2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)
April 3-7, 2020, Iowa City, Iowa, USA

STITCHING METHODOLOGY FOR WHOLE SLIDE LOW-COST ROBOTIC MICROSCOPE BASED ON A SMARTPHONE

Juan E. Ortuño[1,2], Lin Lin[2,3], María del Pilar Ortega[3], Jaime García-Villena[3], Daniel Cuadrado[3], María Linares[4,5], Andrés Santos[2,1], María J. Ledesma-Carbayo[2,1], Miguel Luengo-Oroz[3]

[1] Consorcio Centro de Investigación Biomédica en Red (CIBER-BBN), Madrid, Spain.
[2] Biomedical Image Technology Lab, Universidad Politécnica de Madrid, Spain.
[3] Spotlab, Madrid, Spain. [4] Hospital Universitario 12 de Octubre, Madrid, Spain.
[5] Biochemistry and Molecular Biology, Pharmacy School, Universidad Complutense de Madrid, Spain.

ABSTRACT

This work is framed within the general objective of helping to reduce the cost of telepathology in developing countries and *rural* areas with no access to automated whole slide imaging (WSI) scanners. We present an automated software pipeline to the problem of mosaicing images acquired with a smartphone, attached to a portable, low-cost, robotic microscopic scanner fabricated using 3D printing technology. To achieve this goal, we propose a robust and automatic workflow, which solves all necessary steps to obtain a stitched image, covering the area of interest, from a set of initial 2D grid of overlapping image*s*, including vignetting correction, lens distortion correction, registration and blending. Optimized solutions, like Voronoi cells and Laplacian blending strategies, are adapted to the low-cost optics and scanner device, and solve imperfections caused using smartphone camera optics. The presented solution can obtain histopathological virtual slides with diagnostic value using a low-cost portable device.

Index Terms— Microscopy, whole slide imaging, stitching, smartphone camera

1. INTRODUCTION

Acquisition of whole slide imaging (WSI) is routinely done in histopathological microscopy using automated scanners to get a large-field image composed of multiple stitched image tiles. High-quality optics and precision mechanics facilitate the process of obtaining a virtual slide.

As an alternative to high cost WSI scanners, stitching solutions have also been proposed to create virtual slides from a set of overlapping images acquired on conventional microscopes with manual or motorized stages [1-3]. Stitching tools also have been extended to 3D volumes in light-sheet microscopy [4] and X-ray microphotography [5].

Low-cost solutions based on smartphones attached to a microscope have been proposed, taking advantage of the increasing processing capacity of mobile devices [6]. Replicability in such systems is a limitation, and the need for a conventional microscope and an attached motorized or manual stage limits the ability of the system to be portable.

Current stitching methods in microscopy are mainly focused on the registration step, which corrects the limited precision of the mechanical scan, preferably using phase correlation algorithms for pairwise registration between adjacent tiles [1,4,5] as well as feature-based matching methods [2]. However, images need to be corrected from vignetting and distortion when these effects are present. Vignetting is the lack of homogeneous intensity due to uneven background illumination, resulting in the usual darkening of the corners of the image [7]. The most common approach to correct the vignetting effect consists of selecting a previously acquired image of the background and use it for intensity normalization [4,5]. On the other hand, the distortion introduced by lens is not considered in reviewed stitching methods used in light microscopy imaging, and this effect is not negligible when a smartphone is used [6]. After registration, the overlapping regions are blended linearly or by an average value [1,5] to get a seamless result. Composition methods based on Voronoi diagrams would reduce potential artifacts [8], but are not used in reviewed works on image stitching in microscopy.

Our proposal to get low-cost WSI automatic scanning replaces the conventional microscope by a low-cost portable device (shown in Fig. 1), fabricated using 3D printing technology, leveraging the camera and computing power of a generic smartphone. A software app installed on the smartphone controls the acquisition process of an image grid and further transmission to the cloud where the virtual slides are obtained. Once the virtual slides are reconstructed in the cloud, experts can analyze them from anywhere using a web interface.

