

Data bases and ontologies

VANO: a volume-object image annotation system

Hanchuan Peng*, Fuhui Long and Eugene W. Myers

Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia, VA, USA

Received on December 9, 2008; revised on January 14, 2009; accepted on January 16, 2009

Advance Access publication February 2, 2009

Associate Editor: Jonathan Wren

ABSTRACT

Volume-object annotation system (VANO) is a cross-platform image annotation system that enables one to conveniently visualize and annotate 3D volume objects including nuclei and cells. An application of VANO typically starts with an initial collection of objects produced by a segmentation computation. The objects can then be labeled, categorized, deleted, added, split, merged and redefined. VANO has been used to build high-resolution digital atlases of the nuclei of *Caenorhabditis elegans* at the L1 stage and the nuclei of *Drosophila melanogaster*'s ventral nerve cord at the late embryonic stage.

Availability : Platform independent executables of VANO, a sample dataset, and a detailed description of both its design and usage are available at research.janelia.org/peng/proj/vano. VANO is open-source for co-development.

Contact: pengh@janelia.hhmi.org

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

Image content annotation is a basic problem in the analysis of 3D high-resolution cellular and molecular images (Peng, 2008). Many methods have been developed for determining annotations computationally such as categorizing gene expression patterns (e.g. Zhou and Peng, 2007) or predicting cell identities (e.g. Long *et al.*, 2008a). Critical to their development is the availability of a corpus of curated training data, and since none of these methods is perfect, their application in the field is benefited by having the ability to manually curate the results they produce. VANO, short for volume-object annotation system, was developed specifically to allow one to produce annotations manually and to correct or refine the output of good, but not perfect, automated annotation methods. This tool has and continues to play a critical role in our recent work in building 3D digital atlases of the L1 stage of *Caenorhabditis elegans*, and the late embryo ventral nerve cord and adult brain of *Drosophila melanogaster*.

VANO is a cross-platform 3D annotator that provides a spreadsheet of all 3D image objects that is linked to both 3D view of the raw image and a segmentation 'mask' that specifies the voxels that belong to each object (Fig. 1). For a given object, one can label it and set any number of user-defined attributes either in the spreadsheet or from the 3D views. Moreover, objects can be created, deleted, split, merged or redefined by commands that directly manipulate the segmentation mask. In typical use scenarios,

VANO is started (i) on just a raw image with no objects; (ii) with a segmentation mask produced by a computation (e.g. Long *et al.*, 2007); or (iii) with both a segmentation mask and an initial annotation (i.e. label and attributes), produced again by an automated computation (e.g. Long *et al.*, 2008a) or other means.

Image visualization tools such as Amira (amiravis.com) provide for flexible painting or refinement of a segmentation mask but do not support the concept of text annotations thereof. Existing image-tagging tools such as the annotator of the Open Microscope Environment (www.openmicroscopy.org) permit one to input text descriptions, but only for the entire stack and not for single objects. VANO allows one to define and annotate an arbitrary collection of 3D image objects in an interactive and efficient way.

Due to space limitations, we briefly describe the design, usage and some applications of VANO. The Supplementary Material, downloadable at the VANO website, provides detailed documentation.

2 DESIGN

VANO is designed to view, select, annotate and modify image objects with a simple and intuitive GUI which consists of two main panels—a *3D image viewer* and an *annotation table viewer*—and a number of auxiliary popup dialogs (see Supplementary Material).

The image viewer consists of two *tri-views*, one for the raw image above and another for the segmentation mask, along with a control panel at right providing display options and widget-based control of these views (Fig. 1a). The three images of a tri-view display the *XY*, *ZY* and *XZ* planes passing through a *current focus point* that is indicated by the intersection of dashed lines superimposed on each image. Both tri-views use the same focus point and thus always show the same planes and all views change in correspondence to the movement of this focus point whose location is adjusted by clicking in an image or moving a slider in the control panel.

