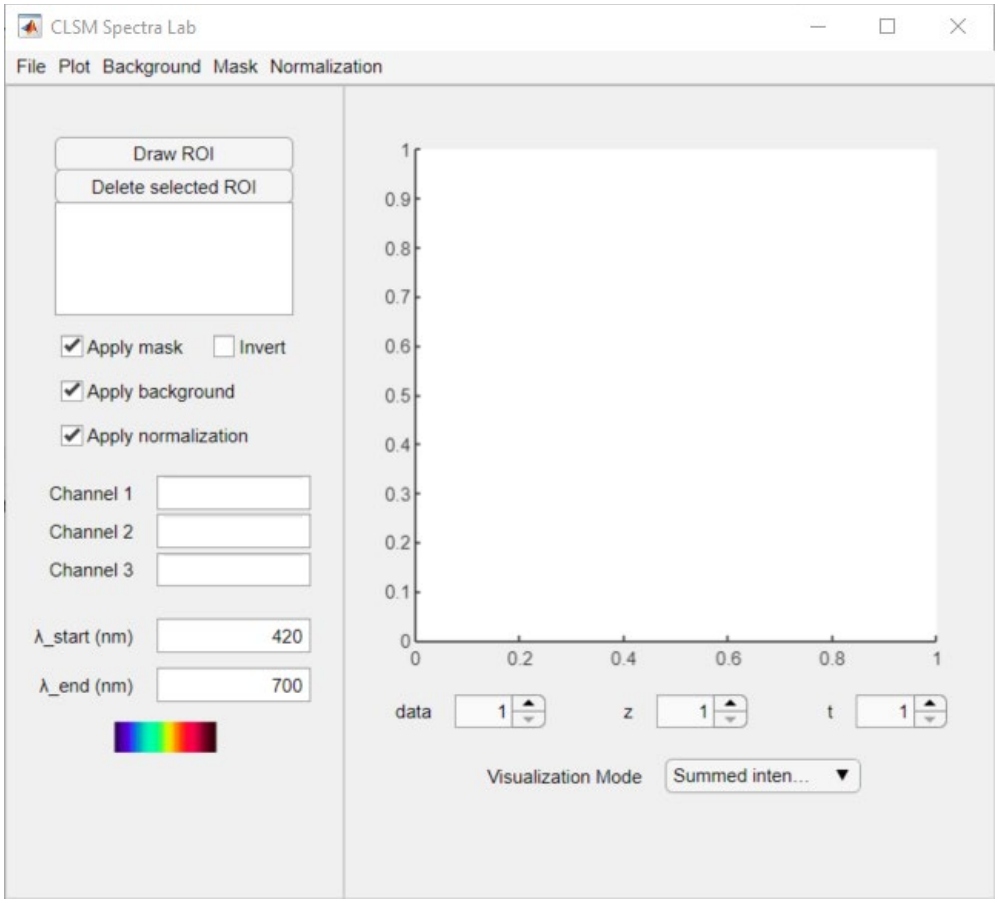
Instruction CLSM Spectra Lab for PCA + clustering in XANES Wizard

