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 whose application in digital pathology is 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 is to determine the colorimetrical uncertainty of such a reference HIMS that is designed to assess the color performance of WSI devices. A set of optical filters are used for that purpose. The color truth, in terms of spectral transmittance in the visible band, of the optical filters is measured by a reference spectroradiometer. The spectral transmittance is 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 are evaluated using the CIE 1976 color difference formulas.

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, i.e. using an eyepiece, 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.

2.   Material and method

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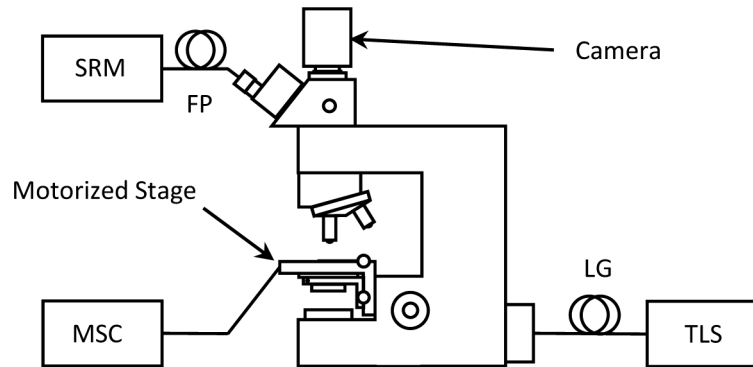

Fig. 1. Schematic layout of the hyperspectral microscope equipped with a PR730 spectroradiometer; SRM: spectroradiometer; FP: fiber probe; MSC: motorized stage control; LG: light guide; TLS: tunable light source

2.1.   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 $\lambda = $ 380 nm to 780 nm with adjustable bandwidth (BW). In this study, λ is spanned in steps of 10 nm with BW = 10 nm as the 41 spectra shown in Fig. 2. A liquid light guide directs the light to a $f = $ 40 mm collector lens (MCWHL5-C4, Thorlabs, Newton, NJ, USA) that is followed by a condenser (Achromatic-aplanatic NA = 0.9, 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 NA = 0.8, 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 676×844 pixels (2.49×3.11 mm^2) 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 the Kohler illumination adjustment are achieved in broadband illumination, i.e. using the total illumination spectrum provided by the xenon light source of the OL490. 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.

### 2.2.   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.
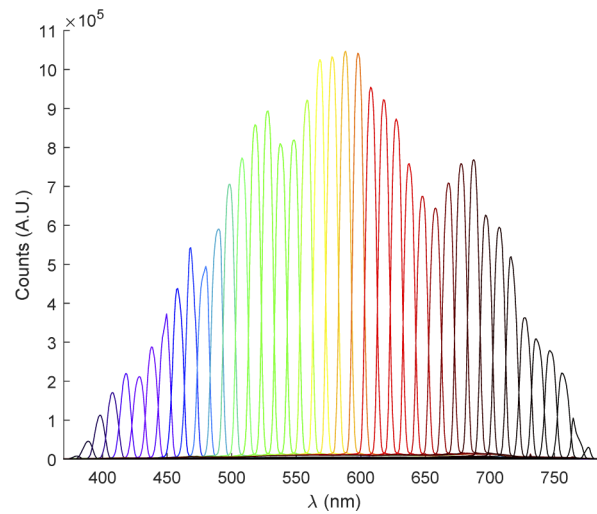
Fig. 2. OL490 light engine output of the 41 spectral bands (scaled to the same light intensity parameter) used to obtain the hyperspectral images.

2.3. Samples

Both standard and non-standard transmittance targets are used. The standard targets include Kodak Wratten (KW) gelatin neutral density (ND) filters with optical density $OD = [0.1, 0.2, 0.3, 0.6, 1.0, 2.0]$ 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 4.76 mm) are punched on a 1 mm-thick supporting cardboard slab using a dot puncher. A thin cardboard slab with 6.35 mm 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, 4.76 mm-punched holes is glued on top (Fig. 3).
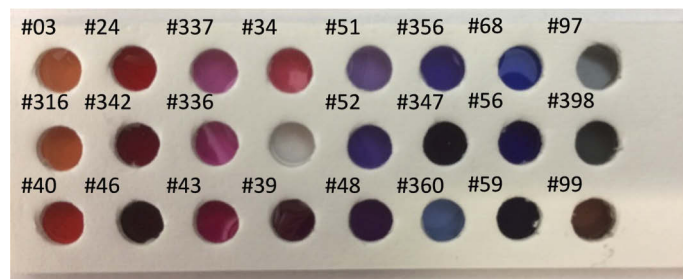
Fig. 3. Color phantom with 23 Roscoloux color filter dots glued on a cardboard slab set with an adequate series of punched holes. The empty position is the 100% transmittance slot.

