# PyHoleFinder v1.0 User's manual

Tohru Terada

Department of Biotechnology, The University of Tokyo

tterada@bi.a.u-tokyo.ac.jp

# PyHoleFinder

- PyHoleFinder is a python program to find holes in holey carbon films on a cryo-electron microscope (cryo-EM) grid.
- This program can work with SerialEM.
- In this manual, I explain how to install this program and how to use it with SerialEM.

# System requirements

- This program was developed and tested on a Windows-10 PC with Anaconda 3.18.9 (Python 3.7.4).
- If you do not have Anaconda, first install it.

# Install (1)

1. Open Anaconda Prompt and install following packages by running:

```
> pip install opencv-python opencv-contrib-python mrcfile PySimpleGUI
```

2. Download PyHoleFinder-main.zip from https://github.com/tterada-utokyo/PyHoleFinder

3. Unzip it to a folder you choose.

4. Change the folder name to PyHoleFinder.

# Install (2)

## 5. Modify script.txt in the PyHoleFinder folder as follows:

```
ScriptName PyHoleFinder

ReportIfNavOpen
If $reportedValue1 != 2
    Exit
Endif
ReportNavFile 1
navfile = $reportedValue1$reportedValue2
navdir = $reportedValue3
SetDirectory $navdir

RunInShell C:¥Users¥tterada¥Anaconda3¥python.exe
C:¥Users¥tterada¥Desktop¥cryoEM¥PyHoleFinder¥PyholeFinder.py -conf config.txt -nav $navfile
-out output.nav

MergeNavFile output.nav
```

Change the folder name to that where you installed Anaconda.

Do not break the line.

Do not break the line.

Change the folder name to that where you installed PyHoleFinder.

# Install (3)

7. Start SerialEM.
8. Select "Script"→"Edit"→"Edit 8" (select unused slot).
9. Copy the contents of script.txt and paste them into the editor.
10. Click "OK"

# Usage (1)

1. Open a SerialEM Navigator file by selecting "Navigator" →"Read & Open…" from the SerialEM menu.
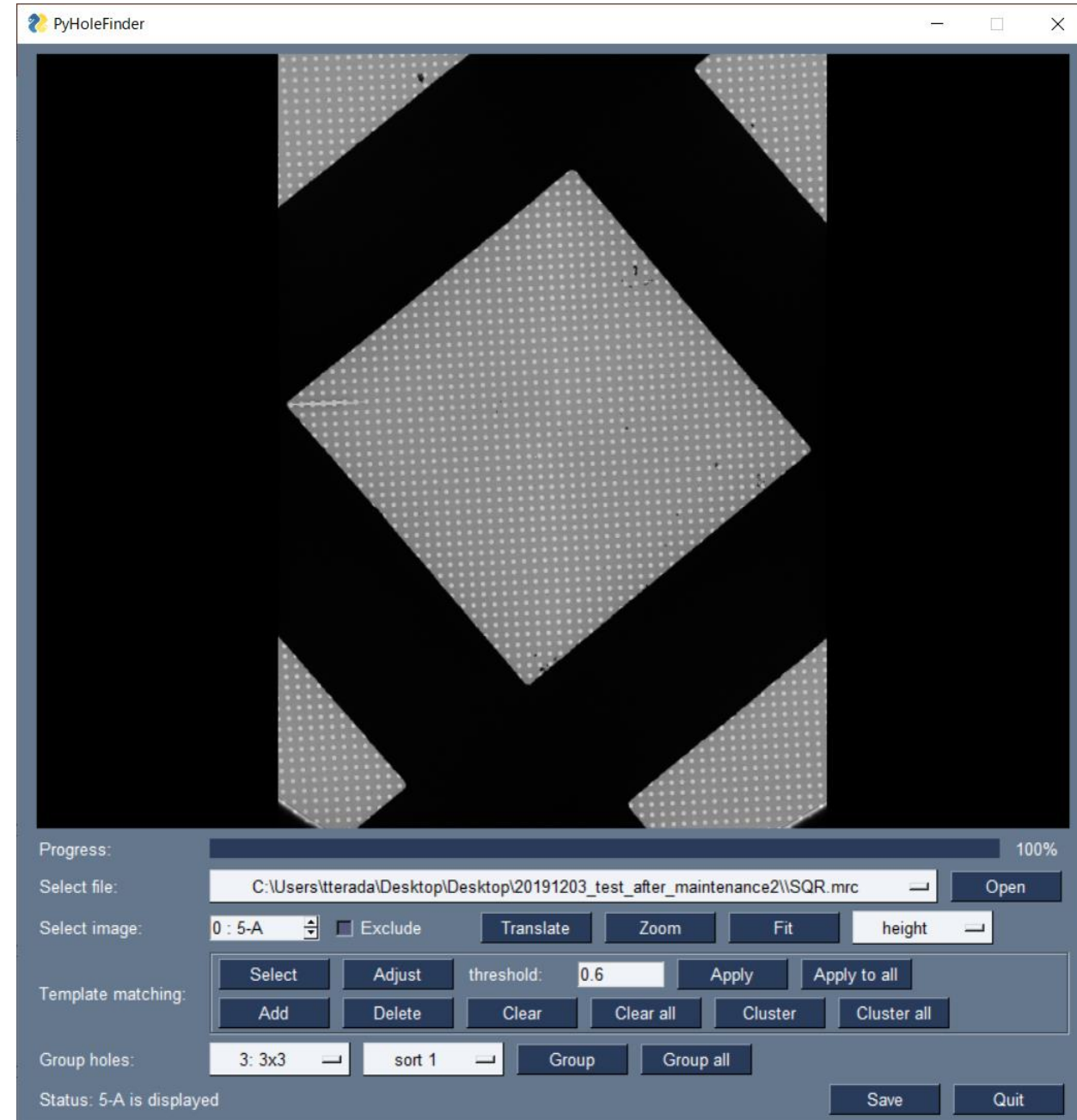
# Usage (2)

