

Im2graph User's Manual

image2graph@gmail.com

December 25th, 2016

Version 1.21



About im2graph

im2graph is a digitizing software that converts graph-images to graph-data, that is to numbers. With behind the scenes fast and efficient image processing algorithms, im2graph effortlessly converts images to graphs with a click of a button. For more complex use cases, a highly tuned user-assisted GUI is available, enabling fast and effortless conversions from graph images to values. Using im2graph is intuitive, addictive and fun and requires very little knowledge or effort on your part. im2graph is the OCR equivalent of graphs— instead of converting scanned images to text, im2graph converts scanned graphs to data.

Im2graph Free and Premium

im2graph comes in two versions: a Free version and a Premium version.

im2graph Premium offers enhanced functionality compared with the Free version, including:

- Optical Character Recognition (OCR)
- Multiple graph support
- Unlimited undo/redo capability
- Export options for multi-graph support
- Free updates

im2graph is available for Windows and Linux.

License Agreement – im2graph

im2graph comes in two versions: a Free version and a Premium (non-free) version. The free version of im2graph is free for personal use. If you use im2graph for non-personal use, or if you would like the additional enhanced functionality of the premium version, you should purchase the software online.

See the website (<http://www.im2graph.co.il>) for price and terms and conditions.

im2graph (both free and premium versions) are provided with absolutely no guarantee whatsoever; nor fitness for any particular task; use it at your own risk.

im2graph is protected by copyright laws. The software copyrights are owned by Shai Vaingast, the author. You may not transfer the software without proper written consent from the author.

No modification to im2graph is allowed.

Copyright (C) 2014- Shai Vaingast

website: <http://www.im2graph.co.il>

Email: shai@im2graph.co.il

All rights reserved.

Installing im2graph

Windows

Download and execute the installation binary **im2graph*-win32.exe**; follow the installation instructions.

Note: Some anti-virus applications interfere with the installation file. Try to temporarily disable them if that happens.

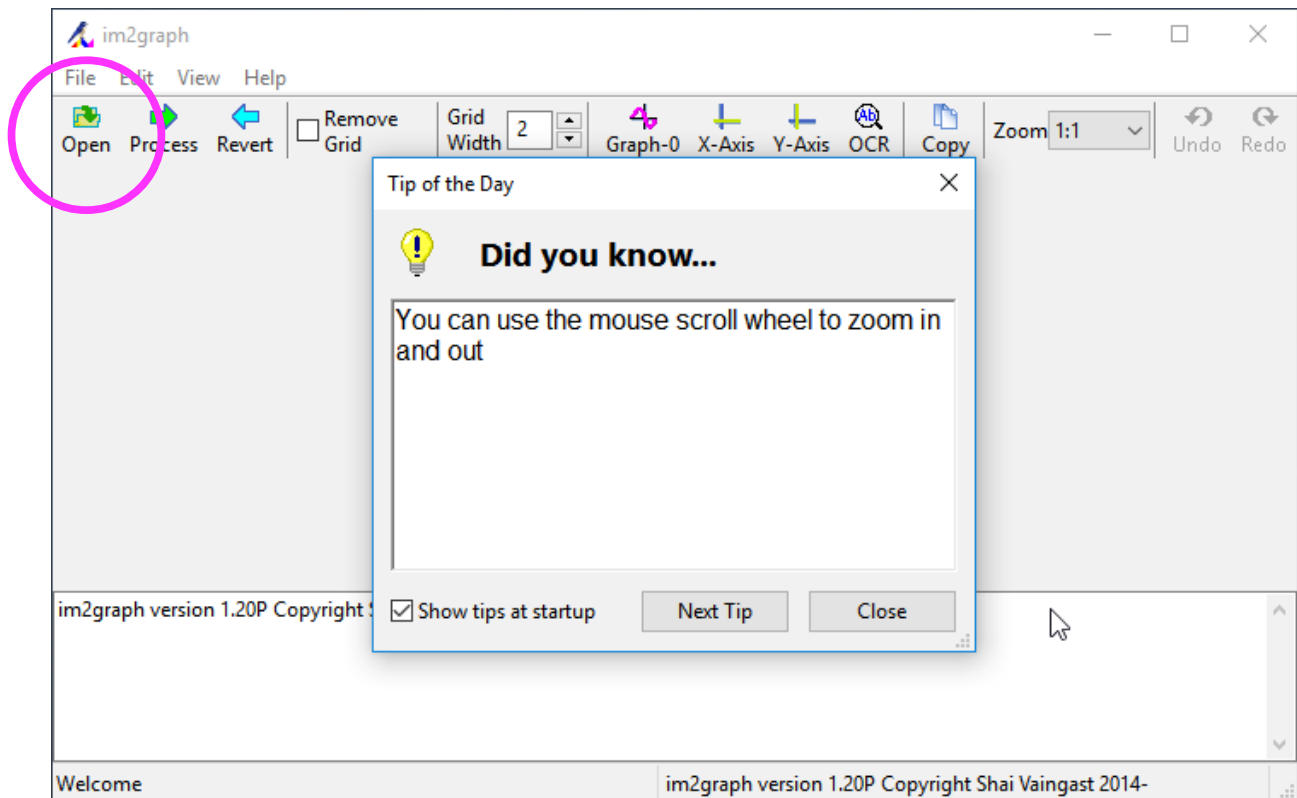
Linux

1. For the Premium version only
 - a. Install wxWidgets and tesseract packages:
Debian/Ubuntu: **apt-get install libwxgtk3.0 tesseract-ocr**
Fedora: **dnf install compat-wxGTK3-gtk2 tesseract**
2. For the Free version only
 - a. Install only the wxWidgets package:
Debian/Ubuntu: **apt-get install libwxgtk3.0**
Fedora: **dnf install compat-wxGTK3-gtk2**
3. Download **im2graph*.tar.gz**. Be sure to choose your OS's distribution.
4. Unpack **im2graph*.tar.gz** to a directory of your choice:
tar zxvf im2graph*.gz.
5. Run im2graph in the newly created directory.

Getting Started

Loading an Image

Your first step is to load an image into im2graph. This can be done by clicking on the Open button from the application's toolbar:

