

# Evaluating whole-slide imaging viewers used in digital pathology

Wei-Chung Cheng, Samuel Lam, Qi Gong, and Paul Lemaillet, Food and Drug Administration, Silver Spring, Maryland, USA

## Abstract

*As a Class II medical device, the whole-slide imaging (WSI) system is emerging to supplement light microscopes used by pathologists in the past decades by digitalizing histological tissue slides into millions of pixels saved in a WSI file. Unlike the standard image file formats such as JPEG or TIFF, a WSI file usually consists of hundreds of compressed images of different magnification levels and focal planes that need to be decompressed, stitched, scaled, and color-managed to reproduce the view demanded by the user with zooming and panning operations. Currently, most WSI files are stored in proprietary file formats, due to the lack of adopting a standard WSI file format, which hinders the development of third-party WSI viewers by making it difficult to interpret WSI files faithfully. To examine the fidelity of third-party WSI viewers, in this study, three freely available viewers, Sedeen, QuPath, and ASAP, were compared with the factory viewer, NDP, at the pixel level. A software tool was developed to register and calculate the 1976 CIE color difference for each pixel between two viewers. The average color differences were found as 1.30, 18.69, and 18.79  $\Delta E$  for Sedeen, QuPath, and ASAP, respectively.*

## Introduction

The whole-slide imaging (WSI) system is the digital paradigm emerging to supplement the conventional light microscope used by pathologists to read histological tissue slides in the past decades. A WSI system usually consists of three components – scanner, viewer, and display. The WSI scanner digitalizes a tissue slide and stores hundreds of tiled images of various magnification levels in a single file, frequently encoded in a proprietary file format. The WSI viewer opens the WSI file, which could be transmitted from a remote server, and reconstructs the tiled images as a unified view on the designated display (Figure 1) [1-3]. Although these WSI devices are costly and accessible by medical professionals only, third-party WSI viewers are freely available to open the WSI files generated by these devices. However, it is unclear whether the third-party viewers reproduce the identical image as the factory viewer does.

The goal of this study was to determine whether different WSI viewers generate identical images for the same input file. To be precise, we needed a bench test method to quantitatively report the CIE color difference  $\Delta E$  for each pixel between two viewers. To obtain pixel-wise differences, the image pair needs to be properly registered, which is complicated by the various user interfaces of zooming and panning provided by different WSI viewers. To minimize human-induced errors, we needed to develop a software tool to interact with the graphical user interface within each individual viewer.

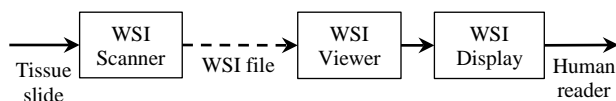


Figure 1: The three components of a WSI system – scanner, viewer, and display. The WSI file is the interface between the scanner and the viewer.

## Materials and Methods

### Target WSI

In this study, we considered the scenario that a tissue slide was scanned by a popular WSI scanner followed by the resultant WSI file shared by the end users, who used different WSI viewers to review the same WSI file.

The Hamamatsu NanoZoomer was chosen as the reference system, which included the reference scanner, viewer, and display [4].

Instead of using a scanner to scan a tissue slide, we used a publicly available WSI file produced by the reference scanner [5]. According to the embedded metadata, the test WSI file “CMU-1.ndpi” was scanned by a Hamamatsu NanoZoomer scanner with a 20x objective on December 31, 2009. The whole-slide image had 51,200\*38,144 pixels at a resolution of 455 nm/pixel covering an area of 23.4\*17.4 mm<sup>2</sup>. The file size was 188 MB after JPEG compressed from 5.45 GB. In this study, the pixels within the rectangle area enclosed by (3994, 29346) and (5912, 30424) in NDP.view2 at 20.1x magnification were used as the ROI.

The reference WSI viewer was NDP.view2, Free Edition, version 2.7.39, dated 2/19/2019, which was freely available from the manufacturer’s website. Figure 2 shows screenshots of the reference viewer displaying the test WSI file.

