High performance vessel analysis tools: 3D vessel reconstruction and analysis tool using whole-slide images.

Google Summer of Code 2015 Proposal

Sayan Maity
PhD Student, University of Miami
sayanmaity.10@gmail.com

Abstract

Histopathological analysis is a common clinical procedure for diagnosing the presence, type, and progression of diseases such as cancer. However, manual examination and decision-making using tissue slides that may potentially contain millions of cells can be time-consuming and subjective. Approximately 80% of the 1 million biopsies performed in the US every year are benign; this suggests that pathologists are spending 80% of their time sieving through benign tissue which would otherwise be utilized is treating patients. Thus high performance vessel analysis tools have an enormous importance to automate dependable clinical decision support.

Contents

I. INTRODUCTION

Monitoring structural changes and spatial relationship of any diseased organ is of high importance in disease diagnosis and to better understand the disease's pathological evolutions and progressions. For example, in case of liver disease tracking the structural variation of liver vessels helps in prognosis and treatment. Compared with radiology image modalities, digital histopathology slides provide a more comprehensive view of disease and its effect on tissues, since the preparation process preserves the underlying tissue architecture. As such, some disease characteristics, e.g., lymphocytic infiltration of cancer, may be deduced only from a histopathology image. Additionally, the diagnosis from a histopathology image remains the 'gold standard' in diagnosing considerable number of diseases including almost all types of cancer. However, histopathology image analysis tools haven't been vastly investigated compared with their counterpart such as radiology images, Computed Tomography (CT) etc., leaving lots of improvement scope.

To address this problem, extensive research [1, 2] been performed; resulted in a framework to study 3D whole-slide microscopy image datasets and reconstruct 3D vessel structure to analyze serial liver slices. This proposed standalone framework [1] perform all the sequence of modules of, image registration, segmentation, vessel cross-section association, interpolation, and volumetric rendering. The simulation based research for [1, 2] been developed in Matlab®. To make this cutting-edge novel idea applicable to large dataset acquired from the patients biopsy result in daily basis, in this project under Google Summer of Code 2015, an efficient and scalable version of HPC implementation will be done using C++ programming language leveraging the powerful and rich open source libraries such as openCV, openSlide, VTK etc.

The remainder of this project proposal is organized as follows: Section II describes a short literature review about the histopathological images, whole-slide microscopic image also the basic steps and the development of WSI image analysis so far. Section III presents the proposed approach of the HPC implementation. Conclusions and the timeline of completion are discussed in Section 4. In section 5, a paragraph is included related to my educational and professional experience and the questionnaires needed for evaluation purpose. Section 6 contains the references.

II. Background & Motivation

Before implementing the HPC framework in C++, to have an idea what is the problem it will solve and knowing the theoretical background is of utmost importance. Thus, I perform a brief literature survey to have an in details knowledge of histopathological image data and what are

the basic processes included in the workflow to automate the trustworthy histopathological (WSI) image analysis. The current state-of-the-art development is also included.

II.I. Histopathological Images

Whole slide imaging (WSI) [3] is creation of a single, high magnification digital image of an entire microscopic slide. An automated microscope scans an entire slide at one or more resolutions. Later the consecutive small images are digitally knitted together into a single large image and stored as the virtual slide ready for performing image analysis. WSI images provide a higher resolution cellular level description on a whole organ scale. However, whole slide imaging is a modality providing only 2D images. Therefore, the challenge is to use stacks of serial sections from which to reconstruct the 3D vessel structures.
Types of Whole slide imaging in use are: fluorescence WSI and multispectral WSI. There are two approaches used to scan slides while acquiring data, tiling and line scanning given in Figure 1.

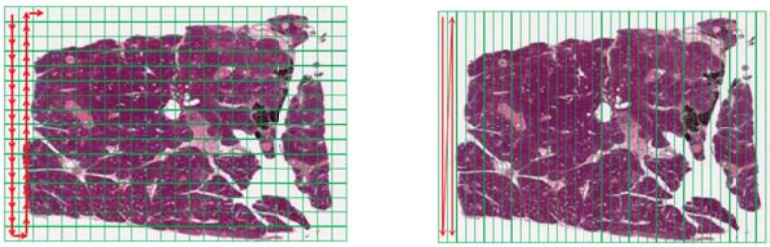

Figure 1: Scanning approaches, tiling (left) & line scanning (right)

