Reports

TLM-Converter: reorganization of long time-lapse microscopy datasets for downstream image analysis

Wee Choo Puah[1], Leong Poh Cheok[1], Maté Biro[1,2], Wee Thong Ng[1], and Martin Wasser[1,3]
*[1]Bioinformatics Institute, A*STAR, 30 Biopolis Street, Singapore, [2]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany, and [3]Department of Biological Sciences, National University of Singapore, Singapore*

BioTechniques 51:49-54 (July 2011) doi 10.2144/000113704
Keywords: Multidimensional imaging; confocal laser scanning microscopy; data reorganization; time-lapse imaging

Automated microscopy enables in vivo studies in developmental biology over long periods of time. Time-lapse recordings in three or more dimensions to study the dynamics of developmental processes can produce huge data sets that extend into the terabyte range. However, depending on the available computational resources and software design, downstream processing of very large image data sets can become highly inefficient, if not impossible. To address the lack of available open source and commercial software tools to efficiently reorganize time-lapse data on a desktop computer with limited system resources, we developed TLM-Converter. The software either fragments oversized files or concatenates multiple files representing single time frames and saves the output files in open standard formats. Our application is undemanding on system resources as it does not require the whole data set to be loaded into the system memory. We tested our tool on time-lapse data sets of live *Drosophila* specimens recorded by laser scanning confocal microscopy. Image data reorganization dramatically enhances the productivity of time-lapse data processing and allows the use of downstream image analysis software that is unable to handle large data sets of ≥2 GB. In addition, saving the outputs in open standard image file formats enables data sharing between independently developed software tools.

Time-lapse imaging using automated microscopy can capture developmental processes in live plants and animals at high spatial and temporal resolution (1,2). The multidimensional image acquisition of live specimens over several hours or days, and often involving multiple fluorophores, can produce sizeable data sets that extend into the terabyte range. Subsequent analysis of 5-D (3-D, time, color) image data involves 3-D reconstruction, visualization, segmentation and tracking (3). As these tasks are CPU and memory intensive, software tools for microscopic image analysis require dedicated high-end graphics workstations (4,5) or computer clusters (6,7) for optimal performance. However, depending on available (suboptimal) hardware resources and software design, downstream processing of very large image data sets can become highly inefficient, if not impossible. For instance, tools developed for 32-bit operating systems are often unable to handle data sets beyond a size of 2 GB. An obvious solution is to fragment large image data sets and to perform analysis in smaller subsets. Although file fragmentation can be achieved using existing software, its execution is often not productive enough for routine use. Some tools, like the Zeiss (Jena, Germany) LSM Image Browser or Zen Light Edition require loading of the entire image data set before data fragmentation is permitted. Other tools like the LOCI plug-in within ImageJ (8) can open and save subsets. However, image subsets cannot be larger than the available memory allocated to ImageJ.

While oversized image files can pose challenges, overfragmentation of data (e.g., saving one file per time frame) can equally slow down data analysis. Motorized microscopes allow the imaging of live samples in multiple locations (www.nature.com/nmeth/journal/v3/n11/full/nmeth971.html) (9), further increasing the amount of data generated per experiment. Specialized acquisition software tools, such as the Zeiss Multiple-Time-Series macro, usually save one file per time point and location. After completion of image acquisition, temporary files are concatenated. The problem is that concatenation can be time consuming and may take several hours for data spanning several gigabytes. If the microscope belongs to a shared facility, precious acquisition time will be lost for the next user. Moreover, in our experience, concatenation may simply abort if the acquisition involves images from multiple locations. An obvious solution to this problem would be a standalone tool that can perform time-lapse concatenation. To our knowledge, such a concatenation function is unavailable in most free image processing tools (10).

