



PROCEEDINGS

Open Access

An open-source, MATLAB based annotation tool for virtual slides

Riku Turkki^{1,2*}, Margarita Walliander¹, Ville Ojansivu¹, Nina Linder¹, Mikael Lundin¹, Johan Lundin^{1,3}

From 11th European Congress on Telepathology and 5th International Congress on Virtual Microscopy Venice, Italy. 6-9 June 2012

Background

Computer aided analysis of virtual slide images has become an increasingly popular research topic and area of development. Novel applications are aimed for instance at automation of certain stages in sample assessment and assay readout (e.g. segmentation, detection, etc.) [2] or for reproducible measurement of a sample's visual appearance (e.g. grading, morphological classes etc.) [3].

Statistical learning methods are one of the key algorithms used in building these image analysis applications. Especially supervised algorithms (e.g. support vector machines (SVM), AdaBoost, k nearest neighbour (kNN), etc.) offer a way to train a classifier that can perform complex quantification tasks. The process requires annotation of labelled examples in order to describe the task and learn a model. Recently, image acquisition techniques for large-scale digitization of tissue samples have become common and require new methods to perform the annotation and associated labelling [4].

In this study we present an annotation tool that combines 1) direct interaction with a remote slide collection (a web based virtual microscopy application), 2) an interface to annotate points of interest in the virtual slides and 3) fast transition from annotation to development of image analysis methods.

Methods

The annotation tool is written in MATLAB (matrix laboratory) that is a cross-platform numerical computing environment (MathWorks, Natick, MA). MATLAB is a tool with a wide range of application areas and adopted by the image analysis community because of its versatile features. The programming language and its computing

environment offers tools for testing and implementation of methods from a number manipulation to creating interactive graphical user interface (GUI) applications.

The annotation tool is based on exploiting two important properties of the current platform: 1) a *random code-stream access* featured image format (e.g. ECW or JPEG 2000) and 2) a *compression streamlining* protocol. The random code-stream compression makes it possible to extract sub-images from a large virtual slide file. Thus, the user can download only a field of interest, instead of loading the whole image file, which might be of gigapixel size. The other required property, the compression streamlining protocol, basically implements this sub-image extraction over standard hypertext transfer protocol (HTTP) in an efficient way.

With the advantages of the above-mentioned properties that a slide collection should have, it is reasonably straightforward to implement a MATLAB based tool to operate a remote database. The overall structure of the tool is illustrated in Figure 1. In a first phase the tool reads the metadata of a virtual slide file and deduces its height and width. Using a server's streamlining protocol, uniform resource locators (URL) are defined to extract the tiles from the pre-calculated coordinates within the virtual slide. In the annotation phase, the tiles are extracted and downloaded one at the time from the server and displayed to an annotator and requested for annotations. The locations and labels of the annotations are saved locally in MAT files.

Results and discussion

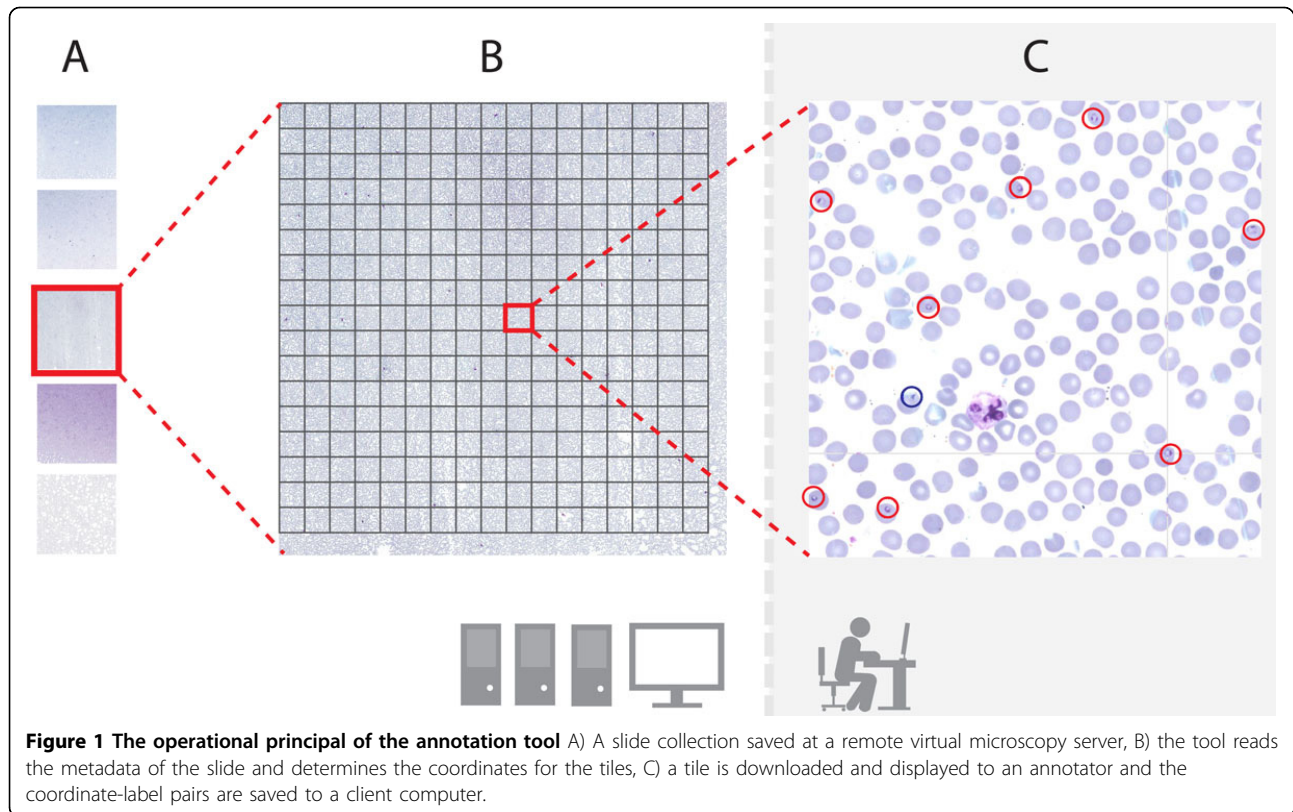
We implemented a MATLAB based annotation program capable of accessing a remote slide collection. The implementation is demonstrated within a previously described