The annotation table viewer is an editable spreadsheet (Fig. 1b) where each row contains the annotation information for an object and each column contains a particular annotation attribute. A typical annotation entry has 'standard' columns, including object order, label, standard name, user comment, 3D center coordinates, volume and the average, standard deviation, peak voxel intensity and mass (= average intensity × volume) of the object. The later attributes starting with the center are computed directly from the mask image. Note that mass in effect gives the level of gene expression or protein abundance in application that require it. In addition, customized annotation columns can be added via a schema file mechanism. Annotations can be entered directly in this spreadsheet or can be

*To whom correspondence should be addressed.

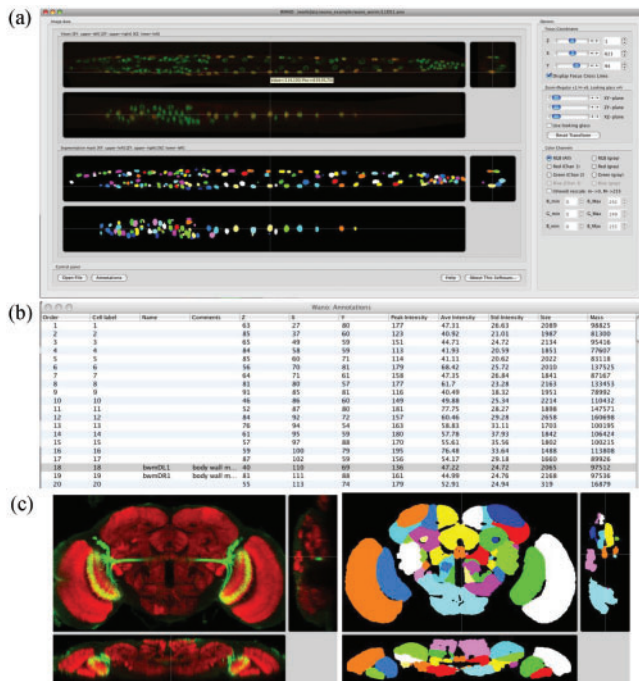


Fig. 1. GUI and applications of VANO in annotating 3D *C.elegans* and *D.melanogaster* image stacks. (a) The 3D image viewer: a tri-view for the subject image (true color, upper left), a tri-view for the mask image (in pseudo-color, lower left) and control panel (right). The displayed images are confocal stacks of a *C.elegans* larva (data by Xiao Liu and Stuart Kim). The worm in this stack was straightened using the algorithm of Peng *et al.* (2008). (b) The annotation table viewer: a spreadsheet where each row is an annotation entry for an image object and each column holds a particular attribute of an object. In this example each object is a nucleus. (c) Application of VANO to annotating neuronal patterns in the brain of *D.melanogaster*. Left: the subject image where NC82 (red) stains synaptic density and GFP (green) stains a target set of neurons (data by Julie Simpson). Right: the mask image where each object, in a distinct pseudo-color, is a glia-delineated compartment of the brain (mask by Arnim Jenett).

entered in dialogs that pop-up for the object under the current focus when an ‘E’ is typed. VANO uses the model-view-controller paradigm (Reenskaug, 1979) so changes made in one view are immediately reflected to all other views of the underlying data.

VANO is launched with either a subject image stack, in which case the mask and annotations are assumed to be empty, or a previously saved or manually prepared ‘.ano’ linker file that contains (derived) names of files containing the subject image, the mask image, the annotation and optionally the attribute schema. The image files are stored in TIFF format, and the annotation is stored in simple comma separated values format so it can easily be parsed and used by other programs such as Excel.

VANO is coded in C++ and uses the platform-independent QT (trolltech.com) GUI library so that it runs under all major operating systems, e.g. OS X, Windows NT/XP and most variants of Unix.

3 USAGE

VANO can be used to (i) create masks and annotations for objects from scratch; (ii) correct the segmentation of an existing mask; or

(iii) enter, edit or correct the annotation of the set of objects in the current mask. The combination of a spreadsheet that lists all objects and pop-ups for individual objects in an image view makes it easy to survey annotation and yet keep track of hundreds of objects as one progresses with the markup of such a collection.