To overcome these above-mentioned bottlenecks in handling large time-lapse microscope data sets, we developed the TLM (time-lapse microscopy)–Converter. The TLM-Converter reorganizes time-lapse microscopy files based on user specifications and saves the output in open standard file formats that are essential when the images and their associated metadata are to be

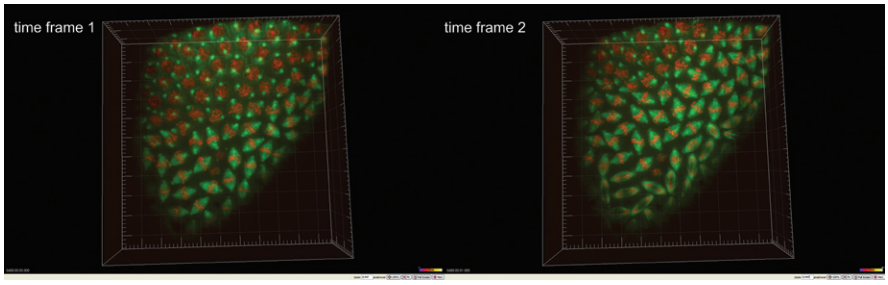

Figure 1. Volume views of 3-D image stacks of live *Drosophila* embryos acquired by confocal laser scanning microscopy. Automated microscopes produce multidimensional data that capture the dynamics of living cells in 3-D. The example shows two frames of a *Drosophila* embryo expressing the live reporters histone-mRFP (red) and tubulin-GFP (green) that highlight the progression through the cell cycle. 100 frames of these images amount to 14 GB of raw data, which exceeds the memory of most desktop computers. In addition, many image processing tools may fail to load subsets of data. Applications developed for 32-bit operating systems can usually only load the first 2 GB of a large data set. Volume rendering was performed using Imaris X64.

shared between independently developed software tools (11). This data reorganization capability complements numerous image analysis software packages that are not designed to handle large volumes of data. The TLM-Converter contains two main modules, one for fragmentation of oversized files and another one for the concatenation of multiple stacks representing single time points.

Materials and methods

The TLM-Converter was developed to facilitate the reorganization of time-lapse data using two modules with associated user interfaces: one for file fragmentation and one for concatenation. To test the TLM-Converter, we recorded time-lapse data of live *Drosophila* embryos (Figure 1) and pupae using the Zeiss LSM 5 Live laser scanning microscope (9) as previously described (12).

The TLM-Converter software was written in C++ with a Java based graphical user interface (GUI) and utilizes the libtiff (www.libtiff.org) and libics (http://libics.sourceforge.net) libraries. It runs on the Windows platform (XP, Vista, and Windows 7) and supports Carl Zeiss confocal laser scanning microscopy (LSM) proprietary files, Image Cytometry Standard (ICS) and Tagged Image File (TIF) formats.

The software and two time-lapse data sets can be downloaded from http://web.bii.a-star.edu.sg/archive/TLM-Converter/User_License_DownloadInstructions_TLM_rev01.htm. The software contains a help manual.

Image data fragmentation

The data fragmentation module (Figure 2A) serves to divide very large multidimensional data sets, especially those with file sizes exceeding the capacity of the system resources (i.e., memory, operating system), into user-defined subsets for more efficient downstream image processing. It is useful when limited system resources prevent the loading of complete data sets or a required image analysis tool lacks the ability to load subsets of data. Another purpose of fragmentation is to facilitate the parallel processing of time-consuming tasks like 3-D deconvolution and 3-D segmentation. The software reads the metadata whenever available and displays the parameters of the input file. Based on dimensionality and size of the input file, the user can specify the subset of data to be stored in the output file in terms of range of time frames, optical sections and color channels. During execution, image data are read from the source and written into the target file one 2-D image at a time. As such, the memory consumption is determined by the size of an individual 2-D image and not the data subset that can span several gigabytes.

To facilitate data sharing between independently developed software, output files are saved in open standard ICS and TIF file formats. ICS was chosen as the preferred output format as it can store images of any dimensionality and data type, alongside the microscopy acquisition parameters (i.e., metadata) (13). The supported input and output formats are summarized in Table 1. The ICS format can be read and written by a variety of commercial and open-source image processing tools. Our software supports both ICS versions 1.0 and 2.0. To save files >2 GB, we recommend the use of ICS version 1.0. In batch fragmentation mode, the input file can be fragmented into multiple output files. The user can choose the number of time frames to save in each output file. This option of batch fragmentation relieves the user from multiple execution of the same sample input.

