MTIMBS (Microtubule Intensity Measurement with Background Subtraction) Documentation

Jon Fernandes and Joseph Slivka
July 13, 2022

Email: jslivka@berkeley.edu

MTIMBS Description:

Take tif image and find average intensity of labeled MTs subtracted by background (found by looking at lowest mean background of the MT area shifted up, down, left, right and diagonals by a specified number of spaces sp [default = 10])

Altogether, MTIMBS can quickly and accurately predict MT intensities minus the image background that occupies the same amount of length as the MT for approximately 95% of movies with all the tools described. This makes it much more efficient than ImageJ plugins to evaluate large data banks of binding assays.

Parameters:
crop_area : crop area (optional): [x y width height] where x and y are measured as if the image is in the first quadrant from the lower left corner. Default: no crop whole image

Returns:
corrected_intensities: array of size num_MT with intensities of MTs after background subtraction all saved in a single excel file for the entire folder of analyzed images

Installation

To use this program, download the entire folder of matlab files included in MTIMBS
Make sure you have MATLAB installed. This code has been extensively tested on 2021a, but will probably work in any version after 2019.
You will need to also download the following toolboxes:

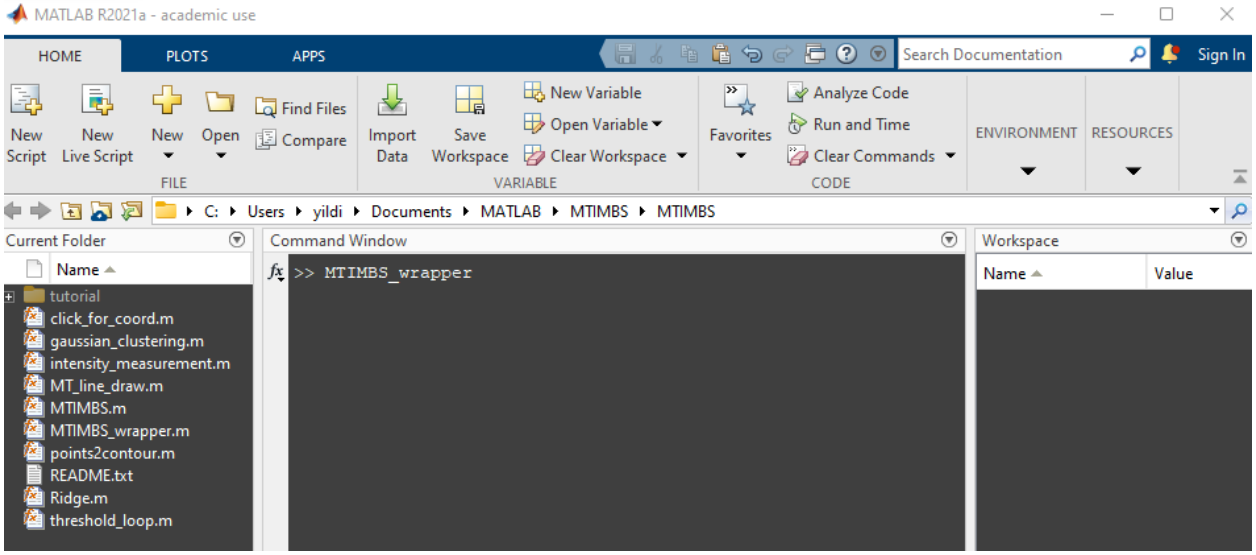
- Audio processing toolbox
- Statistics and Machine Learning Toolbox

There is also an included java macro for use with fiji to crop your images, though the crop_area option takes care of this macro.

A Basic Tutorial

Step 1: Run MTIMBS_wrapper or MTIMBS in MATLAB to preprocess images

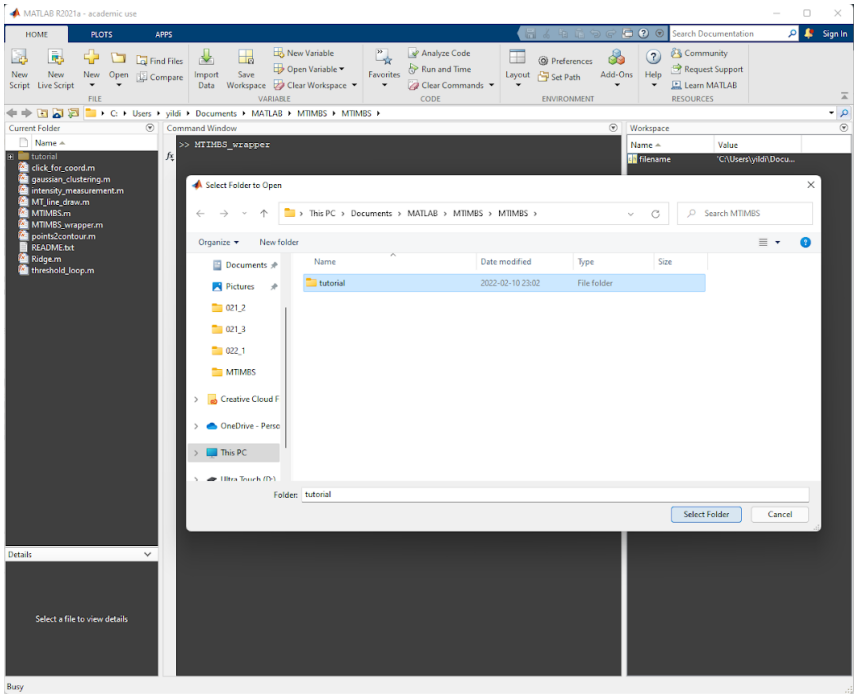
Now open a MATLAB window and make sure you are in the filepath where the files for MTIMBS are. If you just downloaded the software it should look something like this for your current folder:



Now, you will run the function MTIMBS_wrapper either from the function itself or the MATLAB command line.

This will open a window to select the parent directory. The MTIMBS_wrapper function will look to analyze every file with extension *.tif in this directory

- If even a single tif exists, it will immediately launch into the next step and will not do any cropping.
- If no tifs exist, the next level of directories will be checked and cropped according to the crop_area parameter. This is the usual save style for a single condition from the b307 scope.