* Correspondence: riku.turkki@helsinki.fi

¹Institute for Molecular Medicine Finland (FIMM), P.O. Box 20, FI-00014

University of Helsinki, Helsinki, Finland

Full list of author information is available at the end of the article



virtual microscopy environment [1,5] running image web server software (Erdas Inc., Atlanta Georgia), but is modifiable to pair with other platforms as well. The current virtual microscopy platform has proprietary server software (Image Web Server) that accepts both JPEG 2000 and ECW compressed files. The slides are made accessible via an *ImageX* protocol, which is implemented by the server. Other platforms, which have the ability to use JPEG 2000 files only, can utilize the JPIP (JPEG 2000 Interactive Protocol) to extract the metadata and tiles from a virtual slide over HTTP [6]. The code for the proposed tool is freely available [7].

The annotation tool keeps a record of all markings and enables the slides to be annotated in parts: it is possible to continue the annotation process from the beginning of a slide, or alternatively from a tile where the annotation was previously interrupted. The tool also keeps record of tiles that are already displayed to an annotator, which allows the areas that have not been annotated to be excluded from later processing if wanted. All the data are saved locally on the user's computer in standard MAT files.

The tiles are loaded one at the time from the image server and displayed to the annotator. The annotation window of an annotation process in a digitized thin blood smear film is illustrated in Figure 2. All the annotations

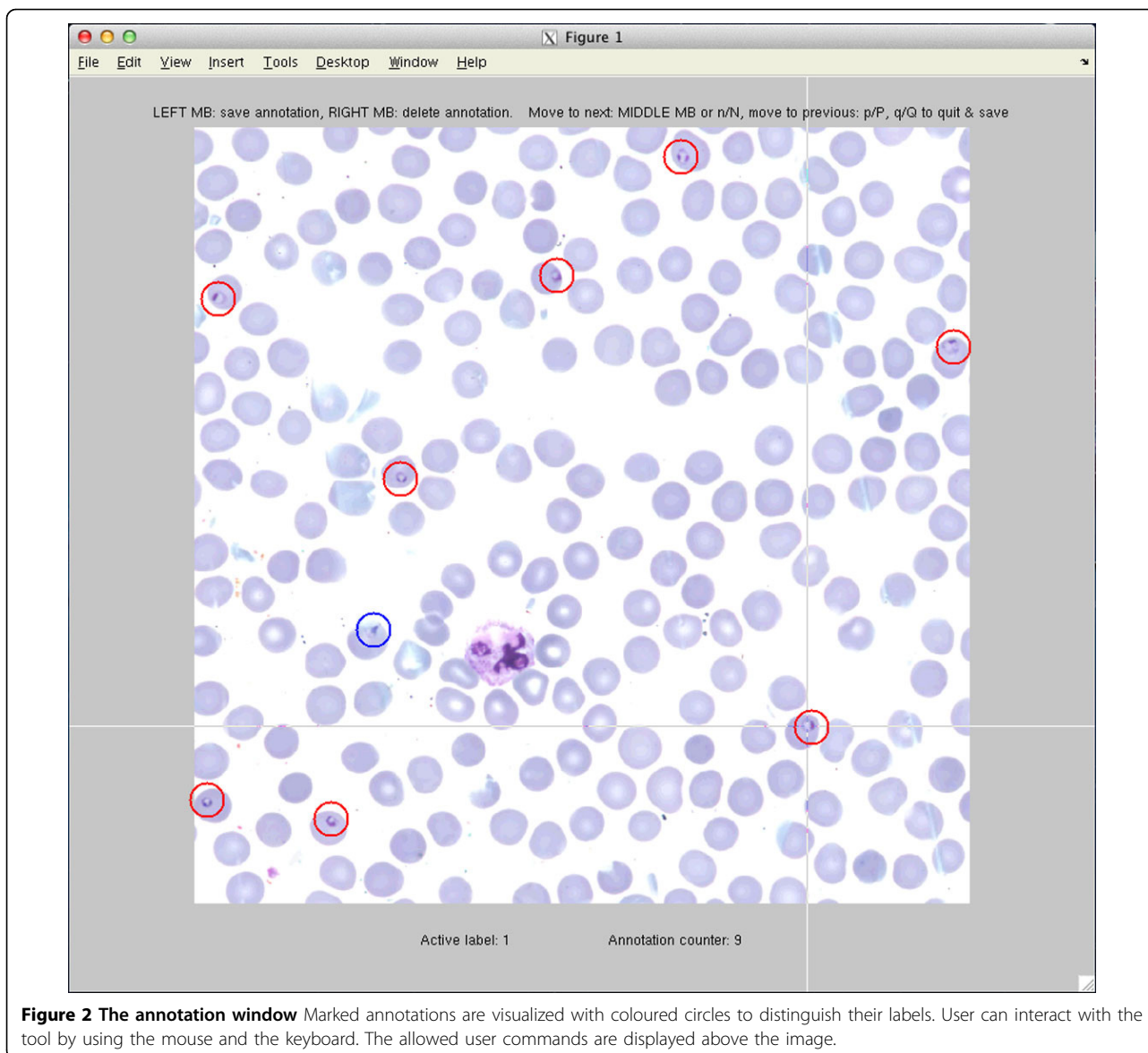
are visualized with circles, which are coloured according to their labels. The tool can be set to handle at the most ten different labels.