In this paper we describe the entire workflow of preprocessing and stitching the 2D grid of acquired images to obtain a final virtual slide, and present preliminary results of histopathological samples. Our work includes novel methods in stitching microscopy images: automatic vignetting correction without reference or background image, integrated distortion correction, and fusion using Voronoi cells and Laplacian blending.

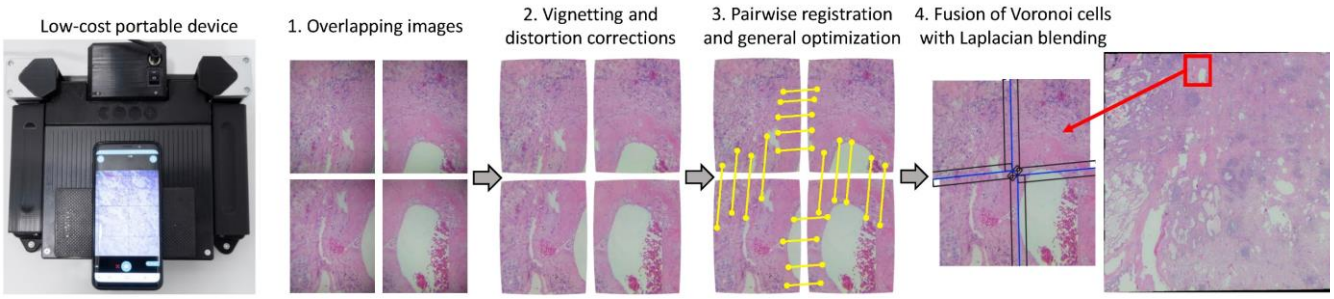


Fig. 1. Left: Top view of the low-cost robotic microscope. Right: Workflow with an array of 2×2 images: 1) acquired overlapping images; (2) images corrected from uneven illumination (vignetting) and distortion; (3) pairwise registration based on matching features (yellow segments) and general optimization; (4) final stitched image using Voronoi cells (blue line segments) and Laplacian blending in regions delimited by back lines. The complete stitched image is shown on the right.

2. METHODS

The workflow to get a virtual slide image from the initial 2D grid of overlapping images consists in the following steps: a) vignetting correction; b) lens distortion correction; c) pairwise registration; d) general registration optimization; e) image fusion or blending. These steps are shown in Fig. 1.

2.1. Vignetting correction

In our device, the smartphone camera is manually attached to the 3D-printed microscopy system using an adapter that allows flexibility so that we do not have to be limited to any model or brand. Therefore, we need a new vignetting calibration at least every time the mobile is attached to the microscope. For greater safety, we perform a different calibration per acquired sample, so it considers variations in illumination due to the placement of the sample.

Our proposal relies on a vignetting correction without reference image to avoid an additional acquisition or the need of background in the sample [7] . We estimate the intensity nonuniformity using a method based of the sparseness of the gradient probability distribution [9]. The images are converted to HSV channels (hue, saturation, value) and the correction is performed over the image with the highest mean value of the V channel. We fit the illumination with a bipolynomial model of 5th degree. We get a robust estimation of the intensity nonuniformity even in the absence of a background image.

However, many histopathological samples contain many background regions surrounding tissue. These regions are automatically detected, using a variance/mean threshold, to create a normalized illumination field which can also be used for vignetting correction. A maximal pixel-wise image is composed to avoid dust or small remaining objects.

2.2. Distortion correction

A polynomial distortion model [10], with radial and tangential distortion is used:

$$\begin{aligned} x_d &= x_u\left(1+k_1 r+k_2 r^2+k_3 r^3\right)+2n_1 x_u y_u+n_2\left(r+2x_u^2\right) \\ y_d &= y_u\left(1+k_1 r+k_2 r^2+k_3 r^3\right)+2n_2 x_u y_u+n_1\left(r+2y_u^2\right) \end{aligned} \quad (1)$$

where the normalized coordinates of the undistorted image (x_u, y_u) are transformed into the distorted coordinates (x_d, y_d) using three radial parameters k and two tangential parameters n, being $r = x_u^2 + y_u^2$ the distance from the principal point. The model is flexible enough to include the distortion induced by the camera optics of a generic smartphone.