If you wish to only run it for one file, you can do this by copying the filepath as below and then running MTIMBS(filename)

```
Command Window
>> filename = 'test2.tif';
>> filename = 'C:\Users\yildi\Documents\MATLAB\MTIMBS\MTIMBS\tutorial\test2.tif';
>> MTIMBS(filename)
```

It is best to use the whole path so that the subdirectory will save without error. So don't use the top, use the bottom (with your current path of course)

After either of these steps, a matlab imshow plot will show what image you will be analyzing. Simply answer 'y' or 'n' to analyze or skip respectively. You can also skip image analysis at any user input by the input 'x'.

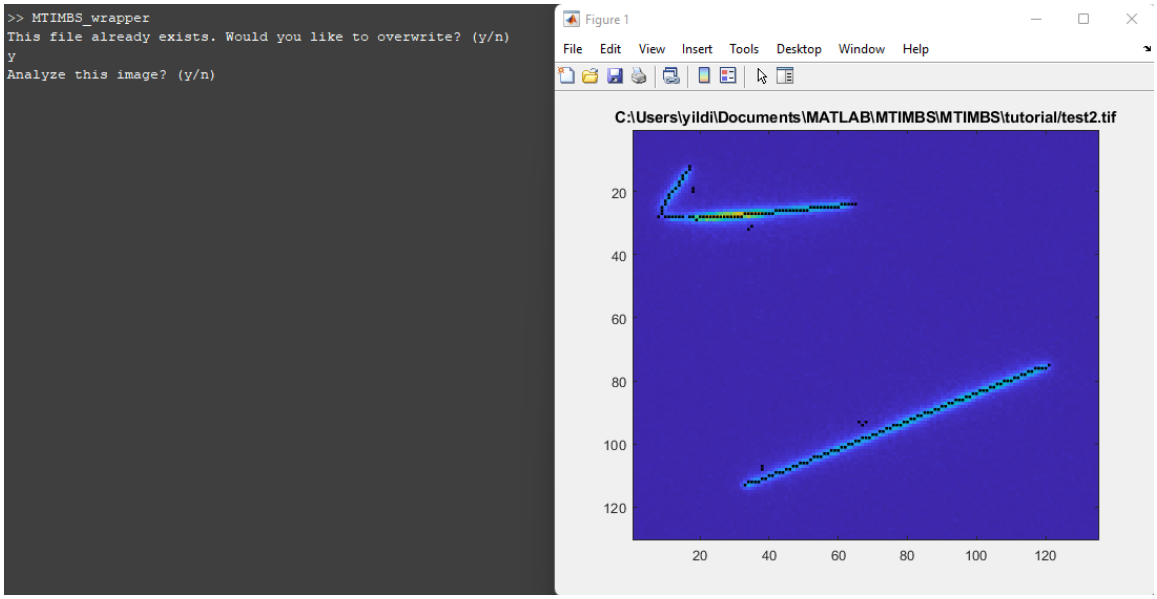
Step 2: Catch if already analyzed

If you have already analyzed this folder, it will send a warning message that this file already exists. Selecting 'n' will cancel the program.

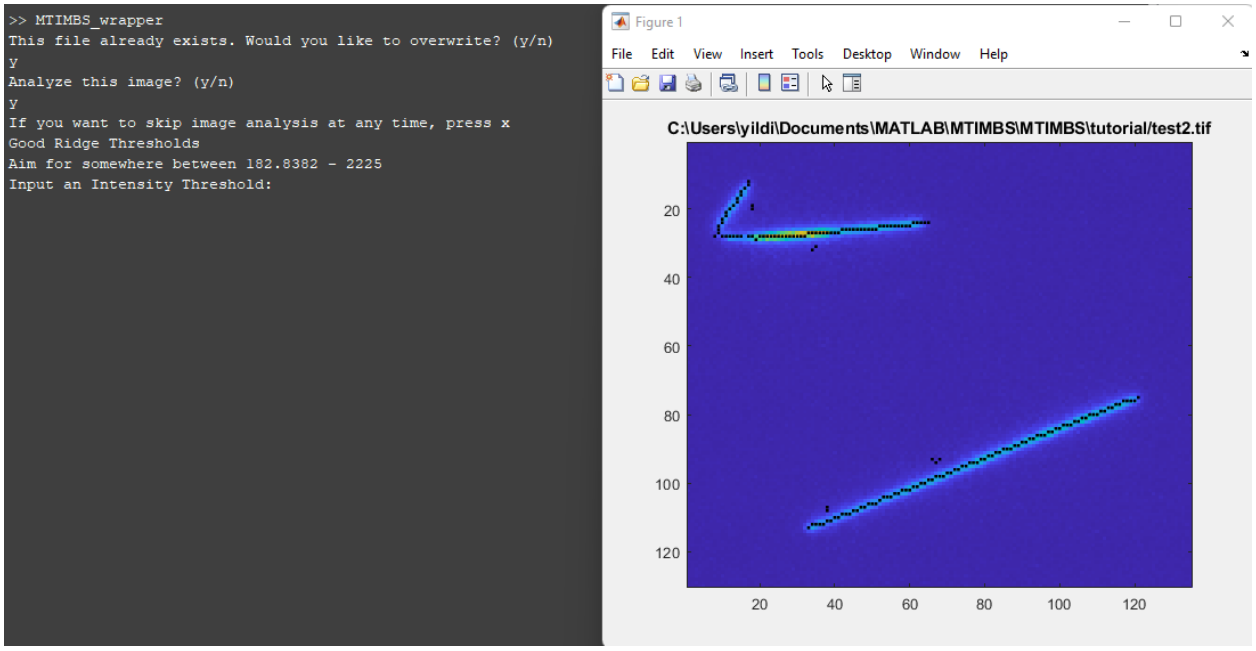
```
>> MTIMBS_wrapper
This file already exists. Would you like to overwrite? (y/n)
```