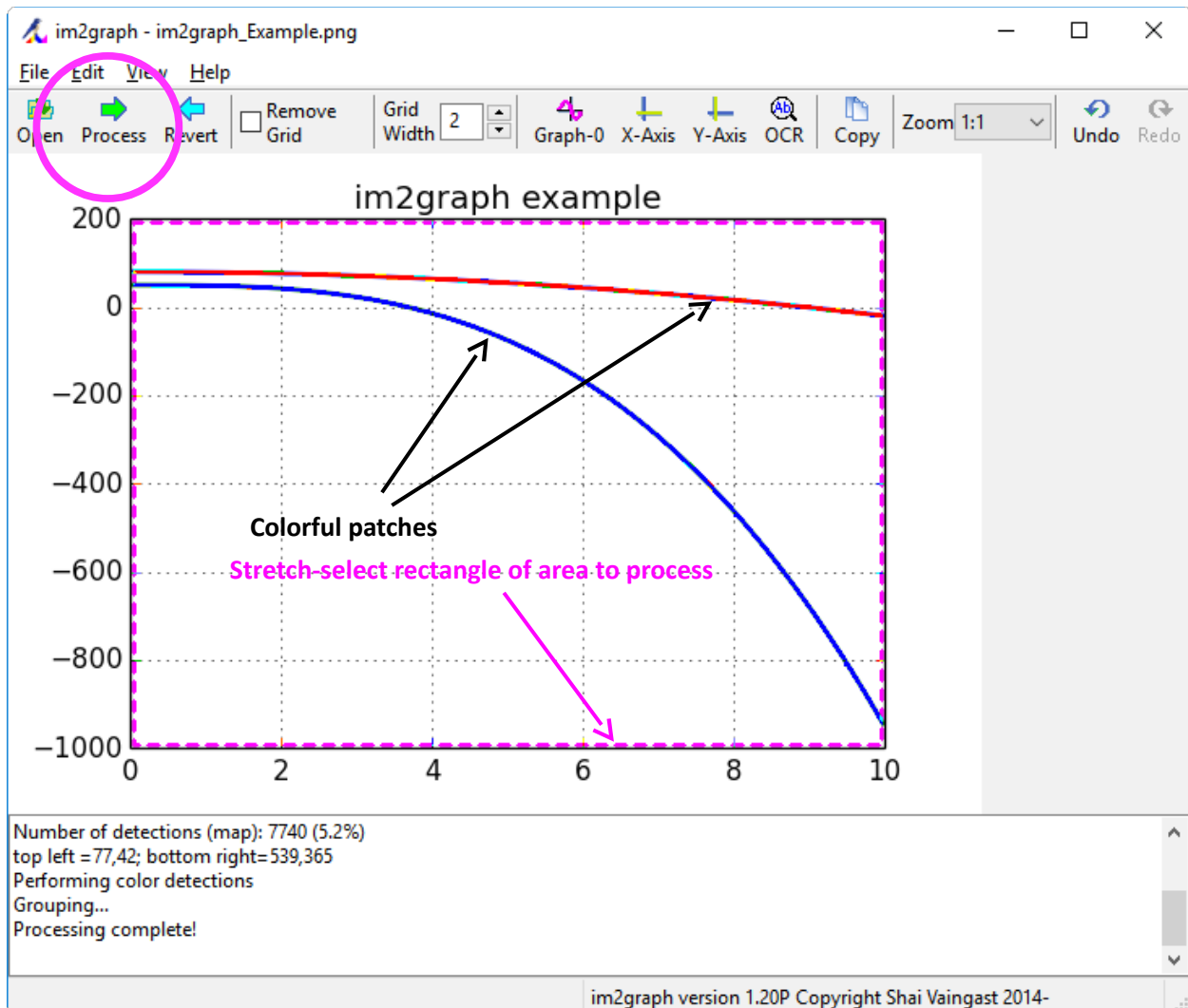


An Open File dialog will appear, enabling the selection of an image from file. A considerable amount of image formats is available including BMP, PNG, JPG, TIF, GIF and many more.

You can also load an image from the menu: File | Open or clicking on Ctrl+o.

Processing an Image (F2)

Once you've loaded an image, the next step is to process the image. This can be done by clicking on the Process button from the applications' toolbar or the menu; you can also press the F2 key. Processing an image does several things. First, a threshold is applied to the image. Next, colorful patches are automatically detected and colored accordingly. And lastly, a stretch selection rectangle is presented representing the area to be processed.



Updating the Area to Process

The next step is to observe the results. First, update the stretch-select rectangle (the magenta dashed rectangle) to include only areas you wish to process and perform the processing again (click Process). When you hover the mouse over the stretch select rectangle the mouse icon will change to indicate you can resize the stretch-select rectangle. Do so until the area you want to process is included within the stretch-select rectangle. You can also change the stretch select area by pressing the CTRL key while clicking on the left mouse and dragging the stretch selection rectangle.

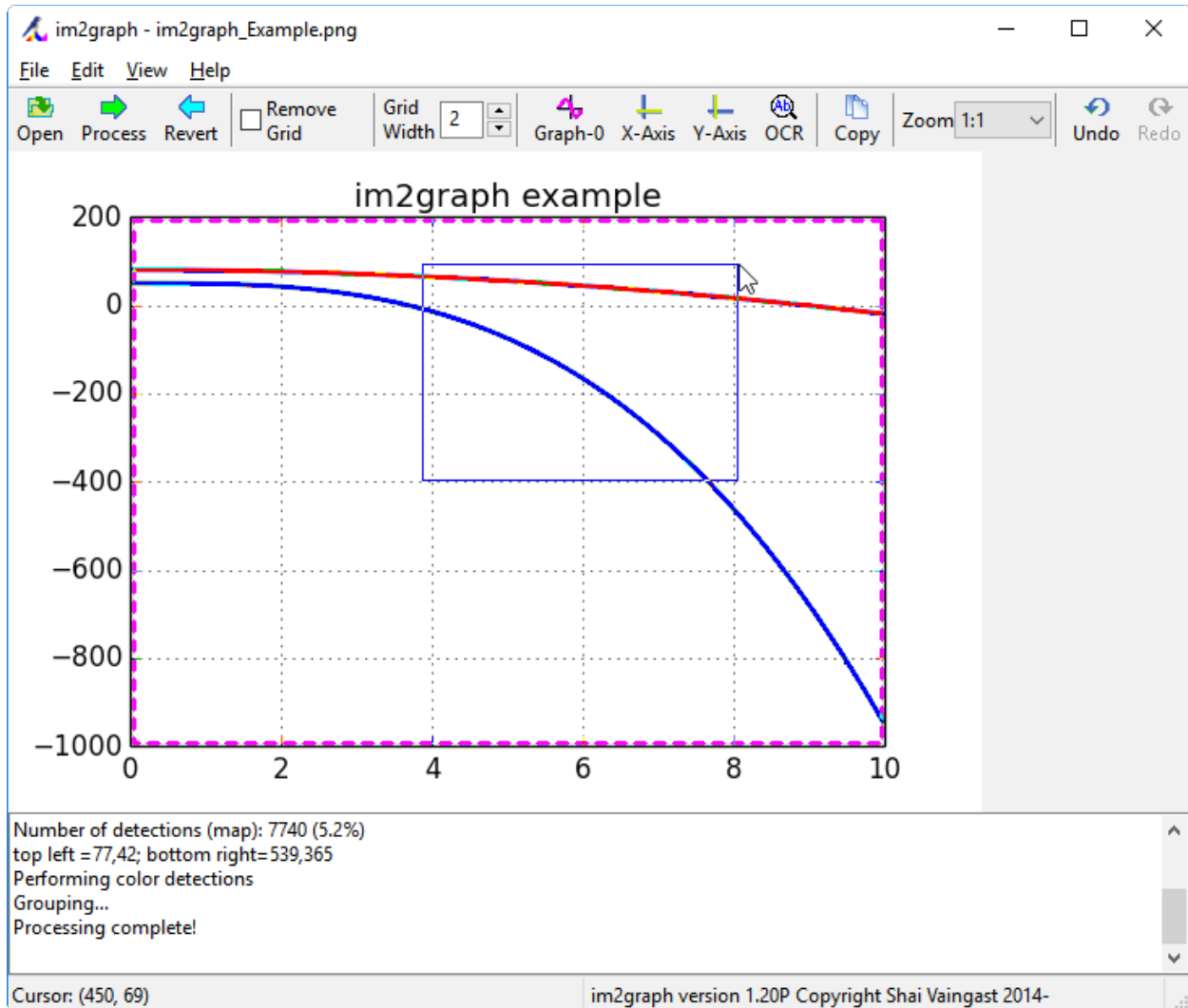
Once you're satisfied with the result, make sure that the graph to be digitized is detected as patches.

Note: You can use the stretch-select rectangle iteratively, that is, select an area to process, process the image, select patches (see below) and then select another area to process and so on.

Selecting the Graph

Once you have the colorful patches after processing, you can select the graph you wish to digitize. There are several ways you can do this: either by left-mouse clicking on a patch, or by creating a selection frame. All the selected patches represent the graph to be digitized. Selected patches change colored in magenta.

When you hover over a patch, the patch changes color and width to a wide magenta patch. This is to inform you that you can select the patch. Once you select the patch, the patch changes to a proper width and colored in magenta.