In the reference WSI system, only the scanner and viewer components were considered in this study. The display component was excluded because the digital-to-optical conversion, the major function of a display, was not part of the evaluation.



Figure 2: Overview of the test slide (top) and a screenshot of the reference viewer.



Figure 3: Screenshot of Sedeen.



Figure 4: Screenshot of ASAP.

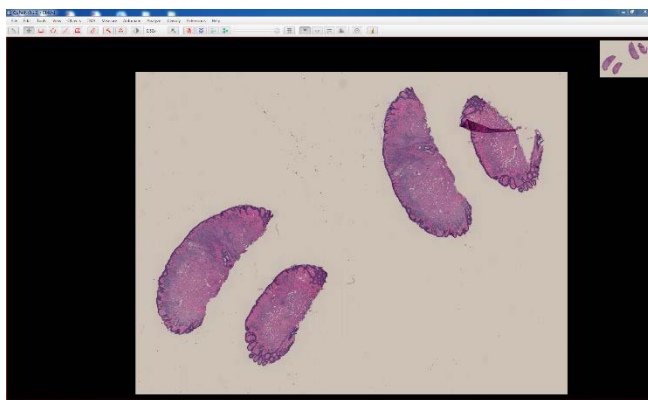


Figure 5: Screenshot of QuPath.

### WSI Viewers under Test

Three freely available third-party viewers were selected in this study. Sedeen (version 5.4.3, build 2019) was freely available from Pathcore, Toronto, Canada as a “universal image viewer designed for digital pathology and/or virtual slide.” It supports various WSI

file formats including Hamamatsu NDPI, NDPIS, VMS, and VMU [6]. Figure 3 shows a screenshot of Sedeen.

ASAP – Automated Slide Analysis Platform (version 1.9.0) was a free and open-source viewer developed by Geert Litjens at the Radboud University Medical Center as an “automated slide analysis platform for fast viewing, annotation and analysis of whole slide histopathology images” [7]. Figure 4 shows a screenshot of ASAP.

QuPath (version 0.1.2, build 2016) was a free and open-source viewer developed for “quantitative pathology and bioimage analysis” at Queen’s University Belfast and University of Edinburgh [8]. Figure 5 shows a screenshot of QuPath.

All four viewers were running in the Microsoft Windows 7 Professional 64-bit environment to open the same WSI file CMU-1.ndpi. The output of each WSI viewer was measured by capturing the pixels shown on the screen by using the Microsoft Snipping Tool, which retrieves the RGB pixel data.

The challenge of comparing the output data from different WSI viewer was the magnification (or zooming) level, which was controlled differently in the four viewers. In the reference viewer NDP.view2, the magnification level was controlled with the computer mouse wheel without any numerical feedback to indicate the magnification level or the pixel coordinate. In Sedeen, the available zoom levels were 0.1x, 0.2x, 0.3x, 0.6x, 1.3x, 2.5x, 5x, 10x, and 20.1x. The pixel coordinate was also provided. In ASAP, the pixel coordinate was provided but there was no numerical feedback of the magnification level. In QuPath, the cursor location was shown in physical distance rather than pixel coordinate. The magnification level could be flexibly entered as a percentage such as 50% or 123.45%. However, the 100% magnification in QuPath is different from the 20.1x magnification in Sedeen. In summary, it was not straight-forward to define the same ROI with the same pixel count between any two of the four different viewers. Defining the same ROI with the same pixel count is imperative because mismatched ROIs would lead to overestimated differences between viewers.

The process of defining the same ROI in different viewers was similar to image registration, which transforms one image to match the other’s coordinates by using different transformation functions. In this study, we assume that the mismatched ROIs between the four viewers can be rectified by two linear transformations – scaling and translation, which are equivalent to the zooming and panning operations, respectively. In other words, to minimize the discrepancies introduced by mismatched ROIs, our strategy was to use the user interface as much as possible to match the ROIs by zooming and panning. The remaining mismatch was then corrected by image registration with a two-dimensional affine transformation.