Editing or entering annotation is simply a matter of editing or entering within the spreadsheet or within a text dialog of a pop-up. Editing segmentation in the general case requires the ability to paint arbitrary sets of pixels. This is extremely tedious and not really scalable to our scenario. Fortunately, the objects of interest to us are always globular and sufficiently modeled by a sphere. So, while VANO supports an arbitrary mask produced by a segmentation program, the simple manual edits supported are (i) deleting an object; (ii) merging an object adjacent to another, (iii) adding a spherical object of some radius; (iv) splitting an object by replacing it with two or more spheres; and (v) redefining an object by replacing it with a sphere. The restriction to spheres has not been found to be limiting and if somewhat more tailored masks are desired we find combining 2–4 spheres generally suffices to give a ‘snug’ mask.

4 APPLICATIONS

VANO allows one to conveniently create or efficiently edit a segmentation and annotation of a collection of globular objects. As such it played a critical support role in building a 3D digital nuclei atlas for *C.elegans* (Fig. 1a and b; collaboration with Stuart Kim lab at Stanford) and a 3D digital atlas of the nuclei of the ventral nerve cord of late stage *D.melanogaster* embryos (collaboration with Chris Doe lab at Oregon). For these two applications, VANO was used to produce the training sets for our automated algorithms (Long *et al.*, 2007, 2008a), to assess the quality of the 3D segmentations and annotations of nuclei produced by these algorithms, and to correct errors wherever needed to produce near perfect reference results.

VANO can also be used to annotate images in contexts where the image objects have an irregular shape. For example, we have been using VANO to annotate the enervation pattern of target neurons in defined compartments of the adult brain of *D.melanogaster* (collaborations with Gerry Rubin and Julie Simpson laboratories at Janelia Farm Research Campus). In Figure 1c, each glia-isolated compartment is defined as a separate object with a complicated 3D shape. VANO enables us to conveniently interact with these 3D objects and annotate the neuronal distribution pattern within each compartment. VANO can be used in many situations in biomedical imaging where the annotation of automatically segmented images is desired.

In addition, VANO can be used to collect the prior information (e.g. 3D location) and statistics (e.g. mean/peak intensity) of image objects. The information is useful for developing 3D segmentation methods of cells or intracellular organelles, or related applications.

ACKNOWLEDGEMENTS

We thank Xiao Liu and Stuart Kim for feedback in annotating *C.elegans* data; Mike Layden, Ellie Heckscher and Chris Doe for feedback in annotating fruit fly embryo data; Arnim Jenett, Gerry Rubin and Julie Simpson for feedback on annotating fruit fly adult brain data; and Frank Midgley for feedback on the design of VANO. We thank Zongcai Ruan for assistance in writing batch-compiling scripts and Margaret Jefferies for editorial review of the manuscript.

Conflict of Interest: none declared.

REFERENCES

- Long,F. *et al.* (2007) Automatic segmentation of nuclei in 3D microscopy images of C. elegans. In *Proceedings of the IEEE ISBI'2007*, pp. 536–539.
- Long,F. *et al.* (2008a) Automatic recognition of cells (ARC) for 3D images of C. elegans. *Lecture Notes in Computer Science: Research in Comp. Mol. Biology*, Vol. 4955, pp. 128–139. Springer, Berlin, Heidelberg.
- Long,F. *et al.* (2008b) A 3D digital atlas of C. elegans and its application to single-cell expression analysis. *HHMI JFRC Technical Report*.
- Peng,H. (2008) Bioimage informatics: a new area of engineering biology. *Bioinformatics*, **24**, 1827–1836.
- Peng,H. *et al.* (2008) Straightening Caenorhabditis elegans images. *Bioinformatics*, **24**, 234–242.
- Reenskaug,T. (1979) THING-MODEL-VIEW-EDITOR: an Example from a planning system. Xerox Parc technical note. 1978–1979.
- Zhou,J. and Peng,H. (2007) Automatic recognition and annotation of gene expression patterns of fly embryos. *Bioinformatics*, **23**, 589–596.