

Software Guide

(February 9, 2012)

Abstract

This material is supplementary to the publication by Zawadzki et. al.
Building on that paper, this document provides information on usage of
our software.

Contents

1	Introduction	1
2	File Information	1
3	System Requirements	2
3.1	The graphical user interface - GUI	2
4	Disclaimer	3

1 Introduction

We report a morphology-based approach for the automatic identification of outlier neurons, as well as its application to the *NeuroMorpho* database, with more than 5,000 neurons. Each neuron in a given analysis is represented by a feature vector composed of 20 measurements, which are then projected into a two-dimensional space by applying principal component analysis. Bivariate kernel density estimation is then used to obtain the probability distribution for the group of cells, so that the cells with highest probabilities are understood as archetypes while those with the smallest probabilities are classified as outliers. The potential of the methodology is illustrated with respect to several cases involving uniform cell types as well as cell types for specific animal species. The results provided several insights regarding the distribution of cells, yielding single and multi-variate clusters, as well as suggesting that outlier cells tend to be more planar and tortuous. The proposed methodology can be used in several situations involving one or more categories of cells, as well as for detection of new categories and possible artifacts.

2 File Information

The following files are supplied:

<code>readme.pdf</code>	this file
<code>outliers.fig</code>	interface of GUI-version
<code>outliers.m</code>	main file of GUI-version
<code>initialize.m</code>	routine to initialize the main parameters
<code>main.m</code>	main routine to archetype and outliers detection
<code>kde2d.m</code>	routine for bivariate kernel estimate (see credits inside)
<code>showvar.m</code>	routine to show a specific neuron measurement
<code>*.txt</code>	Offline <i>NeuroMorpho</i> database.

3 System Requirements

This version requires Matlab (Mathworks Inc, Natick, USA) and allows the user to apply the workflow using a graphical user interface (GUI, Fig. 1).

3.1 The graphical user interface - GUI

Figure 1 shows the user interface where we featured the main components, which are described below.

- 1 - **List box 1:** *cell types*.
- 2 - **List box 2:** *animal specie*.
- 3 - **List box 2:** *brain region*.
- 4 - **Run button:** execute the archetype and outlier detection.
- 5 - **Text:** statistics of the current analyses.
- 6 - **Graphic:** 2D PCA projections where the outliers are shown in red.
- 7 - **Graphic:** estimated probability density function.
- 8 - **List:** list of available measurements.
- 9 - **Graphic:** Average value of the chosen measurement as a function of the region *id*.
- 10 - **Text:** basic statistics related to the chosen measurement.
- 11 - **Download button:** download the images of archetypes and outliers cells from *NeuroMorpho.org*.
- 12 - **List:** list with the names of all archetype cells.
- 13 - **List:** list with the names of all outliers cells.
- 14 - **Graphic:** where the image of the chosen archetype cell is shown.
- 15 - **Graphic:** where the image of the chosen outlier cell is shown.
- 16 - **Text:** status bar.

To use the GUI, start up Matlab and set the working directory to that containing the main-file `outliers.m`, calling the corresponding function `outliers()`. At this point, the GUI opens and the user must choose a combination of entries in the list boxes (shown by markers 1, 2 and 3 in Figure 1), in order to specify the *cell types*, *animal species* and *brain region*, respectively. Next the user can press the button *Run* to execute the main code. Observe that the minimum number of cells required to properly execute the code is twenty. Otherwise, a warning message is shown in the status bar. The 2D PCA projection as well as the estimated probability density function are shown in the places featured by the markers 6 and 7 in Figure 1. Moreover, a new window is also opened where the user can find the plots of the average values of all measurements as a function of the region *id*. In this window, we show in red every measurement whose absolute Pearson correlation coefficient is greater than 0.85. Additional information about a specific measurement can be obtained by using the list 8 of Figure 1. After choosing a measurement, the plot 9 is updated and some basic statistics are shown in the field marked by 10. These statistics correspond to the Pearson correlation coefficient between the measurement and the region *id*, the relative distance between archetypes and outliers, and the boolean results of the Kolmogorov-Smirnov test. This test returns 0 if the values of the chosen measurement come from the same probability distribution, which means that

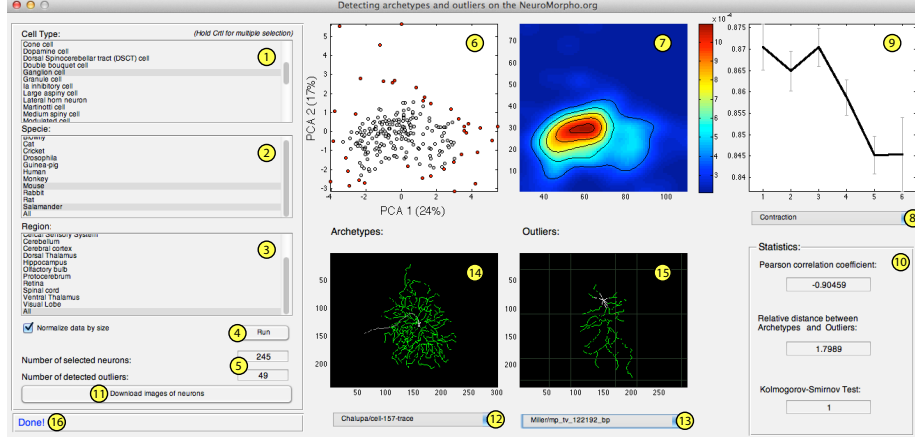


Figure 1: The graphical user interface.

the differences between archetypes and outliers are not significant. In the case the values come from different distributions, the test returns 1.

Images corresponding the the archetypes and outliers cells can be downloaded from NeuroMorpho.org once internet is available. In order to do so, just press the download button (marker 11 in figure 1) and wait for a few minutes for all cells to be downloaded. The images need to be download only once. When the downloading is concluded, the lists marked as 12 and 13 are automatically updated with the name of archetypes and outliers. By choosing a cell from these lists, the neuron image is shown in the plots 14 for archetypes or 15 for outliers. The bar marked by 16 is used to display the current state of the program as well as specific warning messages. The user can also disable the normalization by size by unchecking the check box *Normalize data by size*.

4 Disclaimer

The software and obtained results are not guaranteed regarding correctness and every application of this software is sole responsibility of the reader.