2.4.   Measurements

Measurements of the transmittance are conducted is two distinct steps. First, the sample is measured using the spectroradiometer with a broadband illumination. Secondly, the same sample is measured using the camera with narrow band illumination. In both cases the illumination is provided by the OL490 light engine. The numerical aperture of the spectroradiometer's detector fiber probe averages out over the sample ROI. 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

$$T(\lambda) = \frac{I_S(\lambda) - I_{SBg}(\lambda)}{I_W(\lambda) - I_{WBg}(\lambda)}, \tag{1}$$

where (I_S, I_{SBg}) and (I_W, I_{WBg}) are the intensities of the signal measured with the sample in the light path and with no sample in the light path ($T = 100\%$), respectively. I_{SBg} and I_{WBg} are the background signals obtained setting the intensity of the light engine to zero. They account for both persistent background illumination observed with the light source and camera background noise. For measuring I_S and I_W, 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. For that purpose, we implemented an independent step to define a table of shutter time values. For each wavelength, we acquire images with no sample in the light path, with a maximum light intensity value, starting from the maximum value of the camera shutter time and reducing it until no pixel in the image is saturated. This table defines the camera shutter time at each wavelength. During the regular sample measurements procedure, a preliminary measurement step consists in acquiring images with no sample in the light path and for each wavelength: (i) setting the camera shutter time defined by the camera shutter time table and (ii) reducing the light intensity from its maximum value until no image pixel is saturated. This preliminary step produces a table of maximum light intensity, the light intensity values used in the sample measurements were 90% of the table values to anticipate any light source intensity fluctuations.

The International Commission on Illumination (Commission Internationale de l'Éclairage, CIE) tri-stimulus values are computed from the transmittance spectra $T(\lambda)$ as

$$
\begin{aligned}
X &= K \int T(\lambda)S(\lambda)\bar{x}(\lambda)d\lambda \\
Y &= K \int T(\lambda)S(\lambda)\bar{y}(\lambda)d\lambda, \\
Z &= K \int T(\lambda)S(\lambda)\bar{z}(\lambda)d\lambda
\end{aligned}
\tag{2}
$$

where $S(\lambda)$ is the relative spectral power of one of the CIE standard illuminant, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ are the CIE 1931 color matching functions [8–10], and K is the norlizing factor of $Y(\lambda)$ 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

$$
\begin{aligned}
X &= K \sum_{i=1}^{N} T(\lambda_i)S(\lambda_i)\bar{x}(\lambda_i)\Delta\lambda &= f_1(T) \\
Y &= K \sum_{i=1}^{N} T(\lambda_i)S(\lambda_i)\bar{y}(\lambda_i)\Delta\lambda &= f_2(T) , \\
Z &= K \sum_{i=1}^{N} T(\lambda_i)S(\lambda_i)\bar{z}(\lambda_i)\Delta\lambda &= f_3(T)
\end{aligned}
\tag{3}
$$

Where

$$K = \frac{100}{\sum_{i=1}^{N} S(\lambda_i)\bar{y}(\lambda_i)\Delta\lambda}. \tag{4}$$

The CIELAB (L^*, a^*, b^*) values are then computed as

$$
\begin{aligned}
L^* &= 116 f\left(\frac{Y}{Y_n}\right) - 16 \\
a^* &= 500 \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right), \\
b^* &= 200 \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right)
\end{aligned}
\tag{5}
$$

where

$$
f(t) = \begin{cases} \sqrt[3]{t} & \text{if } t > \delta^3 \\ \frac{t}{3\delta^2} + \frac{4}{29} & \text{otherwise} \end{cases} \quad \text{with } \delta = \frac{6}{29},
\tag{6}
$$

and (X_n, Y_n, Z_n) 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 (L_1^*, a_1^*, b_1^*) and (L_2^*, a_2^*, b_2^*)

$$
\Delta E_{ab}^* = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}}
\tag{7}
$$

where $\Delta L^* = L_2^* - L_1^*$, $\Delta a^* = a_2^* - a_1^*$ and $\Delta b^* = b_2^* - b_1^*$.