Motivation

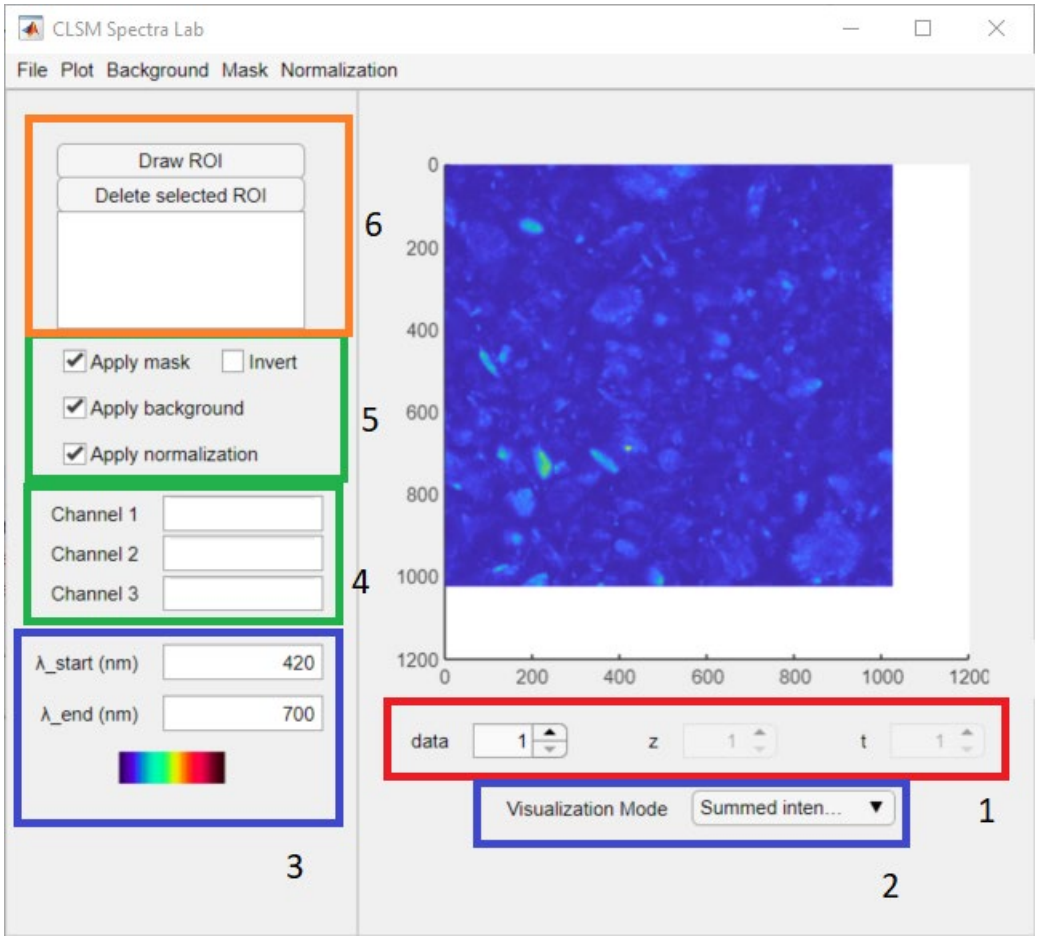
The goal of this document is to provide a practical instruction to load images recorded on a confocal laser scanning microscope (CLSM) in the CLSM Spectra Lab App and prepare them for principal component analysis (PCA) + clustering in the XANES Wizard App. CLSM Spectra Lab is designed for CLSM data with a spectral component. The App has various plotting functions and can perform the computations “background correction”, “masking”, and “normalization”. These pre-treatment steps are essential for reliable clustering result of CLSM data, but are usually not included in the analysis software used in our group, because these were initially designed for other microscopy techniques. Ultimately, I hope that we can converge to a standardized pre-treatment procedure for CLSM data.

Introduction

To start the app, open the CLSM Spectra Lab app via the “APPS” toolstrip. This should look like the figure below when freshly opened.



To load an image file, go to: “File” > “Load”. Multiple files can be selected if they are located in the same folder. If the images have been recorded in vastly different conditions, or have a different number of pixels, it is better to follow the workflow below for each set of ‘similar’ images. The reason is that some manipulations can only be done in the same way for all images in the data set. Even though the App can handle timeseries and z-stacks, it is not yet able export these to XANES Wizard.



The app has various features on the home screen that can be used right away (following the numbers in the figure above):

1. The user can navigate through the data, z-slices, and/or time steps via the buttons. A “please wait” box will appear until the image has been loaded.
2. The user can select the visualisation mode used to display the images in main window. The summed intensity over all spectral bins are displayed by default. The user can change the visualisation via the dropdown menu. “RGB blocks” divides the spectral bins in three equal blocks and displays each from low to high as red, green, and blue. Please note that this order is inverted with respect to the true colour, assuming that the increasing bin numbers represent longer wavelengths. To plot the approximate “true” colours, use the option “True colour”.
3. The wavelength of the first and last bin can be given in this box. These values are used for the “true colour” visualisation and are included in the export for XANES Wizard. The colour bar below displays the colour range of the bins.
4. Certain plot functions require the specification of bins as colour channel. This can be done in these boxes.
5. The mask, background correction, and normalisation can be applied for visualisation and computation. The visualisation is adapted automatically; however, computations have to be recomputed in order for the option to be applied.
6. Regions of interest (ROI) can be drawn and the spectra of the ROIs can be plotted. Click “Draw ROI” and draw the regions in the main visualisation window. Each ROI is assigned a unique number for identification. The corner points of the ROI can be interactively changed. The ROI