2. Select "Script"→"Run"→
   "8: PyHoleFinder" or the slot
   you edited.

   • PyHoleFinder is started.

# Usage (3)

3. Select a MRC file to analyze in "Select file" and "Open"

   - If the program complains the file does not exist, "Quit" the program, save the Navigator file in SerialEM to update the file path, and start again.
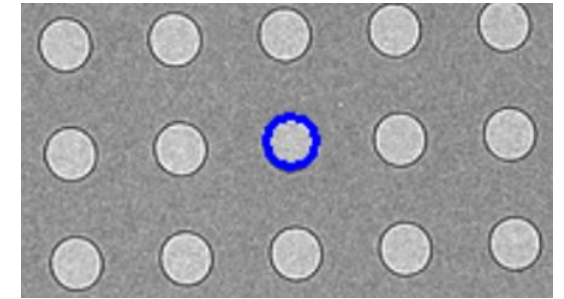
4. Select an image to display in "Select image"

# Usage (4)

5. Click "Select" in "Template matching"
   - The region used as the template is enclosed by a blue square.
6. Click "Adjust" to adjust the region of the template
   - Arrow keys to move
   - + to enlarge and – to shrink
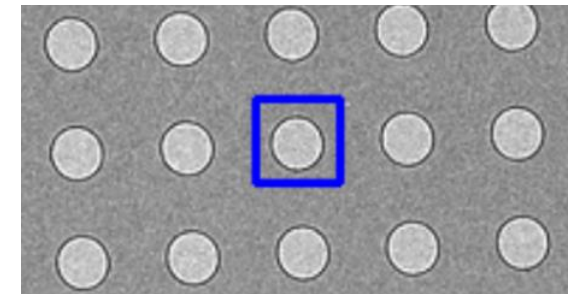   - ESC to set and exit

Click the center of the hole used as the template
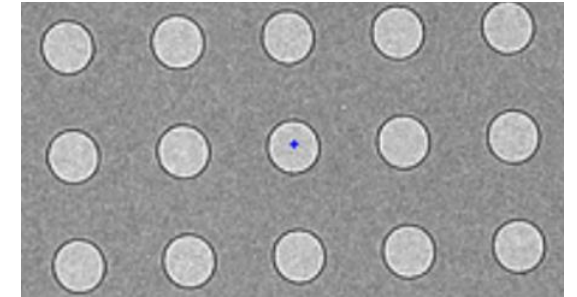