2.5. Uncertainty propagation

The usefulness of measurements results is to a large extend determined by the quality of the statement of uncertainty that accompany them [11]. Since the HIMS is our reference instrument to determine the color truth of reference tissue samples further used to assess the color performance of WSI scanners, it is important to first estimate the uncertainty on the measurements results provided by the HIMS. Hence, 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 $y = f(x) = [f_1(x) f_2(x) \ldots f_p(x)]$ between the output quantities y and the input quantities x about the mean values of x, μ_x [11,12]. In matrix form, assuming that C_x is the covariance matrix of the input quantities, the covariance matrix of the output quantities C_y is

$$
C_y \simeq J C_x J^t
\tag{8}
$$

where J is the Jacobian matrix with $[J]_{ij} = \left.\frac{\partial f_i(x)}{\partial x_j}\right|_{\mu_x}$.

From Eq. (1) and Eq. (8), $C_T = J_1 C_I J_1^t$ at each measurement wavelength λ with

$$
J_1 = \begin{bmatrix} \frac{\partial T}{\partial I_S} & \frac{\partial T}{\partial I_W} & \frac{\partial T}{\partial I_{SBg}} & \frac{\partial T}{\partial I_{WBg}} \end{bmatrix}.
$$

The measurements of the sample, the 100% transmittance and background are independent. Hence, C_I is a diagonal matrix whose elements are the variances on the measured signals, $s_{x_i}^2$, where s_{x_i} is the standard deviation on the measured signals and $x_i = (I_S(\lambda), I_W(\lambda), I_{SBg}(\lambda), I_{WBg}(\lambda))$. In developed form, the resulting variance on the transmittance is

$$
s_{T(\lambda)}{}^2 = \sum_{i=1}^{4} \left(\frac{\partial T(\lambda)}{\partial x_i}\right)^2 s_{x_i}^2.
\tag{9}
$$

From Eq. (3), the tri-stimulus equations are linear functions of the transmittance measurements and Eq. (8) is not an approximation in that case, $C_{CIEXYZ} = J_2 C_T J_2^t$ with

$$
J_2 = \begin{bmatrix} \frac{\partial f_1(T)}{\partial T(\lambda_1)} & \cdots & \frac{\partial f_1(T)}{\partial T(\lambda_N)} \\ \frac{\partial f_2(T)}{\partial T(\lambda_1)} & \cdots & \frac{\partial f_2(T)}{\partial T(\lambda_N)} \\ \frac{\partial f_3(T)}{\partial T(\lambda_1)} & \cdots & \frac{\partial f_3(T)}{\partial T(\lambda_N)} \end{bmatrix} = K \begin{bmatrix} S(\lambda_1)\bar{x}(\lambda_1) & \cdots & S(\lambda_N)\bar{x}(\lambda_N) \\ S(\lambda_1)\bar{y}(\lambda_1) & \cdots & S(\lambda_N)\bar{y}(\lambda_N) \\ S(\lambda_1)\bar{z}(\lambda_1) & \cdots & S(\lambda_N)\bar{z}(\lambda_N) \end{bmatrix},
\tag{10}
$$

where N is the number of measurement wavelengths. Since the transmittance measurements over the wavelengths are independent, C_T is a $N \times N$ diagonal matrix with elements $[s_{T(\lambda_1)}{}^2, \ldots, s_{T(\lambda_N)}{}^2]$, 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 covariance matrix, C_{CIEXYZ}, is 3×3.