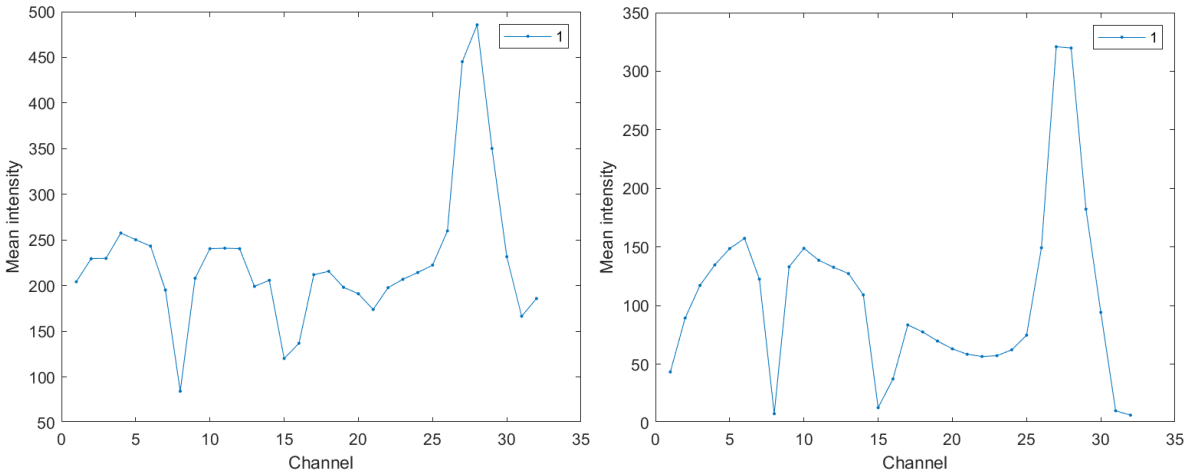
can be deleted by selection of the ROI in the menu and pressing the “Delete selected ROI” button.

The images and their spectra can be plotted in the menu “Plot”.

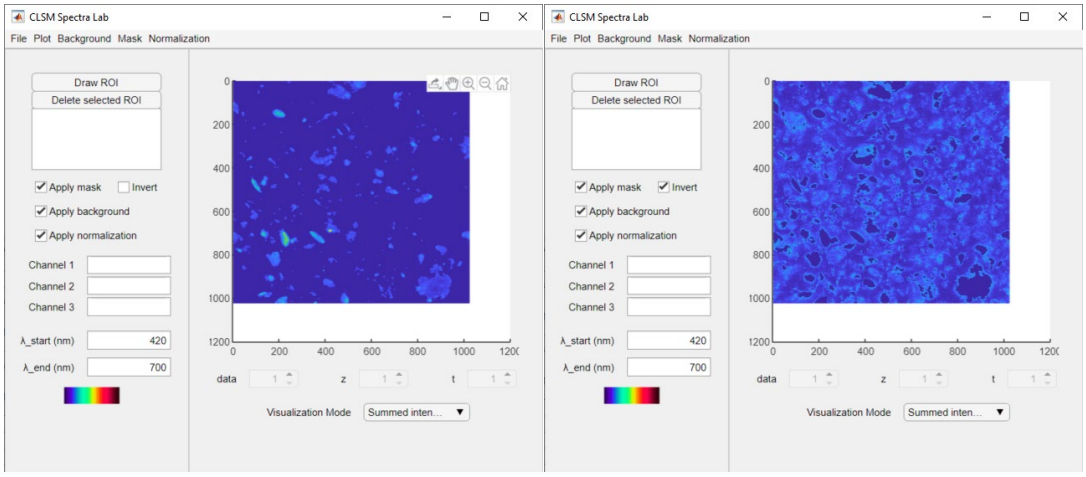
Workflow

In this section, I would like to propose a workflow for the pre-treatment of CLSM data. This workflow should be polished based on user experience. I will explain the steps and how to perform them.