In both cases, the resulting images (tiles or strips) are fitted together into a single large image, i.e. the whole slide image. Once all the image data is acquired, the WSI must be stored on disk. These files are relatively large. Due to the large size, WSI data are usually compressed using lossy algorithms.
WSI files stored in an “image pyramid” fashion, shown in Figure 2. Zoom levels are pre calculated and stored in the file. The image at each zoom level is broken into small tiles (e.g. 256 x 256 px). The small tiles are stored in the image file (or directory). Each additional channel or z-planes (3D WSI images) are also present in the pyramid. WSI’s are usually stored both image data and metadata in a dedicated server or “Digital slide repository”(DSR). The viewer client requests the area of the image being displayed (green) at a particular zoom level. The DSR then sends only the tiles needed to fulfill the request.
3D reconstruction provides many important anatomical measurement that are either not available, or cannot be accurately measured in 2D. For example, the projected length of a vessel is shorter in the projected views. Torque and the curvature of a vessel are virtually impossible to estimate from 2D views. 3D reconstruction provides better and cleaner visualization allowing people without extensive training to understand vessel geometry. It saves

reviewing time for doctors since 3D reconstruction may be performed by a trained technician, and may also help visualize dynamics of the vessel that were not be demonstrated by previous methods.

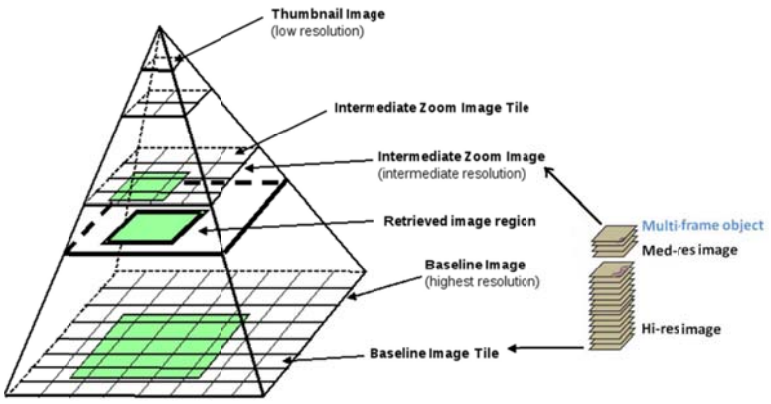

Figure 2: WSI Storage pyramid

II.II. Brief literature review

WSI analysis methods can potentially maximize the amount of information extracted from tissue slides for decision-making and maximize the objectivity and reproducibility of analysis. The importance of quantitative and objective analysis of tissue biopsy WSI has led to several commercial software tools for WSI analysis including GENIE (Aperio, Vista, California, USA), HALO (Indica Labs, Corrales, New Mexico, USA), AQUA Analysis (HistoRx, Branford, Connecticut, USA), and Visiopharm (Hoersholm, Denmark). However, all of these tools provide limited image processing capabilities. In most cases, pathologists manually select the regions of interest (ROI) and make diagnoses based on feedback from these commercial tools. Usually, an expert user calibrates these systems for each laboratory-specific experimental setup. In Figure 3, a flowchart including the basic steps of the histopathological image analysis framework and different available techniques to perform them. As mentioned in different articles [3], none of these tools provides complete data analysis for clinical decision-making that includes all of the steps illustrated in Figure 3.

Researchers in the Virtual Liver reconstruction area, develop methods to digitally reconstruct the organ in 3D based on serial tissue sections. The main challenges for this are deformations of the tissue which are introduced during the cutting process and the placement of the section on the slide. Thus consecutive slides cannot simply be stacked on top of each other to regain the 3D object. In [1, 4], a two-step sequential image-based registration process consists of rigid and non-rigid steps is used to handle the problem. Images are first aligned rigidly (ignoring scale) by

subsampling the virtual slides by a specific factor. This result is then used as input to a non-rigid registration method, that divides the image into a set of regularly spaced individually aligned square patches and perform the registration using cubic B-Spline transformation.