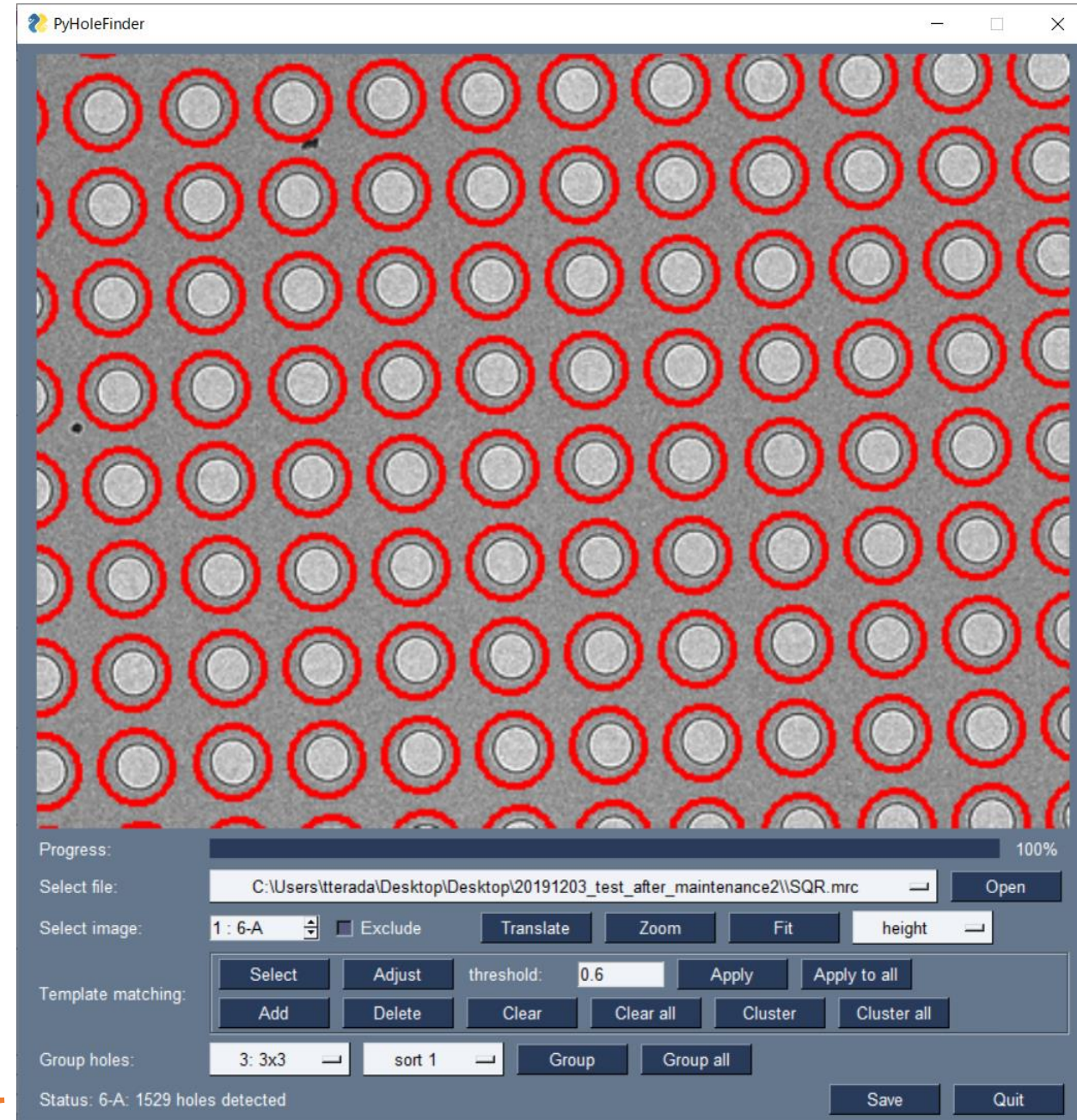Drag until the circle overlaps the rim of the hole



Template region is determined

# Usage (5)

7. Adjust the threshold value (0.4–0.6) and click "Apply" or "Apply to all"

- Detected holes are enclosed by red circles.
- Holes near grid bars are ignored.
- Checking the "Exclude" checkbox excludes the image(s) from the analysis.
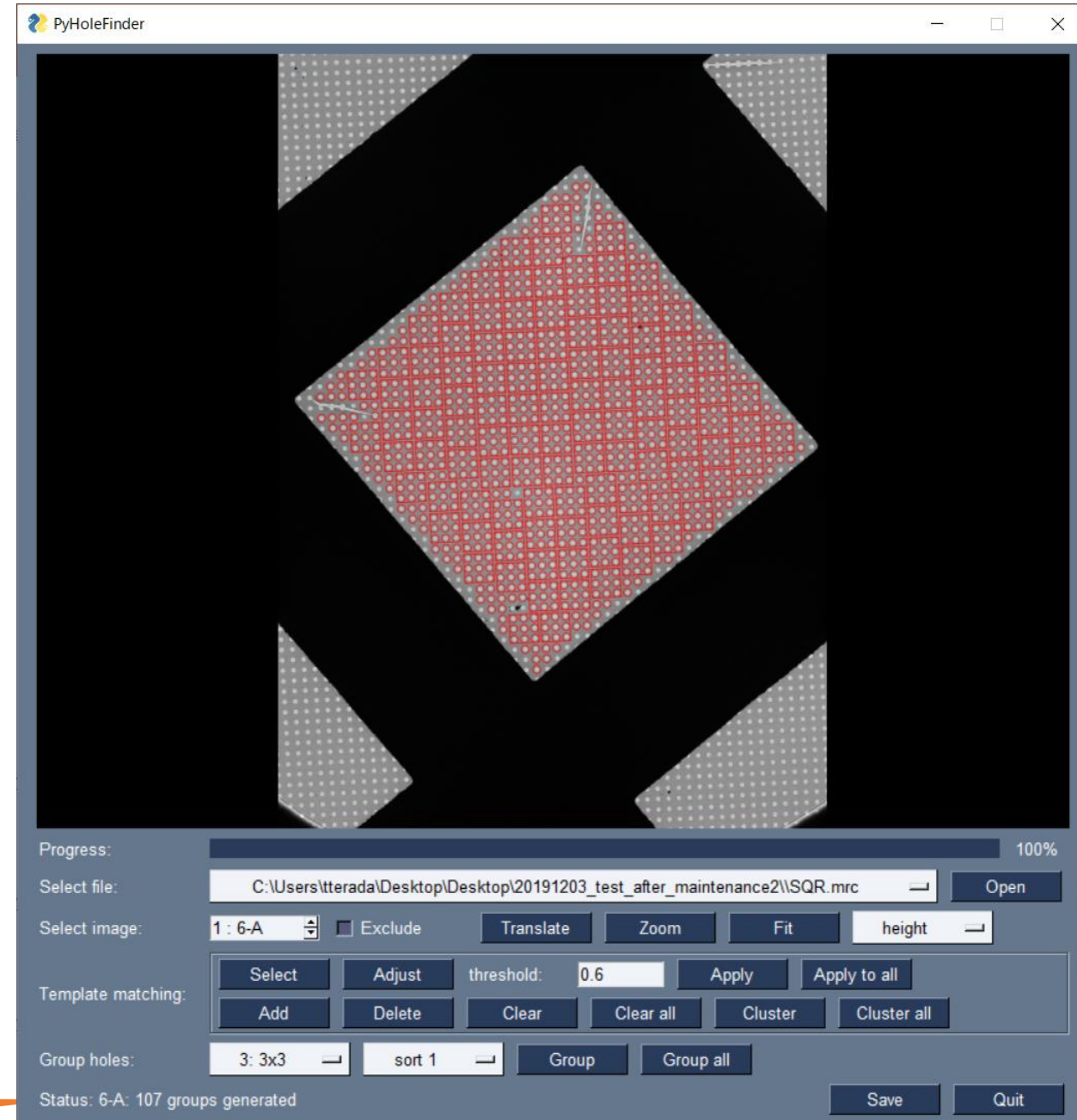
The number of holes is shown.

