Colorimetrical uncertainty of a hyperspectral imaging microscopy system for assessing whole-slide imaging devices

Paul Lemaillet and Wei-Chung Cheng\*

Division of Imaging, Diagnostics, and Software Reliability, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993, USA

\*Wei-Chung.Cheng@fda.hhs.gov

**Abstract:**

A whole-slide imaging (WSI) device is a color medical imaging system used in digital pathology to digitalize stained tissue samples into electronic images for pathologists to diagnose without using a conventional light microscope. Testing the color performance of a WSI device usually implies a color target with known truth that is compared with the device output to estimate color differences. Using stained tissue samples as color targets is challenging because the cellular features cannot be measured with ordinary spectroradiometers unless a hyperspectral imaging microscopy system (HIMS) is used. The goal of this study was to determine the colorimetrical uncertainty of such a reference HIMS that was designed to assess the color performance of WSI devices. A set of optical filters were used for that purpose. The color truth, in terms of spectral transmittance in the visible band, of the optical filters was measured by a reference spectroradiometer. The spectral transmittance was combined with a standard illuminant to generate colorimetrical measures using the CIEXYZ and CIELAB formulas. The differences between the reference HIMS and the reference spectroradiometer were evaluated using the CIE 1976 color difference formulas.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](https://doi.org/10.1364/OA_License_v1)

1. Introduction

A whole slide imaging (WSI) system is an automated digital slide creation, viewing, and management system intended as an aid to the pathologist to review and interpret digital images of surgical pathology slides. The system generates digital images that would otherwise be appropriate for manual visualization by conventional light microscopy [1]. As a medical imaging device, the technical characteristics of a WSI system, such as spatial resolution and focusing accuracy, need to be comparable to the conventional light microscopy that has been used by pathologists for decades. Among the essential technical characteristics of WSI systems, color performance is fundamental because histology is based on staining techniques to color and expose invisible cellular structures [2]. In conventional light microscopy color is transmitted purely in the optical domain whereas for WSI systems digital conversion usually leads to color discrepancies between the original scene and the device output [3, 4], making color performance unique for WSI systems. Characterizing the color behavior of a WSI system requires color measurement of the scene, which is challenging because the cellular structures of tissue samples are too fine to be measured by ordinary spectroradiometers. To assess the color performance of WSI devices, a hyperspectral imaging microscopy system (HIMS) developed to measure the color truth of tissue samples at the pixel level was previously presented [5]. However, the measurement accuracy of the HIMS itself was not reported.

In this study, a test method is presented to determine the measurement accuracy and uncertainty of such a HIMS. The results include transmittance measurements of neutral density filters as well as of color transmittance filters that are compared to reference measurements of the same region of interest (ROI) by a spectroradiometer equipped with a fiber probe. The CIELAB color space coordinates and their uncertainties are derived from the transmittance measurements and the CIE 1976 color difference in the CIELAB color space [6] is used as a color performance assessment metric.

1. Material and method

Figure 1(a) presents the experimental setup used to estimate the color performance of a HIMS.

* 1. Device under Test

The hyperspectral imaging microscope system is based on an upright light microscope (AxioPhot 2, Carl Zeiss Microscopy, White Plains, NY, USA) in bright-field illumination mode. The original lamp housing is replaced with a tunable light source (OL490, Gooch and Housego, TX, USA) using a xenon lamp to provide illumination from to with adjustable band width. In this study, is spanned in steps of with as the 41 spectra shown in Fig. 1(b). A liquid light guide directs the light to a collector lens (MCWHL5-C4, Thorlabs, Newton, NJ, USA) that is followed by a condenser (Achromatic-aplanatic , Carl Zeiss Microscopy, White Plains, NY, USA). The sample is set on a motorized stage system (MAC 6000, Ludl Electronic Products Ltd., Hawthorne, NY, USA) and is imaged using a 20x objective (Plan-Apochromat 20x , Carl Zeiss Microscopy, White Plains, NY, USA). The image of the sample is acquired by a camera (Grasshopper3 9.1 MP Mono USB3 Vision, Point Grey Research Inc., BC, Canada) with a monochrome sensor (ICX814, Sony Electronics, Newton, NJ, USA). The image format used in this study is a pixels area situated in the center of the sensor. The camera shutter time and light source intensity are optimized to prevent detector saturation while maximizing the detected signal for each measurement wavelength. Both the focusing and Kohler illumination are achieved in white light, . The tunable light source, motorized stage, and camera are all controlled by programs written in Matlab (Mathworks, MA, USA) running in the Microsoft Windows 7 Professional 64-bit environment.

