The Journal of Open Source Software

LMAPper – Where Scanning Probe Microscopy and Molecular Visualisation Meet

Luís M. A. Perdigão[1]

1 Independent Researcher

DOI: 10.21105/joss.02553

Software

- Review
- Repository
- Archive

Editor: Pending Editor

Submitted: 05 August 2020
Published: 06 August 2020

Summary

Microscopy, as a means to observe natural phenomena in smaller scales than human eyes can achieve, has undeniably led to enormous advances in science, thanks to a constant struggle to analyse and interpret microscopy images. With scanning probe microscopy images, scientists have developed several tools and methods but, despite the progress, these tools have been limited to manipulate image data and molecule models separately.

This article presents a new software application called LMAPper, that helps combine high resolution SPM images and molecular models into a single environment, making fitting image details to molecular models a simple straightforward task. This is also a much faster alternative than using other office and graphics productivity applications. This application has a friendly user interface and allows ample room for improvement by being open source.

Statement of need

For a number of decades scanning probe microscopy (SPM) has allowed scientists to directly observe phenomena at nanoscale. In surface science studies it has enabled us to improve our understanding of surfaces and catalysis. Molecular visualisation has allowed to give direct evidence of molecular structures and to observe self-assembly. SPM has played a central role as a technical tool in nanotechnology and developments. With evolution of SPM technologies, software that controls and helps analyse the images has gradually evolved, making it more accessible.

When working with data collected with a SPM instrument, there are important aspects in analysing experimental SPM images: image processing, metrology, and modelling using atomistic, molecular and crystal models. The simple process of overlaying molecular models to the images is often done. Although several commercial and free applications exist that can help deal with image manipulation and molecular modelling separately, there is a complete absence of tools that combine both within the same environment. Many scientists often use applications that are not originally optimised for working with both scanning tunnelling microscopy images and molecular models, nevertheless this hugely important because it helps interpreting the features of the images.

Unfortunately, most SPM images do not provide elemental information and the precise location of atoms is often not possible, unless tips are functionalised in situ for sharper contrast (Gross, Mohn, Moll, Liljeroth, & Meyer, 2009 ; Lawrence et al., 2020 ; Mohn, Schuler, Gross, & Meyer, 2013 ; Temirov, Soubatch, Neucheva, Lassise, & Tautz, 2008). In scanning tunnelling microscopy (STM) the images are the height combined with the local density of states (convoluted with the LDOS of the tip), and these can be delocalised over the molecule, resulting in diminished contrast and difficult identification of the atoms. Modelling

and interpretation of SPM images is often guess work which is done by overlaying models on the experimental images.

A typical workflow of overlaying molecular models on an STM image is to (see also Figure 1):

1. Adjust the image in a SPM analysis software (such as Gwyddion(Nečas & Klapetek, 2011) or WSxM (Horcas et al., 2007))
2. Import the image to a drawing program (example Inkscape ("Draw Freely Inkscape," n.d.) or PowerPoint("Microsoft PowerPoint, Slide Presentation Software, PPT," n.d.))
3. Resize SPM image to the correct dimensions.
4. Separately create a molecular model in a molecular modelling program, that may have been optimised or undergone through several calculation algorithms.
5. Create a bitmap of the molecule and import this image onto the vectorial drawing program
6. Resize molecular model to the correct dimensions.
7. Overlap molecular models to SPM image to help interpret contrast features.
8. Correct if needed.