# Usage (6)

8. Divide the holes into 3x3 or 5x5 groups by "Group" or "Group all" in "Group holes"

   - Each group is enclosed by a red square.
   - Checking the "Exclude" checkbox excludes the image(s) from the analysis.
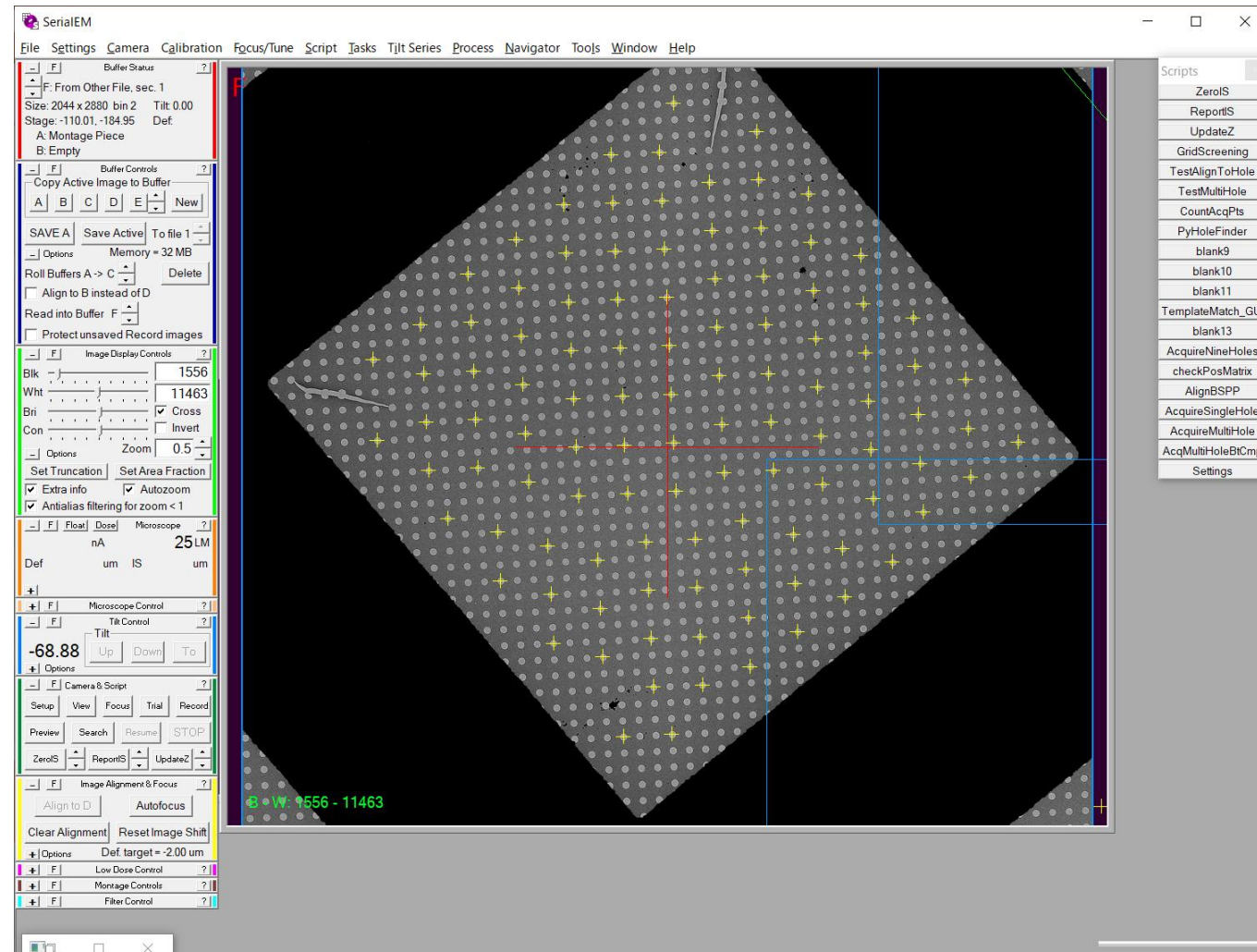
The number of groups is shown.