|  |  |
| --- | --- |
| (b)  OL490 light engine  PR730 spectroradiometer  XYZ motorized stage  (a)  (a) |  |

Fig. 1. (a) Hyperspectral microscope equipped with a PR730 spectroradiometer; (b) OL490 light engine output of the 41 spectral bands (scaled to the same light intensity parameter) used to obtain the hyperspectral images.

* 1. Reference Instrument

A spectroradiometer (PR-730 with fiber probe FP-730, Photo Research, Syracuse, NY, U.S.A.) is used as the reference instrument. The distal end of the fiber probe is positioned in the eyepiece tube to measure the ROI as observed by the hyperspectral imaging microscope system.

* 1. Samples

Both standard and non-standard transmittance targets are used. The standard targets include Kodak Warren (KW) gelatin neutral density (ND) filters with optical density and color gelatin filters #12 (yellow), #25 (red), #32 (magenta), #47 (deep blue), and #58 (green) (Edmunds Optics, Barrington, NJ, USA). One should note that in order to prevent potential interference patterns due to air gap between film and glass[7], the KW filters are held laterally without setting them on a glass slide. To better represent the color gamut of the hematoxylin-and-eosin (H&E) stained tissue samples, we designed a color phantom populated with a choice of Roscolux color filters (Rosco Laboratories Inc., Stamford, CT, USA) as a set of non-standard transmittance targets. 24 holes (diameter ) are punched on a 1 mm-thick supporting cardboard slab using a dot puncher. A thin cardboard slab with holes is glued on the supporting slab and 23 locations are filled with a filter dot glued on the supporting cardboard. The additional hole is left empty for measuring the 100% transmittance. A thin covering slab with 24, -punched holes is glued on top (Fig. 2).



#40

#316

#03

#46

#342

#24

#43

#336

#337

#39

#34

#48

#52

#51

#360

#347

#356

#59

#56

#68

#99

#398

#97

Fig. 2: Color phantom with 23 Roscoloux color filter dots glued on a cardboard slab set with an adequate series of punched holes. The empty position is the 100% transmittance slot.

* 1. Measurements

Measurements of the transmittance are conducted at first using the spectroradiometer with a broadband illumination and, secondly using the camera with narrow band illumination. The numerical aperture of the spectroradiometer’s detector fiber probe averages out over the sample ROI and for comparison purposes, images of the same ROI captured by the camera are spatially averaged as well. Repeated measurements lead to signal intensities from which a mean value and a standard deviation can be computed. The transmittance of the sample is then expressed as

|  |  |
| --- | --- |
|  | (1) |

where and are the intensities of the signal measured with the sample the light path and with no sample in the light path (), respectively. and are the background signals obtained setting the intensity of the light engine to zero and account for both persistent background illumination observed with the light source and camera background noise. For measuring and , the intensity of the light source is optimized to obtain the higher signal output without detector saturation. Additionally, for measurements with the camera, both the light source intensity value and the camera shutter time are optimized wavelength by wavelength. The International Commission on Illumination (Commission Internationale de l’Éclairage, CIE) tri-stimulus values are then computed as

|  |  |
| --- | --- |
| , | (2) |

where is the relative spectral power of one of the CIE standard illuminant, , and are the CIE 1931 color matching functions [8-10], and is the normalizing factor of corresponding to 100 for a perfectly transmitting sample. In this study, the CIE D65 standard illuminant is used. Converting the integral to summations, we have