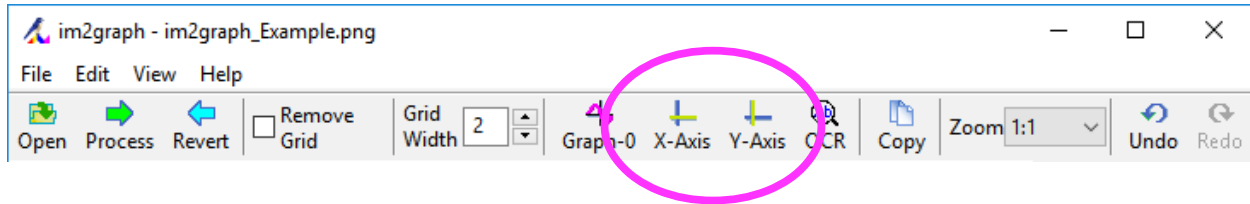
The selection frame is especially useful when you have many small patches to select. To create a selection frame, left mouse click on an empty area of the image (one that does not include colorful patches) and drag it to create a selection frame. A thin blue rectangle will appear. All patches that are fully or partly inside the selection frame will be selected (and turn magenta).

Deselecting Patches

If you'd like to deselect a patch, right mouse click on that patch. You can also deselect patches by creating a deselection frame. This is done by right mouse clicking on an area with no patches and dragging the mouse to change the frame (thin red frame). All patches that are entirely or partially in the deselection frame will be deselected and return to their previous color.

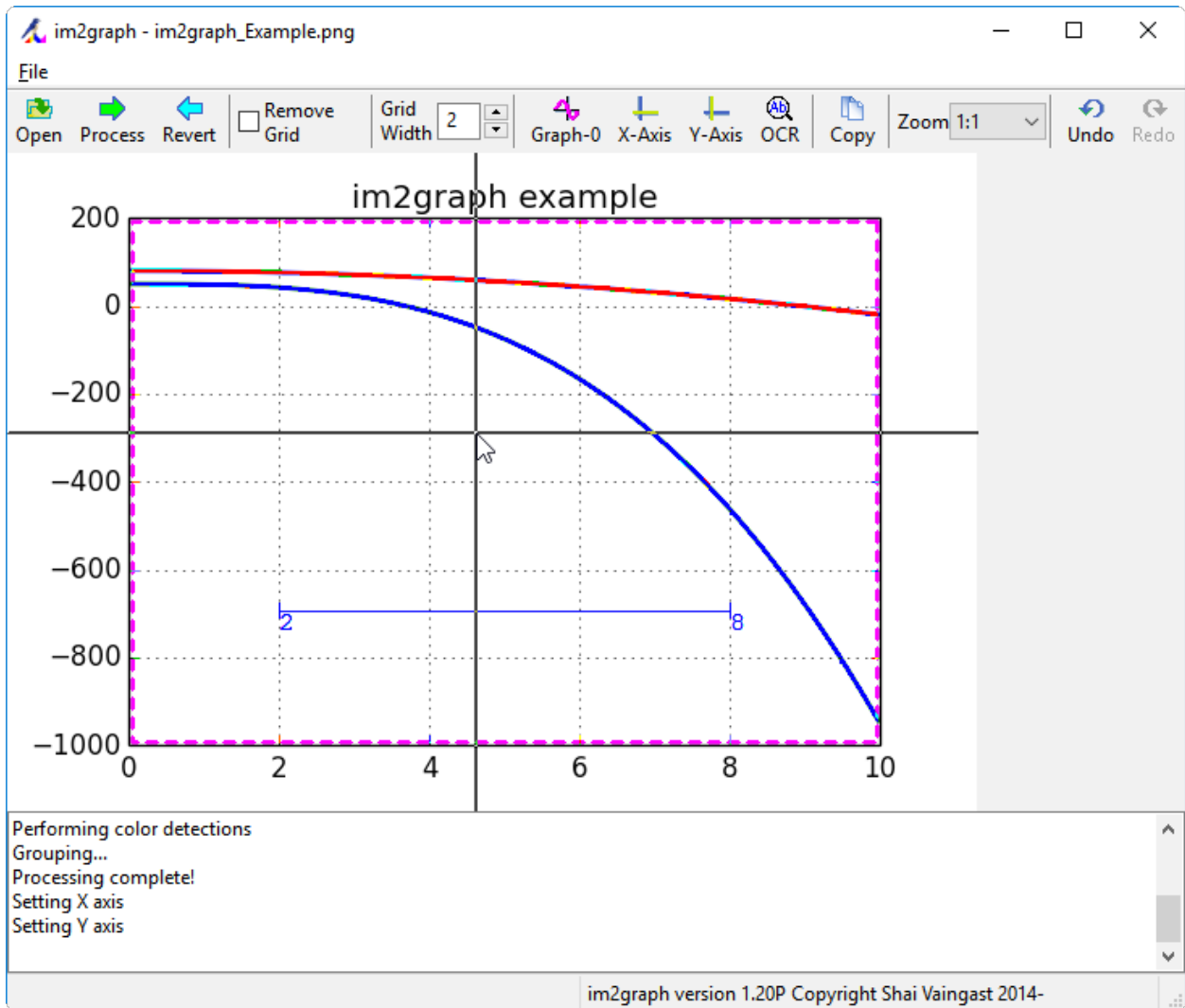
Setting the Scales (optional)

Before digitizing the graph, you should provide the X and Y scales of the graph. Note: this is optional, if you do not provide the scales, the values digitized will correspond to the pixel location. So if you require a scale, it's best to define it.



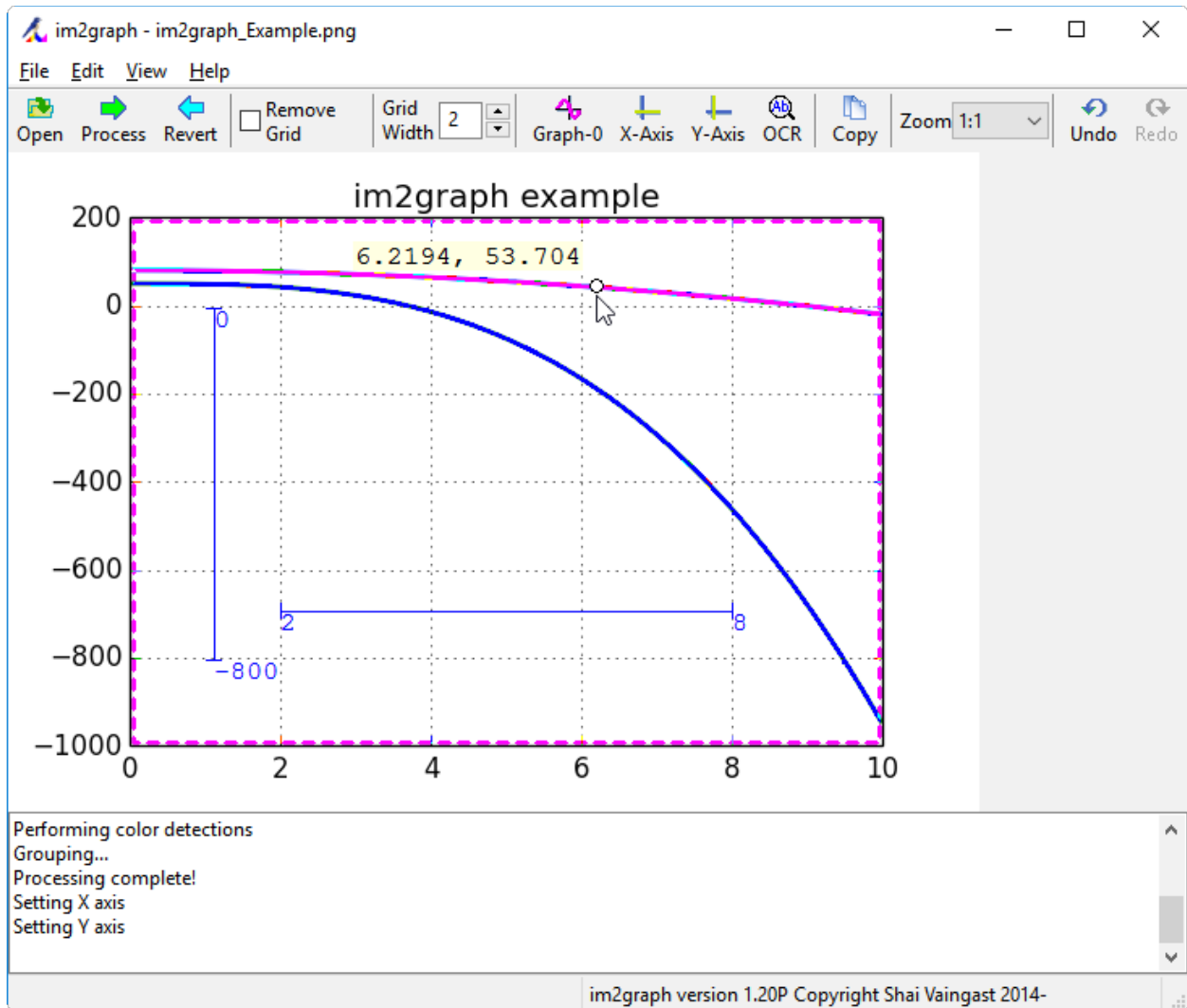
Setting the X-Axis scale is done by clicking on the X-Axis button from the toolbar (or pressing the x key). Once you click on it, a crosshair will appear on the image. Select an arbitrary point for which you know its X-value (hence: X_0) and left-mouse click. A dialog will appear asking for the value of X_0 . Enter the value of X_0 . A new crosshair will appear, this time it will be limited in motion to horizontal movement. Select another arbitrary point for which you know its X-value (hence: X_1) and again left mouse-click and enter the value of X_1 into the dialog. Once you're done, a scale will appear in blue, on the image with the values X_0 and X_1 . To abort setting scales operation, press the Esc key at any time.

