Statistical approaches to signal and image processing in astronomy

Jean-François Cardoso

French-Vietnamese Summer School Mathematics Of Image Processing
Institute of Mathematics. Hanoi. Sept 10-18, 2004

About these lectures

The aim of these lectures is to introduce some simple or not so simple *statistical* methods of image processing. We will do so by following a common thread: their development in the context of an exciting mega-application: processing data from a space-based astrophysical mission dedicated to constructing the *oldest image in the Universe*.

We start with the astrophysical context: the Big Bang theory of the formation of the Universe and the data processing tool-chain as it is currently planned: from satellite observations, to mapping the (very) early Universe, to measuring its age via spectral estimation.

We explain some of the statistical techniques involved in data processing, focusing on some specific aspects which can be categorized as follows

- Exploiting statistical coherence
- Exploiting non Gaussianity
- Exploiting independence

In this process, we hope to offer insights about some key ideas of statistical inference.

1 Cosmology and data analysis

1.1 Latest news from the sky

We will start by giving some background about modern cosmology and why statistical image processing techniques are relevant to it.

- The standard cosmological model: the Big Bang, an expanding Universe.
- The cosmic microwave background (CMB) is a fossil light, permeating the Universe, once in thermal equilibrium with matter, released (uncoupled from matter) at an early stage in the formation of the Universe. The CMB is one of the pillars of the Big Bang model. It provides us with a "snapshot" of the Universe in its infancy.
- The CMB is a gold mine for cosmologists; its study reveals many important features of our Universe on cosmological scales. It is the subject of many , ground-based, airborne or spatial experiments. Recently, NASA has launched the W-MAP satellite which has already sent one-year worth of high quality data; the European Spatial Agency (ESA) will launch an even more powerful satellite in 2007. These experiments and many others are yielding tremendous amounts of data which have to be processed in new and interesting ways to deliver their scientific content.
- The most direct way in which CMB data will be useful is along the following lines: the sky is scanned in (redundant) circles by the instrument; an image of the apparent temperature of the cosmic radiation is formed (map making); the power spectrum of this image is estimated; cosmological parameters (such as the Hubble constant, the density of matter in the Universe, etc...) are estimated by fitting the empirical spectrum to theoretical cosmological models.

1.2 Some issues in CMB data analysis

We have just roughly outlined a connection between cosmology and image processing and ultimately spectrum estimation. The program sounds simple enough in principle but many challenges are faced by the people in charge of extracting science from data. Some of the relevant (to these lectures) are the following.

- **Multiple detectors, multiple channels**. The cosmic radiation is measured by arrays of detectors (bolometers) operating at various frequencies (or channels). These various channels offer different 'points of view' (so to speak) and this redundancy is an important feature of modern experiments. However, how best to combine the images acquired by the channels of a multi- or hyper-spectral data sets is an open question. We will examine how these channels can be used to combat noise.

- **Background and foregrounds, component separation**. The measurements of the cosmic background radiation are not only severely corrupted by noise. They also suffer from contamination by various other emissions, globally referred to as *foregrounds* (as opposed to the cosmic background). These foreground emissions originate from the cosmos, from our own galaxy or even from the instrument itself. All these components are present in all channels. We will examine how they can be disentangled by resorting to independent component analysis.

- **Working on the sphere; the harmonic spectrum**. One of the unusual aspects of image processing data with CMB data is that data are collected data from the whole celestial sphere, that a spherical image is to be estimated, that a spherical harmonic spectrum is to be estimated. Hence, spherical geometry is central to CMB studies. This raises many interesting questions such as how to adapt the familiar Euclidean concepts and methods of image processing to the sphere. These lectures will address some of these points. In particular, we shall introduce spherical harmonics and related ideas.

- **Point source detection and extraction**. The distribution of the CMB map follows very closely a Gaussian model but most of the foregrounds do not appear to be Gaussian distributed at all. In particular, the CMB map must be contaminated by a large number of point-like sources which are in fact images of remote galaxy clusters. Detecting and extracting these point sources is a very important task in two respects. First, the point sources have a significant impact on the spectrum of the CMB at high frequencies, hence ignoring point sources lead to overestimation of the CMB power at small scales. Second, the point sources also have a cosmological interest in their own right: they provide us with a sample of galaxy clusters large enough to perform statistical studies. The success of any procedure depends on taking into account the non Gaussian nature of the object to be detected. We shall examine some statistical aspects of this problem.

2 Exploiting coherence

This section examines how to take advantage of spatial and spectral coherence to build optimal linear filters for processing noisy images.

2.1 General Setting

Noisy observations.

Filters and linear filters.

Objective function: minimum mean-square error.

Wiener filter and conditional expectation.

Maximum likelihood estimation.

2.2 Exploiting spectral coherence

Model of Gaussian stationary random fields.

The Whittle approximation.

Likelihood and spectral matching.

Discussion: optimality and unbiasedness.

2.3 Exploiting spatial coherence

Multivariate observations (several channels for instance).

Conditional expectation.

Direct minimization of the mean square error.

Connection with spectral coherence.

Signal subspace, spatial whitening, projection.

Connection with the case of stationary processes.

Minimum variance beamformer. Beamforming.

Processing without statistical assumption on the image of interest.

Minimizing variance under a linear constraint.

Relative efficiency w.r.t. optimal (statistical) filter.

2.4 Optimal linear filter for mapping the cosmic background

Combining the above two filters to play with spatial and spectral diversity.

From 1D sequences to filtering images in harmonic space.

3 Exploiting non Gaussianity

Taking advantage of non Gaussian features in the distribution of images. Going beyond second order (correlation) descriptors.

3.1 Maximum likelihood and MAP estimates.

Score function for non Gaussian distributions.

Estimating equations.

Convexity.

Soft thresholding.

Asymptotic analysis. Gaussian signals are worst case.

3.2 Detection of point sources

Maximum entropy and maximum likelihood

Application to detection and estimation of galaxy clusters in the cosmic background.

4 Exploiting statistical independence

Images may be corrupted by noise or by other images. This section deals with (blind) component separation: multi-channel images may be modeled as linear superpositions (mixtures) of independent components. By expressing strongly enough the assumption of statistical independence, it is possible to retrieve (disentangle) all underlying components.

- Finding independent components in images.
- Likelihood for stationary random fields.
- Whittle approximation and spectral matching.
- The EM algorithm for maximizing the likelihood.

A Prerequisites

- Basic univariate statistics
- Basic (2nd order) multivariate statistics.
- Basic linear algebra (matrices, eigen-vectors, pseudo-inverse)
- Basic optimization (Lagrange multiplier, gradient descent)