**Third-party whole-slide image viewers do not produce the same image**

**Samuel Lam1, Qi Gong2, Wei-Chung Cheng2**

1*Department of Computer Science, University of Maryland, College Park;* **2***FDA/CDRH/OSEL/DIDSR*

**Description of Purpose**

Do different WSI viewers generate identical images for the same WSI file? Three freely available third-party whole-slide image viewers were compared with the factory viewer on the pixel level. Experiment results show that some viewers generated excessive errors.

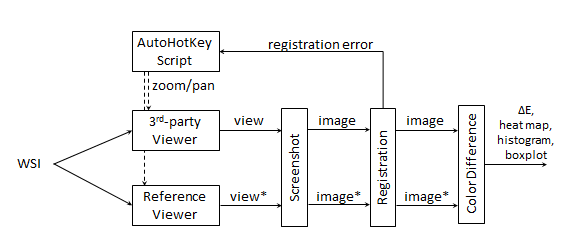
**Abstract**

Whole slide imaging (WSI) involves scanning microscope slides and converting them into high-resolution digital files. In this way, image viewing is more flexible and image transfer is much more efficient. A WSI system consists of the scanner, viewer, and display components. A scanner scans slides and converts them into a standard file format, which can be opened by viewer software and its data converted to color values, which are then sent to the display for pathologists to examine. Currently, only two whole-system WSI devices have been cleared by the FDA. Recently, third-party vendors submitted independent WSI viewer software as alternatives to replace factory viewer components in preapproved devices, claiming to provide more functionality, better integration with the other medical databases, or more economical options. The basis of these submissions is that the third-party viewer software should be interoperable with the cleared devices – since the output of a scanner is a standard file format and the input to the display is color values, as long as a viewer can convert the data in the file to color values, it should be able to replace any other viewer, since they are not fundamentally different. However, the results of this study show that this is clearly not the case. Comparing the color data between viewers on the pixel level shows significant color differences between certain viewers.

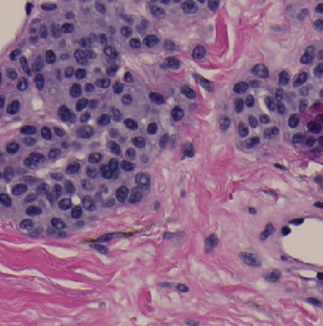
**Methodology**

To compare viewers, the following procedure was devised. We chose a WSI file generated by a Hamamatsu scanner as the target. Hamamatsu scanners are not cleared by the FDA for primary diagnosis, but are very popular for research and have been widely used all over the world for many years. This same file was opened in four separate viewers, including the factory Hamamatsu viewer, NDP.view 2, as well as Sedeen from Pathcore, and two open-source viewers, QuPath from Queen’s University Belfast, and ASAP, from Radboud University. After adjusting the field of view in each viewer to show the region of interest, screenshots of each viewer were taken using Microsoft Snipping Tool. Snipping Tool captures the digital color data sent to the display, without involving the display, which was verified by a previous study [1]. The screenshots were registered, then scanned pixel by pixel to calculate color difference, as defined by the International Commission on Illumination (CIE). The metric, ΔE, is based on human perception of color. Very roughly speaking, human vision can discern the difference of about 2-4 ΔE, and 100 ΔE is the difference between white and black.

**Data flow chart**

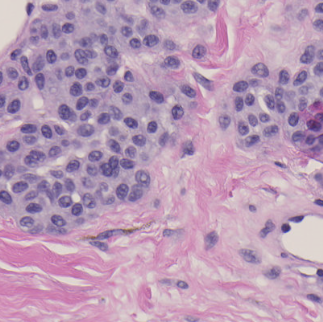


The following registered samples were compared in the study:



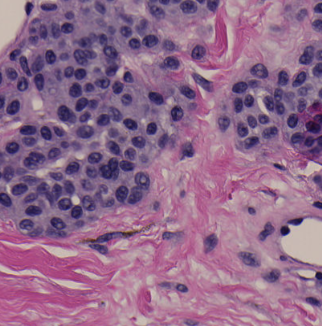
ASAP

Radboud U



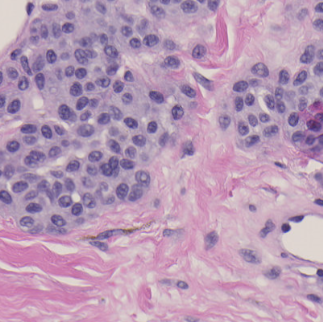
NDP.view2

Hamamatsu



QuPath

Queen’s U



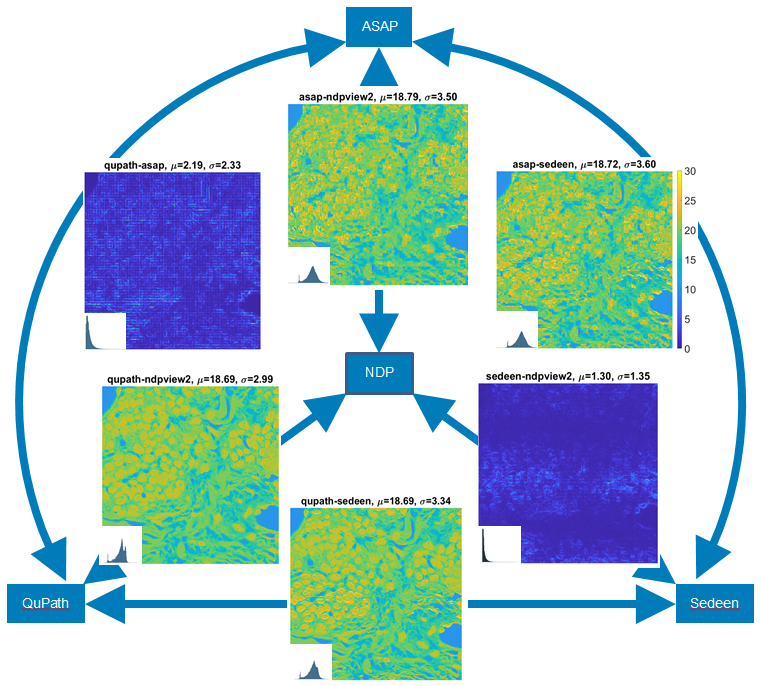
Sedeen

PathCore



**Registered Screenshots**

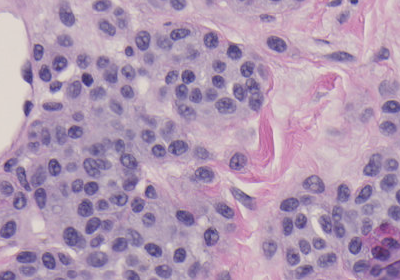
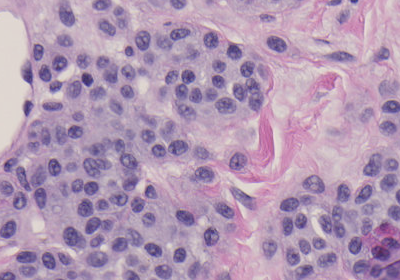
Visually, we can already tell that one looks similar to the reference while the other two look much darker. However, visual inspection is subjective and inadequate for making regulatory decisions. Therefore, we propose a quantitative approach in this study. ΔE information was presented in the form of heat maps and histograms. Average ΔE and standard deviation were also associated with each comparison. A box plot was also generated.



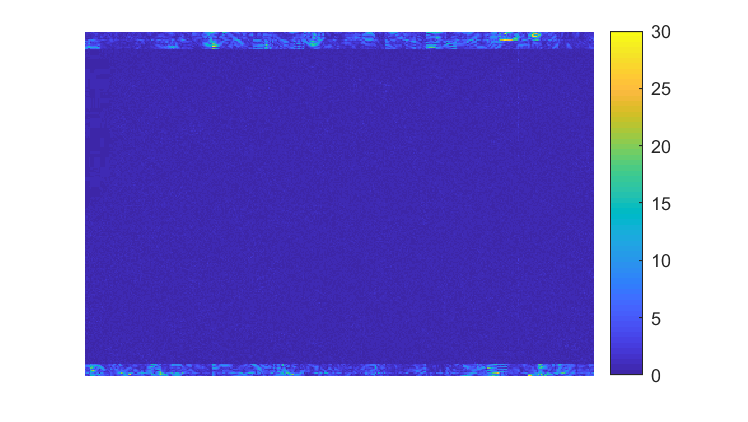
**ΔE from Comparisons between Viewers**



**Distribution of ΔE of Pixels**

****

**NDP.view 2 Sedeen**

****

**ΔE Heat Map Comparing Above Two Screenshots Showing Viewer Stitching Error**

**Results**

Six comparisons are made between the four images. The color difference for each pixel is shown as one dot in these heat maps. The title of each heat map shows the viewer names of the pair. We used the same color bar between 0 and 30 ΔE for all 6 heat maps. Blue areas have low ΔE, meaning similar colors. Yellow areas have high ΔE, meaning different colors. We see that two heat maps are blue and the other four are mostly yellowish green.

The results of the study show that there are significant differences in color between the reference Hamamatsu NDP.view 2 viewer and QuPath and between NDP.view 2 and ASAP, with high average ΔE values of 18.69 and 18.79, respectively. Comparing between Sedeen and QuPath and between Sedeen and ASAP give similar results. These color differences seem to correlate with tissue structure, with nuclei, stroma, and background sections having different levels of ΔE. The cause of these large color differences needs to be determined but is very likely to be the color profile. Sedeen was much closer in color to NDP.view 2, with an average ΔE of 1.30, but there were still slight differences almost imperceptible to the human eye. We can infer that Sedeen uses different mechanism to decode the WSI file, which may call for additional tests. ASAP was also problematic; when comparing ASAP with other viewers, noticeable "pixelized" ΔE patterns show. This could be a sign of differences in decompression.

The most surprising finding was the existence of stitching errors in the Sedeen viewer. This is most noticeable when comparing between Sedeen and NDP since they are the closest in color, but can also be readily seen in comparisons between Sedeen and the other two viewers. Choosing a ROI that straddles a stitch and only registering within one tile clearly shows almost 0 ΔE for pixels in the registered portion, but major color differences across the stitch boundary. This viewer stitching error requires further investigation, because it poses serious challenges to the validity of the viewer decompression method.

**Conclusion**

For the same input file, four different WSI viewers generated four different images, which confirm that it is a fallacy to assume that any WSI viewer can reproduce digital images identically. The concept of interoperability between WSI components needs to be revisited to include image integrity on the pixel level. Before a WSI file format is standardized, third-party viewer vendors should work with the scanner manufacturer instead of relying on untested free libraries. More adequate bench testing data are needed for 510(k) WSI viewer submissions. In the future, more work needs to be done to determine the sources of these observed differences, and the experiments should be repeated with the two FDA-cleared WSI devices included. Ultimately, bench test methods and acceptable criteria for determining substantial equivalence in terms of color performance among viewers should be established.

**References**

Calvin Sun and Wei-Chung Cheng, *How Much Can Bad Display Calibration Cost?* 2017 FDA Summer Student Poster Day.

**Acknowledgements**

This study was supported by ORISE. 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. We are grateful to Jonathan Boswell for helping in setting up this study.