# Usage (7)

9. Click "Save" and "Quit" to incorporate the results in SerialEM.
   - Yellow cross marks indicate the centers of the holes at the centers of the groups.
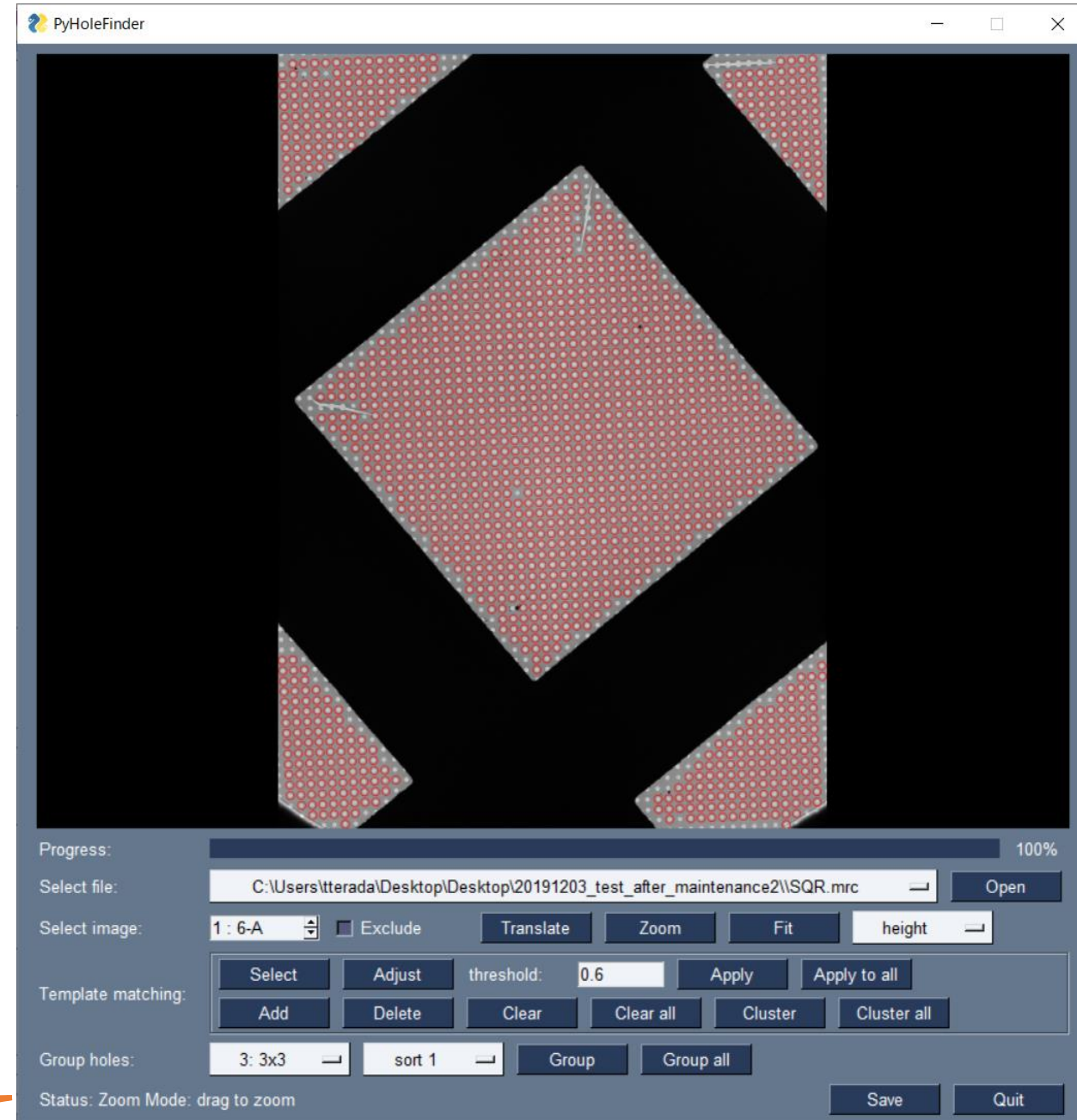
# Functions (1)

- Detection of the holes near grid bars are avoided by applying a mask to the image.

- The mask is created by binarization of the image (A).

  - Noises in the square regions are removed by applying dilate function (B).

  - The bar region is extended by applying the erode function several times (C).