In order to precisely reproduce the experiments with consistent quantitative results, instead of manually controlling the graphical user interface (GUI), a program-driven approach was used in this study. Our method consists of iterations of two steps as shown in Figure 6. In the first step, a keyboard/mouse event generator,

AutoHotKey, was used to create scripts to interact with the GUIs of the four WSI viewers. The interaction included finding the region of interest, performing zooming and panning operations to match the ROI, and capturing the pixel data in the view.

In the second step, a Matlab program was written to calculate the best registration between the two ROIs and return the optimal two-

dimensional affine transformation. The affine transformation included the scaling and translation transformations, which were used as feedback to guide the zooming and panning operations, respectively, in the first step.

The registration quality was measured by the average correlation coefficient over all pixels within the ROI and used as the stop condition for the loop. The final views were then captured to calculate their discrepancies.

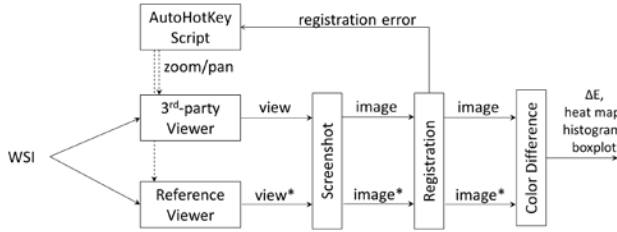


Figure 6: Dataflow of the test method. The same WSI file was opened by the third-party viewer and the reference viewer. Scripts written in AutoHotKey were used to inject user operations such as zooming and panning. The pixel data captured in screenshots were analyzed by MATLAB programs to confirm registration accuracy. Finally, the per-pixel color differences were reported as a heatmap.

## Evaluating discrepancies

The difference between each pixel pair was evaluated with the 1976 CIE color difference  $\Delta E$ . Each captured pixel had the red, green, and blue (RGB) components. First, by assuming that the RGB data were based on the sRGB color space with a white point of CIE D65, the RGB data were converted into the CIELAB color space by using the Matlab `rgb2lab` function. Next, the 1976 CIE color difference was obtained by calculating the Euclidian distance between the two colors in the CIELAB color space.

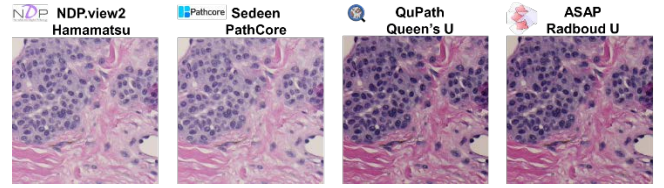


Figure 7: Images generated by four viewers for the same WSI file. It is obvious that the right two images are darker than the left two. Without a bench test method, it is hard to tell whether the left two (or the right two) images are identical.