Step 3: Find Appropriate Ridge Threshold

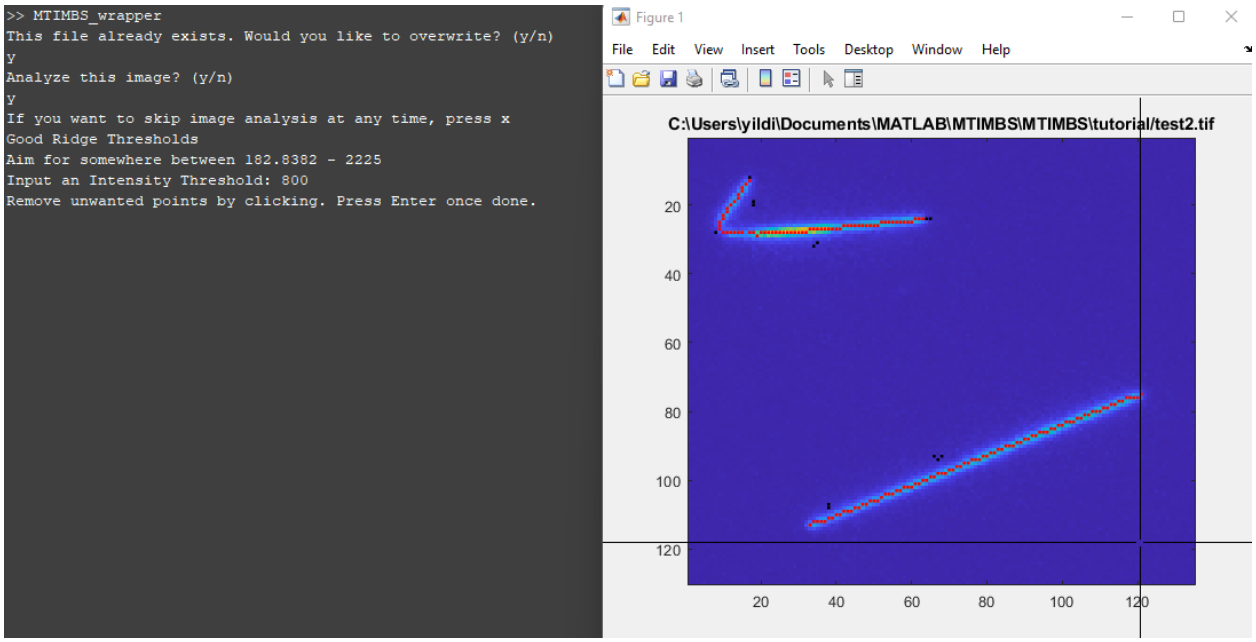
Decide whether you want to analyze a particular image. After this, if at any time you want to cancel analyzing an image, just enter ‘x’.



The wrapper will try and apply the previous threshold to your new image. If there are no points above threshold or no previous selected threshold, you will be prompted to set a ridge threshold. Enter a number. Suggested values are given, and other good probes would be on this scale: [150, 250, 400, 600, 1000, 1400]. Once entered, this will show the values of ridge on the plot (in black) and those above threshold (in red)

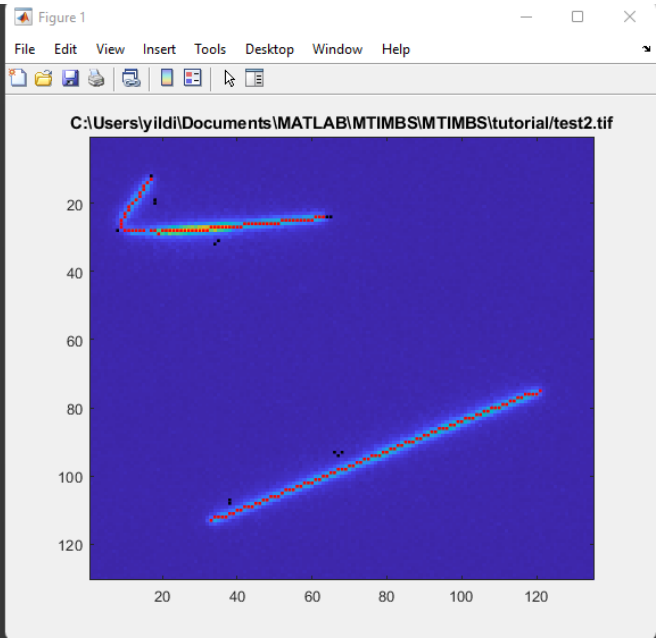


Aim to have key points of the MT in red as shown below. You have the option to delete outliers by clicking and it will delete the nearest red point. Once it looks good, press enter. Here, no outliers seem to need to be deleted so just press enter.



And not something like this:

```
>> MTIMBS_wrapper
This file already exists. Would you like to overwrite? (y/n)
y
Analyze this image? (y/n)
y
If you want to skip image analysis at any time, press x
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 800
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
n
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 600
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
n
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 800
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
```

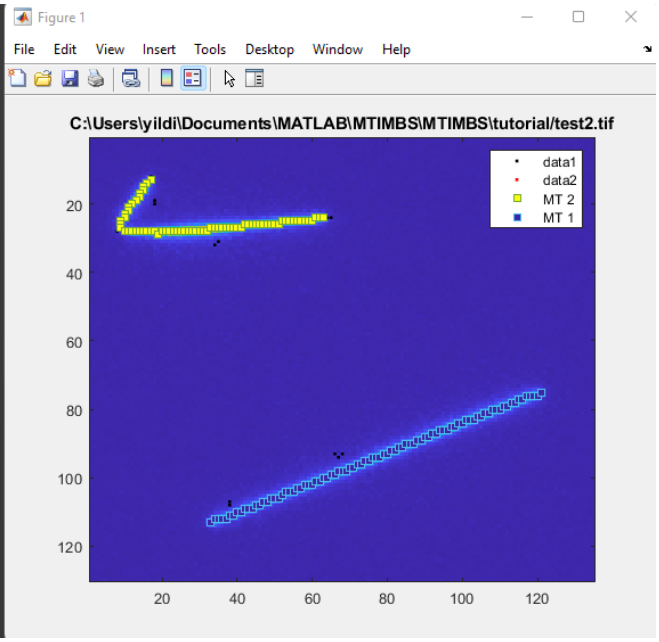


Notice that this loops so you can play with thresholds as much as desired. As soon as you find a reasonable threshold, enter ‘y’ and move onto the next part. If the threshold did not meet expectations, simply enter ‘n’ and try again