The distortion parameters are optimized using a target grid, as shown in Fig. 2, whose coordinates in the distorted image (x_d, y_d) are automatically detected and matched with their ideal positions (x_u, y_u) . The parameters of the model of (1) are optimized using the Levenberg-Marquardt method over the set of matching correspondence pairs. The matching algorithm is totally automatic and robust as it does not require a complete segmentation of the grid, which could be occluded due to dust or blurred in the extremes of the image.

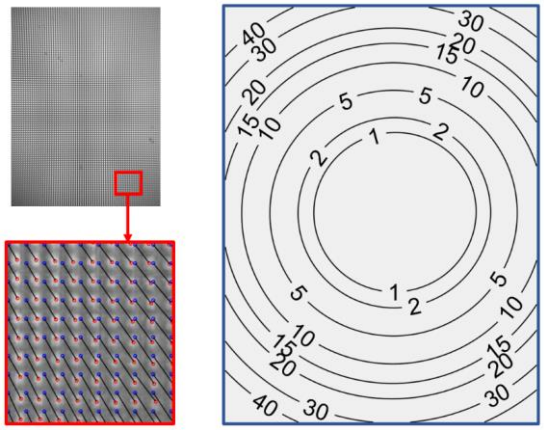
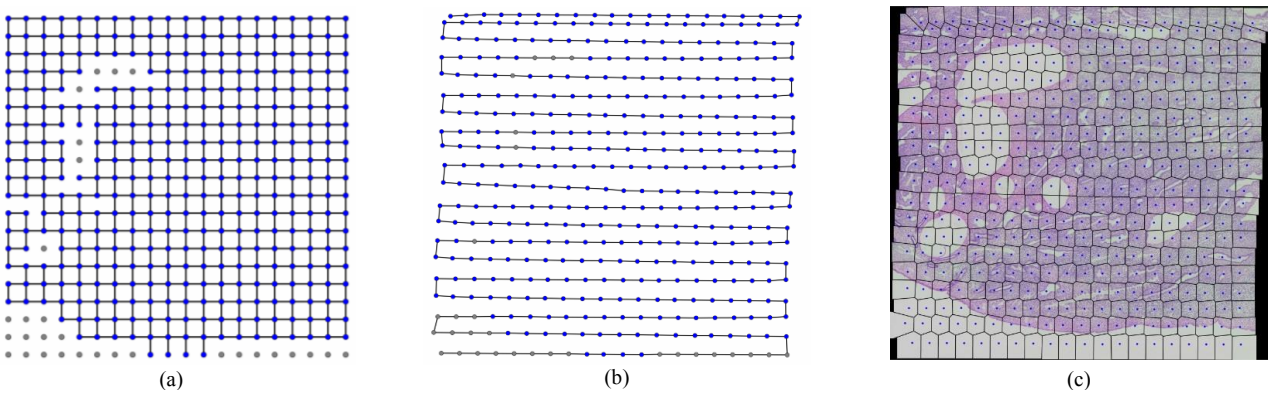


Fig. 2. Left: Distortion correction using a target grid, showing detail with detected centers (in red) matched to ideal coordinates (in blue). Right: Contour plot of the absolute value of distortion in pixels.



(a) (b) (c)

Fig. 3. (a) Array of 20×20 images where segments indicate successful registration between adjacent tiles. Unregistered images are represented ad gray dots; (b) estimated positions showing scanner path. The unregistered images (gray dots) are interpolated from the grid of registered ones (blue dots); c) Voronoi diagram of in superimposed stitched result. Unregistered images correspond to zones with background without significant features.