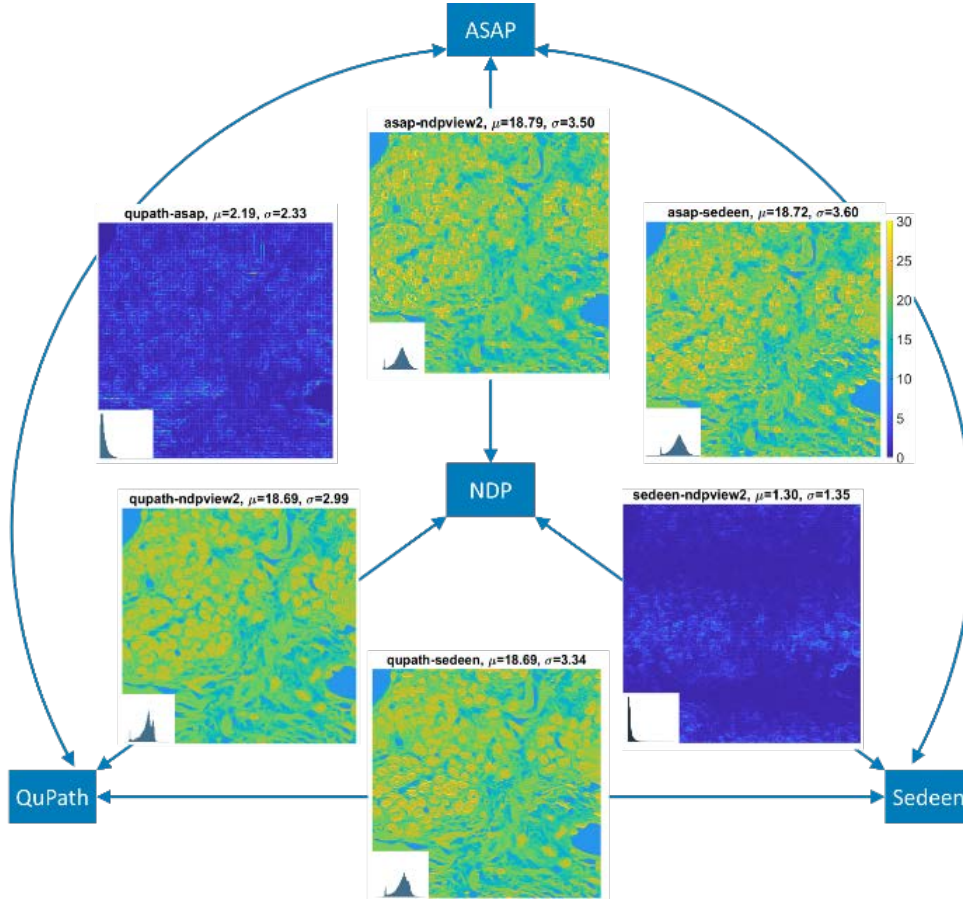


Figure 8: The color differences between the six image pairs as heatmaps. The histogram of color differences is inset in each heatmap.

## Results

The registered images generated by the four viewers are shown in Figure 7. Although the differences between them are visually obvious – two images (QuPath and ASAP) are darker than the others (NDP and Sedeen), quantitative data were demanded in this study. The six pairs of four images were compared at the pixel level as shown in Figure 8. The per-pixel color differences are shown as heatmaps with mean and standard deviation in the titles. The corresponding boxplot is shown in Figure 9.

Considering the mean  $\Delta E$ , none of the three third-party viewers generated an image identical to the factory one. The closest one was generated by Sedeen, which still carried a mean  $\Delta E$  of 1.30. ASAP and QuPath generated 18.79 and 18.69  $\Delta E$ , respectively. They both generated darker images, which turned out to be similar to each other (2.33  $\Delta E$ ).

On the heatmaps, we found that ASAP and QuPath generated  $\Delta E$  values correlated to the tissue structures. For example, nuclei tended to have higher  $\Delta E$  (yellowish in the heatmaps) and stroma lower (bluish). In other words, the color differences seemed to be a function of the color correlates. However, such a trend was not observed in Sedeen because the  $\Delta E$  depended not only on the color but also on the location. After further investigation, we found that the ROI included portions from two adjacent tiles and Sedeen stitched them differently than the factory viewer. After adjusting the registration method to match only the lower part of the ROI, the stitched seam can be easily seen as shown in Figure 10.

Comparing ASAP and QuPath, we observed that the color differences formed a micro-block pattern. It seemed to be related to JPEG compression and consistent with the minor pixelization exhibited in ASAP.

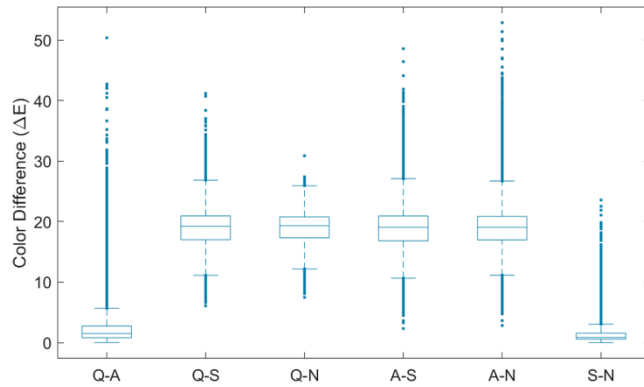


Figure 9: Boxplot of the color differences between the six image pairs. N: NDP. A: ASAP. Q: QuPath. S: Sedeen.