Setting the Y-value is similar (shortcut: pressing the y key). First click the Y-Axis button and follow the process entering values Y_0 and Y_1 (note that now the second crosshair will be limited in motion to horizontal movement).



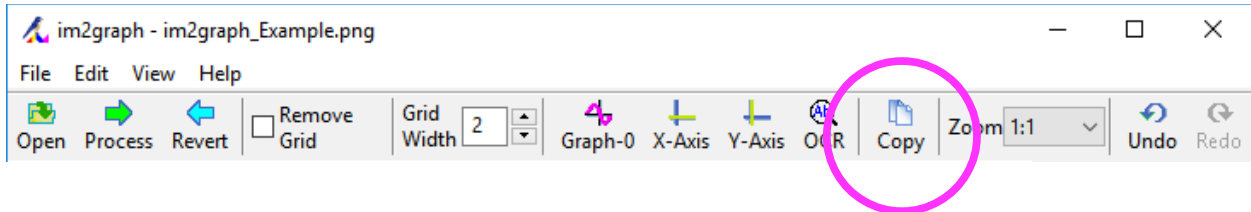
Display of Graph Values (Tooltips)

Once you place your mouse over a selected patch, a tooltip will display the graph values. In case the x-axis and y-axis values are set, the tooltip will use these to display the location in graph coordinates. If either one of these values is set and not the other, the unset value will be displayed as a question mark. If both are unset, the displayed values will be the image coordinates (X and Y).



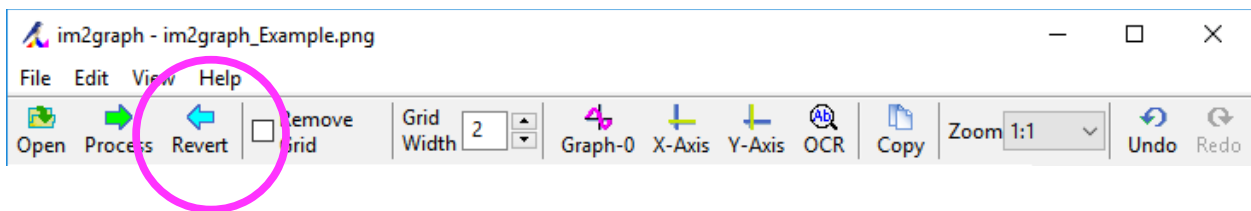
Digitizing the Graph

Finally, transfer the digitized data to file or clipboard. To transfer to clipboard, simply click on Copy in the toolbar. To save the file as a CSV file, click on File|Export or click Ctrl+s. Please note that copying to clipboard is available in the Premium version only. The File|Export option is available for both the Premium and Free version.



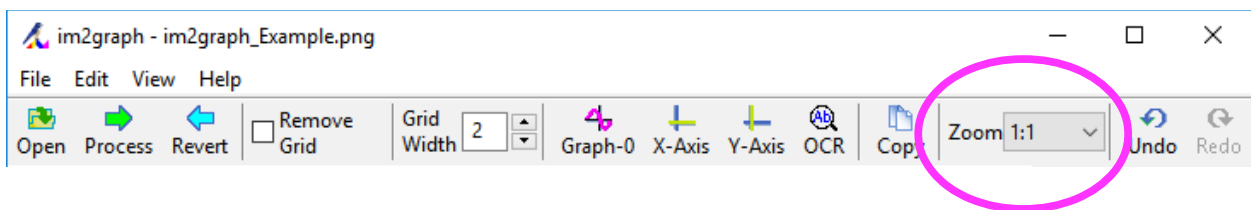
Revert (F9)

You can always revert to the original image, discarding everything you've done by clicking on this button.



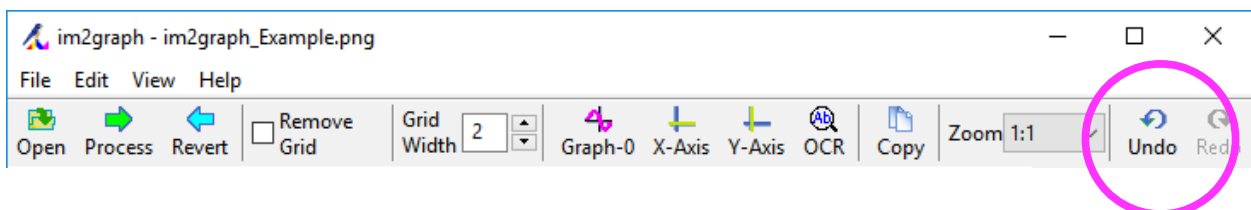
Zooming In and Out

You can zoom in and out for a finer control of the patch selection or to better view graph details. This is done with the mouse scroll wheel or using the zoom control in the toolbar. When you zoom with the mouse wheel, the zoom algorithm zooms around the mouse current location if possible.



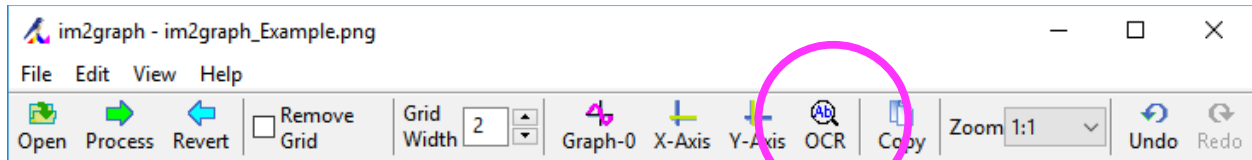
Undo/Redo (Premium only)

You can undo and redo your selections and operations with the Undo and Redo buttons available at the toolbar or use Ctrl-z for undo and Shift-Ctrl-Z for redo.