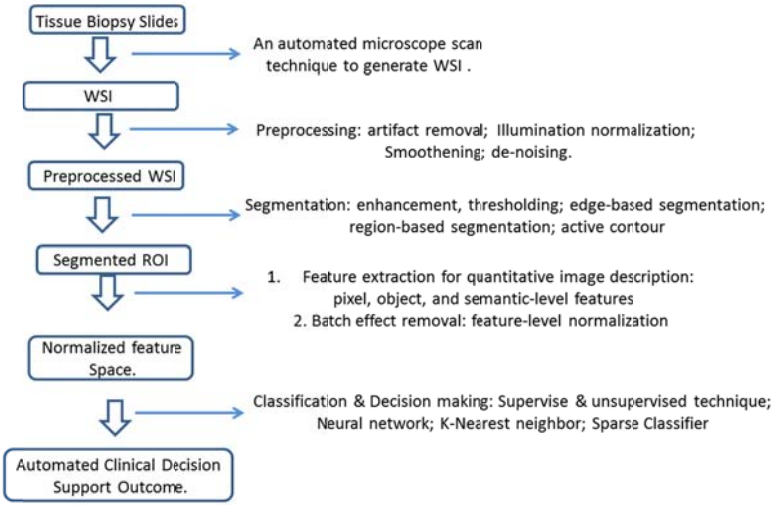

Figure 3: Histopathological Image analysis steps

Followed by the registration reconstruction of 3D Vessel Structures from WSI is proposed in [1]. After registration, the resulting image stack is used for the vessel extraction procedure. By applying color deconvolution technique immunohistochemistry (IHC) stain components are deconvoluted. Later using hysteresis thresholding and morphological reconstruction operation on the deconvoluted color channels the blood vessels are segmented from nuclei, bile ducts and the background. Fourier shape descriptors are used to distinct vessel association in the successive frames. A vessel object-based analysis is proposed in [1] to find the optimal vessel association using entropy-based relaxed integer programming (eRIP). Finally, a B-Spline interpolation performed between associated vessel objects, the fitted curved surface used for 3D volumetric rendering using volumetric tetrahedral mesh generation algorithm. The proposed 3D vessel reconstruction algorithm in [1] is at per according to quantitative and qualitative results compared with laborious human annotations.

III. PROPOSED WORK

C++ is a general purpose language, so one can express any algorithm in it, and it is most definitely better for algorithms that cannot be expressed using arrays, and need to use complex and highly dynamic data structures such as graphs, mesh generators, symbolic manipulation and so on. To implement the standalone 3D vessel reconstruction HPC framework in C++ we

will use different open source libraries such as openCV, openSlide, VTK etc. OpenCV (Open Source Computer Vision Library) includes several hundreds of image processing and computer vision algorithms starting form pixel level to feature level analysis and having multi-threading and re-enterable capability. OpenSlide is a C library that provides a simple interface to read whole-slide images that is the input modality (.svs) of our application. Visualization Toolkit (VTK) is an open-source object-oriented library helpful 3D graphics visualization and rendering application.

