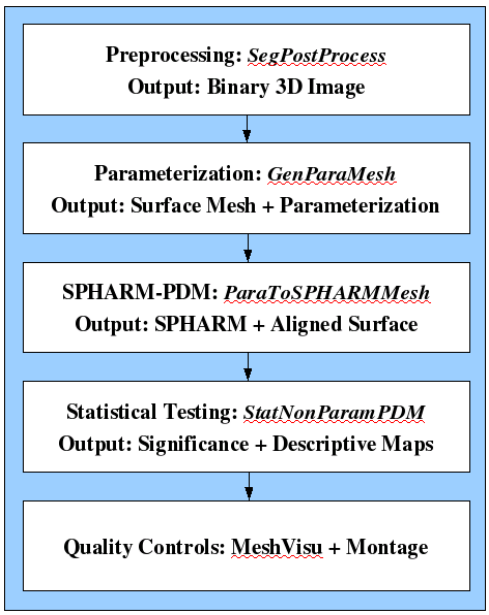
SPHARM-PDM based Shape Analysis Tool

The Shape Analysis Tool is a cross-platform software making the SPHARM-PDM pipeline easily and intuitively usable and giving complete results from a set of parameters and data. I will describe a typical use of the tool. The Shape Analysis Tool can be used with a graphic user interface, but also with command lines.
This software executes a BatchMake script and runs several tools as threads, with the possibility to process datasets locally or on a distributed environment using Condor.
In this document, after a short reminder of the SPHARM pipeline, i will present a overview of the shape analysis tool and then give a tutorial to use it.

1 The Shape Analysis pipeline

The shape analysis tool is guided by a pipeline. Each step of this pipeline provides outputs used by the following step. This pipeline contains five steps.

The original data are a set of binary 3D images. This data are first preprocessed using SegPostProcess in order to get smoother and filled images. Then, using GenParaMesh, we obtain surface meshes and spherical parameterizations. Aligned SPHARM surfaces are then created with ParaToSPHARMMesh. The fourth step is the statistical testing using StatNonParamPDM. Finally a set of Quality Controls is created to give an overview of the result to the user.



1.1 Processing of Binary Labels: SegPostProcess

The application of the SegPostProcess program is optional but its use is recommended, as it ensure spherical topology of the segmentation and offers a series of services:

- Extraction of a single label or a label range: This is especially helpful if the label dataset contains multiple labels.

- Resampling of the label data: Segmentation data from MRI or other modalities are often stored at rather coarse resolution or at non-isotropic resolution. The input to GenParaMesh has to be of isotropic resolution and a relative fine resolution is recommended.

- Spherical topology

1.2 Spherical Parametrization: GenParaMesh

GenParaMesh first extracts the surface of the input label segmentation by following the 'cracks' between the voxels of the foreground (label) and the background. The resulting mesh coordinates are thus off by minus half a voxel-width from coordinate systems that place the origin at the center of the voxels. This surface mesh is mapped to a sphere.

1.3 SPHARM-PDM: ParaToSPHARMMesh

This tool computes the SPHARM-PDM representation and resolves issues of correspondence and alignment.
The input is a surface mesh with spherical parametrization. As a first step, the raw spherical harmonic coefficients are computed only up to the first degree and the first order ellipsoid is determined. The spherical parametrization is then rotated such that the poles of the first order ellipsoid are coinciding with the poles of the parametrization. The spherical harmonic coefficients are recomputed up to the degree specified on the command line option.
The two main parameters of this tool are the maximal degree for the SPHARM computation and the subdivision level for the icosahedron subdivision. In our experience, the SPHARM maximal degree should be chosen between 12 (hippocampus) to 15 (lateral ventricle, caudate) for brain structures. If the degree is chosen too high the reconstructed SPHARM surface will often show signs of voxelization. By using the '- paraOut' command line option, the spherical icosahedron subdivision, as well as local phi and theta, attribute files for the quality control visualization.
Correspondence: The SPHARM shape description has an object-inherent correspondence based on the first order ellipsoid. Unfortunately this is not fully unique as the first order ellipsoid can be flipped along any of its axes with the same result. Thus, without additional measures the correspondence is ambiguously defined in regard to flips along these axes. In order to clarify this ambiguity, an additional