|  |  |
| --- | --- |
| *,* | (3) |

where

|  |  |
| --- | --- |
| . | (4) |

The CIELAB (, , ) values are then computed as

|  |  |
| --- | --- |
|  | (5) |

where

|  |  |
| --- | --- |
|  | (6) |

and are the tri-stimulus values for a perfectly transmitting sample. The CIE 1976 color difference in the CIELAB color space can be computed as the Euclidian distance between two color points of coordinates and

|  |  |
| --- | --- |
| , | (7) |

where and .

* 1. Uncertainty propagation

The uncertainty on the transmittance, CIEXYZ coordinates, CIELAB coordinates are estimated using the law of uncertainty propagation which is based on the Taylor expansion of the functional relationship between the output quantities and the input quantities about mean values, [11, 12]. In matrix form, assuming that is the covariance matrix of the input quantities, the covariance matrix of the output quantities is

|  |  |
| --- | --- |
| , | (8) |

where is the Jacobian matrix with .

From Eq. (1) and Eq. (8), the uncertainty on the transmittance at each measurement wavelength is

|  |  |
| --- | --- |
| , | (9) |

since the measurements of the sample, the 100% transmittance and background are independent. Here is the standard deviation on the measured signals.

From Eq. (3), the tri-stimulus equations are linear functions of the transmittance measurements and Eq. (8) is not an approximation in that case, with

|  |  |
| --- | --- |
|  | (10) |

where is the number of measurement wavelengths. Since the transmittance measurements over the wavelengths are independent, is a diagonal matrix with elements , i.e. the variance of the transmittance measurements. The tri-stimulus values are correlated because the color matching functions overlap, and they depend on the transmittance spectrum [13]. The resulting correlation matrix is .

From Eq. (5), the non-linearity of the relationship between the CIELAB coordinates and the CIEXYZ coordinates implies that Eq. (8) is an approximation, with [14]

|  |  |
| --- | --- |
|  | (11) |

where

|  |  |
| --- | --- |
| . | (12) |

The resulting covariance matrix,,is .

From Eq. (7) the normally distributed CIELAB coordinates and can be used to form a Euclidian distance in the CIELAB space, , that is not normally distributed. Hence, the statistical distributions of between both types of measurements are computed by Monte Carlo simulations of the color points positions using the covariance matrices of the CIELAB coordinates, and . The median of the statistical distribution, , is used as metric to estimate the proximity between and .

Here, we limit our estimation to the type A uncertainty (uncertainty evaluated by the statistical analysis of series of observations[12]) by considering: i) the propagation of the uncertainty on a set of measured transmittances under the same measurement conditions, i.e. repeatability of the results, and ii) conducting the experiments several times under changed conditions to account of the reproducibility. The estimated variance, on the results of the reproducibility experiments is added to the square of the uncertainty on the repeatability experiments to compute the total type A variances, . The expanded uncertainty is where is the coverage factor. We used which correspond to a .confidence interval.

1. Results

To assess the linearity of the transmittance measurements with the camera, we compare the spatial average of the transmittance images, , to the spectroradiometer transmittance measurements, , for a set of KW gelatin ND filters with optical density . The uncertainties on and are computed using Eq. (9). Ten reproducibility experiments are conducted for . The estimated variance from the reproducibility experiments, is used for all samples to compute the total type A variances, . Figure 3(a) shows that for most values, and overlap within the error bars over most wavelengths in the spectral range of measurements but that the differences can be significant for . At these wavelengths, the values of the color matching functions , and are small enough that the impact of the transmittance values on the end results CIELAB coordinates is limited. Figure 4 illustrates this assumption by presenting , , and the Relative Cumulative Weight (RCW) of the color matching functions, expressed as . We compute a weighted linear interpolation of versus using the uncertainty on as weight parameters and considering as the ground-truth of the transmittance of the KW ND filters. For a broad range of wavelengths, , there is a linear relationship between . As an example, Fig. 3(b) presents the results of the linear interpolation at , with a slope and an intercept for a root mean square error (rmse) of .

|  |  |
| --- | --- |
| (a)  (b) |  |

Fig. 3. (a) Comparison of , the transmittance spectra measured with a spectroradiometer, and , the transmittance measured with the camera (spatial average over the image) for a set of KW gelatin neutral density filters with . and overlap within the error bars (coverage factor ) for for most ND filters; (b) versus fitted with a linear model for .

