

## EMAN2: Software for Image Analysis and Single Particle Reconstruction

Grant Tang\*, Liwei Peng\*, Deepy Mann\*, Chao Yang\*\*\*, Pawel A. Penczek\*\*, Grant Goodyear\*\*, Philip R. Baldwin\*\*, Wen Jiang\*\*\*\*, Steven J. Ludtke\*.

\* National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, TX, 77030

\*\* The University of Texas – Houston Medical School, Department of Biochemistry, 6431 Fannin, Houston, TX 77030.

\*\*\* Computational Research Div., Lawrence Berkeley National Laboratory, Berkeley, CA 94720.

\*\*\*\* Department of Biological Sciences, Purdue Univ., West Lafayette, IN 47907.

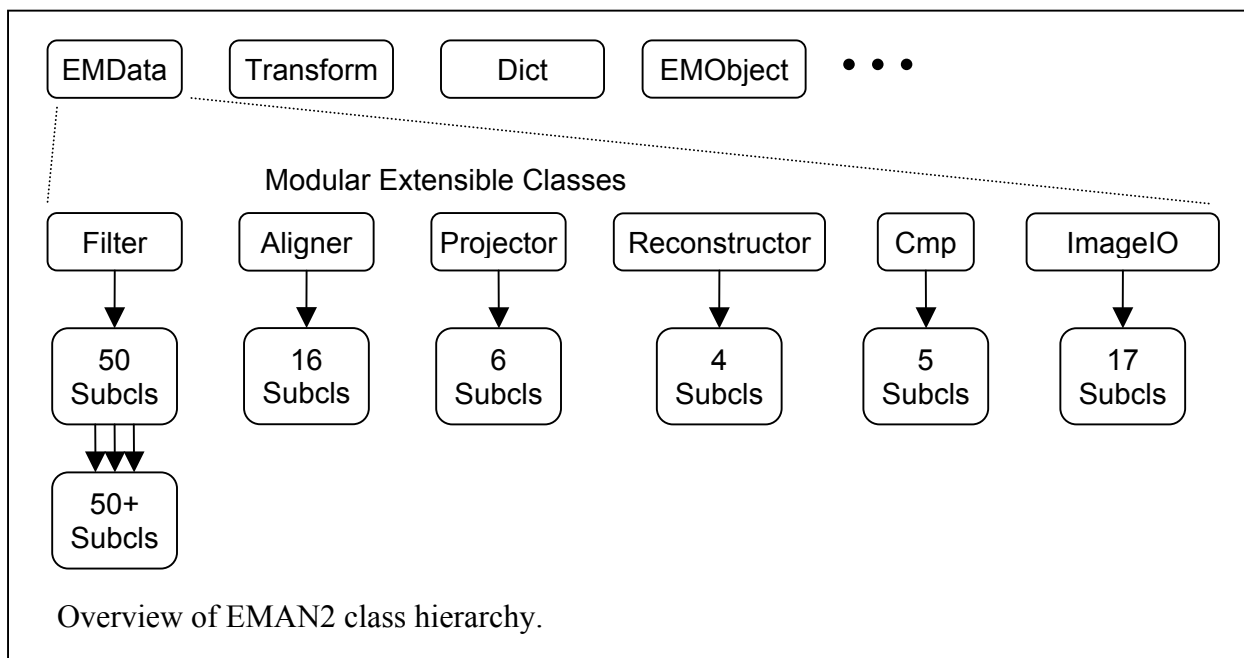
EMAN (Electron Microscopy ANalysis) [3] is an open-source image processing suite in use by hundreds of researchers around the world [1-2]. We present EMAN2, a completely refactored version of this popular package. While it is primarily aimed at analysis of cryo-EM data with a particular focus on single particle reconstruction, its extensive image processing library is also applicable to other modalities such as AFM, SEM, traditional TEM, and any other imaging technique involving 1 to 3-dimensional floating point (or integer) datasets. The suite consists of a core C++ library with complete bindings to the Python scripting language. GUI tools are implemented using the cross-platform compatible WX toolkit. All computationally intensive image processing operations are implemented in C++, with user-level programs implemented as Python scripts. This language offers a vast number of features convenient for high-level scientific programming. As a very readable scripting language, this offers sophisticated end users the ability to subtly alter the functionality of any of the high level code without requiring any recompilation or C++ programming skills. Since the low level operations are performed in C++, there is virtually no performance penalty for using Python.

Extensibility is an explicit part of the new library design. Rather than taking the deep object hierarchy approach used in packages like ITK, EMAN2 uses a set of simple extensible classes for specific categories of operations. Templates are provided for each type of process including: filtration, 3D reconstruction, projection, registration, comparison, etc. Adding a new algorithm requires simply defining the parameters required by your routine and inserting the actual image processing code into the template. After a simple 'make', the new algorithm is immediately available from both the C++ and Python interfaces. The modular structure integrates documentation into the code and supports full introspection for integration with GUI tools. The current library contains the expected image processing algorithms, such as Fourier and real-space filtration, various registration algorithms in addition to sophisticated new methods such as 3-D reconstruction using the gridding algorithm [4].

The core library of EMAN2 is still under active development. The basic GUI widgets have been implemented (using the cross-platform WXPYthon toolkit), but EMAN2 does not yet have the rich end-user functionality of EMAN1. The EMAN2 core is also being used to implement SPARX, an effort to extend the PHENIX crystallography suite. EMAN2 is available for download with full source from <http://ncmi.bcm.tmc.edu>. Contributions by outside developers are welcome, and CVS access to the repository is available (SLudtke@bcm.edu).

## References

- [1] Ludtke, S.J., Serysheva, I.I., Hamilton, S.L. and Chiu, W. (2005) The pore structure of the closed RyR1 channel. *Structure*. 8:1203-11.
- [2] Ludtke, S.J., Chen, D.H., Song, J.L., Chuang, D.T., and Chiu, W. (2004). Seeing GroEL at 6 Å resolution by single particle electron cryomicroscopy. *Structure (Camb)* 12, 1129-1136.
- [3] Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128, 82-97.
- [4] Penczek, P.A., Renka, R. and Schomberg, H.: Gridding-based direct Fourier inversion of the three-dimensional ray transform. *J. Opt. Soc. Am. A*. 21:499-509, 2004.
- [5] Supported by the NIH (P41RR02250, P01GM064692 (Penczek, Ludtke), R0160635 (Penczek), and DK-26758), Welch Foundation (I-1286), and the Agouron Institute.



```

from EMAN2 import *
images=EMData.read_images("orig.spi")
for img in images:
    img.process("LowpassGauss", {"lowpass":14.0})
    img.process("NormalizeEdgeMean")
    img.write_image("filtered.hdf")

```

A sample EMAN v2 Python script which reads in a set of images in the SPIDER file format, applies a Gaussian low-pass filter to each, normalizes each such that the mean value around the edge of the image is 0.0 and the standard deviation is 1.0, then writes each to a HDF5 format stack file.