SPHARM object (described by SPHARM coefficients) can be provided as a flip-template. This flipTemplate is used to test all possible flips of the parametrization along the first order ellipsoid axis and select the one whose reconstruction has minimal distance to the flip-template. Commonly, the first object is computed without providing a fliptemplate and then serves for all subsequent objects as the flip-template.
Alignment: The ParaToSPHARMMesh tool provides two types of alignment. The first type performs an alignment of the normalized first order ellipsoid axis to the unit-axis with the ellipsoid center located at the origin. The second alignment type is a rigid-body Procrustes alignment, i.e. Procrustes alignment only with translation and rotation. Whereas the first alignment is object inherent, the Procrustes alignment is in need of a template, which can optionally be supplied.

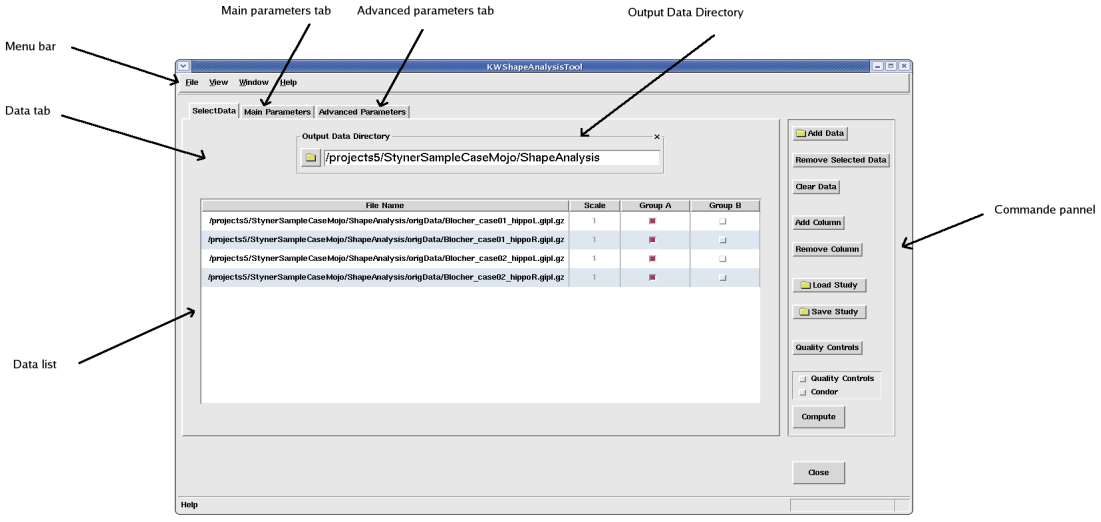
1.4 Statistics: StatNonParamTestPDM

Once all SPHARM-PDM meshes have been computed, group differences of the local surface point distributions can be assessed using StatNonParamTestPDM. There are two main types of results from the statistical testing step:
a) descriptive group statistics, such as the mean and covariance information, and
b) group mean difference hypothesis testing.
The descriptive statistics are used for quality control, sanity check, as well as these are necessary to make sense of the computed significance maps (”Is that significant region enlarged in the patient population?”). The significance maps show the regions of significant difference of the Hotelling T2 metric with 3 different maps: the raw p-value map, the FDR corrected p-value map, and the permutation based corrected p-value map. The value of the Hotelling T2 metric is also useful to visualize the full range of the effect size. The current version of the tool also computes a single global shape difference by averaging the local Hotelling T2 metric across the whole surface. The p-value is computed the same way than for the local tests using non-parametric permutation tests. No correction is needed for the global shape difference as only a single test only is performed. Additional other global shape difference measures will be implemented, such as percentile measures of the Hotelling T2 representing stable estimations of extremal difference, such as the 68% and 95% percentiles.
