Automated area calculation of histopathologic features using SIVQ

Jason Hipp[a], Jerome Cheng[a], Stephanie Daignault[a], Jefferey Sica[a], Michael C. Dugan[d,e], David Lucas[a], Yukako Yagi[c], Stephen Hewitt[b] and Ulysses J. Balis[a,*]

[a]*University of Michigan Health System, Department of Pathology, Medical Science I, Ann Arbor, MI, USA*
[b]*National Institutes of Health, National Cancer Institute, Laboratory of Pathology, Advanced Technology Center, Gaithersburg, MD, USA*
[c]*MGH Pathology Imaging & Communication, Technology (PICT) Center, Boston, MA, USA*
[d]*Esoterix Genetic Laboratories, LLC, Pathology Services, Los Angeles, CA, USA*
[e]*Roche molecular systems, Inc., Pleasanton, CA*

Abstract. Recently, with the advent of the 7th edition of the AJCC Cancer Staging manual, at least one set of criteria (e.g. breast) were modified to now require the measurement of maximal depth of stromal invasion. With the current manual interpretive morphological approaches typically employed by surgical pathologists to assess tumor extent, the specialty now potentially has stumbled upon a crossroads of practice, where the diagnostic criteria have exceeded the capabilities of our commonly available tools. While whole slide imaging (WSI) technology holds the potential to offer many improvements in clinical workflow over conventional slide microscopy including unambiguous utility for facilitating quantitative diagnostic tasks with one important example being the determination of both linear dimension and surface area. However, the availability of histology data in digital form is of little utility if time-consuming and cumbersome manual workflow steps are necessarily imposed upon the pathologist in order to generate such measurements, especially as encountered with the complex and ill-defined shapes inherent to infiltrative tumors. In this communication, we demonstrate the utility of the recently described SIVQ algorithm to serve as the basis of a highly accurate, precise and semi-automated tool for direct surface area measurement of tumor infiltration from WSI data sets. By anticipating the current trend in cancer staging that emphasizes increasingly precise feature characterization, as witnessed by the recent publication of AJCC's 7th edition of the Cancer Staging Manual, this tool holds promise to will be of value to pathologists for clinical utility.

Keywords: SIVQ, area calculation, WSI, digital slide

1. Introduction

The increasingly prevalent ability to create digital whole slide imaging (WSI) data sets (digital slides) from glass slides offers new opportunities for the practice of pathology such as improving workflow, facilitating consultation and accessing archived material. WSI also enables the use of image analysis algorithms that can add further value by providing additional quality assurance and consistency in pre-screening and/or post-screening confirmatory roles.

Measurement of histopathologic features, such as maximum tumor diameter (pathologic Tumor size, pT), is necessary for accurate staging of cancers such as breast and non small cell lung cancer (NSCLC) (2, 3). Furthermore, the new edition (7th) of the American Joint Committee on Cancer (AJCC) Cancer

*Corresponding author: Ulysses J. Balis, MD, Department of Pathology, University of Michigan Health System, M4233A Medical Science I, 1301 Catherine, Ann Arbor, MI 48109-0602, USA. Tel.: +1 734 615 5727; Fax: +1 603 250 3139; E-mail: ulysses@ umich.edu.

Staging Manual for breast cancer has updated criteria stipulating that the maximal depth of tumor invasion into the stroma should be recorded to the nearest millimeter [3].

Working with digital slides offers the ability to annotate and precisely measure distance with "digital rulers." Such tools are very helpful, for example, in measuring: 1) the distance between tumor and margin and 2) maximal tumor dimension. In addition, programs such as Aperio Technology's ImageScope[TM] viewer have intrinsic annotation tools with pre-defined shapes, as well as freehand drawing tools, both of which enable simple types of area determination. However, with these tools alone, area calculation can be time consuming and technically demanding when challenged with the need for measuring ill-defined shapes, such as infiltrative tumors. This reality is further compounded where there is a need for a trained pathologist to perform such activity.

Here we present an evolved area calculation tool, using a high-throughput pattern recognition image analysis algorithm, Spatially Invariant Vector Quantization (SIVQ), which addresses the limitations described above. Briefly, SIVQ is a pattern recognition algorithm that can identify morphologic and architectural features on H&E and special stained digital slides [4, 6]. Initially, one (or a small number of) distinct predicate image candidate feature(s) is identified by the user. This feature in turn is utilized to perform an exhaustive search of the entire surface area of the digital slide, resulting in the generation of a statistical probability heatmap of quality of matches. We have previously described a number of potential clinical and research applications enabled by SIVQ, with examples being identification of: 1) specific cell types such as Reed-Sternberg cells in Hodgkin lymphoma, 2) myeloid precursors in bone marrow smears, 3) breast micro-calcifications, 4) micro-organisms such as fungus, 5) different textures of stroma and finally, 6) specific nuclear morphologies (e.g. nuclear grooves) [4, 6]. Introductory video demonstrations on typical uses of the SIVQ algorithm are available online [8].

In this report, we will first demonstrate the spatial validation for the SIVQ-based area calculation tool to two significant digits, and subsequently, apply the technique to germane clinical use cases. In so doing, we hope to show how such application allows for the imparting of a rigorous quantitative component to surgical pathology, while additionally realizing "added value" with the use of digital slides. To be clear,

the determination of precise, morphologically defines areas, such a tumor cross section, represents a current and growing unmet clinical need. This reality is made all the more urgent by the recent AJCC 7th edition publication, which requires accuracy to a degree that pathologists alone are no longer able to easily generate, owing to time pressures and unavoidable inter-observer variability. Use of SIVQ, as described in this communication, addresses this need with an immediately available tool that is suitable for pathologists in a busy clinical setting.