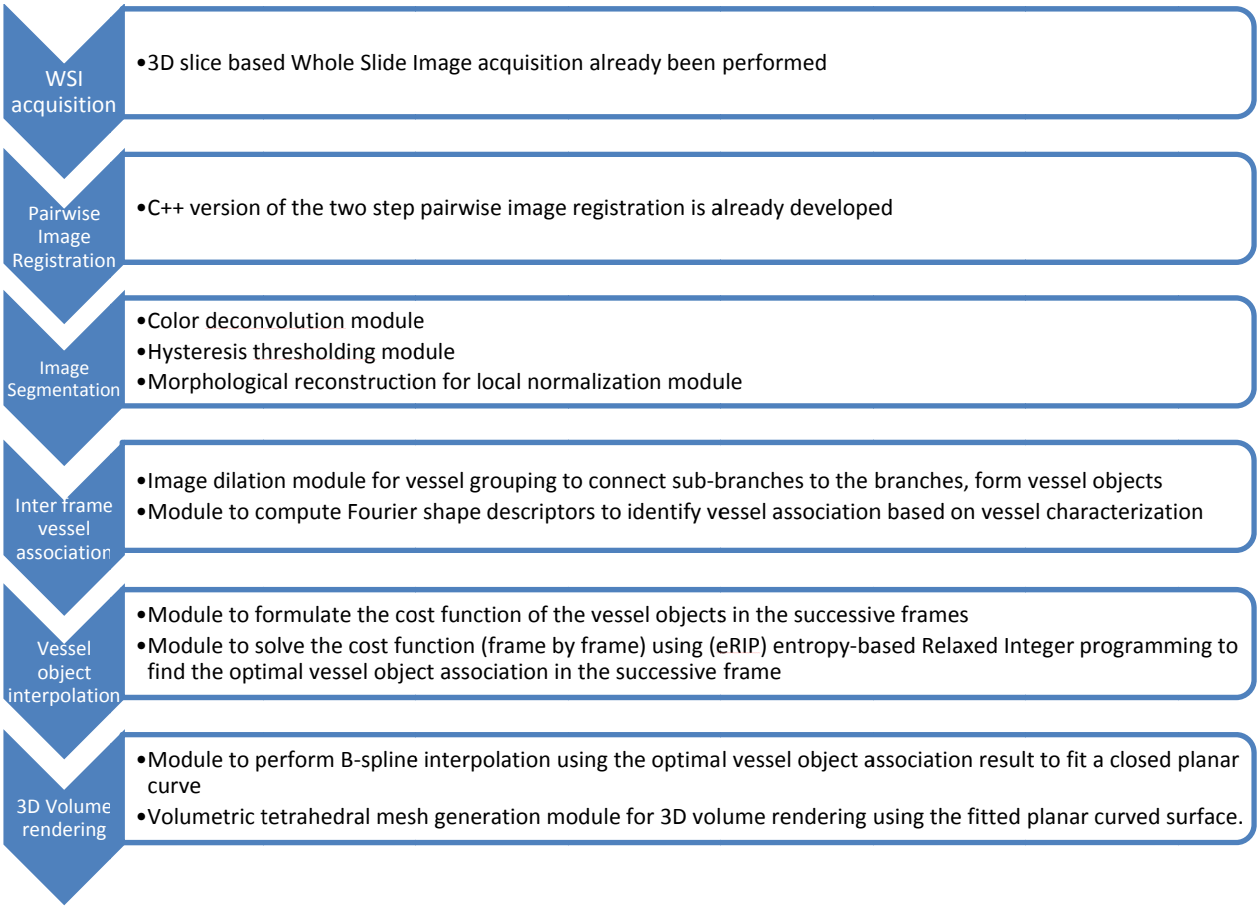

Figure 4: Modular/ Functional approach of building the HPC framework

In Figure 4 the modular approach, i.e. the functions to be implement for a specific task is tabulated. The stand-alone HPC framework will be able to accept registered whole slide images and perform the tasks segmentation, inter frame vessel association, vessel object interpolation and finally 3D volume rendering of the blood vessels.

IV. MILESTONES AND CONCLUSIONS

In this project, we will develop HPC implementation (in C++ language) of the 3D vessel reconstruction proposed in [1]. First with continuous interaction with the mentor/ organization, in-depth review of the application and a vivid description of the system requirement will be prepared. Later following the modular task approach mentioned in Figure 4 the functionalities will be developed. Finally testing will be done in the evaluation and final product demonstration period within the suggested 'pencils down' date mentioned in GSoC 2015 timeline.

TABLE I.
PROPOSED MILESTONES AND THE CORRESPONDING TIMELINE TO COMPLETE THE PROJECT

Research & Development Task / Timeline	3/27-5/25	5/25-6/15	6/15-7/6	7/6-7/20	7/20-8/3	8/3-8/17
T1: In depth review and define system requirement	X					
T2: Image segmentation Modules		X	X			
T3: Inter Frame Vessel Association Modules		X	X	X		
T4: Vessel Object Interpolation Modules				X	X	
T5: 3D Volume rendering Modules					X	X
T6: Evaluation and Final Product Demonstration						X

VI. REFERENCES

[1] Liang, Y, Kong, J, Treanor D, Magee D., Wang, F, et al. “Liver Whole Slide Image Analysis for 3D Vessel Reconstruction”, Proc. International Symposium on Biomedical Imaging , 2015 (Accepted)

[2] Liang, Y, Wang, F, Treanor, D, Magee, D, Teodoro, G, Zhu, Y, Kong, J, “A Framework for 3D Vessel Analysis with Whole Slide Images of Serial Liver Sections”, Proc. of International Conference on Intelligent Biology and Medicine (ICIBM 2014), 2014

[3] Kothari S, Phan JH, Stokes TH, Wang MD, “Pathology imaging informatics for quantitative analysis of whole-slide images”, Journal of the American Medical Informatics Association : JAMIA. 2013;20(6):1099-1108.

[4] Roberts, N., Magee, D., Song, Y., Brabazon, K., Shires, M., Crellin, D., Orsi, N.M., Quirke, R., Quirke, P. and Treanor, D., “Toward routine use of 3D histopathology as a research tool,” Am. J. of Path., 180(5), pp.1835-1842, 2012.