Original image (A) (B)

(C) Masked image
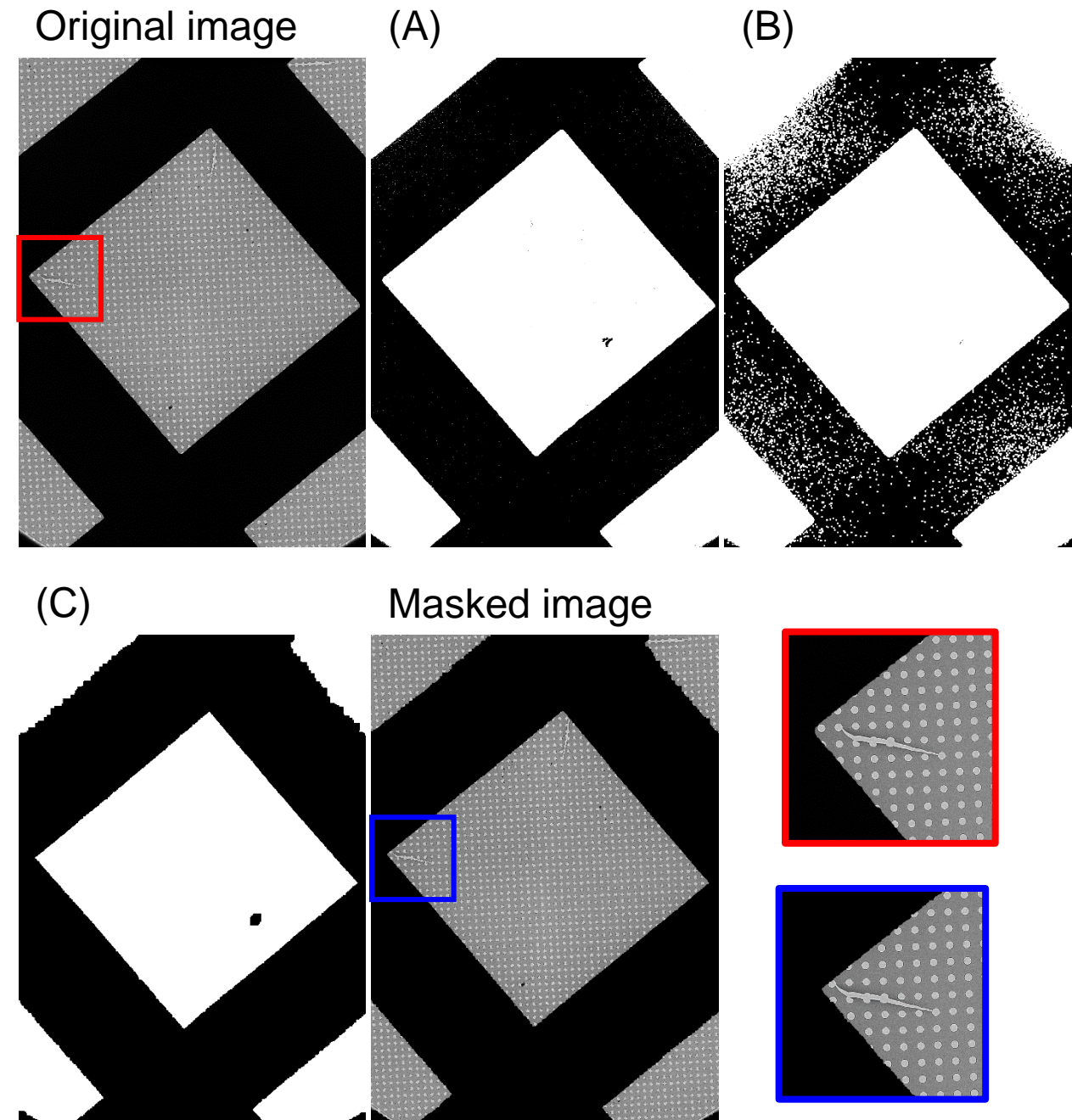
# Functions (2)

- Drag the mouse after clicking "Translate" or pressing T to translate the image.

- Drag the mouse after clicking "Zoom" or pressing Z to zoom the image.

- Turning the mouse wheel also zooms the image.

- Clicking "Fit" fits the height or the width of the image to the canvas.

Zoom mode.