2.3. Pairwise registration

The optical system scans the sample following a horizontal continuous movement with vertical jumps at the ends of the rows, as illustrated in Fig. 3(b), The number of rows *N* and columns *M* is determined in advance to get a grid of *N*×*M* images, with a predetermined approximated overlap between adjacent tiles. Mechanical distress makes the real overlap not coincide with the theoretical one, therefore a registration between images is necessary to correctly estimate the movement of the smartphone when scanning the sample.

We test four registrations by every image (up, down, left and right adjacent images) except in the extremes of the grid, getting *M*×*(N-1)* horizontal registrations and *N*×*(M-1)* vertical registrations. With enough overlap, the common area guarantees a successful registration, but more images need to be acquired to cover the same total area.

We use a feature-based registration method based on matchings of multiscale AKAZE features [11], estimating a translation transformation with a random sample consensus (RANSAC) algorithm to reject outliers. A minimum number of features must be reached in order to label the registration as valid.

2.4. General optimization

The pairwise registration process produces relative translations in *x* and *y* between pairs of adjacent images *a* and *b*, denoted as $\Delta x_{a,b} = x_a - x_b$, $\Delta y_{a,b} = y_a - y_b$. There are *N(M-1) + M(N-1)* ≈ *2MN* pairs of images registered in a grid of *MN* images.

We can express the relative translations in the *x* dimension in a vector-matrix form $\mathbf{Wx} = \mathbf{\Delta x}$, being $\mathbf{x}$ the column vector of absolute locations, and $\mathbf{\Delta x}$ the column vector of relative translations $\Delta x_{a,b}$. The sparse matrix $\mathbf{W}$ has a size of *MN* columns and ~*2MN* rows, fulfilling the $\Delta x_{a,b} = x_a - x_b$ relations. We compose a similar equation for the y coordinates: $\mathbf{Wy} = \mathbf{\Delta y}$.

In the most favorable case, when all possible adjacent images have been successfully registered, the systems of equations are overdetermined and can be solved by least squares to get the absolute positions **x** and **y**.

When some pairwise registration are not successful, mainly because there are homogeneous regions such as a background in the overlapping area, the relative shifts are not defined, and those pairs will not be considered to solve the least squares problem. If a background zone completely separates two tissue zones, the remaining equation is undetermined and cannot be solved.

In our solution, we check the subsets of images that are interconnected through registration pairs, solving the equations $\mathbf{W}_n\mathbf{x}_n = \mathbf{\Delta x}_n$ where $\mathbf{x}_n$ are the coordinates $\mathbf{x}$ of the subset *n* of connected images. Once the sample's movements are identified inside each subset, we estimate the location of unmatched images by extrapolating the registered coordinates row by row. It should be noted that we do not need great accuracy in this estimation of remaining images, because they are expected to be background areas with no significant features. Once we have an estimation of absolute positions **x** and **y,** the fusion process can be performed to obtain the final stitching.

2.5. Image fusion

To construct the stitched image, we define a Voronoi diagram with Euclidean distance using the centers of images as the seeds of the cells. The Voronoi diagram is the optimal division of the stitched image from its original images if it is assumed that the image degradation, (e.g., out-of-focus and chromatic aberrations) increases with the distance to the center of the image [12]. We get polygonal cells as shown in Fig 3(c).

Image degradation, as well as possible registration imperfections, recommend the use of a blending scheme to smooth the result around the line segments between Voronoi cells. We use Laplacian pyramids [13], a multi-scale transform based method, for image blending. This process is applied to the rectangular regions around line segments between two adjacent Voronoi cells, as shown in Fig. 4.

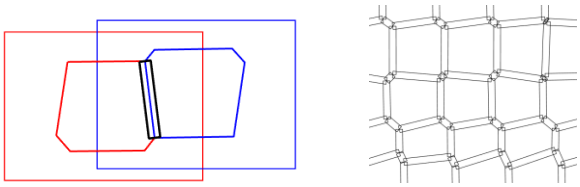


Fig. 4. Laplacian blending is applied to rectangular regions surrounding the segments of the Voronoi diagram. Left: Two image positions and their Voronoi cells with its common blending rectangle. Right: Blending rectangles between image tiles.