Table 1. Input and output file formats supported by TLM-Converter.

Conversion Task	Input Format	Output format
Fragmentation	Single ICS or LSM	Single ICS Multiple TIFF files
Batch Fragmentation	Single ICS or LSM	Multiple ICS
Time-lapse Concatenation	Multiple LSM (1 per time point)	Single ICS
Concatenation with editing of metadata	Multiple TIFFs	Single ICS

Time-lapse concatenation

The Zeiss Multiple-Time-Series (MTS) macro records up to six dimensions involving multiple locations, channels, and time frames, and saves the time-lapse data as individual Z-stacks. The concatenation module of the TLM-Converter (Figure 2B) concatenates the individual files into larger user-defined subsets of the time-lapse experiment that are more convenient for further image analysis. The inputs are LSM files that are acquired from a time-lapse experiment, stored in the same folder, and whose names contain consecutive numbers that represent the single time frames. The user selects one of the LSM files and specifies a template, using question marks as wild cards for the time frame number. The TLM-converter reads the input image parameters and automatically counts the files of the time-lapse sequence that matches the template. The tool provides users with the flexibility to specify a subset of data to incorporate into the reorganized time-lapse file, including the range of Z-sections, time frames, start time, time intervals, and color-channels.

Furthermore, we implemented a utility that can concatenate multiple TIFF files into a single ICS file. Input filenames must match a pattern (*ch##t####z###.tif) that specifies color channel (ch), time point (t) and optical section (z) of each 2-D image. Prior to concatenation, the user can edit the metadata of the output ICS file to be saved.

Results and discussion

We designed TLM-converter to facilitate the handling of large time-lapse microscopy data sets. In our laboratory practice, we routinely