Table 1 presents the CIELAB coordinates results for the set of neutral density filters, their uncertainties and the median value of the statistical distributions of , , obtained for each sample. One should note that despite the proximity of the values issued from and , there is not always an overlap within the error bars at . For , despite the relative proximity of the CIELAB coordinates mean values their large uncertainties explain the larger obtained for this sample for which the and values are close to zero over the detection wavelengths.

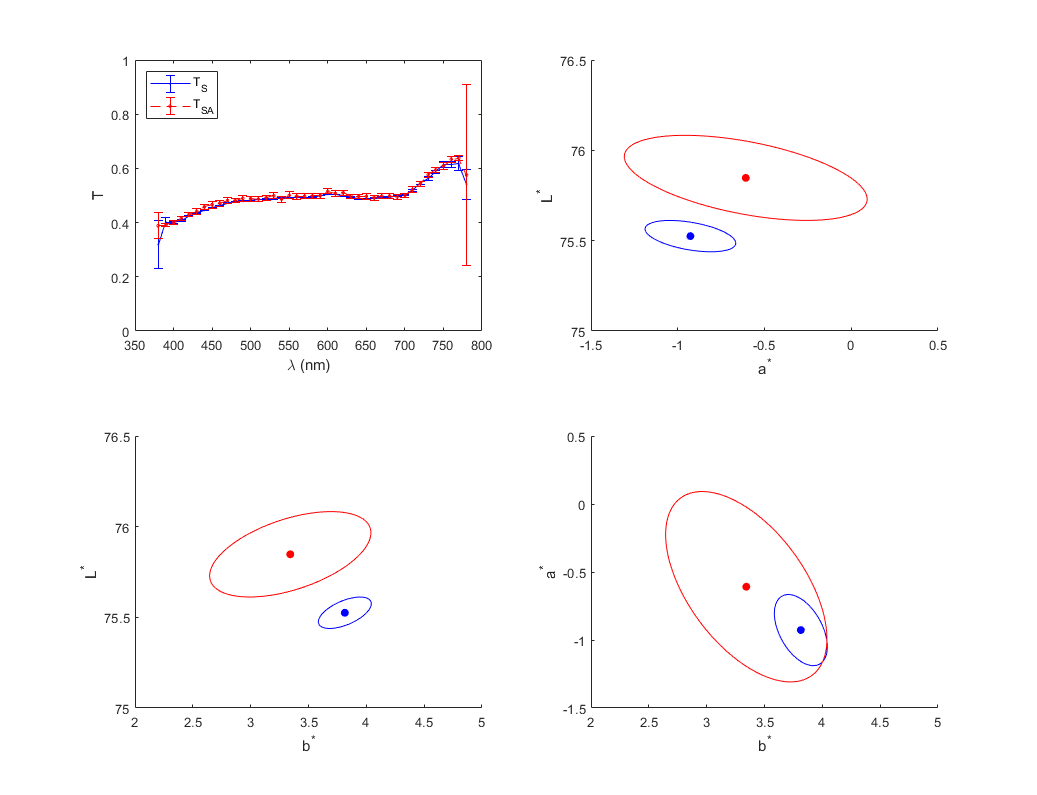


Fig. 4. The CIE 1931 color matching functions , and the relative cumulative weight.