Step 4: Categorize MTs

Once the points of interest have been identified, the program needs to sort out clusters of MTs. Using a gaussian mixture model algorithm to cluster the MTs (described in the section “Why are the MTs not categorized properly?”), the program will attempt to find the optimal number of clusters which we call MTs.

```
>> MTIMBS_wrapper
This file already exists. Would you like to overwrite? (y/n)
y
Analyze this image? (y/n)
y
If you want to skip image analysis at any time, press x
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 800
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
n
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 600
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
n
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 800
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
y
Optimal MT # 2
Do these MTs look correct? (y/n)
```

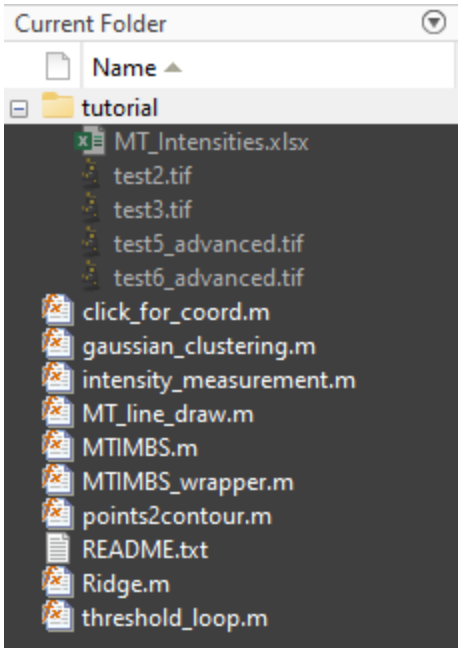


Finally, make a last visual check to see if the MTs look reasonable. If they are acceptable, select ‘y’. You should always check that the figure shows a reasonable MT fit and points for the intensity measurements.

If you ran MTIMBS_wrapper, it will automatically save the results and move on to the next file. If you ran MTIMBS, the program will exit, saving the following array and printing it in terminal.

```
>> MTIMBS(filename, 500, 2)
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225

ans =

   1.0e+03 *

    1.7807    2.0285
```

The answer is the intensity of the MTs with background subtraction specified by the shift array (labeled mod). You should also see at this point a new file labeled MT_intensities. This is a csv file storing all intensities found from the program that you agree to save.

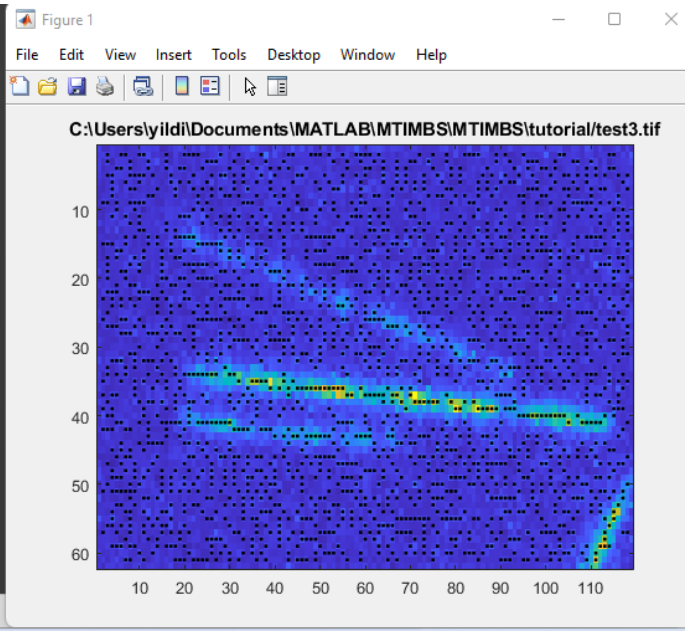


Step 5: Do it again - A more intricate example

Continue to do this for all of your MTs.

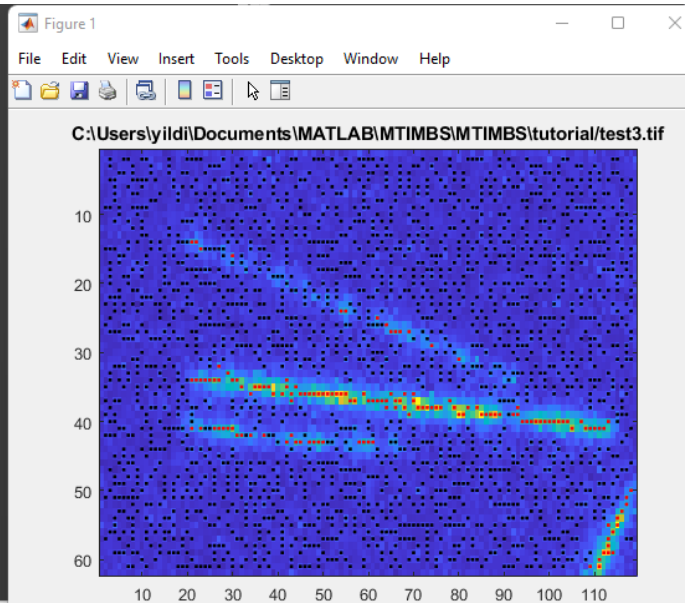
Step 3: Find Ridge Threshold
Step 4: Verify MTs are found

```
Operation terminated by user during MTIMBS (line 32)


In MTIMBS_wrapper (line 42)
    AllI = [AllI, MTIMBS(fname_w_path)];


>> MTIMBS_wrapper
Analyze this image? (y/n)
y
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 800
Use this threshold (y) or try again (n)? (y/n)
y
How many MTs are in the image?
2
Do these MTs look correct? (y/n)
y
Analyze this image? (y/n)
y
Good Ridge Thresholds
Aim for somewhere between 120.1149 - 650
Input an Intensity Threshold: 500
Use this threshold (y) or try again (n)? (y/n)
```