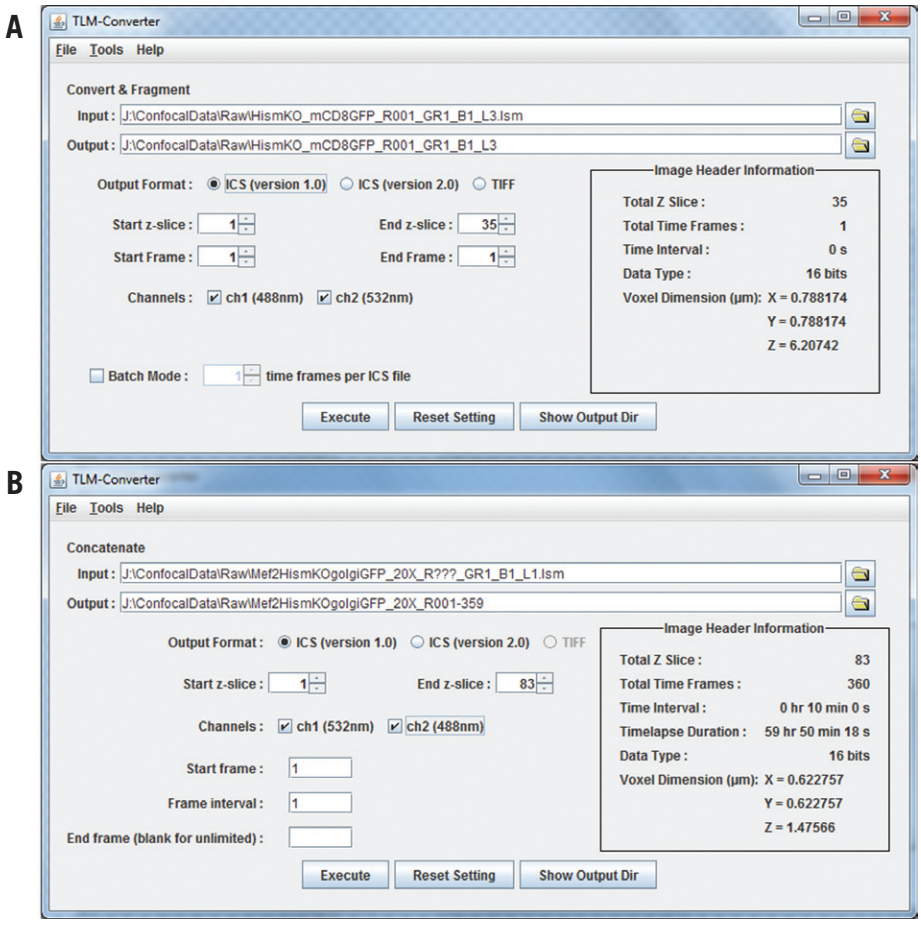

Figure 2. Graphical user interface (GUI) of the TLM-Converter. (A) In the data fragmentation module, the user can specify ranges of optical sections, time frames and color channels to be included into the output files. In batch mode, a single input file can be fragmented into multiple files (e.g., 100 time frames can be broken down into 10 files of 10 frames each). (B) The concatenation module assembles multiple multidimensional files corresponding to single time frames (up to four dimensions) into single output files. Question marks in the input filename indicate the frame number. Users can choose the set of colors, range of time frames and optical sections to incorporate into the output file. The frame interval section allows the skipping of frames.

generate 3-D time-lapse image files of *Drosophila* embryogenesis (Figure 1) or metamorphosis (12) that range in size between 12 GB and 64 GB. However, hardware constraints such as insufficient memory or restrictions imposed by the software—such as the inability of 32-bit applications to handle addresses >2 GB—can severely affect our ability to analyze and visualize those image files. TLM-Converter helps to overcome this bottleneck by fragmenting oversized image data sets into more manageable portions. We tested the fragmentation module of the TLM-Converter with a 66-GB dual-channel 3-D time-lapse image file containing 480 time frames. Each time frame consisted of 35 z-sections of 1024 × 1024 pixels in the *x*- and *y*-axes. On a Desktop PC with an Intel core i7 processor (3.07 GHz) and 12 GB RAM, it took 48 min to convert the single LSM file into six ICS files of 80 frames each. The integrity of the output files was confirmed using ImageJ, Huygens Professional (Scientific Volume Imaging, Hilversum, The Netherlands), and Imaris (Bitplane, Zurich, Switzerland). Apart from overcoming hardware and software constraints, a further benefit of fragmentation is the ability to distribute image processing tasks that can be parallelized (e.g., deconvolution or 3-D segmentation) across multiple CPUs. Subsampling of multidimensional image data can also be accomplished with ImageJ, LSM image browser, or Zen Light Edition, although in a less-productive manner. In ImageJ, the whole subset needs to be loaded into system memory before saving to hard disk. LSM Image browser and Zen Light also have the option to load images exceeding the size of the system memory by using the system's allocated virtual memory. As the TLM-Converter iteratively reads one 2-D frame at a time into the system memory and writes it to a new file on the hard disk, the subsampling is carried out more efficiently. In addition, saving multiple data subsets using ImageJ, LSM image browser, or Zen Light Edition is more time-consuming since these tools lack batch-processing capabilities.

Automated motorized microscopes have the ability to perform the type of 5-D imaging described above in multiple locations. Recorded image data are usually saved as temporary files, one for each time point and location. After completion of image acquisition, the temporary files are concatenated into a single file [e.g., using Zeiss's Multiple Time Series Macro (MTS) macro]. Since users might decide not to concatenate on the microscopy workstation [it can take several hours or, as we encountered, concatenation may simply abort when the recorded image data reaches a considerable size (≥100 GB)], we implemented a module for time-lapse concatenation of LSM files. To test the time-lapse concatenation module, we used a time-lapse dual-channel 8-bit data set consisting of 300 image stacks that each contained 38 optical sections acquired using the MTS macro mode. The software took 11 min to concatenate 300 LSM files into a single ICS file of 23 GB. The only other free software tool that can accomplish this concatenation task is Zen 2009 Light Edition. However, Zen