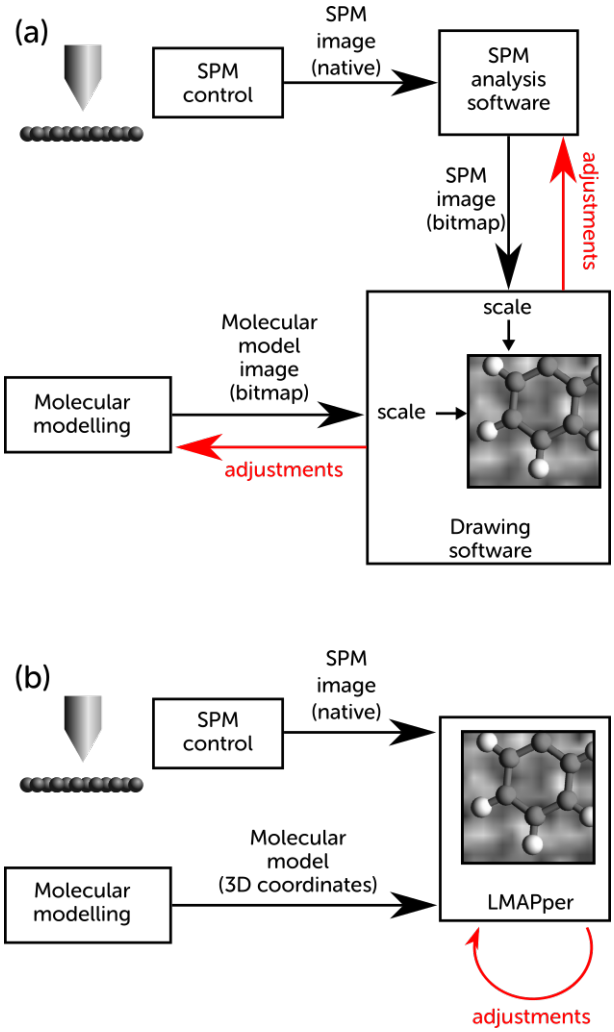

Figure 1: (a) Traditional workflow in analysing experimental SPM images and fitting to molecular models. SPM images and molecular models are handled separately and imported as bitmaps to drawing software. This presents a major problems in that scaling can generate errors, and that any adjustments require user to several steps back in the workflow, hence being time-consuming. (b) Typical workflow when using LMAPper that can import SPM images and molecular models in native format and scale them automatically.

Although this methodology works, the whole process is quite time-consuming and very impratical. The bitmap of the molecular model that was generated and the SPM image must be correctly sized in the vectorial drawing program. Typically, measurements of nanometers can be converted to units of pixels or cm and careful adjustment to the image size in the same units. This is necessary in order to have the molecular model and the SPM image scaled to each other correctly. Also, if the imported bitmap of the molecular model was screenshotted at the wrong angle, the user is often required to go back to the molecular modelling software, fix the model or shooting angle and recreate the correct bitmap. In certain cases, a large number of repeating molecular models are intentionally overlaid in SPM image, and this often results in large memory consumption and a large file which leads to computer hanging or slow response. The large number of steps involved also increase the probability of errors.

Making the correct calibration between the SPM images and the molecular model is very important as incorrect calibrations can lead to wrong conclusions. A typical error is when molecular geometry models are proposed for experimental images obtained but these do not

fit correctly if the correct scaling had been used. The author and colleagues have found a number of calibration errors like this in the literature, turning a published article practically useless. Situations like this are quite unfortunate as it degrades the quality of science, that could be avoided if appropriate verification was in place. Therefore tools to help with analysis and verification would be useful.

The software proposed in this article introduces a new approach in modelling and fitting SPM images to molecular models. It takes advantage of the fact that: (i) 3D molecular structure file formats contain information of the size and location of the atoms as well of connectivity, and (ii) that SPM images are stored with scanning conditions such as scan size. Combining this information together is what the LMAPper software was designed to do. This is a significant advantage compared with alternative software such as the most commonly used vectorial drawing applications, that do not support three dimensional models as molecule are. More recent versions of Power Point now support 3D objects, however in order to use these capabilities to draw molecules in 3D from a file is still far from straightforward.

In computer graphics, existing OpenGL or DirectX libraries and higher-level frameworks are now well matured to make easier to combine 3D molecules with an SPM image. This can be done in the following way. The SPM image is modelled as a flat rectangle, with a texture corresponding to the surface intensity plot of the SPM image. The molecules can be simply represented with three-dimensional spheres as atoms and cylinders as bonds (ball and stick) or by another three-dimensional representation, in the same way many molecular visualisation programs do. The principle of combining both is conceptually simple, but its implementation had never been realised. This was the principle used in developing the software presented here, LMAPper. It was developed to facilitate scientists to combine SPM images with molecular models in 3D, accelerating image analysis and workflow. This application borne out from the frustration of trying to fit molecular models to SPM images in order to determine molecular assembly geometries. The application imports native SPM data formats, and can import also several molecular model formats. Then it represents the molecules in 3D overlaid on the SPM image. Great care has been taken to make this application very easy to use.