The stitching algorithm is deployed using container technology (AWS ECR) and is executed using cloud computing services (AWS Batch jobs) using 2 virtual CPUs, representing a portion of a shared physical CPU (custom Intel Xeon processors with an all core Turbo frequency of 3.6GHz) and 4GB of memory.

3. EXPERIMENTS AND RESULTS

A total of 87 independent slides were acquired and the corresponding stitchings were performed with samples from 9 different pathologies, including breast carcinoma, colorectal carcinoma, endometrial carcinoma, and lung carcinoma, covering different areas depending on the magnification, overlap (ranging from 10% to 40%) and the corresponding variation in the number of images used, i.e., a 20×20 tiles acquisition with magnification 40x and 30% obtains a 7.2×7.2 mm^2 image. Besides, we used different models of mobile phones: BQ X2, OnePlus3T, Samsung S9 and Samsung A40. The stitching was successful for all 85 images acquired with an overlap greater than 20%.

To adjust a robust overlapping value, a histopathological sample was acquired with different overlapping ratios (from a 10% to 40% overlap) with a 10×10 grid. The percentage of failed pairwise registrations is shown in Table 1. The total number of pairwise registrations was 180.

Table 1. Percentage of successful pairwise registrations.

Overlap (%)	10	15	20	25	30	35	40
Failed regs (%)	40.5%	25%	22.2%	1.6%	1.1%	0%	0%

The mean squared error (MSE) of the optimization method gives an estimation of the stitching accuracy. 9 stitchings of lung carcinoma acquired from 20×20 images of 1080×1440 pixels, with 30% overlap were tested. The MSE measured in pixels without lens distortion correction was 14.87 (SD 4.14) and was reduced to 1.43 (SD 0.79).

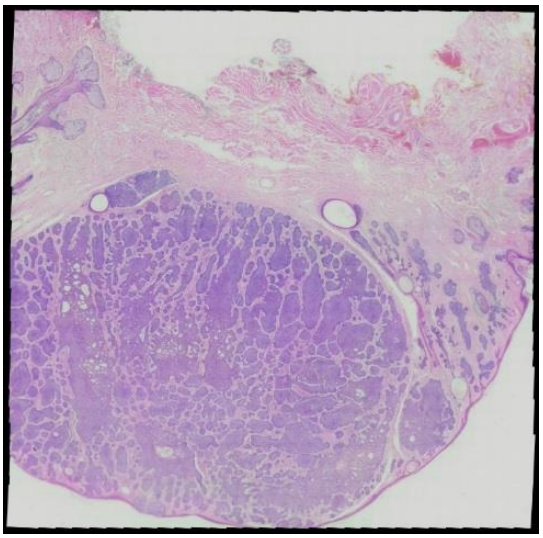


Fig. 5. Endometrial polyp of 6×6 mm^2 stitched from a stack of 30×30 image tiles with 30% overlap at magnification 55x.

4. DISCUSSION

In this work we present a complete workflow for mosaicing histopathological samples, designed to be robust using low-cost optics but completely automatic as is going to be operated by non-trained specialists in remote or rural areas.

Our proposal is aimed to work with a low-cost WSI scanner manufactured using a 3D printer, which leverages the optics and the computing power of a generic smartphone. This device needs a robust method of correction, registration and blending to get a mosaic image. This solution can help to reduce the cost of telepathology in developing countries and rural areas with no access to automated WSI scanners.

Comparing with the existing low-cost solutions, we can stitch the mosaic even though some image tiles associated to homogeneous regions are not registered, extrapolating the approximate missing locations from its adjacent images. Besides, we correct the image distortion, the uneven color, and the seam line, creating a high-resolution virtual slide image. The operational implementation of the proposed solution runs in the cloud within 15 min for a field of view of 7.2x7.2mm^2, creating an image of approximately 15000^2 pixels. The proposed size allows to keep focus and provides a relevant piece of tissue for pathology diagnosis.

