In Silico simulation of smFISH images (How to use)

Before start

- Download the repository or the content of this folder.
- unzip compressed `.tif` image files.

Descriptions

This simulation pipeline is divided into 2 parts:

- 1st part: generation of experimental PSF/mRNA images repository.
- 2nd part: generation of synthetic smFISH images at different mRNA concentrations.

Requirements:

- Tested in MATLAB v R2018b or higher
  - Required toolboxes:
    Image Processing Toolbox
    Statistics and Machine Learning Toolbox

Set up:

Copy the scripts provided (`./scripts_simulation`) to the MATLAB folder (usually located in Documents directory).

Important:

Here, two simulation approaches (**app**) are described:

- Approach 1:
  Simulation is performed by placing experimental mRNA images (PSF) cropped from smFISH images.

- Approach 2:
  Simulation is performed by placing an average experimental mRNA image (average PSF) with the intensities of experimental mRNA images.

  - First, an average experimental PSF image is obtained and normalized.
  - Second, the intensity value of experimental mRNA images is multiplied by the normalized PSF and then placed.

Data provided

Generation of experimental PSF/mRNA images repository

- Approach 1 (location: `./simulation_inputs/mRNA_repositories/spots_for_approach1`):

  - mRNA positions table:

    ```
    .csv file with columns: "Image", "Pos_Y", "Pos_X", "Pos_Z", "Y_det", "X_det", "Z_det" (i.e.: obtained using
    FISH-quant).
    % Image: names should match images where mRNAs are extracted
    % Pos_Y, _X, and _Z: mRNA positions in nm
    % Y_, X_, and Z_det: mRNA position in px index
    ```

  - smFISH Images

- Approach 2 (location: `./simulation_inputs/mRNA_repositories/spots_for_approach2`):

  - mRNA positions table:

    ```
    .csv file with columns: "Image", "Pos_Y", "Pos_X", "Pos_Z", "Y_det", "X_det", "Z_det" (i.e.: obtained using
    FISH-quant).
    % Image: names should match images where mRNAs are extracted
    % Pos_Y, _X, and _Z: mRNA positions in nm
    % Y_, X_, and Z_det: mRNA position in px index
    ```

  - smFISH Images

Generation of synthetic smFISH images at different mRNA concentrations

- Approach 1

- - Binary image of an oocyte where mRNAs will be placed: "oocyte_to_model_16bits.tif" (location: ./simulation_inputs)
    - Repository of mRNA images: "mRNA_repository_approach1.mat" (location: ./simulation_inputs/mRNA_repositories)

- Approach 2

    - Binary image of an oocyte where mRNAs will be placed: "oocyte_to_model_16bits.tif" (location: ./simulation_inputs)
    - Average experimental mRNA image (average PSF): "experimental_single_mol_PSF_8x6_mRNA_AVG_ns" (location: ./simulation_inputs)
    - Repository of mRNA images: "mRNA_repository_approach2.mat" (location: ./simulation_inputs/mRNA_repositories)

User guide

1st part: generation of experimental PSF/mRNA images repository

Procedure:

1. Open MATLAB.
2. In the command window:

Approach 1	Approach 2
Type Script_to_generate_mRNA_repository_approach1 and then press enter	Type Script_to_generate_mRNA_repository_approach2 and then press enter

3. Select the folder containing the mRNA positions table.

Approach 1	Approach 2
The script will look for a .csv file with the identifier 'dissolved' in its file name. To change this: - Type in the command window: edit Script_to_generate_mRNA_repository_approach1 - Press enter. - Then, in the line # 9 change: 'dissolved' for anything you want. Example data location: ./simulation_inputs/mRNA_repositories/spots_for_approach1 Important: Image names should match between the table of positions and the image files.	The script will look for a .csv file with the identifier 'L44440' in its file name. To change this: - Type in the command window: edit Script_to_generate_mRNA_repository_approach2 - Press enter. - Then, in the line # 9 change: 'L44440' for anything you want. Example data location: ./simulation_inputs/mRNA_repositories/spots_for_approach2 Important: Image names should match between the table of positions and the image files.