1. Press “Background” > “Compute background”. An input dialog opens: Compute per pixel (1=true/0=false) and fill in 1. A file selection dialog opens. The user should select a so-called “dark image”, which has been recorded with the shutter off or microscope to “AUX”. Please make sure that the dark image is recorded with the same settings as the image, that is same dwell time, HV, and pixels. This option subtracts the values in the dark image from the image. Inspect the result by plotting the spectrum of all pixels via “Plot” > “Plot Spectrum” > “All” both with and without “Apply background” marked “ON” and “OFF”. Before and after background correction:

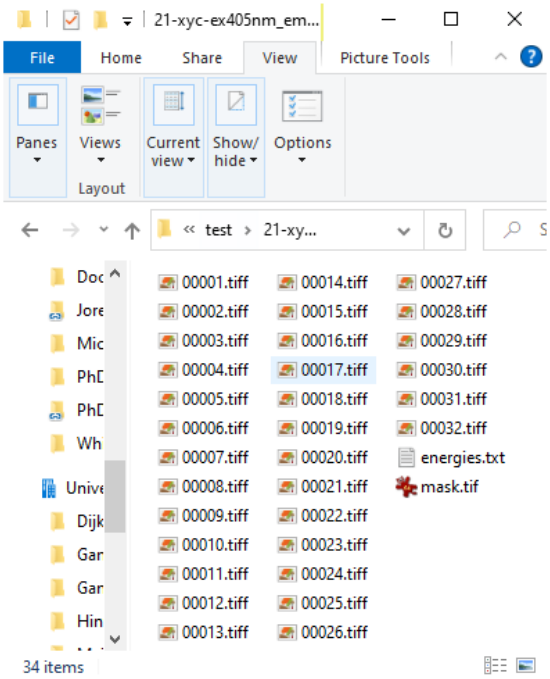


2. Please make sure that the “Apply background” is selected to do the masking with the background correction included. Press “Mask” > “First PCA” > “Compute mask”. A dialog box opens: Enter threshold value (empty for auto). Leave empty and press OK. Now image is displayed with the mask, where the dark blue regions are the background, and the mask can be inverted with the checkbox “Invert”:



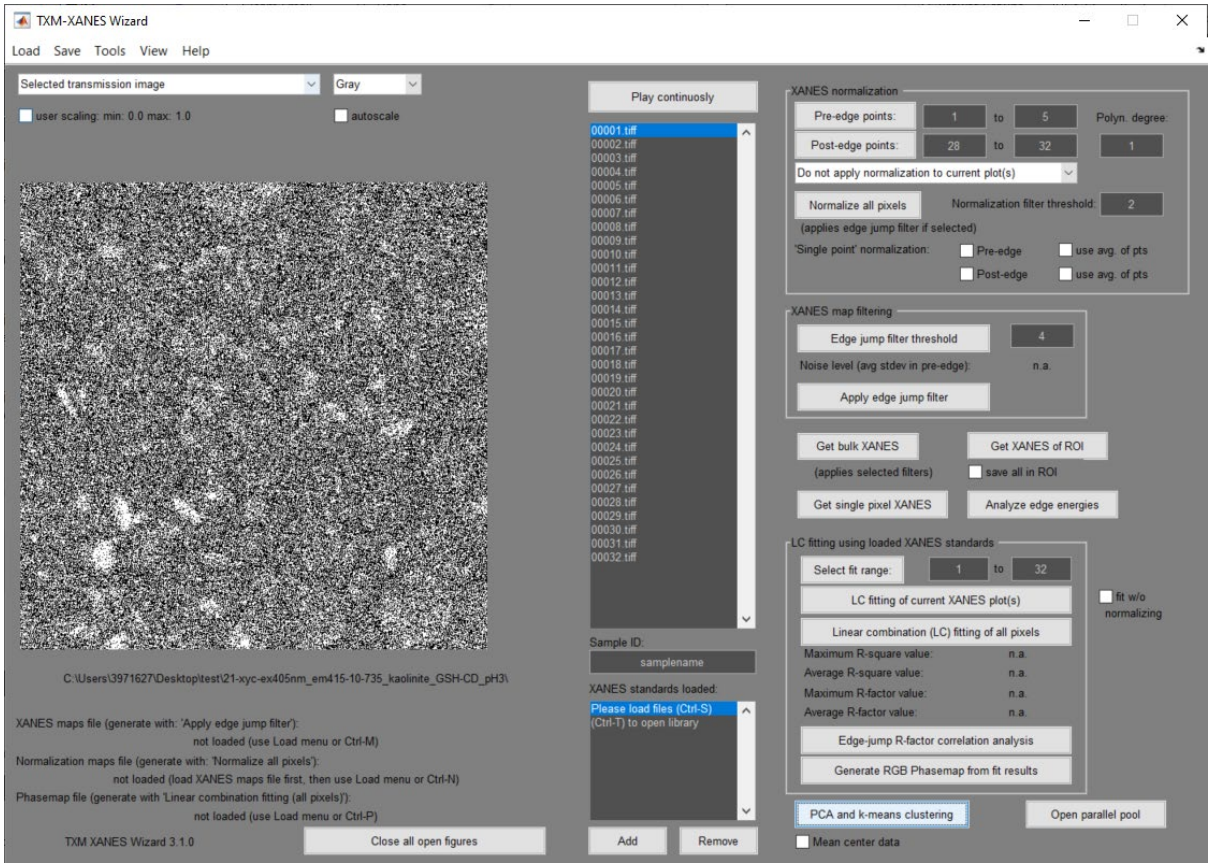
The value can be tuned for each image individually via “Mask” > “First PCA” > “Compute” for current data. The masking is done based on rescaled intensity values. A histogram of all rescaled intensity values can be plotted via “Mask” > “First PCA” > Plot histogram 1st eigen image”, which aids in selecting a good threshold value.

