Bioinformatics, 32(2), 2016, 315–317
doi: 10.1093/bioinformatics/btv559
Advance Access Publication Date: 28 September 2015
Applications Note

Bioimage informatics

XPIWIT—an XML pipeline wrapper for the Insight Toolkit

Andreas Bartschat, Eduard Hübner, Markus Reischl, Ralf Mikut and Johannes Stegmaier*

Institute for Applied Computer Science (IAI), Karlsruhe Institute of Technology, Germany

*To whom correspondence should be addressed.

Associate Editor: Robert Murphy

Received on June 25, 2015; revised on August 31, 2015; accepted on September 20, 2015

Abstract

Summary: The Insight Toolkit offers plenty of features for multidimensional image analysis. Current implementations, however, often suffer either from a lack of flexibility due to hard-coded C++ pipelines for a certain task or by slow execution times, e.g. caused by inefficient implementations or multiple read/write operations for separate filter execution. We present an XML-based wrapper application for the Insight Toolkit that combines the performance of a pure C++ implementation with an easy-to-use graphical setup of dynamic image analysis pipelines. Created XML pipelines can be interpreted and executed by XPIWIT in console mode either locally or on large clusters. We successfully applied the software tool for the automated analysis of terabyte-scale, time-resolved 3D image data of zebrafish embryos.

Availability and implementation: XPIWIT is implemented in C++ using the Insight Toolkit and the Qt SDK. It has been successfully compiled and tested under Windows and Unix-based systems. Software and documentation are distributed under Apache 2.0 license and are publicly available for download at https://bitbucket.org/jstegmaier/xpiwit/downloads/.

Contact: johannes.stegmaier@kit.edu

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

1 Introduction

Automated analysis of large-scale multidimensional image data has become an integral part of current biological research. Particularly, 3D imaging techniques like confocal, light-sheet or electron microscopy easily produce terabytes of image data that cannot be assessed manually (Mikut *et al.* 2013; Tomer *et al.* 2012). Although many comprehensive and user-friendly graphical user interfaces (GUIs) for the automation of image analysis problems have been presented in the past, their applicability to terabyte-scale data remains limited (Eliceiri *et al.* 2012). On the other hand, tools that are usable for large-scale analyses are mostly single-purpose command-line tools with a substantial lack of flexibility. The Insight Toolkit (ITK), for instance, often serves as a basis for such specialized high-performance implementations, as it offers plenty of features for multidimensional image analysis and has an active community in the biomedical field that constantly improves and extents the functionality of ITK (Ibanez *et al.* 2005). However, many available solutions based on ITK require advanced programming skills, are limited to a certain task due to specialized C++ pipelines or suffer from slow execution times, e.g. caused by multiple I/O operations for separate filter execution or inefficient implementations. To overcome these limitations, we developed XPIWIT, an XML-based wrapper application for ITK that allows a graphical setup and rapid prototyping of image analysis pipelines while preserving the performance of a pure C++ implementation (Supplementary Fig. S1). Created XML pipelines can directly be used to interface XPIWIT on large clusters, and the current version of XPIWIT already incorporates about 70 different filters, including I/O filters, preprocessing filters, edge detectors, morphological operators and segmentation filters (Supplementary Table S1). We also integrated our recently published algorithm for the efficient segmentation of fluorescently labeled cellular nuclei as an exemplary pipeline (Stegmaier *et al.* 2014). To ensure

XPIWIT can be constantly extended with new functionality, we created template files that facilitate the implementation of new filters for XPIWIT. Filters that are already implemented as image-to-image filters in ITK can be added to XPIWIT in a matter of minutes and are instantly accessible via the XML interface. Because of the flexible design, literally any additional C++ code can be encapsulated as an XPIWIT filter, as long as the input and output images are appropriately casted to ITK-internal types. The platform-independent open-source software tool was successfully applied to automatically analyze multiple terabytes of time-resolved 3D microscopy image data in embryomics using a Hadoop cluster (Kobitski *et al.* 2015, Supplementary Section S2).

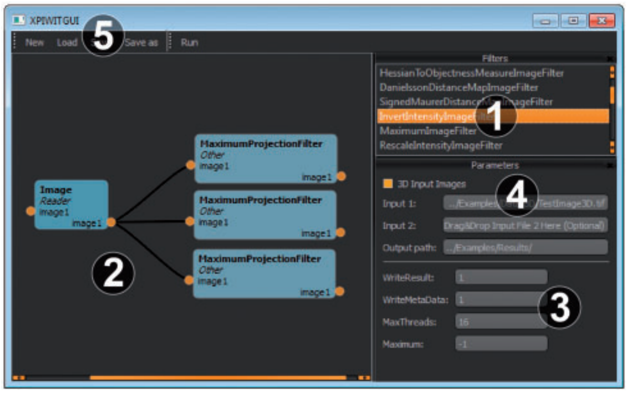
2 Methods

2.1 XML pipeline interface

A general concept of automated image processing, which is also inherent to ITK, is the arrangement of processing operators in a feed-forward pipeline structure. We specifically developed an XML-based pipeline format that allows to create flexible image analysis pipelines using a variety of different processing filters (Supplementary Listing S1). Each of the specified filters has a set of inputs that either point to files on disk or the output data of a preceding filter. In addition to image inputs and outputs, XPIWIT also has an internal metadata handling system that can be used to share meta-information such as extracted object locations or image statistics between processing operators. The optimal execution order of the individual processing operators is internally determined by XPIWIT and ensures that the requested image and meta-inputs are available as soon as a processing operator is executed and that memory is released as soon as the data is not needed anymore. The XML structure allows to adjust all available parameters of each filter using key-value pairs. The XML pipelines are decoupled from the I/O parameters, which facilitates sharing customized pipelines with other users and additionally allows to use a single XML pipeline to process multiple images.