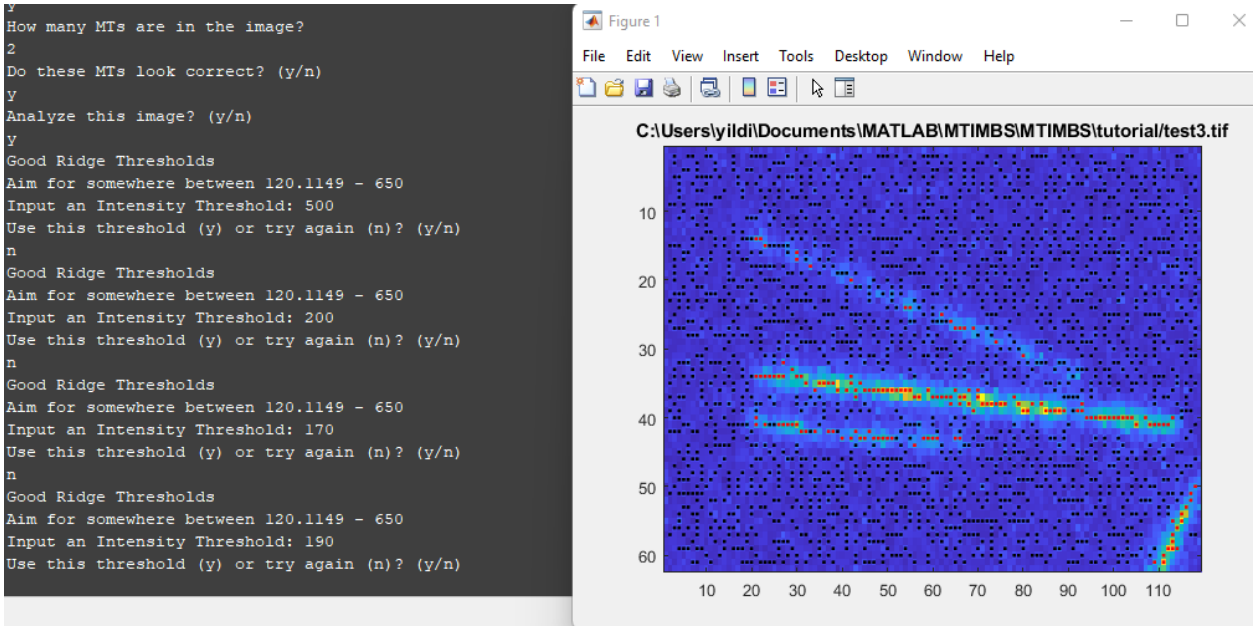


This ridge threshold was too high and clearly did not find any good points. Let's try lowering it to somewhere in the range suggested.

```
    AllI = [AllI, MTIMBS(fname_w_path)];


>> MTIMBS_wrapper
Analyze this image? (y/n)
y
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 800
Use this threshold (y) or try again (n)? (y/n)
y
How many MTs are in the image?
2
Do these MTs look correct? (y/n)
y
Analyze this image? (y/n)
y
Good Ridge Thresholds
Aim for somewhere between 120.1149 - 650
Input an Intensity Threshold: 500
Use this threshold (y) or try again (n)? (y/n)
n
Good Ridge Thresholds
Aim for somewhere between 120.1149 - 650
Input an Intensity Threshold: 200
Use this threshold (y) or try again (n)? (y/n)
```

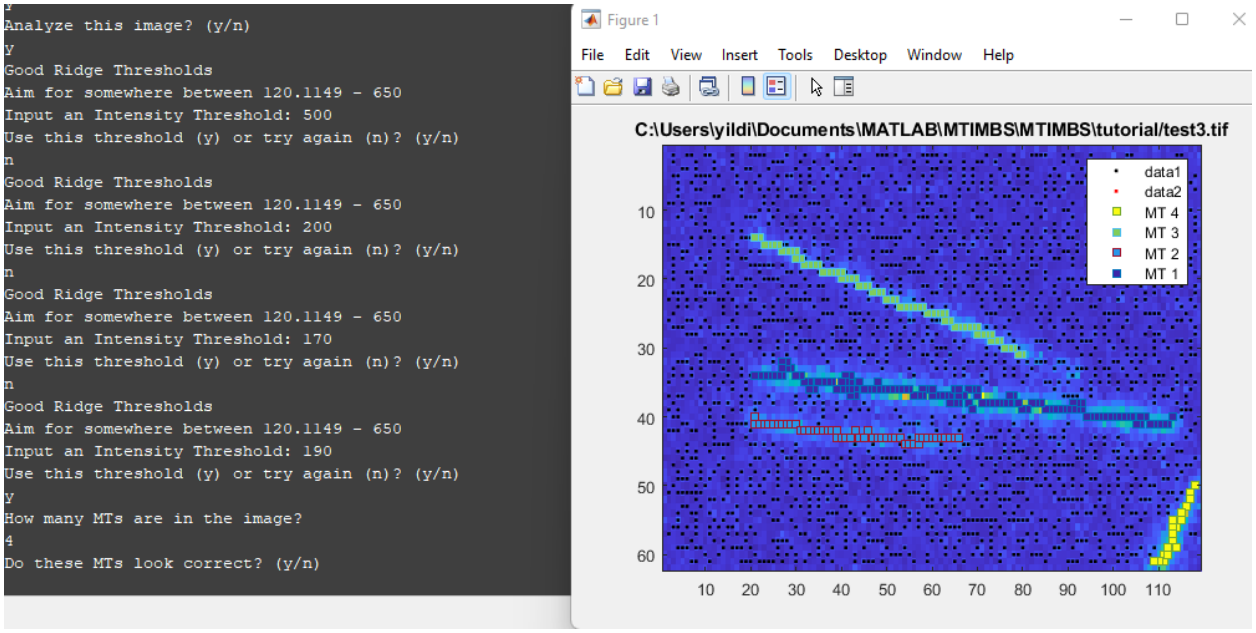


This seems like it found much more reasonable points to trace out the MTs, but we are missing some in between points and edges. Let's decrease it to include a couple more points



Let's try this. Enter 'y'

Step 4: num_MTs
This looks like four MTs.



This looks much better and the values seem much more reasonable.