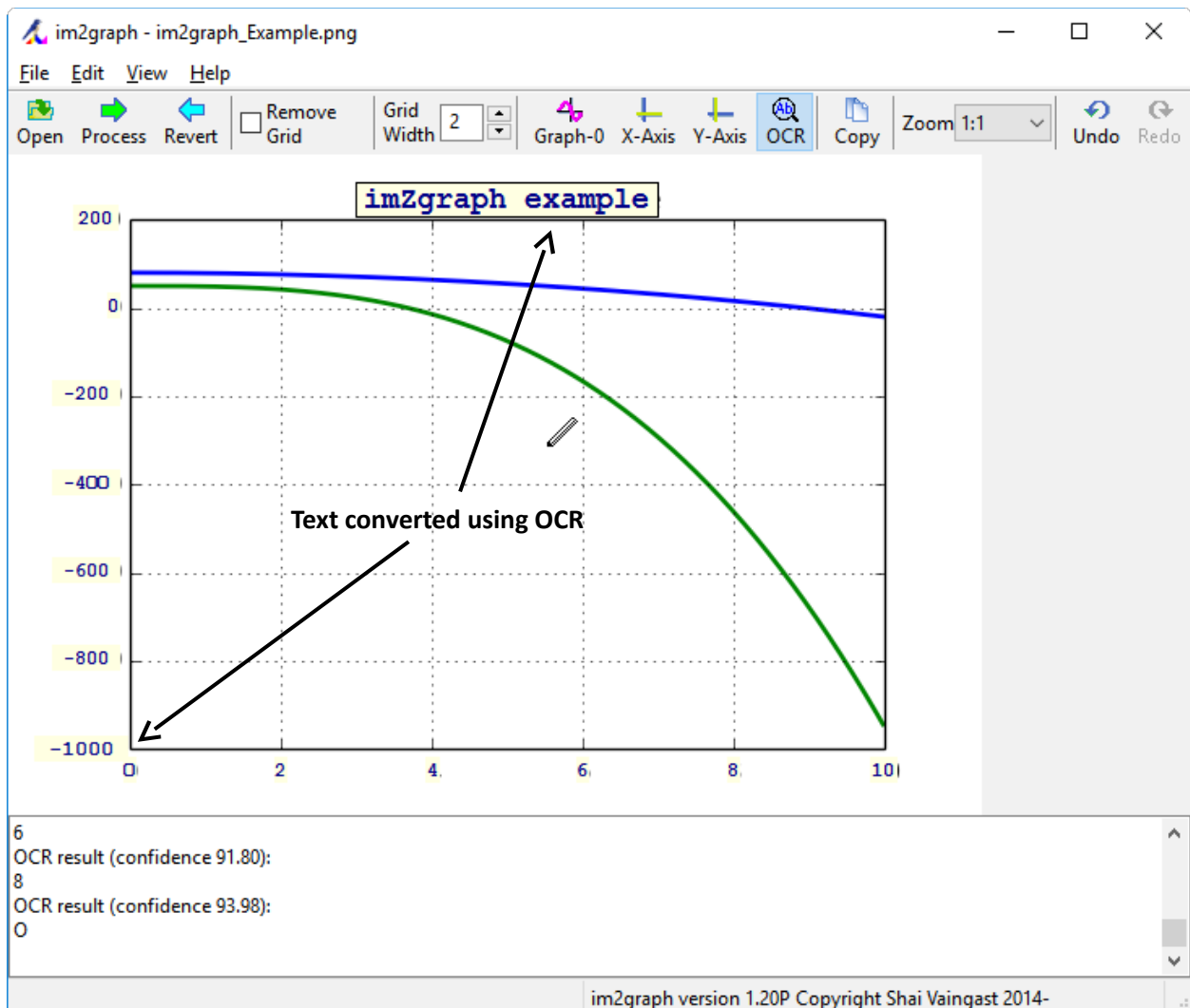


Optical Character Recognition - OCR (Premium only)

The premium version of im2graph also has built-in OCR capabilities. This means that you can convert not only graphs, but text and data too.



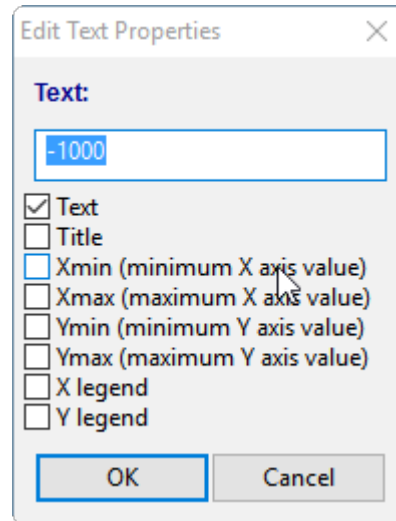
To perform OCR, either select the OCR button from the toolbar, press the T key or click the menu item Edit|OCR Text Area. This will put im2graph in OCR mode. When OCR mode is active, the OCR button will be depressed and the mouse icon will change to a pencil. To exit OCR mode, either press the OCR button in the toolbar, press the T key or use the menu. While in OCR mode, use the mouse to select an area to OCR process by right clicking on the image and dragging a stretch rectangle. Text within that rectangle will then be automatically OCRed. Text that has been detected using OCR will turn yellow and you can edit that text and set its properties.



If you are having issues with converting the text in a specific area, try again with a smaller area.

OCR values will be output to the CSV file or clipboard (see Digitizing the Graph).

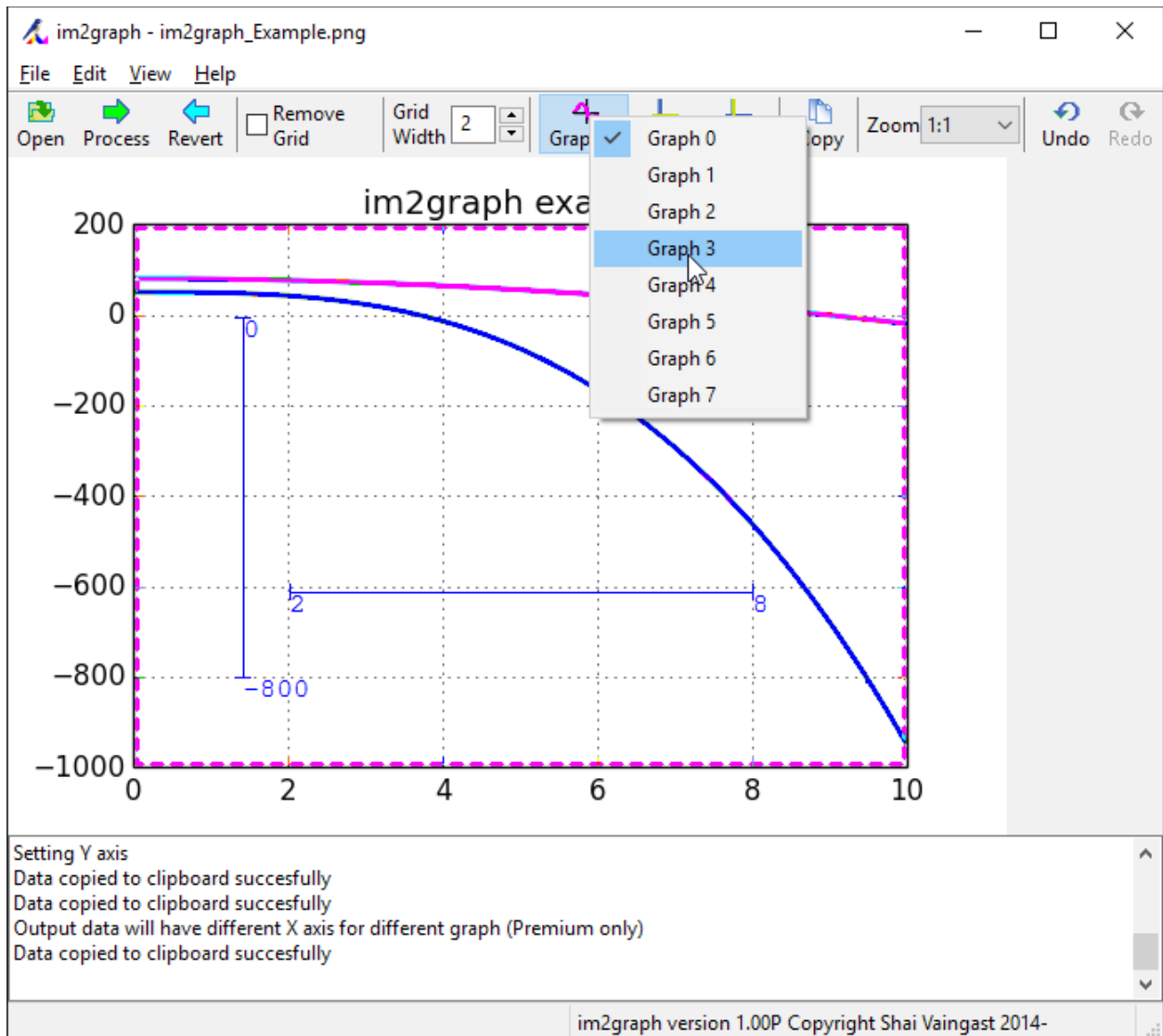
You can select an OCR text by right clicking on it. Once selected, you can either delete the text (for example, if the conversion is wrong) by pressing the DELETE key or you can edit its properties by pressing ENTER or double clicking on it. In the properties dialog box you can edit the text (in case the conversion is inaccurate, for example, the letter O and the digit 0 are prone to replace one another). You can also set the text properties.