2. Methods

2.1. Whole slide images

Colon, breast, and gastric tissue and calibration tools were scanned using an Aperio XT[TM] slide scanner at $20\times$ and/or $40\times$ magnification. The pixel resolutions of the resultant WSI data sets are 0.5 um/pixel ($20\times$) and 0.25 um/pixel ($40\times$), with all primary digital data, including the full WSI files, being available at the University of Michigan Department of Pathology WSI Repository (www.wsirepository.org).

2.2. Calibration tools

3.05 mm diameter, 200 HEXAGONAL mesh copper grids (6228Cu) were purchased from Tousimis Corp. (Rockville, MD) and scanned into digital format. Also, a stage mapping calibration grid, precision glass stage micrometer (IAM-5-S-N-OP) was purchased from Applied Image Inc. (Rochester, NY) and scanned into a digital format.

2.3. Statistical analysis

A TOST ("two one-sided tests") test is a type of equivalence test. An equivalence test describes if two items differ within some predetermined bound. TOST equivalence tests were used to compare the $20\times$ and $40\times$ area measurements of the glass stage micrometer grid to the actual specified areas as provided by the company. We choose a bound to maintain 2 significant digits. The TOST test is chosen because it is a simple test that is commonly used for this question when the data is normally distributed.

2.4. SIVQ

The use of SIVQ image analysis has been previously described elsewhere by Hipp & Cheng et al. [4, 6] Areas to be measured within the digital images were identified and "painted" in an automated fashion, using the SIVQ pattern matching algorithm. For this effort, an additional module was developed added to the SIVQ discovery package, allowing for the quantification of total painted areas identified within the resultant matches. Briefly, this was accomplished by an iterative surface scanning algorithm that exhaustively counted the total number of pixels in the painted regions, and subsequently multiplied this value by a corrective constant (0.0635 for 40× scans and 0.251 for 20× scans), yielding final resultant values in units of square microns.[1]

The minimum systems requirements for effective execution of this software application include any systems with at least Pentium 4/2.9 GHz processor and 2 Gb of ram, running upon Windows XP or later operating system. To validate the ease with which users can quickly operate the tool in independent fashion, the SIVQ suite has been distributed to informatics groups at Rutgers, NIH/NCI Laboratory of Pathology, Harvard/MGH, U.C. Davis, and Cornell, with all groups reporting back success in demonstrating reproducibility. To facilitate facile use of the application, collaborating groups were offered a 20 minute tutorial on the theory and operation of the software, an explanation of the user interface, overview of the vector selection process and a concluding demonstration of several examples. Upon completion of these training sessions, collaborating investigators were able to identify morphologic features and perform their own independent analyses. As a final example of the ease with which the tool can be mastered, pathology residents at the University of Michigan, with no prior image analysis experience, were able to effectively use SIVQ and the additional area calculation feature, after a modest 30 minute introductory session.

---

[1] The conversion factors were derived from taking the square of the length equivalent of each pixel. The length of each pixel on the physical slide was identified in the microns per pixel property of the Aperio Viewport ActiveX Control (0.252 microns for slides scanned at 20× and 0.501 for slides scanned at 40×). The square of these values gave the area of each square pixel at each of the respective magnifications. The area covered by each square pixel was multiplied by the total number of positive pixels, resulting in the total area "painted" from the SIVQ analysis.

3. Results

3.1. Calibration of the SIVQ-based area calculation feature

Since there is no established reference methodology for validating the accuracy and precision of *in silico* measurement tools intended for use on digital slides, the following series of studies were performed. Initially, the commonly used "ruler" intrinsic to Aperio's ImageScope™ viewer was validated against a paper micrometer for determination of linear accuracy. To calibrate the spatial properties of the digital slides against an external reproducible standard (a feature absent with the Aperio approach), we obtained digital scans at 20× and 40× of two "reference constructs," a copper EM grid and a glass stage micrometer. These micro-structured objects were selected as a result of their exhibiting reproducible and measurable features at the micron length-scale, with the added property of such features exhibiting tight dimensional tolerances.

3.1.1. Paper micrometer

We began by assessing the ruler function and the measurement grid within Aperio's ImageScope™. A paper micrometer was scanned into a digital slide. Figure 1A depicts a representative low power field from the paper micrometer. Figure 1B depicts a representative high power field of view of the same micrometer, with the dark blue lines in the inferior portion representing 20 micron intervals and the juxtaposing data rendered by ImageScope™ viewer's measurement grid in the superior portion. Both the micrometer and the measurement grids visually corresponded.