Implementation

The application LMAPper is a single executable file that runs under windows. The software was programmed in Visual Basic Language (.NET), using Visual Studio 2017 ("Visual Studio IDE, Code Editor, Azure DevOps, & App Center," n.d.). The API used was the Windows Presentation Framework, because it offers support for 3D modelling using the Viewport3D control. Also, the easy data binding, simplified handling of memory, debugging and profiling tools make it easier to develop and maintain.

The application opens a single window with menus, buttons and other widgets offering common tasks. Most of the operations can be executed without using the menus. For example, instead of importing SPM files or 3D molecule description files by using the open command in the menu, this can be achieved by simply dragging and dropping a file from the file explorer.

The left side of the window is dedicated to the (single) SPM image, and the right side is dedicated to the multiple molecular models. This right-left separation corresponding with the molecule and SPM image is one of the foundations of how this user interface works. It naturally separates the processing of each of the data types. In the middle is the viewport which shows both the SPM image and the molecular models.

SPM images

Currently, the application can read the following binary SPM image formats:

1. Nanonis SXM
2. Nottingham NTI
3. Createc DAT
4. WSxM binary STP

With the binary SPM image, the software displays it as a height colour map. The following image adjustments operations are available: Raw, Global Plane or Flatten. Two color schemes can be used, but potentially more in the future.

Not all SPM image formats are supported because there are too many types, and it would require a great amount of work to be able to write code to support all of them. Also, the software offers limited image analysis tools, as it intended to be simple. Nonetheless, the application can import bitmap formats such as PNG, JPG and BMP, as if they are SPM images. The information about the size (in nanometers) is not stored in these bitmap file formats so this information needs to be provided by a bitmap-loving-human. A dialog prompts the user for this additional information. Alternatively, some bitmap images that are imported have a scale bar to indicate the size, so the import dialog offers a method to calibrate the bitmap image from this scale bar provided. This is a quite useful feature as this software can be used to verify images from published articles whereby only the scale bar is provided. Bitmap images can also be imported from clipboard using the paste shortcut Ctrl+V.

It is often assumed that calibration of the SPM instruments is never perfect. Piezo voltages calibrations can vary on a daily basis and are strongly dependent on their temperature. Experimental measurements are conducted in different conditions which may need slight adjustments in spatial measurements. These additional adjustments are normally estimated by user and are not stored in the SPM file formats. Additionally, SPM scanning experiments often suffer from creep and drift effects, which is often noticeable in the slow scan direction and as a skewing of the image. Skewing can also occur when the scanner piezos are not perfectly designed with the X and Y electrodes at 90° degrees. To account for these potential scaling and skew adjustments, the software offers a very accessible way to adjust the SPM image, with optional parameters to scale in the X or Y direction and skew angle correction. Internally, these corrections are applied by modifying the rectangle shape to a parallelogram, which is 'filled' with the SPM image as a bitmap texture. This is a great advantage compared to other SPM software that does these corrections by either pixel/data interpolation or by internal parameter scaling.

Molecular models

Several molecular file formats are supported:

- Protein database (PDB)
- Chemical Markup Language (CML)
- PubChem conformer (XML)
- Chemical table (MOL or MDL)

The software can also import 3D molecular models directly from clipboard if they are in mime type *chemical/x-mdl-molfile* format, which is the one used by Avogadro (Hanwell et al., 2012) software when models are copied to clipboard. This is a very useful feature that allows user to skip the step of saving the molecular model from the molecular modelling application.

The imported molecule is scaled precisely to the scale of the SPM image. This minimizes the need to worry about scaling the models, as is often done with alternative methods.