4. Select the folder containing the smFISH image files.

Approach 1	Approach 2
The script will look for images with the identifier 'spn4' in their file names. To change this: - Type in the command window: edit Script_to_generate_mRNA_repository_approach1 - Press enter. - Then, in the line # 10 change: 'spn4' for anything you want. Example data location: ./simulation_inputs/mRNA_repositories/spots_for_approach1 Important: Image names should match between the table of positions and the image files.	The script will look for images with the identifier 'distal' in their file names. To change this: - Type in the command window: edit Script_to_generate_mRNA_repository_approach2 - Press enter. - Then, in the line # 10 change: 'distal' for anything you want. Example data location: ./simulation_inputs/mRNA_repositories/spots_for_approach2 Important: Image names should match between the table of positions and the image files.

5. Select the folder where to save the repositories.

The default parameters of the script include:

```
    % ============ voxel size in nm (*)
    pixel_size.xy = 49;
    pixel_size.z = 250;
    % (*)This is the voxel size of the example images.  It is used to compute the locations of the mRNA spots from
 the positions table (see step 3) when the positions are given in nm.
```

To change this:

- Type in the command window:

  ```
  edit Script_to_generate_mRNA_repository_approach1
  ```

  or

  ```
  edit Script_to_generate_mRNA_repository_approach2
  ```
- Press enter.
- Then, in the line # 5 and # 6, default values can be changed.

2nd part: generation of synthetic smFISH images at different mRNA concentrations

Procedure:

Note: by default, the script requires a lot of time as it generates 25 different concentrations with 3 replicates in a whole oocyte. To test the script see the **Recommended parameters to test** section below.

6. Open MATLAB
7. In the command window:

Approach 1	Approach 2
Type `Script_to_generate_synthetic_images_approach1` and then press enter	Type `Script_to_generate_synthetic_images_approach2` and then press enter

8. Select the file of the binary image where random placement will be performed.

   - Provided image example:
     `oocyte_to_model_16bits.tif`

     location: `./simulation_inputs`

9. Select the file of average PSF image (**only for approach 2**)

Approach 1	Approach 2
Step not performed	Provided image example: `experimental_single_mol_PSF_8x6_mRNA_AVG_ns` location: `./simulation_inputs`

10. Select the file of the mRNA images repository.

Approach 1	Approach 2
Provided example: `mRNA_repository_approach1.mat` location: `./simulation_inputs/mRNA_repositories`	Provided example: `mRNA_repository_approach2.mat` location: `./simulation_inputs/mRNA_repositories`

The default parameters of the script include:

```
% ============= voxel size
pixel_size.xy = 49;
pixel_size.z = 250;

%============= expected concentrations in mol/um3
concentration_desired = [0.01  0.025:0.025:0.1   0.25:0.25:1  1.25:0.25:5];
% number of replicates by concentration
replicates_simulation = 3;
% reduce PSF
additional_BGD_to_substract_mRNA = 34;
```

Recommended parameters to test

```
% ============= voxel size
pixel_size.xy = 49;
pixel_size.z = 250;
```

```
%============= expected concentrations in mol/um3
concentration_desired = [0.025];
% number of replicates by concentration
replicates_simulation = 1;
% reduce PSF
additional_BGD_to_substract_mRNA = 34;
```

To change default parameters:

- Type in the command window:

  edit Script_to_generate_synthetic_images_approach1

  or

  edit Script_to_generate_synthetic_images_approach2

- Press enter.
- Then, modifications can be made and saved.

Outputs:

In Silico simulated images.

Provided output examples:

Example simulations with an expected concentration of 0.025 mol/μm3 (location: ./simulated_images_examples)

Note: expected concentrations are used as inputs. However, molecules placed very close to the oocyte edges whose PSF does not fit within the oocyte area are removed. Therefore, the actual concentration will be lower when the number of molecules placed is very high. The script will give the actual number of molecules placed.