

Bsoft: Image processing and molecular modeling for electron microscopy

J. Bernard Heymann^{a,*}, David M. Belnap^{a,b}

^a *Laboratory of Structural Biology, National Institute of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892, USA*

^b *Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT 84602, USA*

Received 9 April 2006; received in revised form 7 June 2006; accepted 11 June 2006

Available online 28 June 2006

Abstract

Bsoft is a software package written for image processing of electron micrographs, interpretation of reconstructions, molecular modeling, and general image processing. The code is modularized to allow for rapid testing and deployment of new processing algorithms, while also providing sufficient infrastructure to deal with many file formats and parametric data. The design is deliberately open to allow interchange of information with other image and molecular processing software through a standard parameter file (currently a text-based encoding of parameters in the STAR format) and its support of multiple image and molecular formats. It also allows shell scripting of processes and allows subtasks to be distributed across multiple computers for concurrent processing. Bsoft has undergone many modifications and advancements since its initial release [Heymann, J.B., 2001. Bsoft: image and molecular processing in electron microscopy. *J. Struct. Biol.* 133, 156–169]. Much of the emphasis is on single particle analysis and tomography, and sufficient functionality is available in the package to support most needed operations for these techniques. The key graphical user interface is the program bshow, which displays an image and is used for many interactive purposes such as fitting the contrast transfer function or picking particles. Bsoft also offers various tools to manipulate atomic structures and to refine the fit of a known molecular structure to a density in a reconstruction. Published by Elsevier Inc.

Keywords: Single particle analysis; Tomography; Three-dimensional image reconstruction; Distributed processing; Image processing workflow

1. Introduction

The availability of electron microscopes capable of achieving high-resolution imaging of biological specimens and increased access to large-scale computational resources allows more elaborate data processing algorithms, provided the software capable of handling large datasets exists. This trend requires a change in both software development habits and users' data processing habits. To handle the enormous amounts of data, more efficient data management techniques need to be employed. Bsoft was written with the purpose of providing easy access to image and molecular data in many formats, offering a way to manage

the large numbers of parameters generated and used during processing, and supporting distributed processing for fast processing of large datasets. In addition, Bsoft is intended as a platform for the rapid development of new algorithms.

Bsoft covers image processing for electron microscopy in general, with new programs added as required. It has grown from a relatively small package with limited capabilities (Heymann, 2001), to a much more comprehensive suit of programs dealing with many aspects of electron microscopy in structural biology. The current scope of the package includes single particle analysis (SPA)¹ as well as its

* Corresponding author. Fax: +1 301 480 7629.

E-mail address: Bernard_Heymann@nih.gov (J.B. Heymann).

¹ Abbreviations used: 2D, two-dimensional; 3D, three-dimensional; AFM, atomic force microscopy; CTF, contrast transfer function; EM, electron microscopy; MPA, multiple particle analysis; SPA, single particle analysis; STAR, self-defining text archiving and retrieval.

multi-particle counterpart (MPA) to allow the analysis of heterogeneous sample preparations (Heymann et al., 2004). The second technique covered is tomography, with operations ranging from micrograph alignment and reconstruction to interpretation. The third area of emphasis is molecular manipulation and rigid-body fitting to interpret reconstructions from electron micrographs. In addition, Bsoft offers a number of programs for other processing tasks and simple “housekeeping” chores.

The key aim in Bsoft from the start (Heymann, 2001) has been to enable the user to accomplish as many conceivable processing tasks as possible. Therefore, programs are written with generalization in mind, so that programs can handle any type of image and a variety of processing tasks in some reasonable fashion. This power has its advantages (a user can do things the programmer may not even have anticipated) as well as disadvantages (users do things the programmer never intended and may end up with strange results or crashed programs). Ease-of-use in Bsoft is directly related to understanding the command line and the intended functionality of the individual programs. In some cases this is easy and intuitive; in others, it is more complicated due to the nature of the processing.

Bsoft also evolved as a package within which a developer can test data processing algorithms relatively quickly without having to deal with issues such as image format handling and to some extent parameter handling. This has manifested in the speed with which relatively inexperienced programmers have been able to assess algorithms for specific image processing tasks.

In the Bsoft package, extensive consideration has been given to the handling of parametric data generated and required for image and molecular data processing. The origin of Bsoft lies in the BioImage database prototype project (Carazo et al., 1999), where the effort concentrated on providing a unified approach to all types of data derived from any type of microscopy. Accomplishing such a complex task proved to be very difficult, exposing many issues requiring attention. Perhaps the most important of these is the need to be able to interchange information generated by many disparate software packages. In particular, the problems include differences in conventions, support for image formats and incompatible handling of parametric data associated with processing. An effort is in progress to establish common conventions in the field (Heymann et al., 2005a), as well as common data exchange formats. The difficulty with the latter is that there must be an agreement on the particular items that need to be included in such a format. Efforts are underway via the EM Data Bank and others to remedy this situation.

2. General image processing programs

Many programs in Bsoft deal with rather simple and mundane operations on images, although such operations provide essential tools to manipulate images in various ways (Table 1). The header information of an image

can be quickly checked for features such as the data type, size and pixel size, and the user may set some of these parameters (*bhead*).² Many programs can be used to convert from one format to another simply by reading and writing an image, with the output format determined by the extension (e.g., *map* is the CCP4 3D image format). The program *bimg* can be used to do many simple operations on images, such as inverting or rescaling the density, or translating and changing the dimensions of the image. The program *bint* performs operations requiring some form of interpolation, such as scaling and rotating the image. The options provided are general and the user has full control over the transformations, therefore, care must be taken to specify the correct origin for an image for scaling and rotation. Various filtering operations can be done with the program *bfilter*, including a generalized kernel-based filtering, where the kernel can be constructed as a text file or using the program *bkern*. Pairs of images can be compared in various ways, and difference images calculated (*bop*).

The program *bshow* is the primary visualization tool, allowing display of images based on the five-dimensional internal image model: number of channels (e.g., gray-levels, or red–green–blue [RGB] color, *x* coordinate, *y* coordinate, *z* coordinate, and number of sub-images). It also provides capabilities for interactive manipulations such as modifying the density range, generating a power spectrum, and calculating a histogram. Micrographs can be processed in various ways, such as picking particles for single particle analysis and markers for tomographic series alignment (see later). The associated parameters can be read and written as text files for use by other programs (menu items “Micrograph/Read parameters” and “Micrograph/Write parameters”). Multiple 2D images or slices of a 3D image can be arranged in a montage in *bshow* (menu item “Image/Modify image”) or on the command line (*bmontage*). The image model in Bsoft supports various color models, including bitmap, gray-scale and RGB (red–green–blue), allowing the user to convert from or to color (*bcolour*), and to display colored images in *bshow*.

3. The concept of workflow in Bsoft

Bsoft offers many command-line programs, and only one interactive program, *bshow*. The most compelling advantage of command-line programs is that they are easily incorporated in shell scripts (such as *csh*, *Perl* and *Python*), allowing very complicated ways of processing that are completely up to the user. At the same time, there are several processing workflows that do not vary much from project to project, and they warrant their own programs or shell scripts. The programs therefore

² Relevant Bsoft programs will be mentioned where appropriate (often in brackets).

Download English Version:

<https://daneshyari.com/en/article/2829473>

Download Persian Version:

<https://daneshyari.com/article/2829473>

[Daneshyari.com](https://daneshyari.com)