai Avidan. INRIA Rhone-Alpes, 17th July.

"Non-rigid 3D shape matching" – Kiran Varanasi. INRIA Rhone-Alpes, 31st August.

"Stereo matching under image ambiguity and appearance changes"- Kuk-Jin Yoon.
INRIA Rhone-Alpes, 7th September.

Attended Workshops

Workshop in " Analyse statistique et physique d'images hyperspectrales planétaires"
Laboratoire de Planetologie de Grenoble, CNRS, 31st May 2006

Presentation: " Réduction de dimensionalité et classification d'images hyperspectrales"

Imparted Seminars

" Dimensionality reduction with image data" – 17th February. INRIA