3. “Mask” > “Add sat. pixels”. Input dialog “Threshold counts”, keep the default value 4095. This function removes saturated pixels. A pixel is saturated when it has at least one bin in the spectrum with maximum value set by the user. Default is 4095, which is correct for the 12-bit format the Nikon A1 saves by default. Use 255 for 8-bit data and 65,535 for 16-bit images. The saturated pixels are removed from the current mask. This step is essential because the shape of the spectrum is distorted in the saturated pixels, because their true value is larger than the saved value in the image.
4. Please make sure that the “Apply background” and “Apply mask” are selected and “Invert” is deselected to do the normalisation with the background correction included. “Normalization” > “Single point” > “Compute norm”. A message box opens: “Enter bins for normalisation (multiple for average)”. Provide the bin number(s) that are set to one for the normalisation. If multiple bins are provided, the mean over these bins is taken. Use MATLAB notation to give a range of bins, e.g. 2:5, or individual bins, e.g. 2 4 7 10. Use “Plot” > “Plot Spectrum” > “All” to inspect the result.
5. Please make sure that Apply background”, “Apply mask”, and “Apply normalization” are selected. Then, press “File” > “Export” > “XANES Wizard”, which opens a folder selection dialog box. Please select the folder where you want to save the exported images for XANES Wizard. In this folder, a new folder is created for each image in CLSM Spectra Lab. In this folder, ‘tiff’ images with an increasing number in the format “00001.tiff” are images of the individual spectral bins, “energies.txt” is a file containing the wavelength of the bins, and “mask.tif” is the mask of the image.



Now, we can import the files in XANES Wizard:

6. In XANES Wizard, go to “Load” > “Image stack” and select the numbered tiff images (not the mask!). An error will appear: “Energy could not be determined from the filename! […] List of energies?”. Press “yes” and load the file “energies.txt”. De images have now been loaded.



7. Perform PCA and Clustering without using XANES Wizard’s other functions via the button “PCA and k-means clustering”. A question dialog appears: “No normalization data file found. Do you […] filtered) data?”. Press “Yes, and apply external filter”. Press “save” and in the next window select “mask.tif” to “open”.
8. You’re ready to do the PCA and clustering. For more information about this procedure, please visit the course on the wiki https://solisservices.sharepoint.com/sites/10857/icc_wiki and go to “Knowledge” > “Lectures & Workshops” > “Principal component analysis (PCA) workshop”.

For more information contact Erik Maris at j.j.e.maris@uu.nl.