Edit the text value in the text field or set its type to one of the options available: Text, Title, X min, Y min, X max, Y max, X Legend and Y Legend. If you select either X min, X max, Y min or Y max, these values will be used for determining the axis values. Other types will have different results in the digitized data as their names suggest.

Multiple Graphs (Premium only)

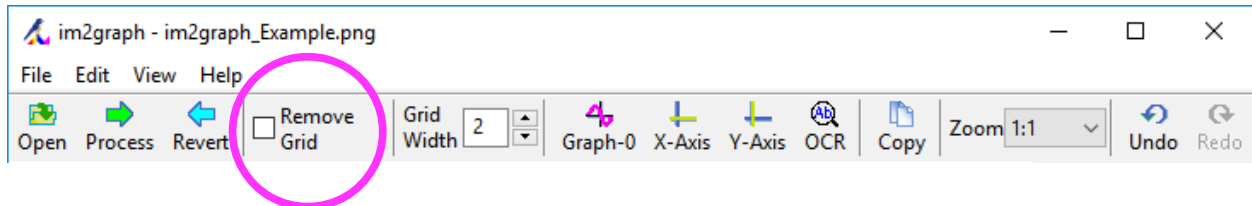
Im2graph supports multiple graph. To select multiple graphs, press on the Graph button and select the Graph number (from 0 to 7). The following is a typical scenario to use multiple graph support. First, select Graph-0, then select the patches from the image corresponding to the first graph. Next, select Graph-1, then select the patches from the image corresponding to the 2nd graph, and so on. Once you're done, you can view each graph separately by choosing that graph from the Graph menu. When you export the data (either to clipboard or to file), you will have separate graph values (see also the advanced option: Use a Common X axis for Multiple Graphs)



Advanced Features

Removing the Grid

At times the graph patch and the grid will merge into a single patch. In this case try degridding (removing the grid) by clicking on the Remove Grid check box. Note that this might cause your graph to break into several patches but you can either manually select the patches or drag select several patches. If the grid is thick, try increasing the Grid Width with the Grid Width spin control.

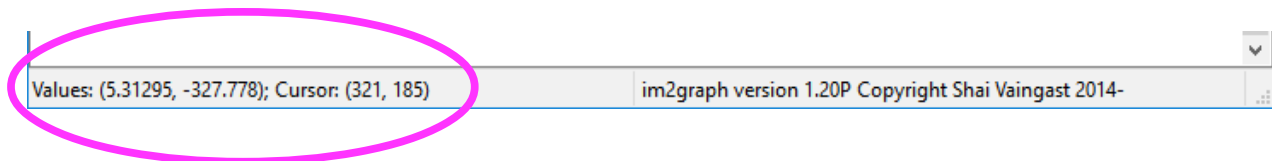


View Fit (Ctrl-=)

To set the size of the window to fit the image size, use the View|Fit menu option or press Ctrl and the key +.

Mouse Location

The cursor location is available at the bottom of the window, in the status bar. Once you set the scales, the graph values are also displayed. In case only one scale is set, the other value will be displayed as "?".



Paste from Clipboard (Ctrl-v)

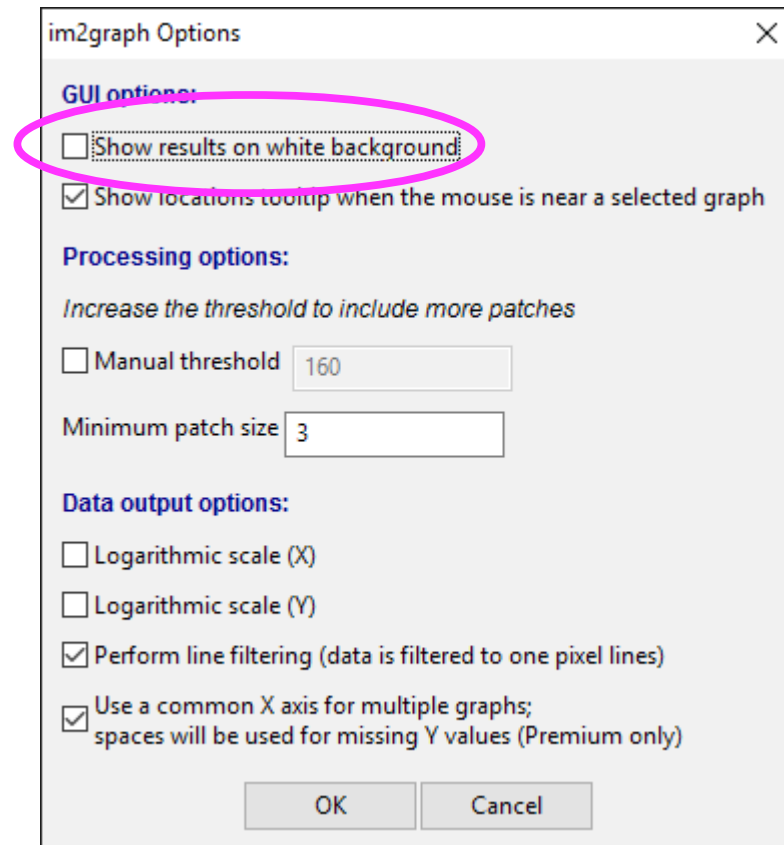
Instead of loading an image from file, you can paste the image directly to im2graph. This is especially useful if you're interested in digitizing an image from an already existing document such as a webpage or a PDF document. To paste an image directly to im2graph, press Ctrl-v or use the menu Edit|Paste.

Advanced Options

Additional im2graph options are available from the Options menu (Edit | Options).

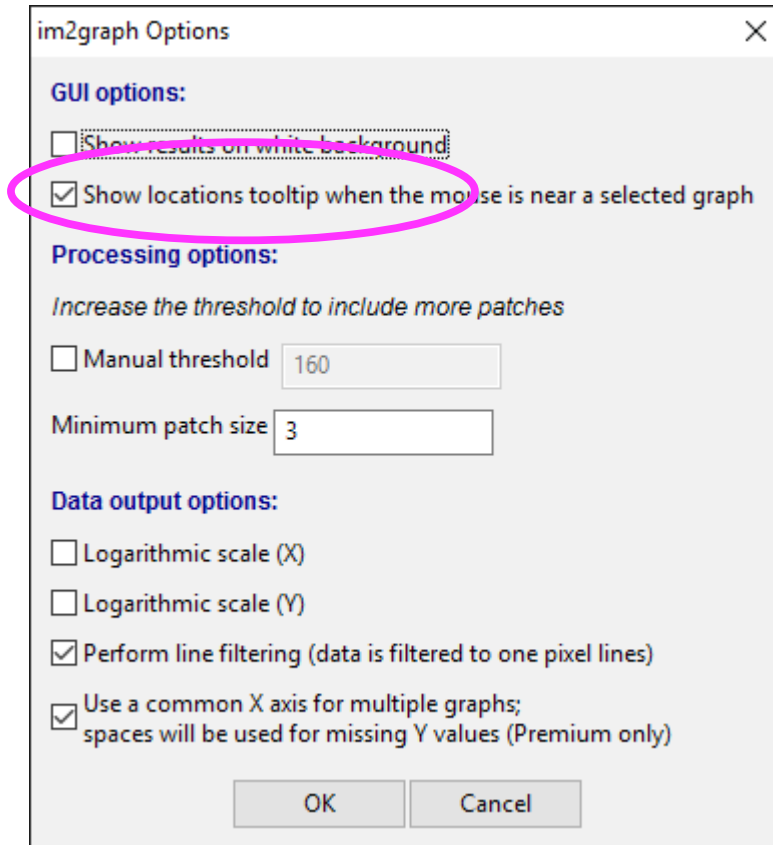
Show Results on White Background

Dealing with colorful patches and a colorful graph to be processed can be confusing. In this case, you can set the processed image to appear as gray while the background is set to white.