Note that once 'y' is selected, the MT_intensities will also change to include the most recent intensities. It will store in columns and automatically sort the data by increasing intensity. It is

important to use MTIMBS_wrapper for similar conditions to not add ambiguity to your output MT intensities. (Clearly, these two test cases are at much different protein concentrations)

	A	B
1	60.93846154	
2	95.28571429	
3	158.9663866	
4	178.5789474	
5	1786.764045	
6	2133.376812	
7		
8		
9		

Congratulations for making it through the beginning tutorial. About 70% of the time, this will be all you need to do to get your MT intensities with an equivalent background subtraction. However, multiple issues can arise from the gmm algorithm that gives us our MT categorizations.

MTIMBS is robust and has functionality to deal with these cases where the basic protocol breaks down. We've found that this added functionality increases the success rate of MTIMBS to around 95%! This is addressed in more detailed in the advanced tutorial which mostly deals with the case when your MTs do not fit properly.

An Advanced Tutorial

Assuming you already have the basics of the programming down, we want to deal with some of the cases where MTIMBS fails to correctly categorize MTs. This will take your data throughput from approximately 70% to 95% of MT tif images.

Why are the MTs not categorized properly?

This is a quick section to describe why MTs may fail to follow what your eye sees. This is of course unnecessary to operate the program but may help to decrease frustration and understand the machinery that is categorizing MTs.

The Machine of Categorization: Gaussian Mixture Models

More of this is described in: [Cluster Using Gaussian Mixture Model - MATLAB & Simulink (mathworks.com)]

The way that MTIMBS determines different linear correlations of points is through a statistical fitting method known as a gaussian mixture model (GMM). This model attempts to describe a collection of points as a 2D statistical distribution. So it iteratively categorizes the space according to the most likely statistical distribution assuming a 2d non-diagonal covariance matrix to describe the gaussian.

The gmm model is initialized with a uniform prior. This means that it initially has no bias where it will converge. This is the best model to determine statistical significance. On the downside this does not always help find MTs that are closely correlated in space, especially in the presence of very long MTs which are hard for a gaussian to want to fit.

We could break the uniform prior by setting initial means so that the gaussian has a starting prior of where to look for its peak. This breaks some fidelity of statistical significance, but this is of little concern to what we are using the gmm to categorize.

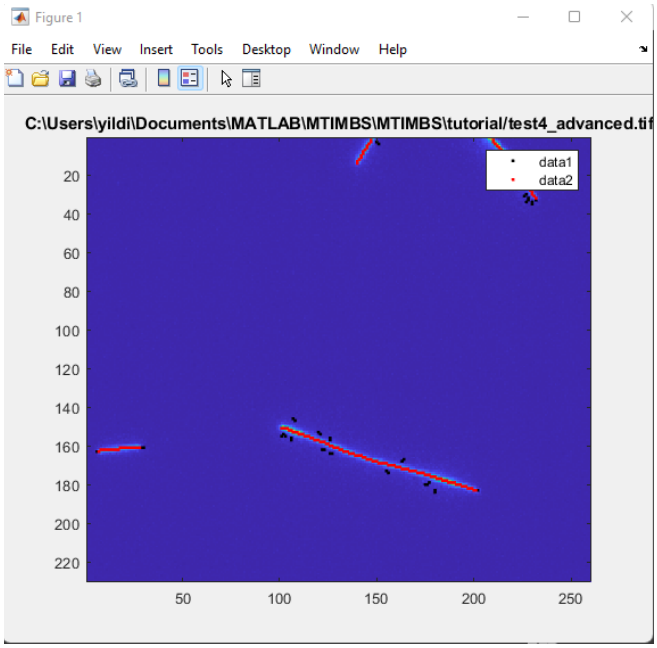
From the prior, the gaussian mixture now makes randomly generated steps and accepts certain moves that increases its statistical likelihood (known as the maximum a posterior criterion). This implies that two runs of the gmm model may not necessarily converge to the same distribution. We could correct this by seeding (having the same steps every time), but this randomness may actually benefit us which will be described later.

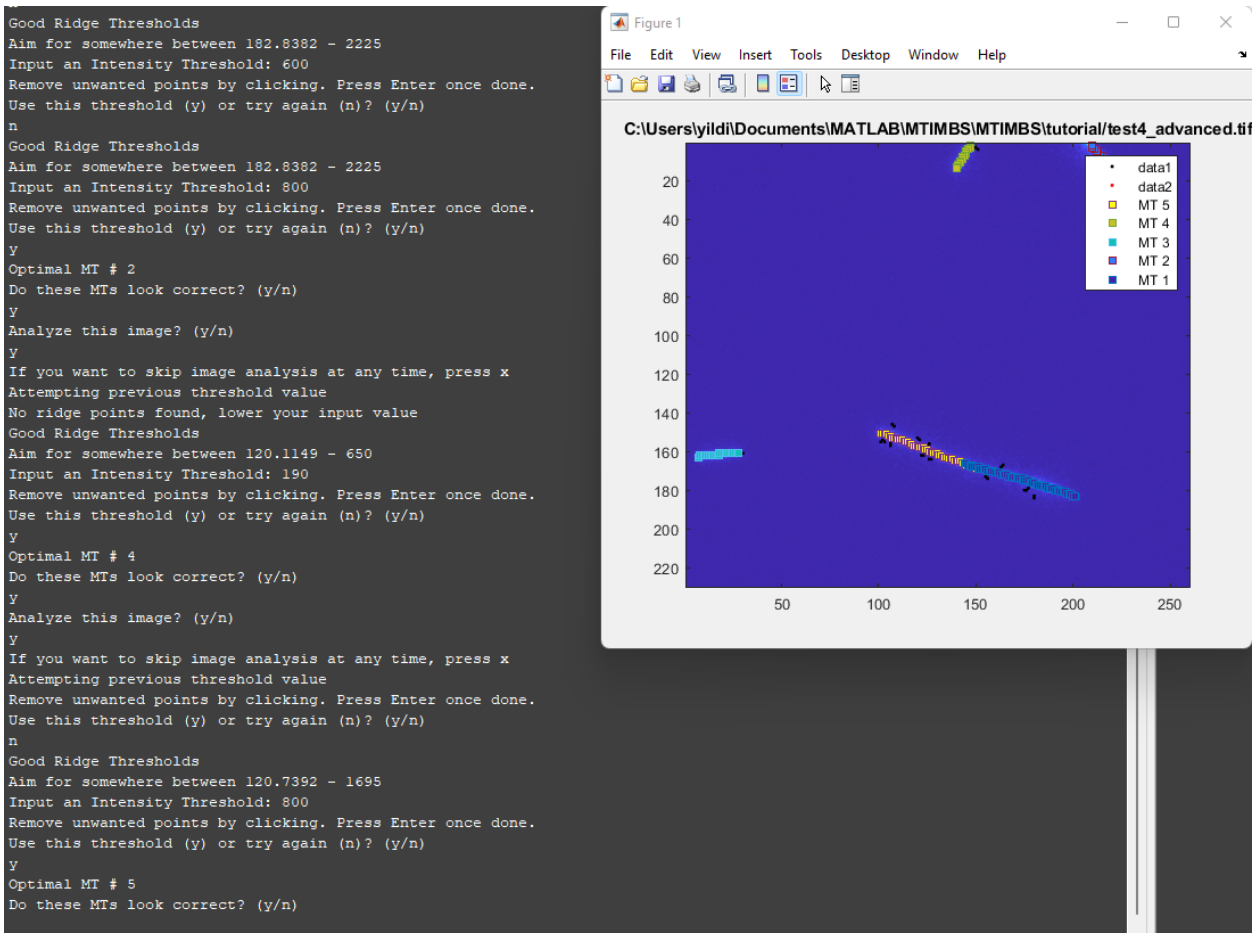
Hope that was a bit worthwhile to understand the gmmfitdist function in MATLAB. Now, we will utilize some of its functionality to make MTIMBS fit even more of our data.