3.1.2. Electron microscopy copper grid

Electron microscopy (EM) copper grids have predefined features that can be measured at the micron level. An EM copper grid was scanned into a digital slide (Fig. 2). Figure 2A shows a low power view of the EM copper grid. A medium power view of the "mesh" is shown in Panel (Fig. 2B) and the hexagon was measured with the ImageScope™ ruler to have a length of 50 microns and a height of 85 microns (a calculated area to be 6375 square microns) which is confirmed by the company's manual. We then created a vector to recognize the white space inside the "mesh" (Fig. 2C), which resulted in the "painted" area in Fig. 2D. The
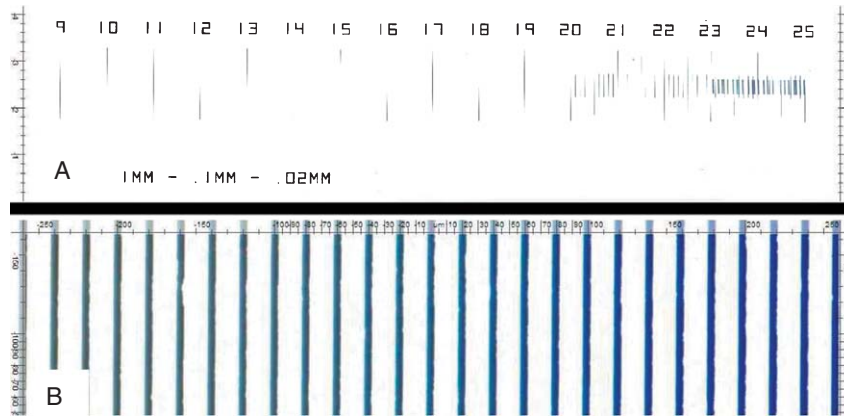
Fig. 1. A WSI (whole slide image) of a micrometer. To assess ImageScope's[TM] measurement "ruler", a micrometer was scanned using the Aperio XT. Panel A shows a low power view, with the measurement dimension key in the bottom left. Panel B is a high power image of the small hash marks on the right side of the ruler. The blue lines are spaced out every 20 microns. On the top of the image, is ImageScope's[TM] ruler, which perfectly corresponds to the measurements on the micrometer.
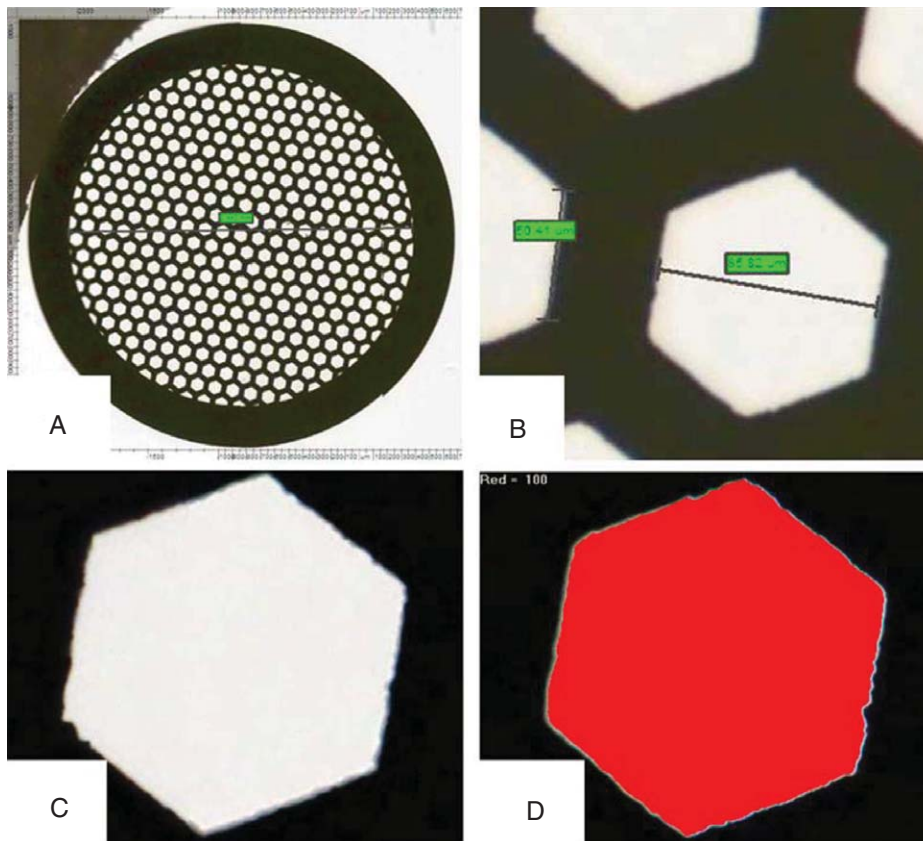

Fig. 2. A WSI of an electron microscopy copper grid. Panel A is low power view of the grid. Panel B is a medium power view of the "mesh" was measured with ImageScope[TM]. The dimensions were Length = 50 microns and height 85 microns. Panel C is a high power view of a single mesh had the pre-calculated area of 6375 um^2. Panel D is the white space of C was "painted" with SIVQ. The area of paint was calculated to be 6375 um^2 and was very accurate.
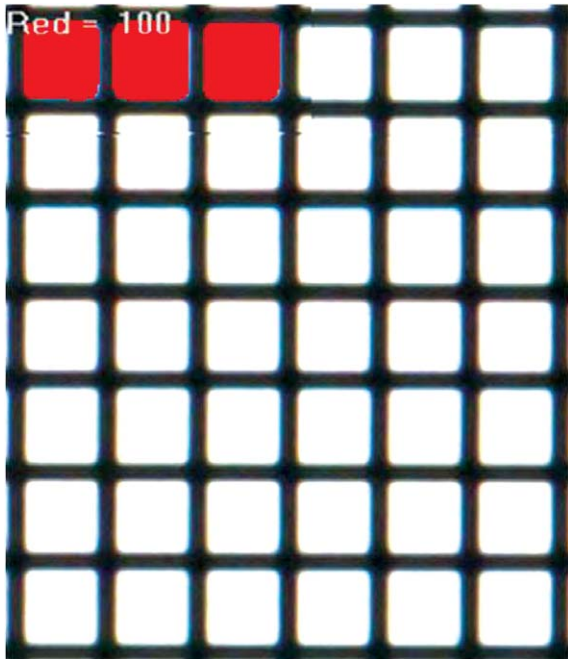
Fig. 3. Glass stage micrometer grid. A glass stage micrometer grid was scanned into a digital slide at 20× and 40×. Ring vectors were created to "paint" and then calculate the area of the white space of the squares. While minor variations in the quality of the "painting" in the box was noted, a visual example of the accuracy of the "painting" at 40× is shown here.