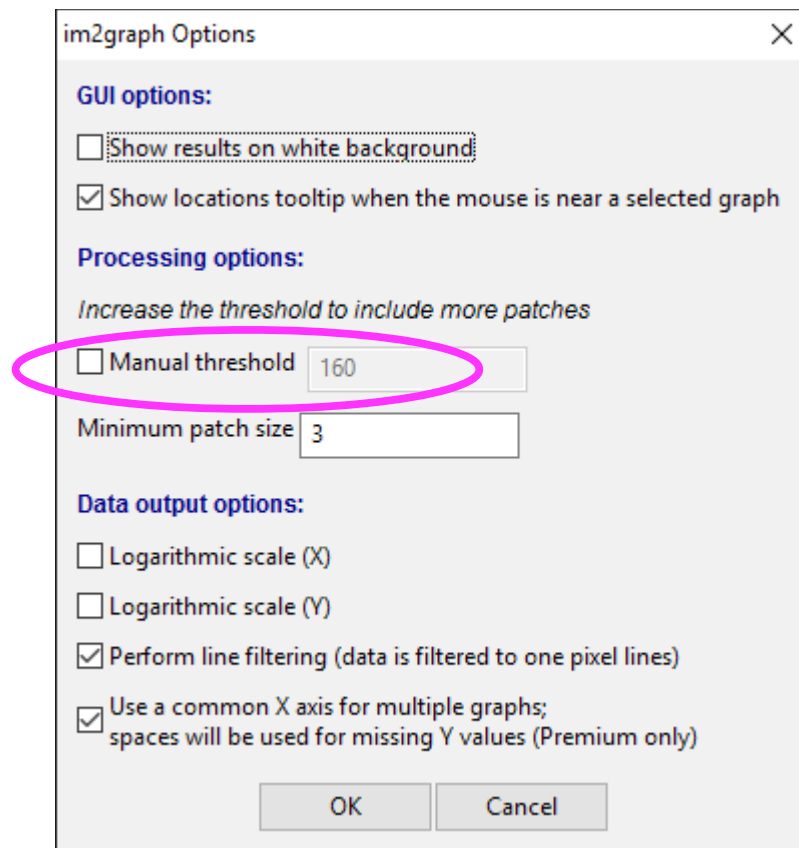
Show Locations Tooltip

This option controls whether or not the location on a selected graph is shown in a tooltip on the graph itself.



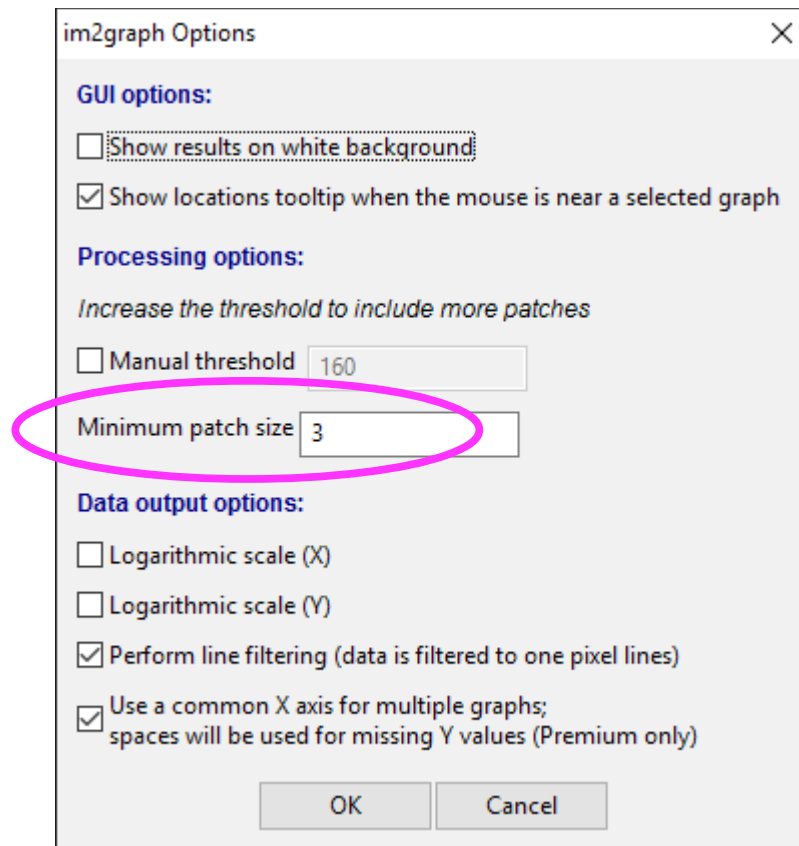
Manual Threshold

In case the graph you wish to digitize is not represented by colorful patches, you can manually set the threshold by selecting the Manual Threshold check button and setting the threshold value (default 160). This will allow for graphs with a lower contrast to be selected as well. As a result, the values that pass the threshold are presented in blue. Once you're satisfied that the graph you want to digitize is detected (becomes blue), click on Process again to show the colorful patches.



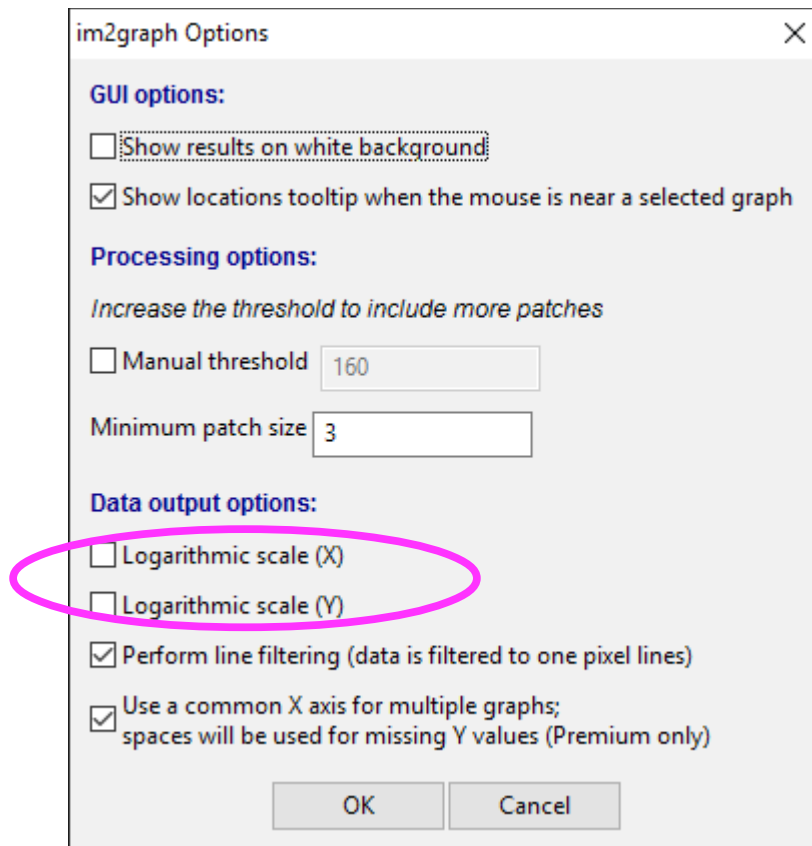
Minimum Patch Size

The minimum patch size value determines the minimum size of selectable patches. The default value is 2 and should work well for many occasions. However, if you'd like to select patches smaller than 2 pixels, change this value to 1. If you'd like to eliminate larger patches from the selection process, increase this value.