2009 Light Edition requires human intervention to manually concatenate one stack at a time until system memory is depleted. This means that concatenation of all 300 LSM files is not possible on our desktop PC with 12 GB of RAM. Users can control the range and temporal sampling. Since time-lapse experiments can involve several hundred time frames recorded simultaneously in multiple locations, our concatenation tool removes another bottleneck in image visualization and analysis.

In summary, TLM-Converter is a user-friendly data reorganization tool for time-lapse data sets larger than system resources, and which runs on single desktop computers. Sizeable multidimensional data sets can be reorganized to match available software and hardware resources, saving the need to invest in new computational equipment. In the future, we intend to include the Bio-Formats library (www.loci.wisc.edu/software/bio-formats), which supports 106 image formats to expand the list of file formats that can be imported into the TLM-converter.

Acknowledgments

We thank Carl Zeiss Singapore for providing the manual containing specifications of the LSM file format.

Competing interests

The authors declare no competing interests.

References

1. **Meijering, E., O. Dzyubachyk, I. Smal, and W.A. van Cappellen.** 2009. Tracking in cell and developmental biology. Semin. Cell Dev. Biol. *20*:894-902.
2. **Schmitz, M.H.A. and D.W. Gerlich.** 2009. Automated live microscopy to study mitotic gene function in fluorescent reporter cell lines. Methods Mol. Biol. *545*:113-134.
3. **Walter, T., D.W. Shattuck, R. Baldock, M.E. Bastin, A.E. Carpenter, S. Duce, J. Ellenberg, A. Fraser, et al.** 2010. Visualization of image data from cells to organisms. Nat. Methods *7*:S26-S41.
4. **Wu, Q., F.A. Merchant, and K.R. Castleman.** 2008. Microscope image processing. Academic Press.
5. **Looney, M.R., E.E. Thornton, D. Sen, W.J. Lamm, R.W. Glenny, and M.F. Krummel.** 2011. Stabilized imaging of immune surveillance in the mouse lung. Nat. Methods *8*:91-96.
6. **Keller, P.J., A.D. Schmidt, J. Wittbrodt, and E.H.K. Stelzer.** 2008. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science *322*:1065-1069.
7. **Carpenter, A.E., T.R. Jones, M.R. Lamprecht, C. Clarke, I.H. Kang, O. Friman, D.A. Guertin, J.H. Chang, et al.** 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. *7*:R100.
8. **Collins, T.J.** 2007. ImageJ for microscopy. BioTechniques *43*:25-30.
9. **Wolleschensky, R., B. Zimmermann, and M. Kempe. 2006. High-speed confocal fluorescence** imaging with a novel line scanning microscope. J. Biomed. Opt. *11*:064011.
10. **Swedlow, J.R. and K.W. Eliceiri.** 2009. Open source bioimage informatics for cell biology. Trends Cell Biol. *19*:656-660.
11. **Linkert, M., C.T. Rueden, C. Allan, J.-M. Burel, W. Moore, A. Patterson, B. Loranger, J. Moore, et al.** 2010. Metadata matters: access to image data in the real world. J. Cell Biol. *189*:777-782.
12. **Du, T. and M. Wasser.** 2009. 3D image stack reconstruction in live cell microscopy of Drosophila muscles and its validation. Cytometry A *75*:329-343.
13. **Dean, P., L. Mascio, D. Ow, D. Sudar, and J. Mullikin.** 1990. Proposed standard for image cytometry data files. Cytometry *11*:561-569.

Received 30 March 2011; accepted 23 May 2011.

Address correspondence to Wee Choo Puah, Bioinformatics Institute, A*STAR, 30 Biopolis Street #07-01 Matrix, Singapore 138671. e-mail: puahwc@bii.a-star.edu.sg