area of the "paint" was calculated by SIVQ to be 6375 square microns. While irregularities in the EM grid's shape were visually observed, a tool that was more consistent in shape was required to determine the precision of our tool. While the EM grid confirmed the accuracy of SIVQ-based area calculation feature, it also elucidated the need for an additional validation process to characterize overall algorithm precision.

3.1.3. Glass stage micrometer grid

A glass stage micrometer grid was purchased and scanned into a digital slide at 20× and 40×. The grid contained numerous boxes with each edge measuring 0.0085 mm. Multiple boxes were analyzed at a time due to the constraints of the viewport window, 6 boxes were analyzed at 20× and 3 boxes at 40× (Fig. 3).

As while there were minor variations in the quality of the "paint" in the box due to potential variations in the grid and for scanning bias, we therefore analyzed the sets of squares in a saw like pattern to obtain a statistical mean. For example, at 40×, 3 squares were analyzed and then the reference frame moved to the

right 1 square and down one square. This analysis was performed 8 times. Then analysis shifted to a northeastern step function of one stop down and 3 squares over (to avoid repeating the analysis on the same squares) then up one. At 20× 6 squares were analyzed, moving up one and to the right one, 7 more times. This was performed 8 times resulting in 32 squares being analyzed. A similar methodology was used at 20×.

At 20×, the 95% Confidence Limit of the Mean was 0.00041 mm^2 and 0.00042 mm^2. The true area was 0.00043 mm^2 which is outside the confidence interval. At 40×, the 95% Confidence Limit is 0.00021 mm^2 and 0.00022 mm^2 with a mean of 0.00021 mm^2. Using TOST equivalence test, the 40× is equivalent to 2 significant digits.

3.2. Clinical use cases

3.2.1. Tumor measurements

In current clinical practice, tumors are measured and often staged by their greatest maximal dimension. This is typically performed by a pathologist marking the edge of the tumor with a dotting pen, and then measuring with a ruler. It usually results in measurements in millimeters. The ocular micrometer employed for Breslow depth measurement results in measurements reported in hundredths of millimeters. The accuracy and precision of such measurements are user dependant and can vary.

3.2.1.1. Well defined tumor nodules. An H&E stained tissue section of a post-imatinib treated gastrointestinal stromal tumor was scanned into a digital slide. The low power view (1×) is shown in Fig. 4A. Using Aperio's ImageScope^{TM}, we calculated the greatest diameters of a viable nodule to be 4.05 mm and 3.84 mm.

With SIVQ, the entire tumor nodule was painted with two vectors (Fig. 4B). Using the area calculation tool, the painted area was calculated to be 11.22 mm^2.

