

## SOP details

Title	Aligning Image Analysis Results With TopoMap
Description	This SOP describes the steps required to aligning image analysis results with TopoMap.
Author	Aliaksei Vasilevich, Miguel Faase; Tim Kuijpers
SOP number	4.3
Version number	3

## 1 Purpose

After CellProfiler Image analysis the user is left with a csv file that contains descriptors for each TopoChip. In this SOP, we append the TopoUnit *FeatureIdx* identifiers to the datasheet. *FeatureIdx* is the unique identifier of each TopoUnit surface.

## 2 Principle

This guide explains the position of the *FeatureIdx* on the TopoChip and explains how it can be linked to results of the screening.

## 3 Important to know before starting

1. TopoChip has 2176 unique surfaces, and each unique design is referred to as a feature, as described in the original TopoChip paper in PNAS (<https://doi.org/10.1073/pnas.1109861108>). They all have a unique identifier, which is called *FeatureIdx* which thus has a range from 1 to 2176. The flat surface has *FeatureIdx* 2177.
2. All surfaces are duplicated in the TopoChip design, flat surface has four replicas.
3. Every *FeatureIdx* has its predefined row and column coordinates, which is stored in file "TopoMap.csv". Four flat surfaces are located on the TopoChip diagonal as shown in figure 1.

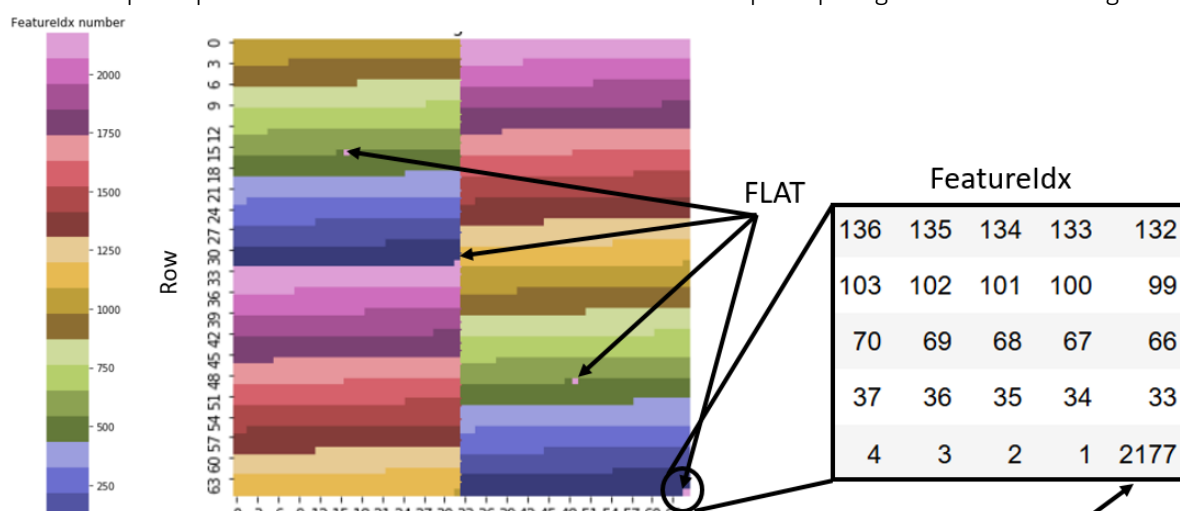


Figure 1 Visual representation of the FeatureIdx on the TopoMap.

4. By default, the TopoChip is imaged such that the Flat surfaces on the edge of the TopoChip, is located at the bottom right. The coordinates of this flat surface will be Row 66, Col 66 (Figure 1, flat surface on the edge). These coordinates can be found in the filename of the corresponding images.

## 4 TopoChip screen analysis

### 4.1 Workplace

This SOP can be performed in the office or home on your Laptop/Desktop.

### 4.2 Requirements

1. **Before you start** the analysis, it is **important** that you have **finished** the **CellProfiler pipeline** (SOP) and have the following files:
  - o Segmentation images
  - o Output file: image.csv
2. *Jupyter Notebook* and *Anaconda Navigator* are installed on your laptop/desktop (see SOP 4.1)

### 4.3 Load and verify image analysis data

1. Launch a Jupyter Notebook session, from Anaconda Navigator. A tab showing your root directory should open in your browser.
2. Navigate to the folder where you unzipped *Data Analysis TopoScreen* and open the *Jupyter Notebook* called *AlignFeatures.ipynb*
3. Next, you need to run every part of the script. The script is divided in blocks of code called *cells*. It is important to run every cell separately, from top to bottom and to verify the output of each *cell*. You can run a cell by selecting it and pressing *Shift + Enter*.
4. Load the "image.csv" data file by running *cell 2.1*. When the file can be located the data is stored in the variable *rawImageData*, otherwise an error message appears. If the file cannot be found, check if you finished the CellProfiler pipeline and validate if you didn't move the file from its original location.
5. *Cell 2.2* checks the number of rows and columns in the data. For instance, the number of rows should be a multiple of 4536. The number of columns depends on the descriptors selected in CellProfiler.
6. Check if the necessary columns are present in your data (*Cell 2.3*).
7. Perform a final check (*Cell 2.4*) to validate if all images have a row and column coordinate. If not, it will estimate it from the *Metadata\_Well* column, which was extracted from image filenames. It uses the following procedure: the last 2 numbers will be interpreted as column id and remain numbers at the beginning of code (1 or 2) as row ids. Ignore the warning produce by this function.

### 4.4 Connect the data with the feature ids in the TopoMap

1. Load the *TopoMap.csv* to align the feature ids to the image data. *TopoMap.csv* is an annotation file located in the folder *"/AnnotationFiles/"*.
2. *Cell 3.2* will ask you if it is necessary to rotate the *TopoMap*. When you rotated the TopoChip during imaging and did not already rotate the image, you have to rotate the *TopoMap*. To perform the rotation of the *TopoMap*, the desired number of times you need to change in the script the rotation coefficient, which is specified in the table below. The rotation of *FeatureIdx* is necessary to match the original *TopoMap* file with the actual position of surfaces on the screen.

Location of the flat	Rotation coefficient
Row 66, Col 66 (default)	0
Row 66, Col 01 (rotated to the left)	-1
Row 01, Col 01 (rotated twice)	-2 or 2
Row 01, Col 66 (rotated to the right)	1

3. *Cell 3.4* will combine the *TopoMap* with the raw image data. This will generate an output similar to the original raw image data, except now the data will contain the columns *FeatureIdx*, *MetadataQuadrant*, and *Metadata\_Duplicate*.
4. Evaluate *Cell 4.1* to save the aligned features to the image data in the output file *ImageWithFIdx.csv*. The notebook will save this file to the */DataAnalysis/* folder and will be used in the *Jupyter notebook 2\_IdentifyOutliers.ipynb*. Therefore, you should not change the location of *ImagewithFIdx.csv*.
5. Finally, save the Jupyter notebook by clicking the 'Save' icon. This way, you will have a document with the different steps you took while generating the data.