The study was approved by the Ethics Committee of Hospital Universitario 12 de Octubre, Madrid, Spain.

6. ACKNOWLEDGMENTS

We thank Dr. José Luis Rodríguez Peralto from Hospital 12 de Octubre for his help with data access. This work has been partially funded by projects TEC2015-66978-R (MINECO/ FEDER, EU), RTI2018-098682-B-I00 (MCIU/AEI/FEDER, EU), CDTI NEOTEC SNEO-20171197,EU SME Instrument Phase 2-881062, and by IND2019/TIC-17167 (Comunidad de Madrid).

6. REFERENCES

[1] J. Chalfoun, M. Majurski, T. Blattner, K. Bhadriraju, W. Keyrouz, P. Bajcsy, and M. Brady, "MIST: Accurate and Scalable Microscopy Image Stitching Tool with Stage Modeling and Error Minimization," *Scientific Reports,* vol. 7, no. 1, pp. 4988, 2017.

[2] A. Gherardi, and A. Bevilacqua, "Manual Stage Acquisition and Interactive Display of Digital Slides in Histopathology," *IEEE Journal of Biomedical and Health Informatics,* vol. 18, no. 4, pp. 1413-1422, 2014.

[3] D. Pradhan, S. E. Monaco, A. V. Parwani, I. Ahmed, J. Duboy, and L. Pantanowitz, "Evaluation of Panoramic Digital Images Using Panoptiq for Frozen Section Diagnosis," *Journal of Pathology Informatics,* vol. 7, no. 26, 2016.

[4] D. Hörl, F. Rojas Rusak, F. Preusser, P. Tillberg, N. Randel, R. K. Chhetri, A. Cardona, P. J. Keller, H. Harz, H. Leonhardt, M. Treier, and S. Preibisch, "BigStitcher: Reconstructing High-Resolution Image Datasets of Cleared and Expanded Samples," *Nature Methods,* vol. 16, no. 9, pp. 870-874, 2019.

[5] A. Miettinen, I. V. Oikonomidis, A. Bonnin, and M. Stampanoni, "NRStitcher: Non-Rigid Stitching of Terapixel-Scale Volumetric Images," *Bioinformatics*, pp. btz423, 2019.

[6] H. Yu, F. Gao, L. Jiang, and S. Ma, "Development of a Whole Slide Imaging System on Smartphones and Evaluation with Frozen Section Samples," *JMIR mHealth and uHealth,* vol. 5, no. 9, pp. e132, 2017.

[7] F. Piccinini, E. Lucarelli, A. Gherardi, and A. Bevilacqua, "Multi-Image Based Method to Correct Vignetting Effect in Light Microscopy Images," *Journal of Microscopy,* vol. 248, no. 1, pp. 6-22, 2012.

[8] J. Pan, M. Wang, D. Ma, Q. Zhou, and J. Li, "Seamline Network Refinement Based on Area Voronoi Diagrams with Overlap," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 52, no. 3, pp. 1658-1666, 2014.

[9] Y. Zheng, M. Grossman, S. P. Awate, and J. C. Gee, "Automatic Correction of Intensity Nonuniformity from Sparseness of Gradient Distribution in Medical Images," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009.* pp. 852-859.

[10] Z. Tang, R. G. von Gioi, P. Monasse, and J.-M. Morel, "A precision Analysis of Camera Distortion Models," *IEEE Transactions on Image Processing,* vol. 26, no. 6, pp. 2694-2704, 2017.

[11] P. F. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 34, no. 7, pp. 1281-1298, 2011.

[12] S. Hsu, H. S. Sawhney, and R. Kumar, "Automated Mosaics Via Topology Inference," *IEEE Computer Graphics and Applications*, no. 2, pp. 44-54, 2002.

[13] P. Burt, and E. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communications,* vol. 31, no. 4, pp. 532-540, 1983.