The main testing procedure parameter is the number of permutations employed for both raw computation and corrected p-values. As a rule of thumb, we are employing ten times the number of surface mesh points, which is a safe estimate and results in quite extensive computation and memory requirements.
Other relevant parameters include the maximal significance threshold (commonly chosen at 5%), as well as the resolution of the significance correction captured by the number of thresholds uniformly spaced between null and the maximal significance threshold (a good value is about 1000). For example, a maximal threshold of 5% with 10 threshold steps results in a correction resolution of 5%/10 = 0.5%. The number of steps parameter does not affect the raw significance map, which is rather determined by the number of permutations.

The StatNonParamTestPDM tool has several options for specifying its input. There are 2 main input types:
a) text file listing all mesh files with corresponding information
b) text file containing raw feature information.
Using the second option, the tool can be applied to many different types of data that are embedded in a linear space, but limited descriptive group statistics are provided. The first option is simpler to use and has a richer set of outputs. Both options use ASCII text files, which reserve a single line per subject. For the mesh-list file, each line contains first the group identifier number (often either 0 or 1), then the scaling factor for an optional scaling normalization, followed by the absolute or relative path to the mesh-file. The scaling normalization is only performed when the appropriate '-scale' command line option is also present, otherwise the scaling information is simply omitted.
In case, data other than meshes are to be tested, the current version of the program allows the testing of univariate, bi and tri-variate data via the '-2DvectorData' and '-3DvectorData' options. Additional information needs to be provided, identifying the column numbers (indexed at 0) for the group identification number and for the first feature, as well as the numbers of features.

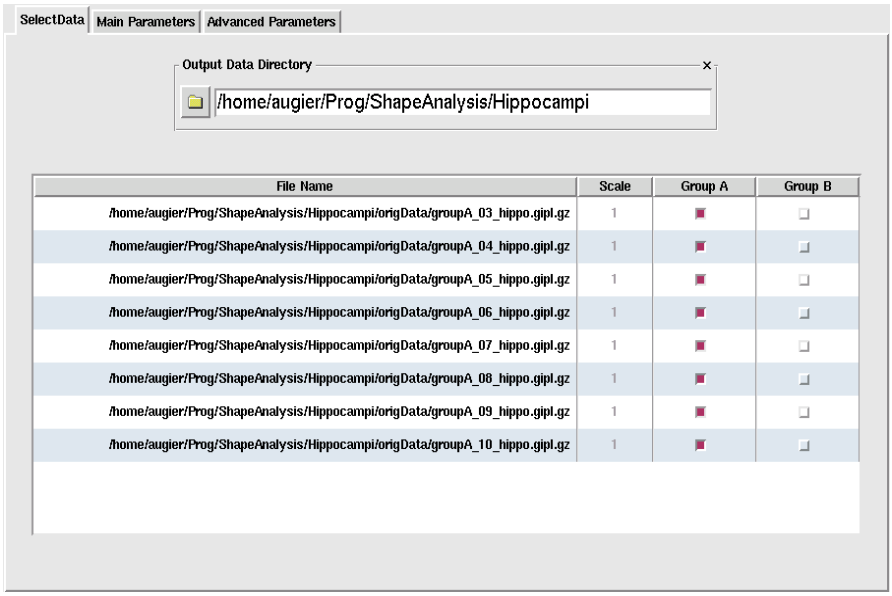
2 Overview

The shape analysis tool is divided in three tabs plus a command panel.