Step 6: Try setting the number of MTs

So your MTs were not properly identified the first time? The program has a random path to find its stationary distribution, and there is a chance that it just got stuck in the wrong minimum and couldn’t get out.

Therefore, the simplest solution is to just try and run it again setting the number of MTs.

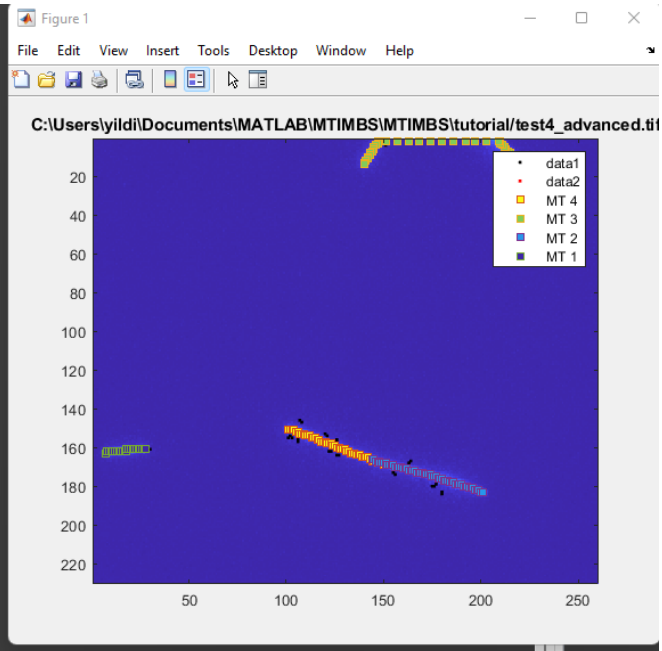




Simply enter ‘n’ when asked if the MTs look good.

Enter ‘y’ to try and run the gaussian mixture model (gmm) again.
This time, enter the number of MTs you want to find to try and help the algorithm

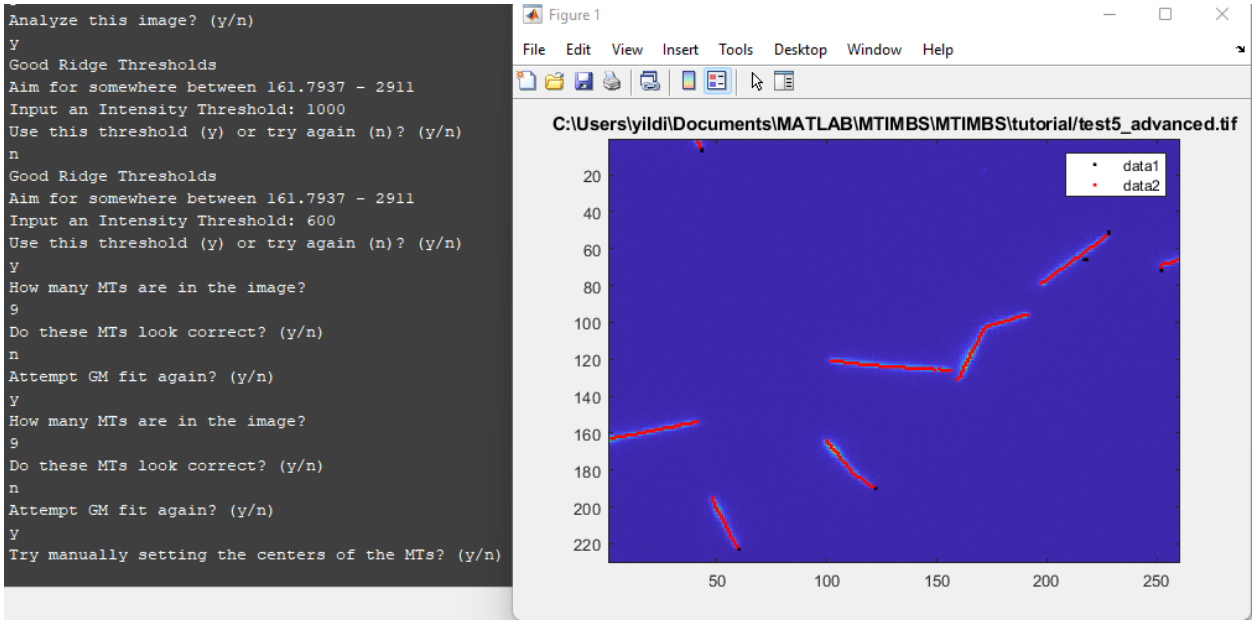
```
Good Ridge Thresholds
Aim for somewhere between 182.8382 - 2225
Input an Intensity Threshold: 800
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
y
Optimal MT # 2
Do these MTs look correct? (y/n)
y
Analyze this image? (y/n)
y
If you want to skip image analysis at any time, press x
Attempting previous threshold value
No ridge points found, lower your input value
Good Ridge Thresholds
Aim for somewhere between 120.1149 - 650
Input an Intensity Threshold: 190
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
y
Optimal MT # 4
Do these MTs look correct? (y/n)
y
Analyze this image? (y/n)
y
If you want to skip image analysis at any time, press x
Attempting previous threshold value
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
n
Good Ridge Thresholds
Aim for somewhere between 120.7392 - 1695
Input an Intensity Threshold: 800
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
y
Optimal MT # 5
Do these MTs look correct? (y/n)
n
Attempt GM fit again? (y/n)
y
How many MTs are in the image?
4
Do these MTs look correct? (y/n)
```