Figure 5(a) presents the transmittance data for and shows that and overlap within the error bars at all wavelengths. For the overlap is explained by the large uncertainty on . Figure 5(b) (c) and (d) present the CIELAB coordinates and 95% confidence regions issued from and measurements and their uncertainties in the (), () and () projection planes, respectively. There is no overlap between the 95% confidence regions apart from in the () projection plane which points to a systematic error on the color coordinates. However, the median value of () is relatively small and the agreement between the color coordinates issued from and is considered reasonable.

Table 1. CIELAB coordinates for the KW gelatin neutral density filters as measured by the spectroradiometer, , and the spatially averaged images acquired by the camera, . The statistical distribution of the Euclidian distance in the CIELAB color space between both types of measurements is computed by Monte Carlo simulations of the color point positions using the covariance matrices of the CIELAB coordinates. The median value of these distributions are reported. The uncertainties are presented with a coverage factor .

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ND | Transmittance |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |



(a)

(b)

(c)

(d)

Fig. 5. (a) Comparison of , the transmittance spectra measured with a spectroradiometer, and , the transmittance measured with the camera (spatial average over the image) for KW gelatin neutral density filter (error bars with a coverage factor ); (b) (c) and (d) CIELAB coordinates and 95% confidence regions from (blue) and (red) in the (), () and () projection planes, respectively.

Since neutral density filters have low chromaticity values, color filters are measured to assess the color performance of the setup. We first measure a set of KW color gelatin filters and then we measure the color phantom. Ten reproducibility experiments are conducted for samples KW #32 and #47. To maximize the resulting uncertainties on the CIELAB coordinates, the estimated variance from the reproducibility experiments, of sample #32 is used for all KW color filters but for sample #47 to compute the total type A variances, on the measured transmittances. Figure 6 shows that for all samples and curves overlap within the error bars at all wavelengths. Again, for the overlap is explained by the large uncertainty on .

The CIELAB coordinates, their uncertainty and the median value of the statistical distributions of obtained for each KW color filters are presented in Table 2. Again, despite the proximity of the values issued from and , there is not always an overlap within the error bars at .