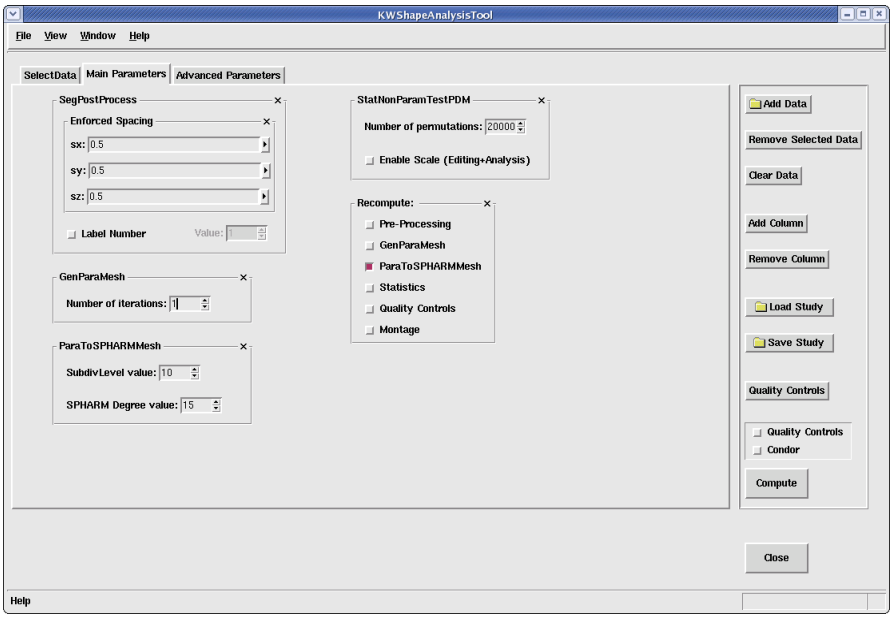
2.1 Data tab

This tab contains a 'select output data directory' section. The directory selected will contain all the outputs generated by each step of the pipeline. The other section is a multi column list showing all the data about to be processed.



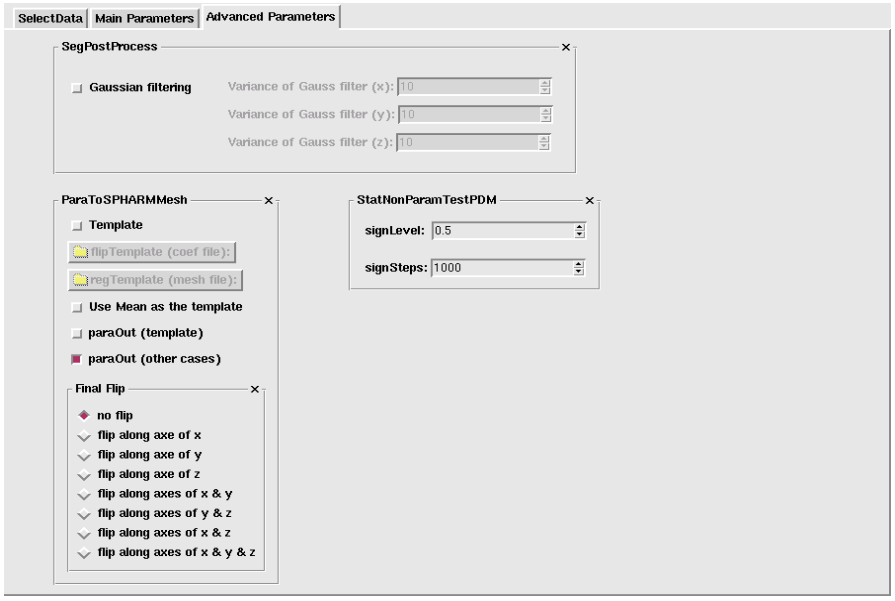
2.2 Main parameters tab

This tab contains the most important parameters separated in several frames corresponding to each step of the shape analysis pipeline. An additional frame allows the user to recompute one or several steps, the selected data being already processed.