3.2.1.2. Infiltrative cancers. Measuring ill defined infiltrative tumors is very challenging and nearly unattainable with current practices. An H&E stained tissue section of a moderately differentiated colonic adenocarcinoma was scanned into a WSI and is shown in Fig. 5A. A vector was selected to identify and paint the malignant glands Fig. 5B. Using the area calculation tool, the painted area was calculated to be 0.39 mm^2.
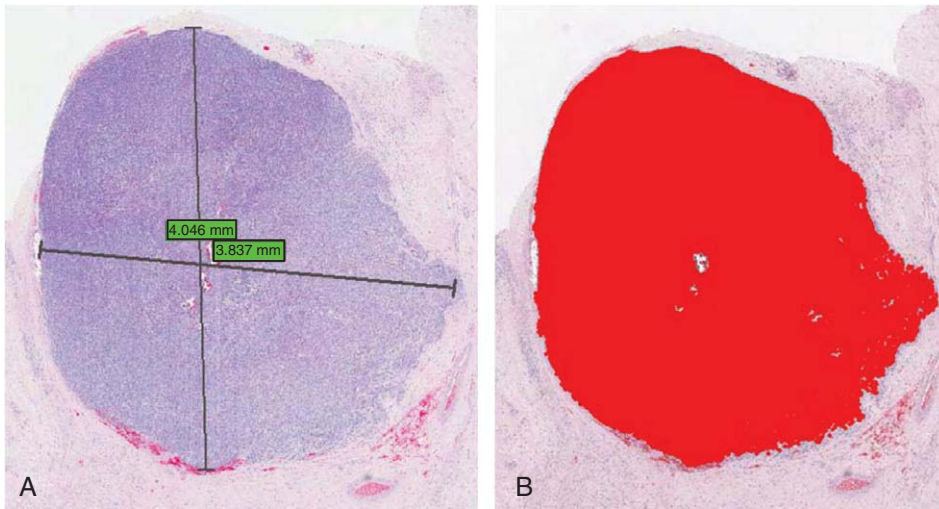
Fig. 4. Gastrointestinal stromal tumor. An H&E stained tissue section of a post-imatinib treated gastrointestinal stroma tumor was scanned into a digital slide at 40×. Using ImageScope's measurement tool, the two greatest maximal dimensions of a viable nodule were calculated to be 4.046 mm and 3.84 mm (Panel A). Two ring vectors were selected to "paint" the entire tumor nodule. The area of the "paint" was calculated to be 11.22 mm^2.
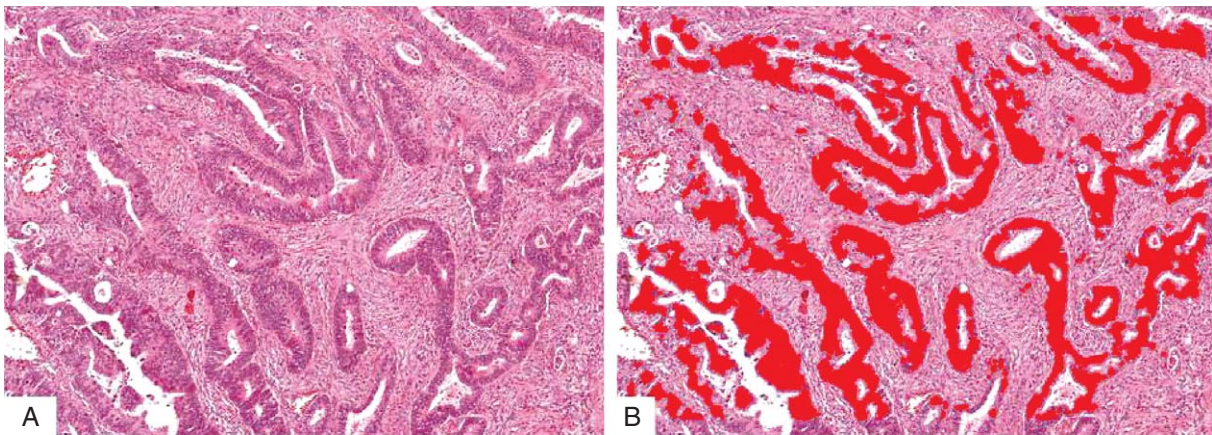

Fig. 5. Example of area estimation of colonic adenocarcinoma. Panel A is a 4× magnification of an H&E tissue section. A ring vector was selected to "paint" the malignant glands and the area was calculated to be 0.39 mm^2.

3.3. Tissue surface area classification and measurement

3.3.1. Breast biopsy

The ability to classify and measure the cell types present in biopsies and resections has potential value for translational and clinical studies. A WSI of a H&E breast core biopsy containing breast epithelium and stroma is shown in Fig. 6A. A vector was selected to recognize and "paint" the breast epithelium Fig. 6B. Using the SIVQ calculation tool, the paint was calculated to be 5.91 mm^2. A vector was then created to select and paint the breast stroma Fig. 6C. The Using the SIVQ calculation tool, the paint was calculated to be 31.84 mm^2. With the above measurements, the percentage of breast epithelium on these 3 core biopsies was calculated to be 15.65% and the percentage of stroma was 84.35%.
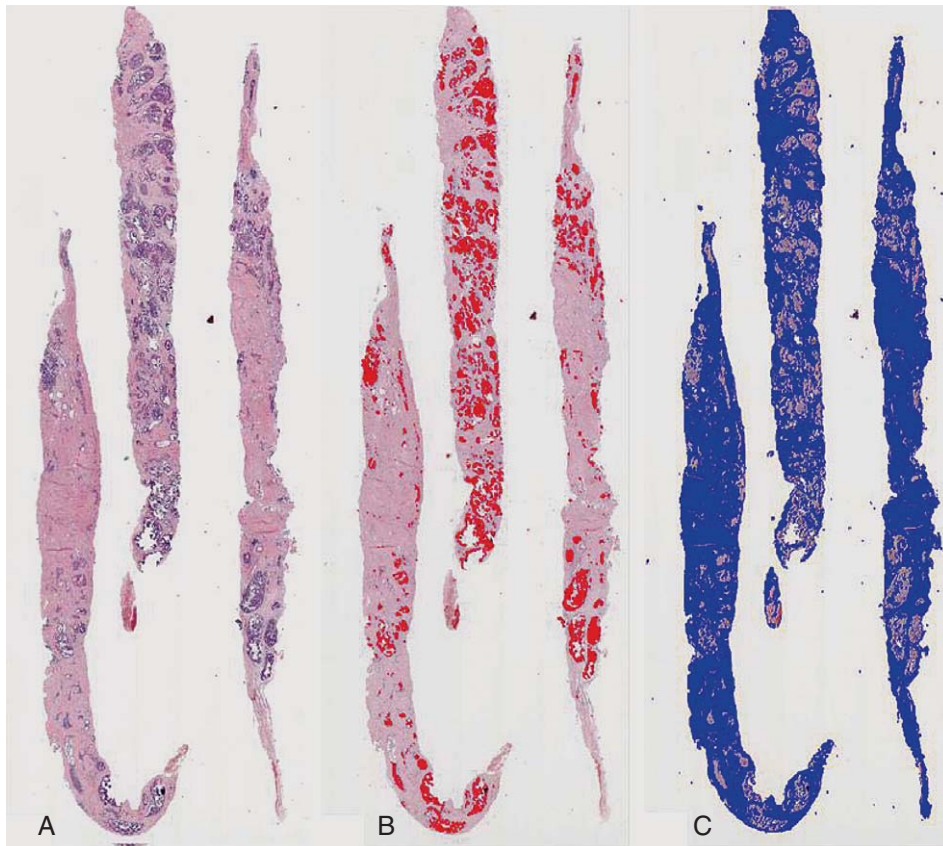
Fig. 6. Breast core biopsy. A low power view ($1\times$ magnification) of a breast core biopsy H&E tissue section, Panel A. A vector was created to recognize and "paint" the breast epithelium. The area of "paint" was calculated to be $5.91\,\text{mm}^2$. A vector was created to recognize and "paint" the breast stroma (Panel C). The area of the "paint" was calculated to be $31.84\,\text{mm}^2$. The percentage of epithelium in the analyzed tissue was 15.65%, and the percentage of stroma was 84.35%.