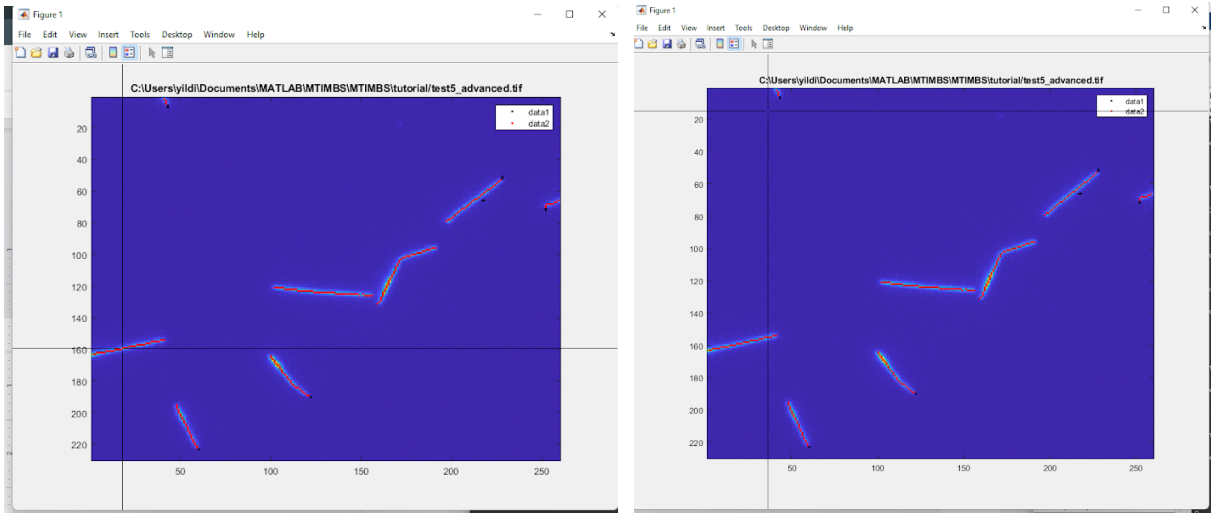
Well this isn’t fitting properly. What else can we do?

Step 7: Set initial MT centers to help break uniform prior

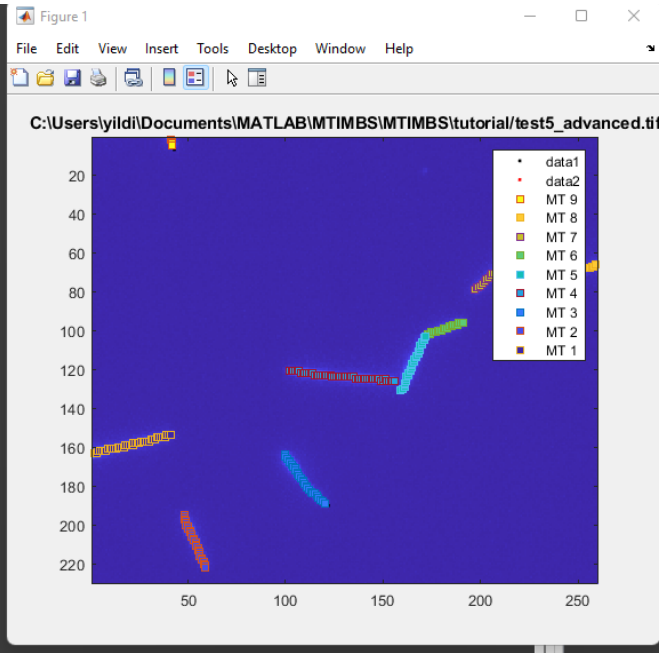
If Step 6 doesn’t work after a couple tries, you will see a new option pop up with the dialogue “Try manually setting the centers of the MTs?“ (currently set to on the third attempt and after, but changeable in the MTIMBS code)



If you enter ‘y’, you will be prompted to enter the number of MTs, but following you will be able to interactively point and click on the figure. Click once for each MT on the distinct MT (try and aim for the center, but just anywhere within the region should be sufficient). This will help the algorithm converge to the appropriate distribution.



```
Input an Intensity Threshold: 800
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
y
Optimal MT # 5
Do these MTs look correct? (y/n)
n
Attempt GM fit again? (y/n)
y
How many MTs are in the image?
4
Do these MTs look correct? (y/n)
n
Attempt GM fit again? (y/n)
y
Try manually setting the centers of the MTs? (y/n)
y
How many MTs are in the image?
4
Do these MTs look correct? (y/n)
y
Analyze this image? (y/n)
y
If you want to skip image analysis at any time, press x
Attempting previous threshold value
Remove unwanted points by clicking. Press Enter once done.
Use this threshold (y) or try again (n)? (y/n)
y
Optimal MT # 11
Do these MTs look correct? (y/n)
n
Attempt GM fit again? (y/n)
y
How many MTs are in the image?
9
Do these MTs look correct? (y/n)
n
Attempt GM fit again? (y/n)
y
Try manually setting the centers of the MTs? (y/n)
y
How many MTs are in the image?
9
Do these MTs look correct? (y/n)
```



With just a couple clicks, the MTs have been found successfully. This will work about 85% of the movies that make it this far. Save the MTs and move on to the last test (which will be left as an exercise for you)