The actual annotation process is kept as simple as possible to provide a fast and easy-to-use interface. The user is able to add and remove an annotation by simple mouse clicks: the left button for adding a new annotation and the right button to delete the nearest detected annotation. The keyboard's number keys are used to set the active category label and by keys N (next) and P (previous) the user can move between the tiles.

The work was motivated by our own need to annotate large virtual slide collections. In addition to the proposed applications, the tool can be applied to any kind of analysis of virtual slides. For instance, the tool offers a way to compare the agreement and accuracy of human annotators to recognize and locate objects, or it can be used to deliver one's annotations to a shared resource.

Conclusions

We have described a simple and effective way to combine a computational environment with a virtual microscopy environment [1] to allow fast transition and iteration between method development and data annotation. The code for the annotation tool is freely available [7] and can be modified to suit different server settings. The use of the



tool is demonstrated on a virtual microscopy platform by providing example slides of thin blood films.

List of abbreviations

ECW: Enhanced Compressed Wavelet; GUI: Graphical User Interface; HTTP: Hypertext Transfer Protocol; kNN: k-Nearest Neighbour; SVM: Support Vector Machine; URL: Uniform Resource Locator

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RT planned and wrote the code for the annotation tool. MW and VO reviewed the implementations and suggested modifications. ML consulted on the use of the virtual microscopy platform. RT drafted the article and NL and JL contributed by supervision of the work and review of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The study was kindly supported by the national Biomedinfra and Biocenter Finland projects.

Authors' details

¹Institute for Molecular Medicine Finland (FIMM), P.O. Box 20, FI-00014 University of Helsinki, Helsinki, Finland. ²Center for Machine Vision Research, Department of Computer Science and Engineering, P.O. Box 4500, FI-90014 University of Oulu, Finland. ³Division of Global Health, Department of Public Sciences, Karolinska Institutet, Stockholm, Sweden.

Published: 30 September 2013

References

1. The WebMicroscope virtual microscopy environment. [<http://www.webmicroscope.net/>].
2. Linder N, Konsti J, Turkki R, Rahtu E, Lundin M, Nordling S, Ahonen T, Pietikainen M, Lundin J: Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. *Diagn Pathol* 2012, **7**:22.

3. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC: **Automatic breast cancer grading of histopathological images.** *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE 2008*, 3052-3055.
4. Kayser K: **Introduction of virtual microscopy in routine surgical pathology-a hypothesis and personal view from Europe.** *Diagn Pathol* 2012, **7**:48.
5. Szymas J, Lundin M: **Five years of experience teaching pathology to dental students using the WebMicroscope.** *Diagn Pathol* 2011, **6**(Suppl 1):S13.
6. **The description of JPIP-protocol.** [<http://www.jpeg.org/jpeg2000/j2kpart9.html>].
7. **The code for the proposed annotation tool.** [<http://firm.webmicroscope.net/Research/Momic/tp2012>].

doi:10.1186/1746-1596-8-S1-S30

Cite this article as: Turkki *et al.*: An open-source, MATLAB based annotation tool for virtual slides. *Diagnostic Pathology* 2013 **8**(Suppl 1):S30.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