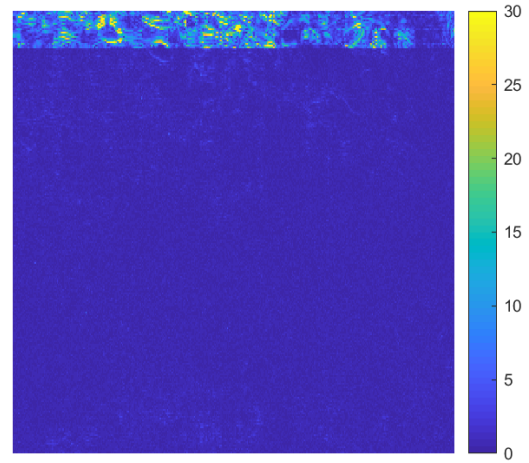


Figure 10: Color differences caused by stitching errors. The original image needs to be reconstructed by stitching two adjacent tiles. Sedeen stitched the two tiles differently compared with the reference viewer. The figure shows that when the image is registered to match the lower tile only, then the residual color differences concentrate in the upper tile only.

## Limitations

The evaluation method in this study is based on colorimetry to quantify discrepancies between WSI viewers and does not incorporate aspects of clinical significance or limitations of the human reader. Only the 1976 CIE color difference was used in the comparison, although the other metrics such as CIEDE2000 can also be used.

## Conclusions

It is commonly perceived that different image viewers should reproduce the same input image file identically. Consequently, the verification of image reproducibility is frequently overlooked. The infidelity of image reproduction may or may not be tolerable for medical imaging since the image viewers are regulated as medical devices. A bench test method was developed to investigate three third-party WSI viewers. The experiment results showed that one viewer reproduced an image very similar to the reference although the stitching method was different. The other two viewers generated images very similar to each other but visually apparently different from the reference. The findings suggest that WSI viewers should not be evaluated by visual comparison only but also bench test methods at the pixel level to determine equivalence between viewers or to validate implementation correctness.

## Acknowledgments

This study was supported by the Critical Path Initiative. The authors thank Drs. Weijie Chen and Aldo Badano for their technical comments and Dr. Jonathan Phillips for his feedback.



## Disclosure

The mention of commercial products herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

## References

- [1] P. Shrestha, and B. Hulsken, "Color accuracy and reproducibility in whole slide imaging scanners," *J Med Imaging* 1(2), 027501 (2014).
- [2] E. L. Clarke, C. Revie, D. Brettell, M. Shires, P. Jackson, R. Cochrane, R. Wilson, C. Mello-Thoms, and D. Treanor, "Development of a novel tissue-mimicking color calibration slide for digital microscopy," *Color Res Appl* 43(2), 184-197 (2018).
- [3] W. C. Cheng, F. Saleheen, and A. Badano, "Assessing color performance of whole-slide imaging scanners for digital pathology," *Color Res Appl* 44(3), 322-334 (2019).
- [4] Hamamatsu, "NDP", <<https://www.hamamatsu.com/us/en/product/type/U12388-01/index.html>> (accessed 10 February 2020).
- [5] OpenSlide project website (<http://openslide.cs.cmu.edu/download/openslide-testdata/Hamamatsu/CMU-1.ndpi>)
- [6] Pathcore, "Sedeen", <<https://pathcore.com/sedeen/>> (accessed 10 February 2020).
- [7] G. Litjens, "Automated Slide Analysis Platform (ASAP)", 2015, <<https://github.com/GeertLitjens/ASAP>> (accessed 10 February 2020).
- [8] Bankhead, P. et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep*. <https://doi.org/10.1038/s41598-017-17204-5> (2017).