

A Wavelet Parallel Code for Structure Detection

A. Pagliaro

Istituto di Astronomia dell'Università di Catania

U. Becciani, V. Antonuccio and M. Gambera

Osservatorio Astrofisico di Catania

Abstract. We describe a parallel code for 3-D structure detection and morphological analysis. The method is based on a multiscale technique, the wavelet transform and on segmentation analysis. The wavelet transform allows us to find substructures at different scales and the segmentation method allows us to make a quantitative analysis of them. The code is based on the message passing programming paradigm and major speed improvements are achieved by using a strategy of domain decomposition.

1. Introduction

This paper presents a new parallel code, that allows a rapid structure detection and morphological analysis in a 3-D set of data points. The code is described in greater detail in Becciani & Pagliaro, 1997. A serial version of this code has been successfully used in the analysis of the Coma cluster (Gambera et al., 1997). However, possible applications of this code are not limited to astrophysics but may also benefit several other fields of science.

Our method of structure detection is based on the wavelet transform (see Grossmann & Morlet 1984, 1987) evaluated at several scales and on segmentation analysis (see also Escalera & Mazure 1992, Lega 1994, Lega et al. 1995).

2. Method Overview

The detection method can be divided into three main steps.

1. Computation of the wavelet matrices on all the scales investigated. These are computed, by means of the “à trous” algorithm, for the data to be analyzed and on a random distribution in the same region of space and on the same grid as the real data. On these latter matrices we calculate the threshold corresponding to a fixed confidence level in the structure detection.

2. Segmentation analysis. The aim of this analysis is to have all the connected pixels with a wavelet coefficients greater than the threshold labeled with an integer number different for every single structure.

3. Computation of a morphological parameter for every structure singled out and of a mean morphological parameter for each scale.

3. The Implementation

Our strategy has been to develop a parallel code that can run both on multiprocessors or MPP systems and on clusters of workstations; the latter are easily available at low cost without a specific investment in supercomputing. Hence, we have adopted a **message passing technique** and subdivided the computational domain into subdomains, assigning the wavelet and segmentation analysis of the subdomain to different processors: each processor executes in a parallel way the analysis with memory usage and computational load decreasing as the number of working processors grow. For the development of our code we have choose to use **PVM** (Parallel Virtual Machine), a set of C and FORTRAN function written by the OACK Ridge National Laboratory.

4. The Parallel Code

The computational domain is made of the region of space that comprises all the data points from the simulation or the catalogue to be analyzed. Recognition of the substructures happens by means of the wavelet coefficient computation and by the segmentation analysis. Our code is based on the programming paradigm **MASTER/SLAVE**. The computational domain is subdivided into M subdomains (proportionally to the weights of the hosts), along a split axis chosen as the longest one. Each subdomain is assigned to a host of the virtual machine. On the wave slaves both a fault tolerance mechanism and a dynamical load balance one have been implemented. All the quantities computed are written on a shared area of the disk, accessible to all the tasks. This is useful for an efficient fault tolerance mechanism.

The label slaves are spawned at the end of the previous jobs. Each slave executes the segmentation analysis on the subdomain assigned to it by the master and recognizes and numbers the substructures inside it. The labels found by each slave are written on a shared area.

The last part of the code is executed only by the master, that reads the results written on the shared area by the slaves and rearranges them. The master reads the labels that each slaves has assigned to the substructures found in its subdomain and makes an analysis of the borders between the subdomains. The task is to recognize those structures that different adjacent hosts have singled out and that are really only one substructure across the border. Finally, the morphological analysis is executed by the master.

References

- Becciani, U., Pagliaro, A., *Comp. Phys. Comms*, submitted
- Escalera, E., Mazure, A., 1992, *ApJ*, 388, 23

- Gambera, M., Pagliaro, A., Antonuccio-Delogu, V., Becciani, U., 1997, *ApJ*, 488, 136
- Grossmann, A., Morlet, J., 1984, *SIAM J. Math.*, 15, 723
- Grossmann, A., Morlet, J., 1987, *Math. & Phys., Lectures on recent results*, ed. L.Streit, World Scientific
- Lega, E., 1994, *These de Doctorat, Université de Nice (L94)*
- Lega, E., Scholl, H., Alimi, J.-M., Bijaoui, A., Bury, P., 1995, *Parallel Computing*, 21, 265