3.4. Immunohistochemical (IHC) staining measurements

3.4.1. Breast biopsy with ER staining

A digital slide of a ER positive breast core biopsy is shown in Fig. 7A. Representative ring vectors were selected to paint the lighter staining nuclei (Panel B) and the darker staining nuclei (Panel C). The dark staining regions had an area of $0.0026\,\text{mm}^2$ and the lighter staining had an area of $0.0014\,\text{mm}^2$.

4. Discussion

Here we describe a new tool for highly accurate, precise and semi-automated surface area measurement (possessing at least two significant figures) from morphologic and textural features from digital slides stained with either H&E or IHC. By anticipating the current trend in cancer staging that emphasizes increasingly precise feature characterization, as witnessed by the recent publication of AJCC's 7th edition of the Cancer Staging Manual, we believe this tool will be of utility have value to pathologists.

While calculating the surface area of the invasive component is not presently an unmet clinical need, the updated AJCC guidelines highlight the need for precise tumor measurement (linear measurement) for staging. Although digital slide technology makes it easier and more precise to measure maximal depth of invasion, a single linear depth measurement for invasion is a statistically underpowered metric, when compared to topologically rich quantitative approaches such as made available by measuring surface area
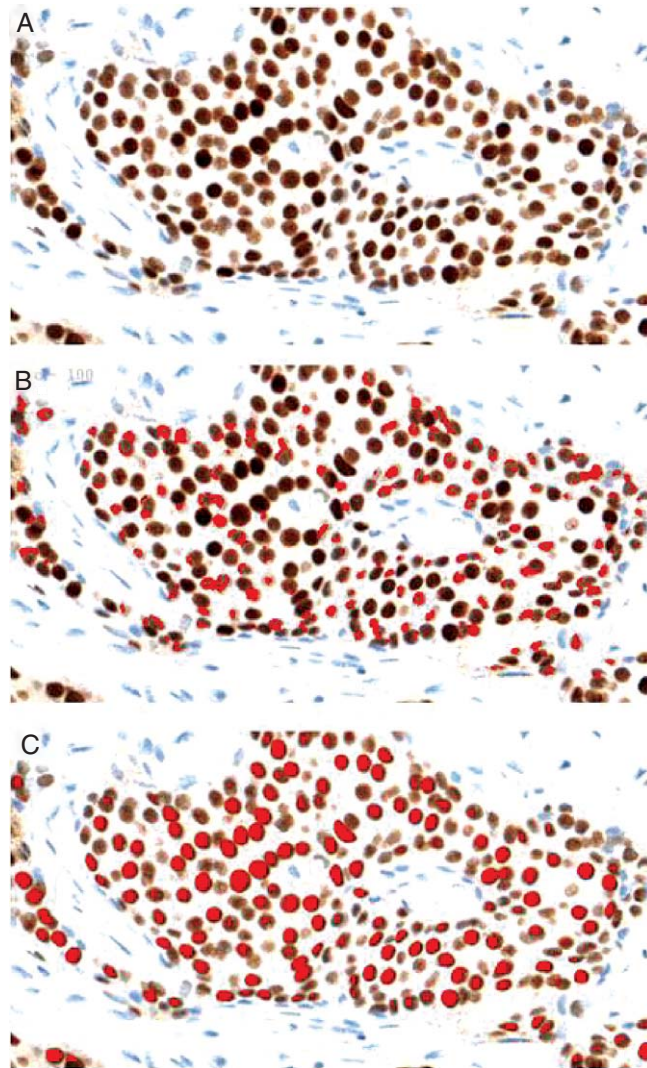
Fig. 7. Breast biopsy with ER staining. A breast core biopsy was stained with ER and scanned into a WSI (Panel A). Ring vectors were selected from lighter stained cells (Panel B) and darker stained cells (C). The dark staining had an area of $0.00255\,mm^2$ and the lighter staining had an area of $0.00137\,mm^2$.

of tumor juxtaposing the invasive margin or measuring the perimeter length of the infiltrative interface margin itself – both constructs that are only approachable via the use of sophisticated numerical analysis of digital pathology data. Consequently, we believe that the collective future unmet clinical needs include the matching of higher-order tools, such as those that determine surface area, with the increased precision called for in current and upcoming AJCC editions.