# Functions (3)

- Click "Add" or press A to add a hole.
- Click "Delete" or press D to delete a hole.
- Click "Clear" to delete holes in the current image.
- Click "Clear all" to delete all holes in all images.
- These functions are enabled after holes are detected.
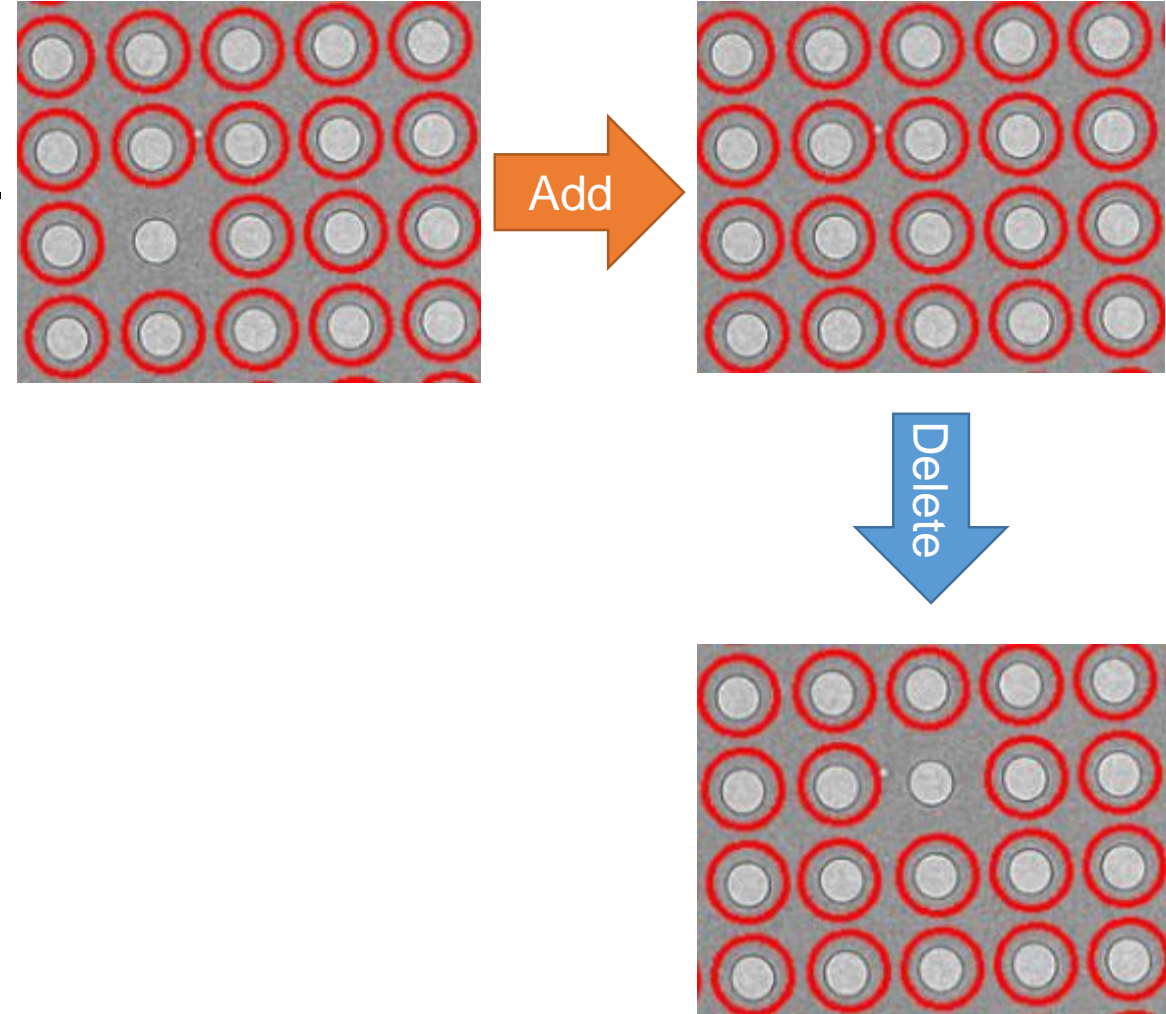
# Functions (4)

- By clicking "Cluster," holes are subjected to clustering and the holes in the largest cluster, which are usually those in the central square, are retained.
- "Cluster all" performs clustering on all the images.
- When "Group" or "Group all" is clicked, clustering and retaining the largest cluster are performed in advance of grouping.

# Functions (5)

- In the grouping, holes are sorted according to their coordinates.
- Groups are formed in the sorted order.
- The sort order is controlled by the option (sort 0–15)
  - Arrows indicate the sort order.
  - In sort 8–15, the sort order along the perpendicular direction of the arrow is reversed.