2.2 Data generation

To apply a predefined XML processing pipeline to a desired set of images, XPIWIT has to be executed from the command prompt with command line input arguments. Alternatively, a configuration text file can be piped to the executable (Supplementary Listing S2). This configuration file needs to contain the output path, one or more input paths, the path to the XML file to be processed and may be customized using further optional parameters as described in the online documentation. This concept allows to heavily parallelize the data processing, e.g. using distributed computing environments such as Apache Hadoop, where one XPIWIT configuration file is generated per image and all jobs are distributed among available processing nodes. In addition to the image output, metadata generated by one of XPIWIT's processing operators is saved. The used CSV format (Comma Separated Value) is able to store large tables in a text file where each column is separated by a specific character (e.g. ',' or ';') and each row by a line break. Many powerful data analysis tools (e.g. Excel, MATLAB, Gait-CAD, SPSS, Knime and RapidMiner) provide functionality to import such CSV files and therefore allow to explore the generated data in detail. All data formats used by XPIWIT are either simple text files or common data formats. This allows third party software to easily automate the creation of input files, to interpret the generated data and to embed XPIWIT into existing image processing pipelines.



■ 3D Input Images

Input 1: ../Examples/ TestImage3D.tif

Input 2: Drag&Drop Input File 2 Here (Optional)

Output path: ../Examples/Results/

WriteResult: 1

WriteMetaData: 1

MaxThreads: 16

Maximum: -1

Fig. 1. Exemplary screenshot of the XPIWIT GUI on Windows. The main control elements of the GUI are the filter list (1), the pipeline drawing area (2), the parameter adjustment panel (3), the input/output definition panel used for pipeline test runs (4) and controls for saving, loading and executing pipelines (5). Generated pipelines are independent from the GUI, e.g. to perform large-scale analyses on computing clusters

2.3 GUI for rapid prototyping

We developed an easy-to-use GUI that allows to create and modify XML pipelines based on the filters compiled into the XPIWIT executable. The GUI is entirely decoupled from the command-line tool and offers a more convenient environment for rapid prototyping and parameter optimization than directly adjusting the XML text files. In Figure 1, the basic layout of the GUI is depicted. Filters can be placed via drag and drop from the filter list (1) to the pipeline drawing area (2). All parameters are adjustable by selecting the respective filter in the drawing area and by setting the parameters in the parameter customization panel (3). Once the input and output parameters (4) are defined, the pipelines can be used to process the specified data and may be saved and re-opened using the menu bar (5).

2.4 Implementation details

XPIWIT is implemented in C++ using ITK (http://www.itk.org) and the Qt SDK (http://www.qt-project.org). Platform-independent project files have been realized using the CMake build tool (http://www.cmake.org). Both XPIWIT and the associated GUI have been successfully compiled and tested under Windows, Linux and Mac OS X. The software is licensed under Apache 2.0, and we host a public repository that encompasses source code, example pipelines and a detailed documentation on https://bitbucket.org/jstegmaier/xpiwit/downloads/. To encourage non-programmers to use the software for their image analysis tasks, precompiled binaries for the three major operating systems are hosted in the repository as well.

3 Conclusions

XPIWIT is a new command-line tool based on ITK that combines the performance of a native C++ application with the ability to create filter pipelines at runtime using a specialized XML format. The functionality is completed with an interactive GUI that simplifies rapid prototyping, parameter optimization and the generation of XML pipelines. XPIWIT is explicitly not meant to replace any of the general purpose GUIs that exist for biomedical imaging but aims at filling the niche of an easy-to-use and flexible tool that is highly efficient, lightweight and platform independent. Both XPIWIT and the GUI open up a wide area of possible biomedical applications ranging from automated analyses of 2D image data on single workstations to large-scale analyses of multidimensional image data on high performance computing clusters. We successfully applied the presented software solution to segment and visualize terabyte-scale 3D+t light-sheet microscopy images of developing embryos. Further work will be put on a shared library-based plugin system to simplify adding new filters to XPIWIT and to combine a step-by-step execution with result previewing for interactive parameter optimization. Moreover, we will try to identify possibilities to integrate XPIWIT with widespread tools like Fiji, ICY or KNIME (Eliceiri *et al*. 2012).

Funding

This work was supported by the Helmholtz Association in the Program BioInterfaces and by the German Research Foundation DFG (grant no. MI 1315/4).

Conflict of Interest: none declared.

References

Eliceiri,K. *et al*. (2012) Biological imaging software tools. *Nat. Methods*, **9**, 697–710.

Ibanez,L. *et al*. (2005) *The ITK Software Guide*. Kitware, Inc. ISBN 1-930934-15-7.

Kobitski,A. *et al*. (2015) An ensemble-averaged, cell density-based digital model of zebrafish embryo development derived from light-sheet microscopy data with single-cell resolution. *Sci. Rep*., **5**, 1–10.

Mikut,R. *et al*. (2013) Automated processing of zebrafish imaging data—a survey. *Zebrafish*, **10**, 401–421.

Stegmaier,J. *et al*. (2014) Fast segmentation of stained nuclei in terabyte-scale, time resolved 3D microscopy image stacks. *PLoS One*, **9**, e90036.

Tomer,R. *et al*. (2012) Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nat. Methods*, **9**, 755–763.