Evaluating Color Performance of Whole-Slide Imaging Devices by Multispectral Imaging of Biological Tissues

Firdous Saleheen[1], Aldo Badano[2], Wei-Chung Cheng[2]

[1]*Temple University, Philadelphia, PA;* [2]*CDRH, U.S. Food and Drug Administration, Silver Spring, MD.*

ABSTRACT

The color reproducibility of two whole-slide imaging (WSI) devices was evaluated with biological tissue slides. Three tissue slides – human colon, skin, and kidney – were used to test a modern and a legacy WSI devices. The color truth of the tissue slides was obtained using a multispectral imaging system. The output WSI images were compared with the color truth to calculate the color difference for each pixel. A psychophysical experiment was also conducted to measure the perceptual color reproducibility (PCR) of the same slides with 4 subjects. The results show that the mean color differences of the modern, legacy, and monochrome WSI devices are 10.94, 22.35, and 42.74 ΔE_{00}, while their mean PCRs are 70.35%, 23.06%, and 0.91%, respectively.

REFERENCES

[1] FDA, *Guidance for WSI devices* (2016).
[2] A Pagi, *Diagnostic Pathology 6* (suppl. 1) (2011).
[3] WC Cheng et al., *SPIE Medical Imaging* (2013).
[4] P Shrestha et al., *J Med Imaging 1.2* (2014).
[5] WC Revie et al., *Anal Cell Pathol* (2014).

INTRODUCTION

Motivation: Color performance is an essential factor when evaluating WSI devices for making regulatory decisions [1].

Challenge: Color truth of biological tissues is difficult to measure because their microscopic structures are too small for color meters.

Existing Methods:

[2] compares slides with images visually – subjective and sensitive to viewing conditions

[3-4] use photographic film-based targets – do not match biological tissues spectrally

[5] uses spectrally matching color targets – lack of spatial uniformity and full color gamut

Our approaches:

- Developed a multispectral imaging system to measure the color truth for each pixel
- Compare WSI devices with an optical microscope to determine perceptual color reproducibility

Correspondence to: Wei-Chung.Cheng@fda.hhs.gov

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.

SPIE Medical Imaging 2017,
Digital Pathology Conference,
13 Feb, 2017. Paper: 10140-26

METHODOLOGY

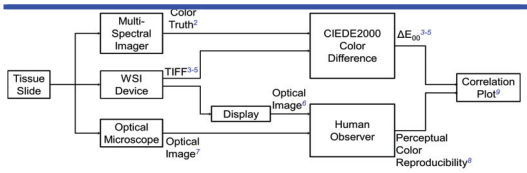


Fig. 1: Methodology workflow. Any tissue slide can be used as the input target. The upper two streams evaluate the per-pixel colorimetrical differences. The lower two streams evaluate the perceptual color differences. The correlation between colorimetrical and perceptual differences is analyzed in the end. The superscript numbers indicate the corresponding figures for each block.

PER-PIXEL COLOR DIFFERENCES (ΔE_{00})

The measured color differences are highlighted in the WSI images.

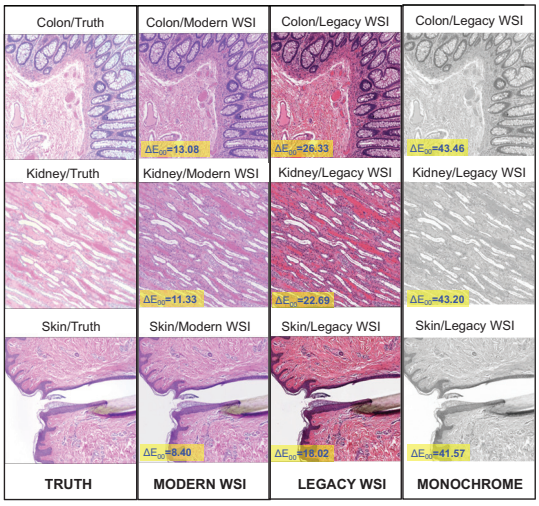


Fig. 2: Reference images from the color truth. *Fig. 3: Images from the modern WSI device.* *Fig. 4: Images from the legacy WSI device.* *Fig. 5: Images from the monochrome WSI device.*