sort 0, 8

sort 7, 15

sort 1, 9

sort 6, 14

sort 2, 10

sort 5, 13

sort 3, 11

sort 4, 12
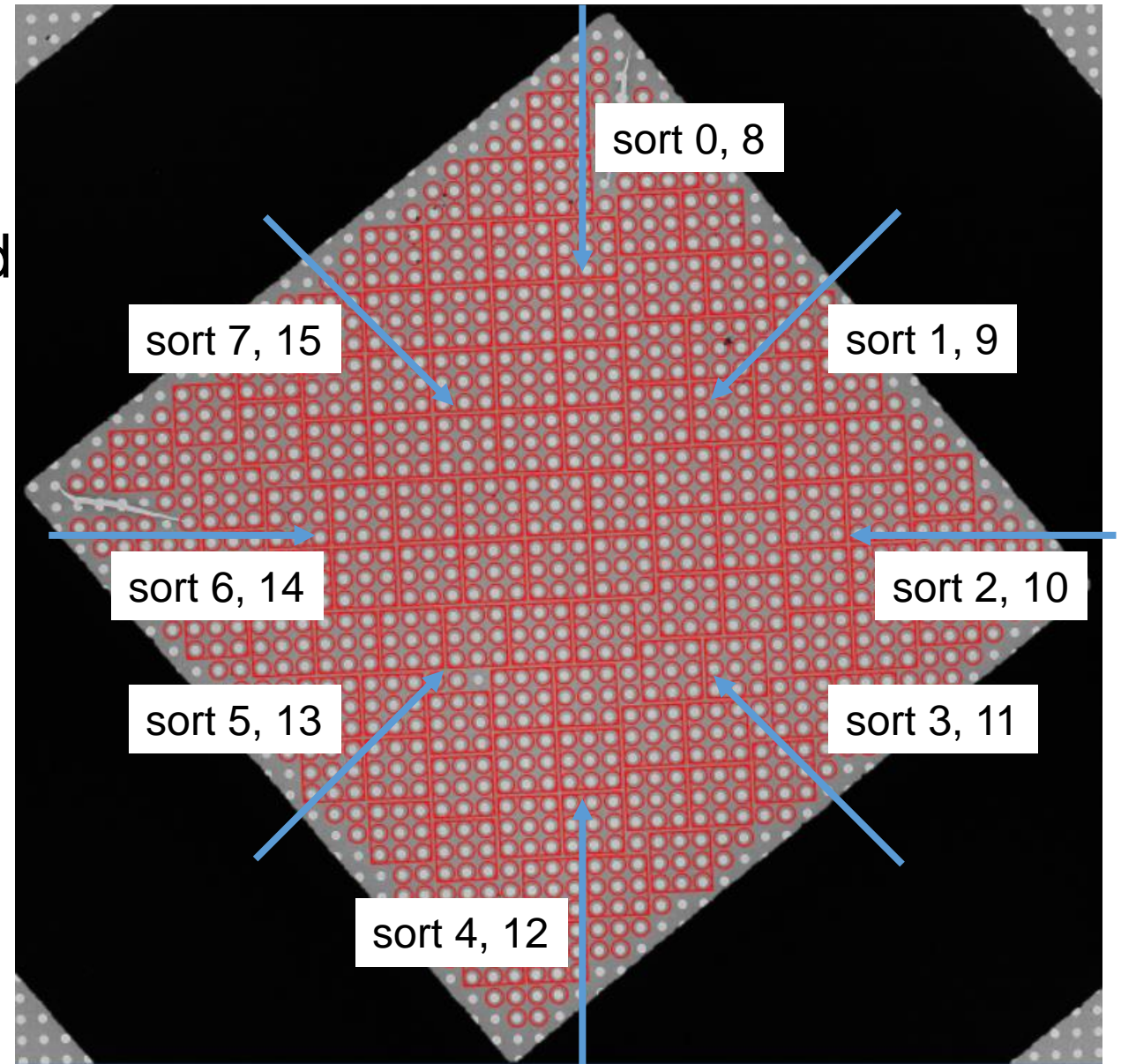
# Functions (6)

- Clicking "Save" after detecting save the center coordinates of all the detected holes.
- Clicking "Save" after clustering save the center coordinates of the holes in the largest clusters.
- Clicking "Save" after grouping save the center coordinates of the holes at the centers of the groups.
- If a file with the output file name exists, this file is renamed by adding ".1" to the end of the file name when the program starts.
  - If the ".1" file exists, the file is renamed by adding ".2", and so on.

# Functions (7)

- Default values of the parameters are described in the file specified by -conf.

  - In script.txt, config.txt is specified.

- If the file specified by –conf does not exist, parameter values described in config_default.txt in the PyHoleFinder folder are used.

# Parameters (1)

| | |
|---|---|
| xsize = 800 | Width of the image canvas |
| ysize = 600 | Height of the image canvas |
| fit_mode = width | Fitting direction |
| MRC_histogram_bins = 2000 | N of bins of brightness histogram |
| MRC_dark_cut = 0.005 | Ratio of dark pixels below cut-off |
| MRC_bright_cut = 0.005 | Ratio of bright pixels above cut-off |
| mask_threshold = 10 | Threshold of grid bar mask |
| mask_kernel_size = 10 | Kernel size of the mask |
| mask_iterations = 5 | N of iterations of applying erode |

# Parameters (2)

| | |
|---|---|
| hole_radius = 10 | Unused |
| hole_padding = 6 | Padding added to hole radius |
| line_width = 2 | Width of lines |
| template_padding = 6 | Padding added to template |
| TM_threshold = 0.6 | Template matching threshold |
| cluster_threshold = 100 | Threshold used in clustering |
| gsize = 3 | Group size (3 or 5) |
| sort = 1 | Sort order |

# Parameters (3)

| | |
|---|---|
| circle_color = 255 0 0 | Color of circles enclosing holes |
| group_color = 255 0 0 | Color of squares enclosing groups |
| template_color = 0 0 255 | Color of circle and sqare enclosing template |
| navi_color = 3 | Color of cross marks in SerialEM |

# Acknowledgements