Several molecules can fill the workspace and these are included in a list, that can be easily navigated. Molecules can be displayed in commonly used formats, ball and stick (covalent or

fixed atom radius) or van der Walls spheres. When working with self-assembly images it is useful to use the duplicate and the mirror tool.

Results and discussion

Figure 2 shows a screenshot of the LMAPper application in action, fitting 3D models of perylene tetracarboxylic diimide (PTCDI) and melamine to the experimental STM image. Molecules are manipulated with the help of the mouse buttons a movement. Left-mouse button to drag in the X-Y plane, right-mouse button to rotate in the Z axis, and middle-mouse button for free rotate. The mouse roller zooms in and out at the location pointed by the mouse. This mode of operation is very similar to the user-interface that many commercial drawing and molecular modelling programs work. In combination with the Shift key, this movement can be done in more discrete steps. The selected molecule(s) is(are) represented with a blue semi-transparent sphere, and multiple selection is possible. Zooming in/out and panning is supported in touchscreen displays.

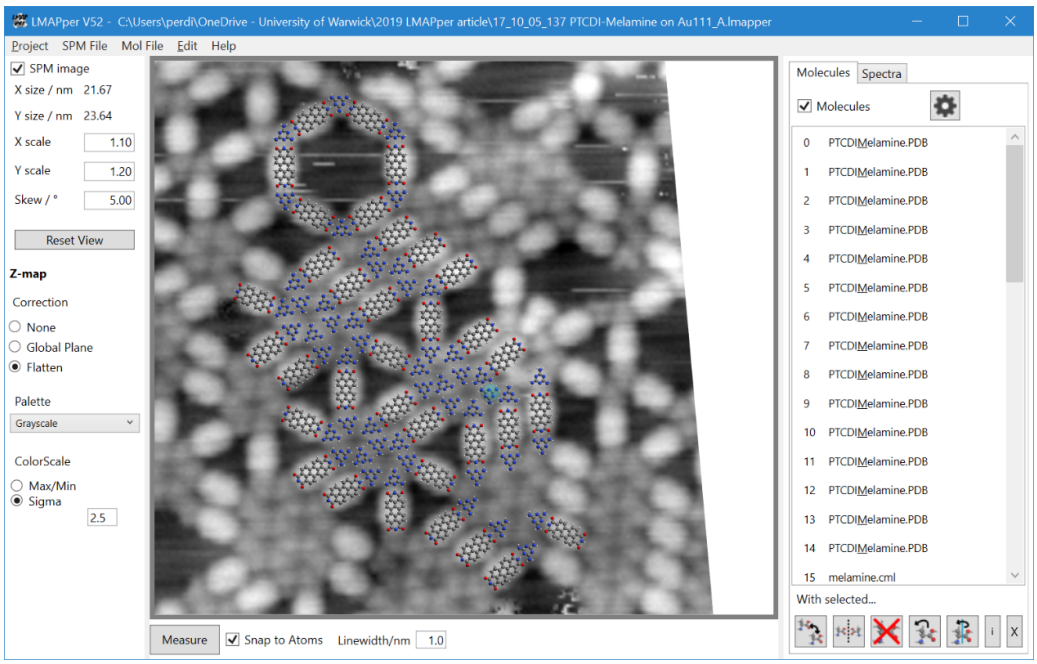

Figure 2: Screenshot of the application window of self-assembled PTCDI-melamine deposited on Au(111).

The SPM image cannot be moved nor rotated, except for scaling and skewing transforms. This restriction is intentional, as this is the image that user wants to fit models with. Only molecules and 'camera' can be manipulated.

Checkboxes for visualising/hiding either the SPM image or the molecule can rapidly hide/unhide each of the layers. This can be used, for example, when user creates a molecular model based in an experimental image, but then wants to publish/present the molecular model only, or if user wants to see the SPM image without the molecular model that conceals details underneath.