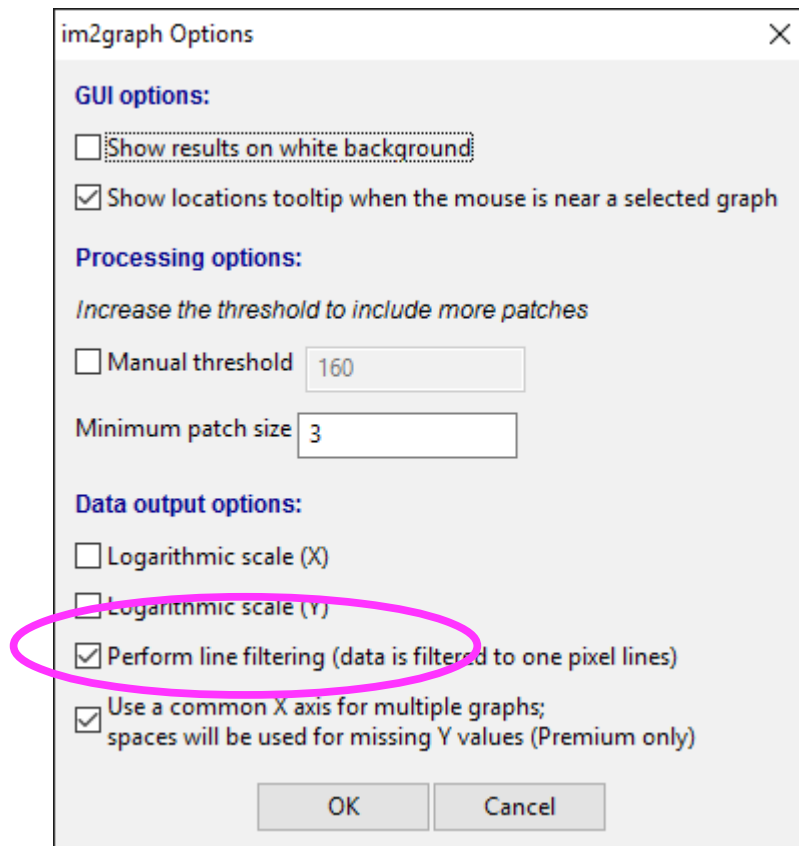
Logarithmic Scales (X and Y)

Use logarithmic scales as you would use normal scales (linear scales) by setting two points (X0-Y0; X1-Y1) per axis. Im2graph will automatically recalculate values once you've selected a logarithmic scale. You can have either both X and Y scales, logarithmic or any mix of linear and logarithmic scale (if the scale is not logarithmic, it is linear).



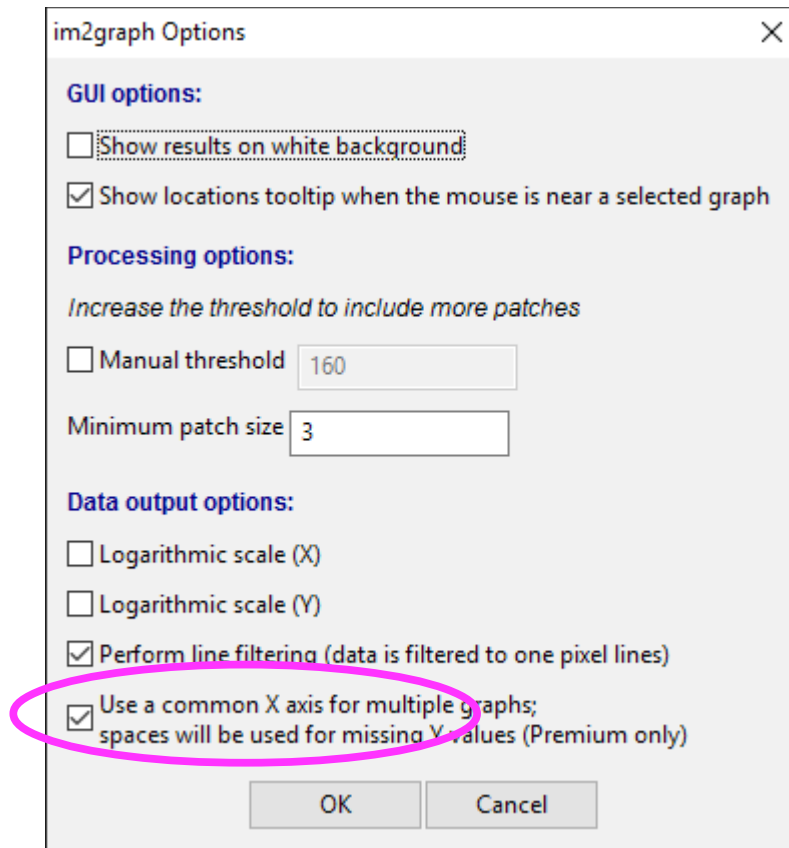
Perform Line Filtering

This option (enabled by default) will perform filtering on the data. The algorithm reduces line width so that for each X value there is a corresponding single Y value. The algorithm takes into account complex shapes including those that include several Y values for a given X value. If, however, the result is not what you expect, de-select this option and observe the results.



Use a Common X axis for Multiple Graphs (Premium only)

When using the multi-graph option, you will end up with two or more graphs. In this case, there are two options to export the data. One uses a common X-axis and the other generates a separate X-axis for each graph, as shown below.



Common X-axis:

notice there's one X-axis and several Y axis (Y_0, Y_1) corresponding to each graph:

| im2graph version | http://www.im2graph.co.il | |
|------------------|---|---------|
| X | Y_0 | Y_1 |
| 0.0397112 | 91.1628 | 61.3953 |
| 0.0613718 | 91.1628 | 61.3953 |
| 0.0830325 | 91.1628 | 61.3953 |
| 0.104693 | 91.1628 | 61.3953 |
| 0.126354 | 91.1628 | 61.3953 |
| 0.148014 | 91.1628 | 61.3953 |
| 0.169675 | 91.1628 | 61.3953 |
| 0.191336 | 91.1628 | 61.3953 |
| 0.212996 | 91.1628 | 61.3953 |
| 0.234657 | 91.1628 | 61.3953 |
| 0.256318 | 91.1628 | 61.3953 |

Without the common X-axis: each graph has its own X-axis values (X_0 and Y_0, X_1 and Y_1, etc.):

| im2graph version | | http://www.im2graph.co.il | | | |
|------------------|---------|---|---------|-----|-----|
| X_0 | Y_0 | X_1 | Y_1 | X_0 | Y_1 |
| 0.0613718 | 91.1628 | 0.0613718 | 61.3953 | | |
| 0.0830325 | 91.1628 | 0.0830325 | 61.3953 | | |
| 0.104693 | 91.1628 | 0.104693 | 61.3953 | | |
| 0.126354 | 91.1628 | 0.126354 | 61.3953 | | |
| 0.148014 | 91.1628 | 0.148014 | 61.3953 | | |
| 0.169675 | 91.1628 | 0.169675 | 61.3953 | | |
| 0.191336 | 91.1628 | 0.191336 | 61.3953 | | |
| 0.212996 | 91.1628 | 0.212996 | 61.3953 | | |
| 0.234657 | 91.1628 | 0.234657 | 61.3953 | | |
| 0.256318 | 91.1628 | 0.256318 | 61.3953 | | |
| 0.277978 | 91.1628 | 0.277978 | 61.3953 | | |
| 0.299639 | 91.1628 | 0.299639 | 61.3953 | | |
| 0.3213 | 91.1628 | 0.3213 | 61.3953 | | |
| 0.34296 | 91.1628 | 0.34296 | 61.3953 | | |
| 0.364621 | 91.1628 | 0.364621 | 61.3953 | | |

What Next?

Data Analysis and Visualization in Python

Now that you have the data in CSV format, what to do with it? You can plot it with Excel and do some processing but for real-world applications, you need a scientific computing tool. Python with its extensive Math libraries including: NumPy, Matplotlib, Scipy and many more are an excellent choice. You can read all about it in my book: [Beginning Python Visualization \(2nd edition\)](http://www.apress.com/9781484200537), from Apress. (Direct link: <http://www.apress.com/9781484200537>).