The use of the current "gold standard", the ocular micrometer, is highly subjective (with typically fewer than even a single significant digit of resultant precision). It also permits for an inexact process, as witnessed by the measurement of only the greatest maximal dimension of a tumor, with this approach greatly subsampling the total actual amount of available information. The net outcome of this process yields loss of much potentially valuable data. Thus, we feel that the use of the ocular micrometer ill fit to the challenge of fully characterizing ill-defined infiltrative lesions.

As is the practice in applying the directives of the contemporary current cancer staging manual, such

poorly defined and infiltrative lesions are graded as being either present or absent, with this binary assertion entirely missing the opportunity to fully categorize the actual spectrum of microscopic tumor presence. The net outcome of this imprecision lies in the potential loss of accuracy, for both staging and consistency of measurements in longitudinal studies. Consequently, the new tool described within this report is intended to address this need, with it adding to the repertoire of "tools for the pathologist."

The goal of using the calibration grid/precision glass stage micrometer grid is to confirm the precision and accuracy in the SIVQ-based area calculation tool. There are precise measurements for the area of the spaces in the grid that are provided. Knowing the precise area, we aimed to show that the SIVQ-based area calculation tool could be accurate and precise to the same number of significant digits. Taking repeated measurements, we compared the measurements to the actual area. The TOST equivalence test confirmed shows that the SIVQ-based measurements were, with 95% confidence, equivalent to two significant digits at $40\times$ to the actual measurements of the precision grid. Thus, we can conclude that the SIVQ-based area calculation tool is accurate and precise in its measurement of area to two significant digits at $40\times$ and to one significant digit at $20\times$.

4.1. "Value added"

Extracting precise area calculations, with low interobserver variability, from glass slides alone can be challenging, if not impossible. Conversely, being able to do this with digital slides represents significant "added value," especially when considering that such effort can be carried out in automated fashion, while maintaining sufficiently rigorous sensitivity and specificity. Tumor nodules are currently measured and graded by their maximal diameter. However, such lesions are not always symmetrical or well-circumscribed. Consequently, total tumor surface area may be a better metric of the lesion's extent, as made possible by use of SIVQ. For example, if a tumor nodule had a greatest diameter of 5 mm, the other dimension could measure 1 mm or 5 mm (depending on the shape), with this variation contributing to resulting areas between of $5\,\text{mm}^2$ vs. $25\,\text{mm}^2$. Compared with the newly recommended levels of measurement precision as stipulated in the AJCC 7th edition Can-

cer Staging Manual (which mandates precise linear measurement of the depth of invasion (to the nearest millimeter). The 500% variance in the above example of glass slide measurement would appear to be an unacceptably large interval. Therefore, application of current grading systems could be missing potential important diagnostic and prognostic information if improperly implemented. Use of SIVQ-based area calculation holds the promise to elevate the current level of clinical practice to the newly encountered degree of accuracy now required to effectively carry out microscopic tumor staging. In addition, use of this tool should allow for much greater precision in assessing post-treatment effect in cancers, a process which currently exhibits a fair degree of subjective bias.

With appropriate numerical modeling, it is a straightforward process to approximate total tumor volume from the successive area measurements of tumor within sequential tissue step sections (when constraining section thickness and number of sections taken between each sampled section to uniform incremental values).

Accurately measuring the depth of tumor invasion with varied stromal borders by hand is nearly impossible on glass slides. Even using the current tools (e.g. digital ruler) available on digital slides is very time consuming and would require a domain expert (such as a pathologist) to manual annotate and circle the specific cells of interest. To classify and calculate the area of all the cells present in a biopsy or tissue resection is also nearly impossible to perform on glass slides alone, and could require many hours per slide in order that a for a pathologist could achieve a high level of accuracy.

4.2. Calibrations of the SIVQ-based area calculation tool

Since there was no standardized way to calibrate the results of a WSI area calculation tool, we began by comparing primary image data to a presently available and widely used measurement tool (ImageScope™). We sought to define the accuracy and precision of its resultant measurements. For our first attempt, we scanned and measured the area of EM copper grids, which have predefined dimensions. However, significant variability in the shape of such grids was observed. A calibration grid/precision glass stage micrometer grid with precise measurements (containing a known number of significant figures) was then considered.

This enabled us to calculate the precision and accuracy at both the $20\times$ (1 significant digit) and $40\times$ magnification (2 significant digits). Confirming the accuracy of SIVQ-based measurement in relation to EM copper grids and micrometer grids allowed us to move on to more challenging applications.

4.3. Advantages and applications

The SIVQ-based area calculation method offers a reliable way to accurately and precisely measure such infiltrative instances of malignancy, while at the same time avoiding the time consuming, technical hurdles inherent in extant contemporary manual methods or current conventional digital imaging technologies. Furthermore, because SIVQ allows for semi-automated workflow (after a suitable vector is selected), area calculation results are available immediately, with this being a critical criterion for potential clinical utility.

4.4. Quality assurance

One of the greatest advantages of using such a tool is in quality assurance. Rather than providing its utility absolute measurements (such as a cell count), the methodology described above can provide a relative percentage of the cells of interest in the context of the amount of cells/tissue examined by a comparison to the total amount of tissue analyzed. Thus, this approach could potentially reduce inter-observer variability in reading selected fields of view.