In certain occasions it is useful to perform distance measurements. A measurement tool lets user pick to points on the surface and it shows the distance. The selection can optionally snap to the centre of atom positions. It should be noted that the distance value given is not

the 3D distance but the projected distance on the two-dimensional plane of the SPM image. This means that, in case two atoms are at different heights, the distance is measured without taking into account their difference in height. In this way the distances measured can be related directly with the distances in the SPM image.

Additional Features

To make working with the modelling user friendly, the following additional features are available

- Save/Load project: The workspace can be saved to a single file, without dependency in original images.
- Undo/Redo: Ctrl+Z or Ctrl+R, that can do up to ten steps.
- Clipboard import (Ctrl+V). SPM images can be imported from clipboard as if loading a bitmap. A dialog will prompt user for the size of the image in nanometers. Molecule models can be imported if they were copied from Avogadro.
- Clipboard export (Ctrl+C). The image in the viewport can be copied to the clipboard so that it can be used in other programs. The image is exported in PNG format with double the resolution that is being displayed, or optionally with a resolution multiplier. Background areas are recorded with transparency (alpha). As a molecular visualiser, this is one of the few applications available that can export molecular models with transparency in such easy way.
- Multiple selection of molecular models.
- Import and manipulate single-point Spectra (under development). Spectra appears with a coloured marker on the SPM image, and a chart presents the measurement.

Discussion

In many occasions, LMAPper has been used successfully to model conjugate polymers arrangements [Lawton et al. (2020); chen_effect_2019], monomolecular (Lawrence et al., 2020) and multicomponent assembly (Pinfold et al., 2020a, 2020b). LMAPper excels in this type of analysis where repetitive units (individual molecules or oligomers) are present and self-assembly is observed without overall periodicity.

Since the first experimental version of this application (which had a lot less features) users quickly realised how image analysis and interpretation with the corresponding molecular models became much easier, then benefiting scientists who have used. New users also quickly learnt how to use it in such a way that it became the default method of analysing SPM images with molecular models, dispensing alternative methodologies that had been previously established.

Perhaps one of the major drawbacks is that this application does not offer a full suite of tools that are available in advanced SPM imaging software or in molecular modelling applications. This problem was minimised by ensuring that importing/exporting data from/to clipboard and that moving data from one application to another, so it became easy to integrate well with these specialised tools. The aim in developing this application is not to be an advanced SPM image analysis or molecular modelling application, but to best combine these two types of data.

One unexpected use that was discovered while testing the application, was that this software could be used to quickly verify images against corresponding molecular models in published

nanoscience literature. Incredibly, the author and colleagues have discovered a number of mis-modelling in several publications. This was done by simply by extracting the images from the articles, and combine in LMAPper with a quick model in drawn Avogadro, and the error becomes apparent. We now support the idea that this application should be used by publishers as an important tool verify the validity of molecular models proposed to explain microscopy images obtained.

The Spectra visualisation feature is under development, but it already offers an advantage in analysing by combining visualisation on the SPM image with the plotted spectra. This feature will be expanded further to support more file formats.

LMAPper could be used for teaching purposes. Many science departments offer ambient STM imaging in their laboratories as part of their curriculum. This application could help students check the calibration and model fitting, and helps to quickly relate the images with the atomistic representation. Where no instrument is available, exercises can be given to students to fit a given molecular model to a particular image, in order to identify its molecular self-assembly preferred arrangement. Overall, LMAPper has the potential to boost accessibility of nanoscience to a wide range of people.

The source code is available to download for free so that users can add features to it. The binary file has also very small size (version 53, $<$500 kB) and is provided as an executable that does not require installation. Although theoretically possible, it is not available in other platforms such as Linux, MacOS, or Android. The author and developer understands that this a major drawback, as scientists are very diverse and demanding (and sometimes critical) with their preferred computer OS choice. Unfortunately, cross-platform development introduces a large number of coding complications such as complete rewrite of the code. Presently, potential API alternatives that have been tested are Qt, Electron (NodeJS) or .Net Core, though there is no imminent plan to implement under these formats. An experimental version LMAPperJS for browsers (using Three.js) already exists but with limited functionality.