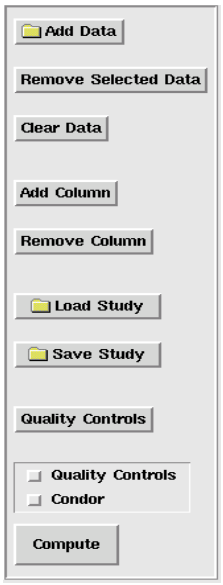


2.3 Advanced parameters tab

This tab contains the secondaries parameters. The user doesn't usually needs to change these parameters, unless a very specific study needs to be computed.



2.4 Command panel



The command panel contains all the necessary buttons to manage the data column list. Two buttons allowing the user to save or load a study and another one showing the generated quality controls are also available in this panel. Eventually, the computation button and a frame containing the computation options are also present in this panel.

3 Tutorial

This tutorial explains how to use the Shape Analysis Tool through the KWWidget graphic user interface.

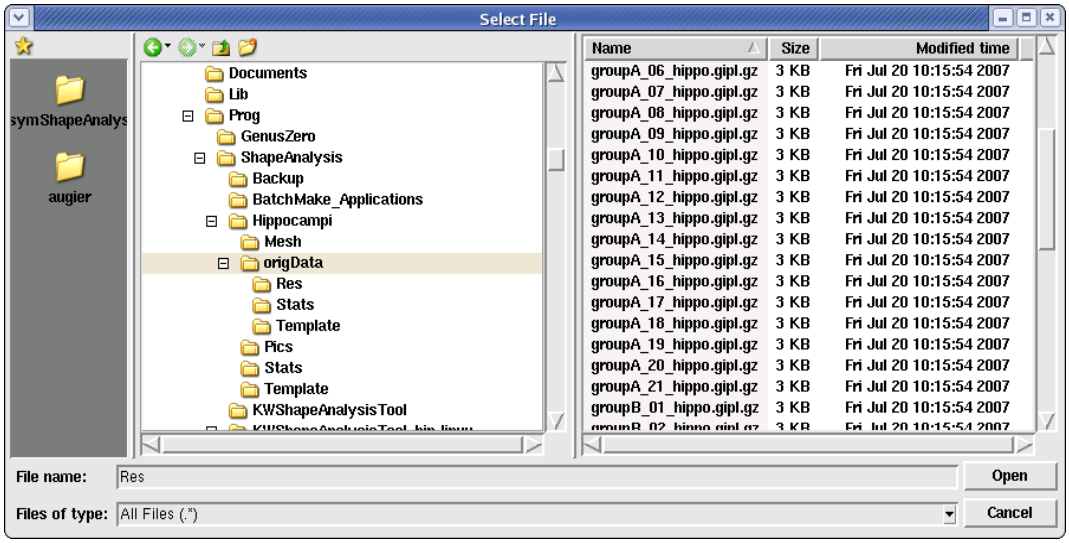
First of all an environment variable called KWShapeAnalysis_Home needs to be created using the following command: setenv KWShapeAnalysis_Home *environment_path*

3.1 Basic use

3.1.1 Select the Data

The user first needs to select the data required for his study.

A file selection dialog window pops up when the user clicks on the 'AddData' button. 



The selected data will then be displayed in a multi column list.

File Name	Scale	Group A	Group B
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_03_hippo.gipl.gz	1	■	□
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_04_hippo.gipl.gz	1	■	□
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_05_hippo.gipl.gz	1	■	□
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_06_hippo.gipl.gz	1	■	□
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_07_hippo.gipl.gz	1	■	□
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_08_hippo.gipl.gz	1	■	□
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_09_hippo.gipl.gz	1	■	□
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_10_hippo.gipl.gz	1	■	□

These data are about to be processed.

3.1.2 Compute

Once the data are selected, the user can directly compute these data with default parameters by clicking on the compute button. 

3.2 Frequently Asked Questions

How to analyze with scaling?