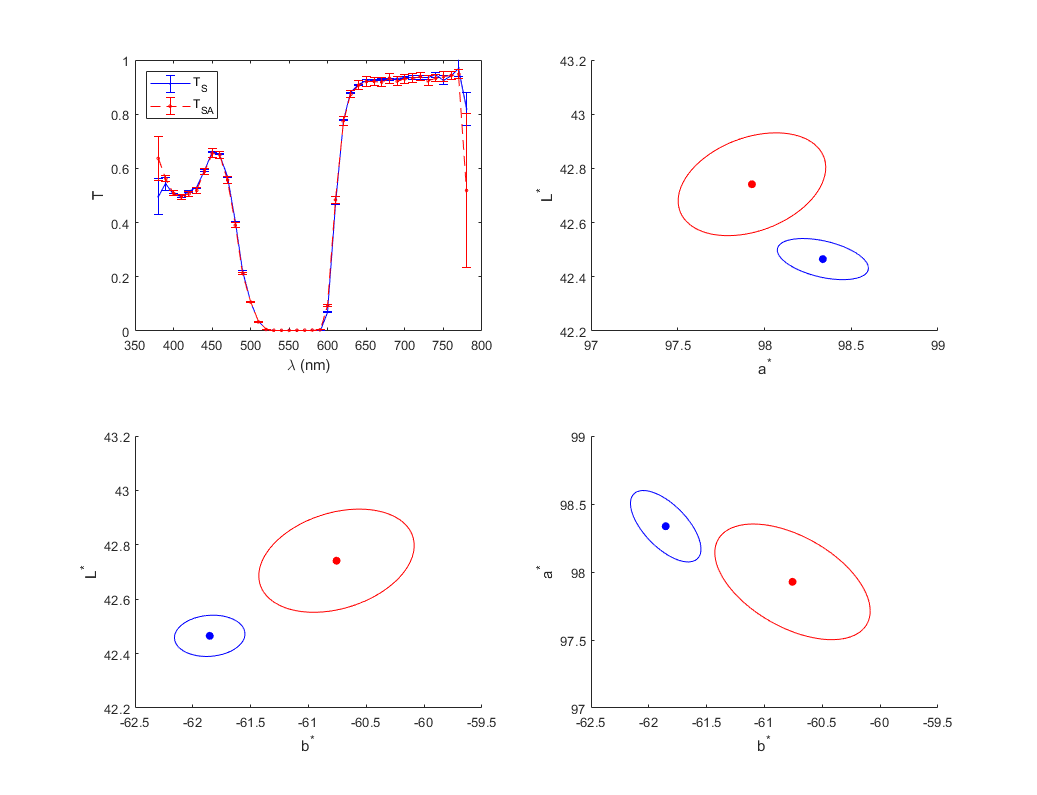


Fig. 6: Transmittance spectra of five KW color gelatin filters (#12: yellow; #25: red: #32: magenta; #47: deep blue: #58: green) measured by the spectroradiometer (, plain) and the camera (, dash). and overlap within the error bars (coverage factor ) for .

Table 2. CIELAB coordinates of five KW color gelatin filters (#12: yellow; #25: red: #32: magenta; #47: deep blue: #58: green) derived from the spectra measured by the spectroradiometer, , and the spatially averaged images acquired by the camera, . The statistical distribution of the Euclidian distance in the CIELAB color space between both types of measurements was computed by Monte Carlo simulations of the color point positions using the covariance matrices of the CIELAB coordinates. The median value of these distributions are reported. The uncertainties are presented with a coverage factor .

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Filter | Transmittance |  |  |  |  |  |
| 12 |  |  |  |  |  |  |
|  |  |  |  |  |
| 25 |  |  |  |  |  |  |
|  |  |  |  |  |
| 32 |  |  |  |  |  |  |
|  |  |  |  |  |
| 47 |  |  |  |  |  |  |
|  |  |  |  |  |
| 58 |  |  |  |  |  |  |
|  |  |  |  |  |

As an illustration of the uncertainty analysis for the KW color filters, Fig. 7(a) presents the and spectra with their uncertainty for filter #32 (magenta) and shows that the largest discrepancies indeed occur at . The 95% confidence regions around the CIELAB coordinates derived from and and their uncertainties presented in Fig. 7(b) (c) and (d) (CIELAB projection planes) do not overlap. Again, this points to a systematic error. For the whole set of KW color filter, ranges from to and the agreement between the color coordinates issued from and is considered reasonable.



(a)

(b)

(c)

(d)

Fig. 7. (a) Comparison of , the transmittance spectra measured with a spectroradiometer, and , the transmittance measured with the camera (spatial average over the image) for KW color gelatin filter #32 (magenta) (error bars with a coverage factor ); (b) (c) and (d) CIELAB coordinates and 95% confidence regions from (blue) and (red) in the (), () and () projection planes, respectively.

The gamut of the color filters composing the color phantom is represented in the CIELAB space along with their color appearance (Fig. 8(a)). Figure 8(b) is a boxplot of the corresponding values computed from , and their uncertainties. Median value of are in the 0.5 to 1.0 range while most outliers are smaller than 3 with some exceptions for a small sub-selection of 5 filters (#24, #342, #46, #347, #59). However, these patches have upper whiskers smaller than 3. Again, there is a good agreement between the color results computed from the and measurements. The phantom will be used provide traceability of the measurements by the hyperspectral microscope.

|  |  |
| --- | --- |
| (b)  (a)  Fig. 8. (a) CIE LAB representation of the 23 patches composing the color filter phantom; (b) Corresponding issued from (spectroradiometer) and (camera) measurements. |  |