PERCEPTUAL COLOR REPRODUCIBILITY (PCR)

- Determine *perceptual color reproducibility* quantitatively with a psychophysical experiment
- Study design: Scale PCR of 4 display images (*truth*, *modern*, *legacy*, and *monochrome WSI devices* with respect to the microscope image (Fig. 6 and 7)
- Psychophysical method: Cross-modality ratio scaling - perceptual color reproducibility vs. perceptual length
- Subject task: For each image, place a mark on a line to represent its perceptual color reproducibility. The left end of the line represents the least PCR, while the right end the most (sample shown in Fig. 8).
- Results from 4 subjects, 6 trials per image

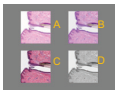


Fig. 6: Display images.

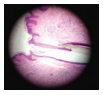


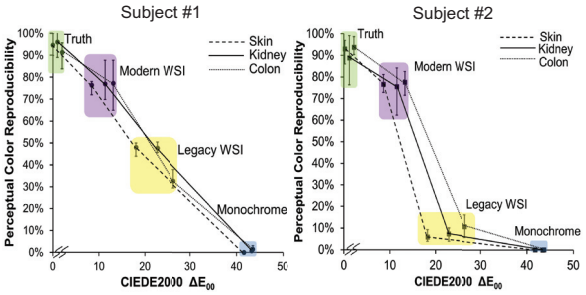
Fig. 7: Microscope image.

Table 1: Perceptual Color Reproducibility

	Perceptual Color Reproducibility (%)			
Tissue	Truth	Modern	Legacy	Monochrome
Colon	89.71	69.08	20.63	0.75
Kidney	89.75	70.82	24.65	0.97
Skin	86.84	71.14	23.89	1.01
Mean	88.77	70.35	23.06	0.91

Fig. 8: Responses to the stimuli in Fig. 6 and 7 from 4 subjects.

CORRELATION BETWEEN ΔE_{00} AND PCR



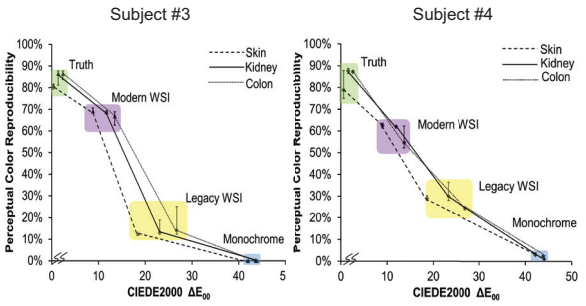


Fig. 9: PCR vs. color difference of 3 tissue types (colon, kidney, and skin) and 4 devices (truth, modern, legacy, and monochrome) from 4 subjects. Each data point represents the mean PCR of 6 trials, while the error bars represent the ranges. For clarity, the data points, especially the "truth" dataset, are slightly staggered in the X-axis direction.

FINDINGS

- Monochrome images have near zero PCR (0.91%) in high consensus.
- Truth images do not have 100% PCR (88.77%) Although ΔE_{00}=0, the images were not perfectly reproduced on the display.
- Modern WSI images have ΔE_{00}=10.94 and PCR=70.35%.
- Legacy WSI images have ΔE_{00}=22.35 and PCR=23.06%.
- Both PCR and ΔE_{00} can perfectly detect legacy vs. modern WSI devices (sensitivity=specificity=100%).
- Inter-reader variability in PCR for legacy WSI is greater than modern WSI.
- Subject #2 and #3 scaled legacy WSI much lower.

CONCLUSION

Non-inferiority of color reproducibility in the modern WSI device is reduced using two quantitative methods.

ACKNOWLEDGEMENT

This study was supported by Critical Path Initiative and ORISE. The authors thank our study participants Yu-Han Cheng, Padraic Flynn, Qi Gong, and Jonathan Rosenberg as well as Katherine Shea and Rodney Rouse, CDER for providing the legacy WSI scans.