For a full shape analysis, we recommend to both analyze the objects in their original scale, as well as normalized for the individual intra-cranial cavity volume (ICV). The ICV is usually acquired in a separate brain tissue segmentation and the scaling factor f is then chosen as $fi = (Mean(ICV)/(ICVi)^{1/3})$

3.2.1 How to correct a not well represented object?

If the object is not well represented, the user may have to increase the SPHARM degree value of ParaToSPHARMMesh in the Main parameters section. However, if the degree is chosen too high the reconstructed SPHARM surface will often show signs of voxelization.

How to improve the spherical parametrization?

To improve a bad spherical parametrization, the user must increase the number of iterations of GenParaMesh in the Main parameters section.

The object looks blocky, what to do?

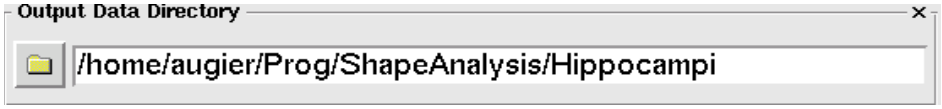
If the object looks dense, this is definitely due to a too low resolution. The user should then lower the enforced spacing of SegPostProcess in the Main parameters section.

4 Reference section

4.1 Data tab

4.1.1 Output directory

This directory will contain all the outputs of the process.



4.1.2 Data

The selected data are displayed in a multi column list.

File Name	Scale	Group A	Group B
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_03_hippo.gipl.gz	1	■	☐
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_04_hippo.gipl.gz	1	■	☐
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_05_hippo.gipl.gz	1	■	☐
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_06_hippo.gipl.gz	1	■	☐
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_07_hippo.gipl.gz	1	■	☐
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_08_hippo.gipl.gz	1	■	☐
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_09_hippo.gipl.gz	1	■	☐
/home/augier/Prog/ShapeAnalysis/Hippocampi/origData/groupA_10_hippo.gipl.gz	1	■	☐

The user can :

add one or several data: 

delete one data: 

clear the entire list: 

4.2 Parameters tabs

Several parameters need to be set before starting the computation. A default parameter set is loaded when the shape analysis tool starts. These parameters can be set manually.

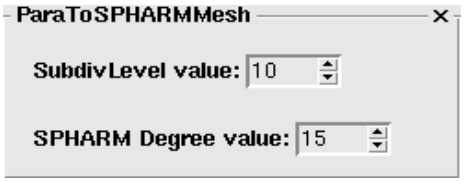
4.2.1 Main parameters

As evoked in the section 1, important parameters need to be set conserning each step of the shape analysis pipeline.

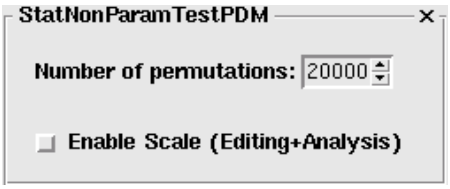
4.2.1.1 Set the SegPostProcess parameters

4.2.1.2 Set the GenParaMesh parameters

4.2.1.3 Set the ParaToSPHARMMesh parameters

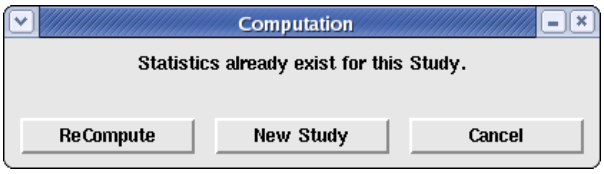


4.2.1.4 Set the StatNonParamTestPDM parameters

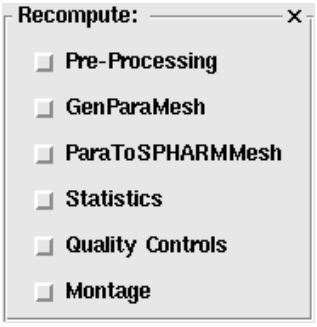


4.2.1.5 Set the recomputing options

If the user click on the Compute button whereas final results are already existing in the output directory, a pop up window appears asking the user what needs to be done :