1. Conclusion

Tissue slides can be used for assessing the color performances of WSI scanners. A hyperspectral microscope was previously developed by our group to measure the color data of such reference tissue slides at the pixel level. The color performances of this microscope are estimated by measuring the transmittance of a set of spatially uniform KW neutral density filters, KW color filters and a color phantom composed of 23 Roscolux color filter dots glued on a cardboard. The results are compared to a reference obtained measuring the same ROI with a spectroradiometer equipped with a fiber probe whose distal end is set in the one of the microscope eyepiece tube. The hyperspectral microscope acquires 10 spatially-averaged images at each acquisition wavelength. Mean values and standard deviation of the corresponding signal intensities when measuring the sample and a 100% transmittance (no sample) with the light source on and off (background) are composed to estimate the transmittance spectra. Similar repeated measurements are conducted with the spectroradiometer. The CIELAB color coordinates and their uncertainties are then estimated and compared by computing a Monte Carlo simulation of the Euclidian distance in the CIELAB space. The 95% confidence region generated in the CIELAB space projection planes generally do not overlap for the samples measured, pointing to some systematic error. However, the median values of are in the range of to for ND filters with an outlier of for for which the transmittance signal is close to zero. For the color filters, are in the 0.5 to 1.0 range while most statistical outliers are smaller than 3. These differences are deemed small enough for the color assessment of WSI scanners using tissue slides measured by the reference HIMS. The color phantom will be used to maintain a traceability on the measurements by the hyperspectral microscope.

Our future work plan toward measuring a selection of tissue slides to estimate their color coordinates and uncertainty pixel by pixel. Subsequently, these reference slides will be used to assess the color performance of a series of WSI scanners.

1. Funding, acknowledgments, and disclosures

5.1 Funding

This study was supported by the Critical Path Initiative.

5.2 Acknowledgments

The authors thank Drs. Anant Agrawal, Ali Afshari, Si Wen, Aldo Badano, Ryan Beams for their technical support and comments.

5.3 Disclosures

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. "Medical Devices; Hematology and Pathology Devices; Classification of the Whole Slide Imaging System. Final order," Federal register **83**, 20 (2018).

2. E. L. Clarke, and D. Treanor, "Colour in digital pathology: a review," Histopathology **70**, 153-163 (2017).

3. E. L. Clarke, C. Revie, D. Brettle, 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 Research & Application **43**, 184-197 (2018).

4. A. Badano, C. Revie, A. Casertano, W.-C. Cheng, P. Green, T. Kimpe, E. Krupinski, C. Sisson, S. Skrøvseth, and D. Treanor, "Consistency and standardization of color in medical imaging: a consensus report," Journal of digital imaging **28**, 41-52 (2015).

5. W. C. Cheng, F. Saleheen, and A. Badano, "Assessing color performance of whole‐slide imaging scanners for digital pathology," Color Research & Application **44**, 322-334 (2019).

6. A. R. Robertson, "The CIE 1976 color‐difference formulae," Color Research & Application **2**, 7-11 (1977).

7. P. Shrestha, and B. Hulsken, "Color accuracy and reproducibility in whole slide imaging scanners," Journal of Medical Imaging **1**, 027501 (2014).

8. M. E. Nadal, E. A. Early, and R. R. Bousquet, "0: 45 Surface Color," NIST Special Publication SP250-71 (2008).

9. "CIE S014-1/E: 2006: Colorimetry - Part I: CIE Standard Colorimetric Observer " (2007).

10. "CIE S014-1/E: 2006: Colorimetry - Part II: CIE Standard Illuminant " (2007).

11. "Guide to the Expression of Uncertainty in Measurement (GUM)–Supplement 1: Numerical Methods for the Propagation of Distributions," International Organization for Standardization (2004).

12. B. N. Taylor, and C. E. Kuyatt, "Guidelines for evaluating and expressing the uncertainty of NIST measurement results," NIST Technical Report 1297 (1994).

13. J. Gardner, and R. Frenkel, "Correlation coefficients for tristimulus response value uncertainties," Metrologica **36**, 477 (1999).

14. G. Wübbeler, J. Campos Acosta, and C. Elster, "Evaluation of uncertainties for CIELAB color coordinates," Color Research & Application **42**, 564-570 (2017).