As this application is source code, it is easy for others to add support to other file formats. Also envisioned is support for protein visualisation, which can be combined with electron microscope images of cells and viruses.

The author would prefer that no more instrument file formats to be created. It would be nice that SPM manufacturers, and molecular modelling developers, create a unified file formats. This would make the developers job much easier so that focus could be diverted to developing better tools to analyse the data.

Conclusion

LMAPper reconciles two separate worlds, one of the SPM imaging and the other of the molecular modelling, into one united workspace, where they combine in perfect harmony. This is what scientists have always envisioned since the discovery of the cells and atoms, that gradually progressed through chemistry and physics, then culminating at this pinnacle of scientific software development.

Acknowledgments

The author wants to thanks to all people in the Giovanni Costantini's group at University of Warwick, who contributed to the development by testing this software extensively with their own data and provided feedback and suggestions. In particular, extra thanks to Harry Pinfold who has heavily tested, and reported bugs whenever these were discovered. The

author also wishes to thank Peter H. Beton for giving permission to use the STM image of PTCDI-melamine for the figure in this article.

The author also wishes to thank Sourceforge website for providing free-project hosting.

References

Draw Freely Inkscape. (n.d.). Retrieved from https://inkscape.org/

Gross, L., Mohn, F., Moll, N., Liljeroth, P., & Meyer, G. (2009). The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy. *Science*, *325*(5944), 1110–1114. doi:10.1126/science.1176210

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, *4*(1), 17. doi:10.1186/1758-2946-4-17

Horcas, I., Fernández, R., Gómez-Rodríguez, J. M., Colchero, J., Gómez-Herrero, J., & Baro, A. M. (2007). WSXM: A software for scanning probe microscopy and a tool for nanotechnology. *Review of Scientific Instruments*, *78*(1), 013705. doi:10.1063/1.2432410

Lawrence, J., Sosso, G. C., Đorđević, L., Pinfold, H., Bonifazi, D., & Costantini, G. (2020). Combining high-resolution scanning tunnelling microscopy and first-principles simulations to identify halogen bonding. *Nature Communications*, *11*(1), 1–7. doi:10.1038/s41467-020-15898-2

Lawton, S. S., Warr, D., Perdigão, L. M. A., Chang, Y., Pron, A., Costantini, G., & Haddleton, D. M. (2020). Determining the sequence and backbone structure of "semi-statistical" copolymers as donor–acceptor polymers in organic solar cells. *Sustainable Energy & Fuels*. doi:10.1039/C9SE01261C

Microsoft PowerPoint, Slide Presentation Software, PPT. (n.d.). Retrieved from https://www.microsoft.com/en-us/microsoft-365/powerpoint

Mohn, F., Schuler, B., Gross, L., & Meyer, G. (2013). Different tips for high-resolution atomic force microscopy and scanning tunneling microscopy of single molecules. *Applied Physics Letters*, *102*(7), 073109. doi:10.1063/1.4793200

Nečas, D., & Klapetek, P. (2011). Gwyddion: An open-source software for SPM data analysis. *Open Physics*, *10*(1), 181–188. doi:10.2478/s11534-011-0096-2

Pinfold, H., Greenland, C., Pattison, G., & Costantini, G. (2020a). Fluorinated carboxylic acids as powerful building blocks for the formation of bimolecular monolayers. *Chemical Communications*, *56*(1), 125–128. doi:10.1039/C9CC08361H

Pinfold, H., Pattison, G., & Costantini, G. (2020b). Fluorination as a route towards unlocking the hydrogen bond donor ability of phenolic compounds in self-assembled monolayers. *CrystEngComm*. doi:10.1039/D0CE00213E

Temirov, R., Soubatch, S., Neucheva, O., Lassise, A. C., & Tautz, F. S. (2008). A novel method achieving ultra-high geometrical resolution in scanning tunnelling microscopy. *New Journal of Physics*, *10*(5), 053012. doi:10.1088/1367-2630/10/5/053012

Visual Studio IDE, Code Editor, Azure DevOps, & App Center. (n.d.). *Visual Studio*. Retrieved from https://visualstudio.microsoft.com