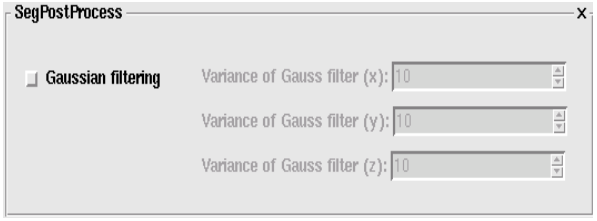


If the user chooses to recompute the study, the selected steps in the following panel will be recomputed for the selected data in the multi column list.

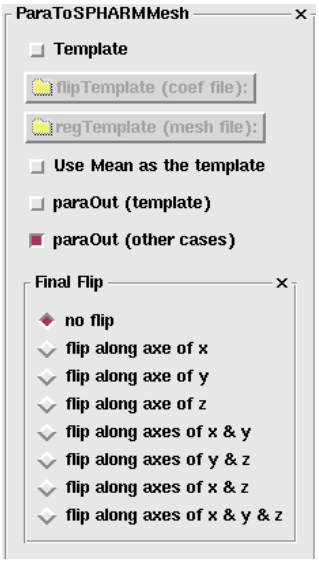


4.2.2 Advanced parameters

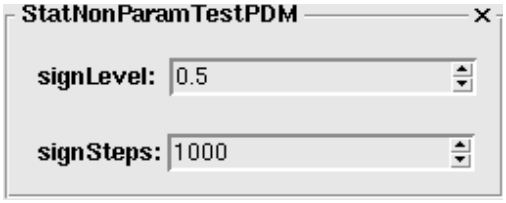
4.2.2.1 Set the SegPostProcess parameters



4.2.2.2 Set the ParaToSPHARMMesh parameters



4.2.2.3 Set the StatNonParamTestPDM parameters



4.3 Computation

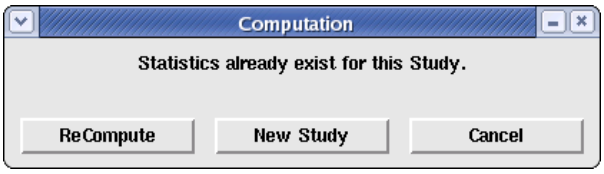
4.3.1 First computation

The user starts the computation by clicking on the compute button. 

Then, all the steps described in the section 1 are processed one by one.

4.3.2 New computaion

If the user clicks on the Compute button whereas final results are already existing in the output directory, a pop up window appears asking the user what needs to be done :



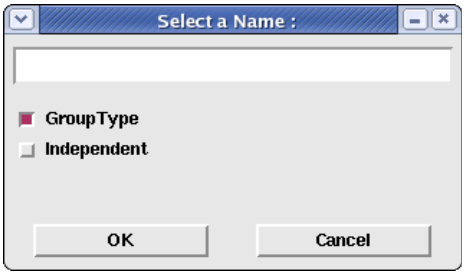
The user can then either recompute the selected data as evoked in section 2.3.1.5, or create a new study of which the data will also be in the specified output directory (e.g the new statistical results will be in the folder Stats_2).

5 Ongoing functionalities

One of the new functionalities currently developed is the opportunity to add other columns to the multi column list.

When the user clicks on the 'Add Column' button, a pop up window appears.

This window allows the user to set the name of the new column. The user needs then to select whether it is a group type or an independent characteristic. The column will be called *Name (GT)* or *Name (I).*



The user can then:

- Add one or several other column
- Remove one column

When the user clicks on the 'Remove Column' button, a pop up window appears letting the user choose the column he wants to remove. The entire name of the column needs to be specified, that is to say the name plus the type id between parenthesis.