4.5. Future applications

In addition to the use cases described above, there are many other potential applications. Area of a tumor may be used to assess specimen adequacy for subsequent evaluation such as a molecular test requiring a certain amount of DNA or it may be used to determine percentage of a tumor that is necrotic following induction chemotherapy. SIVQ could provide an automated approach to such questions. This tool also has utility for assisting with the work flow of tissue banks or biorepositories. When depositing biospecimens, the corresponding WSI could be automatically screened to quantify the amount of cancer present in each tissue block and this could be recorded as meta-data supporting decisions of how best to make use of this tissue.

For clinical trials, SIVQ could be used to measure particular cellular morphologies of interest, before and after treatment, on large number of slides/cases in cohort studies (a number of patients at the same treatment group) or in longitudinal studies (a single patient throughout the treatment protocol). Lastly, the ability to generate vectors to recognize particular IHC intensities (demonstrated here and in Hipp et al. [4]) and to robustly measure their expression intensity per surface area could usher in novel scoring algorithms that are more precise and accurate for common tests such as estrogen receptor (ER), PR, HER2, and Ki67.

4.6. Public availability of WSI

As described by Hipp et al. [5], the public availability of WSI provides many opportunities for the biomedical community such as transparency, validation, and simplified process vehicles for building upon publically available work. While there are no current repositories to host such WSI data sets, we hope to spark the adoption process of such tools and have therefore taken the initiative of making all the WSI used in this manuscript also available on our institution's permalink website repository (www.wsirepository.org). Additionally, this resource implements what we feel to be best practices for both user interface design and long-term data architecture curation, with these features now being offered to the public domain, as a reference for continued open-source development and validation [7]. Furthermore, we also plan to forward these WSI datasets to centrally-sponsored WSI repositories if and when they are established.

The aforemention website enables investigators to analyze and upload the results of their own image analysis area calculation algorithms, enabling performance quality and metrics to be available to other investigators. This will allow for comparison and contrast of algorithms, similarly to what Irizarry *et al.* did in Benchmark with Affymetrix RNA microarray data [1, 9, 10], thus, allowing investigators to identify the optimal algorithm for their study.

Lastly, we have enabled the website to allow for other investigators to create accounts and deposit their WSI content, thus creating a single, centralized database for pathology imaging WSI datasets, as has been described in detail in our recent editorials [5, 7].

In conclusion, we believe that automated, high-throughput tools, such as morphologic feature-based area measurement, as enabled by the use of SIVQ, will "add value" to the use of digital slides in surgical pathology. Similarly, due to the approach's high throughput nature, which allows for more than simply acquiring a single maximal dimension, an approach can be envisioned whereby true tumor dimensional extent could be assessed on multiple levels/sections. This quantitative approach could be further extended to tumor area of all associated digital slides of any given case, thus yielding a far more complete understanding of total tumor burden. While it is beyond the scope of this manuscript to include an outcomes-based side by side comparison of ocular measurement of maximal tumor measurement vs. maximal tumor measurement using a digital ruler vs. SIVQ area based measurement, it is not lost upon the authors that this is a logical next step. And lastly, none of the authors claim any conflict of interest present with the development of this tool, and moreover, uniformly agree with the value of extending it, in the form of a freely-available executable software application, to the general world-wide pathology community. It can be easily accessed from the Whole Slide Imaging Repository Website, [with the specific resource located at: http://www.wsirepository.org/SIVQ/Area_Calculation].

Acknowledgments

We would like to thank Ron Craig for his assistance. Disclosures: None

References

[1] L.M. Cope, R.A. Irizarry, H.A. Jaffee et al., A benchmark for Affymetrix GeneChip expression measures, *Bioinformatics* **20** (2004), 323–331.

[2] F.C. Detterbeck, D.J. Boffa and L.T. Tanoue, The new lung cancer staging system, *Chest* **136** (2009), 260–271.

[3] S.B. Edge and C.C. Compton, The American. Joint Committee on Cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM, *Ann Surg Oncol* **17**, 1471–1474.

[4] J. Hipp, J. Cheng, J.C. Hanson et al., SIVQ-aided laser capture microdissection: A tool for high-throughput expression profiling, *Journal of Pathology Informatics* **2** (2011), 19.

[5] J. Hipp, D. Lucas, M.R. Emmert-Buck et al., Digital slide repositories for publications lessons learned from the microarray community, *American Journal of Surgical Pathology*, In press.

[6] J.D. Hipp, J.Y. Cheng, M. Toner et al., Spatially Invariant Vector Quantization: A pattern matching algorithm for multiple classes of image subject matter- including pathology, *J Pathol Inform* **2** (2010), 13.

[7] J.D. Hipp, J. Sica, B. McKenna et al., The need for the pathology community to sponsor a WSI repository with technical guidance from the pathology informatics community, *Journal of Pathology Informatics*, In Press.

[8] Hipp) UHSUBaJ. Algorithm will help search for cancer cells, 2011. Available at:<http://www.youtube.com/watch?v=O5IJXqkb9SQ&feature=relmfu>.

[9] R.A. Irizarry, B.M. Bolstad, F. Collin et al., Summaries of Affymetrix GeneChip probe level data, *Nucleic Acids Res* **31** (2003), e15.

[10] R.A. Irizarry, Z. Wu and H.A. Jaffee, Comparison of Affymetrix GeneChip expression measures, *Bioinformatics* **22** (2006), 789–794.