From Eq. (5), the non-linearity of the relationship between the CIELAB coordinates and the CIEXYZ coordinates implies that Eq. (8) is an approximation, $C_{CIELAB} \simeq J_3 C_{CIEXYZ} J_3^t$ with [14]

$$J_3 = \begin{bmatrix} 0 & \frac{116}{Y_n} f'\left(\frac{Y}{Y_n}\right) & 0 \\ \frac{500}{X_n} f'\left(\frac{X}{X_n}\right) & -\frac{500}{Y_n} f'\left(\frac{Y}{Y_n}\right) & 0 \\ 0 & \frac{200}{Y_n} f'\left(\frac{Y}{Y_n}\right) & -\frac{200}{Z_n} f'\left(\frac{Z}{Z_n}\right) \end{bmatrix}, \tag{11}$$

Where

$$f'(t) = \begin{cases} \frac{1}{3} t^{-2/3} & \text{if } t > \delta^3 \\ \frac{1}{3\delta^2} & \text{otherwise} \end{cases} \quad with \; \delta = \frac{6}{29} \tag{12}$$

The resulting covariance matrix, C_{CIELAB}, is 3×3.

From Eq. (7) the normally distributed CIELAB coordinates (L_1^*, a_1^*, b_1^*) and (L_2^*, a_2^*, b_2^*) can be used to form a Euclidian distance in the CIELAB space, ΔE_{ab}^*, that is not normally distributed. Hence, the statistical distributions of ΔE_{ab}^* between both types of measurements are computed by Monte Carlo simulations of the color points positions using the covariance matrices of the CIELAB coordinates, C_{CIELAB_1} and C_{CIELAB_2}. The median of the ΔE_{ab}^* statistical distribution, $\text{med}_{\Delta E_{ab}^*}$, is used as metric to estimate the proximity between (L_1^*, a_1^*, b_1^*) and (L_2^*, a_2^*, b_2^*).

Here, we limit our estimation to the type A uncertainty (uncertainty evaluated by the statistical analysis of series of observations [11]) 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, s_R^2, 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, u_A^2. The expanded uncertainty is $U = k u_A$ where k is the coverage factor. We used $k = 2$ which correspond to a 95%-confidence interval.
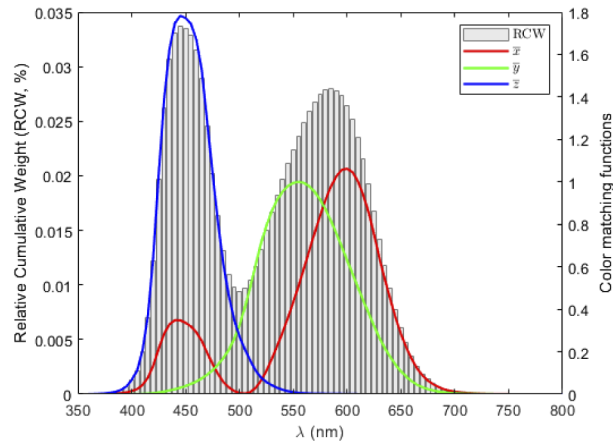
3. Results

To assess the linearity of the transmittance measurements with the camera, we compare the spatial average of the transmittance images, T_{SA}, to the spectroradiometer transmittance measurements, T_S, for a set of KW gelatin ND filters with optical density $OD = [0.1, 0.2, 0.3, 0.6, 1.0, 2.0]$. The uncertainties on T_{SA} and T_S are computed using Eq. (9). Ten reproducibility experiments are conducted for $OD = 0.3$. The estimated variance from the reproducibility experiments, s_R^2, is used for all samples to compute the total type A variances, u_A^2. Figure 4(a) shows that for most OD values, T_{SA} and T_S overlap within the error bars over most wavelengths in the spectral range of measurements but that the differences can be significant for $\lambda = 380$ nm and 780 nm. At these wavelengths, the values of the color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ are small enough that the impact of the transmittance values on the end results CIELAB coordinates is limited. Figure 5 illustrates this assumption by presenting $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ and the Relative Cumulative Weight (RCW) of the color matching functions, expressed as $\frac{\bar{x}(\lambda_i) + \bar{y}(\lambda_i) + \bar{z}(\lambda_i)}{\sum_{i=1}^{N} \bar{x}(\lambda_i) + \bar{y}(\lambda_i) + \bar{z}(\lambda_i)}$. We compute a weighted linear interpolation of T_{SA} versus T_S using the uncertainty on T_{SA} as weight parameters and considering T_S as the ground-truth of the transmittance of the KW ND filters. For a broad range of wavelengths, $\lambda = 390$ nm to 770 nm , there is a linear relationship between T_{SA} and T_S.
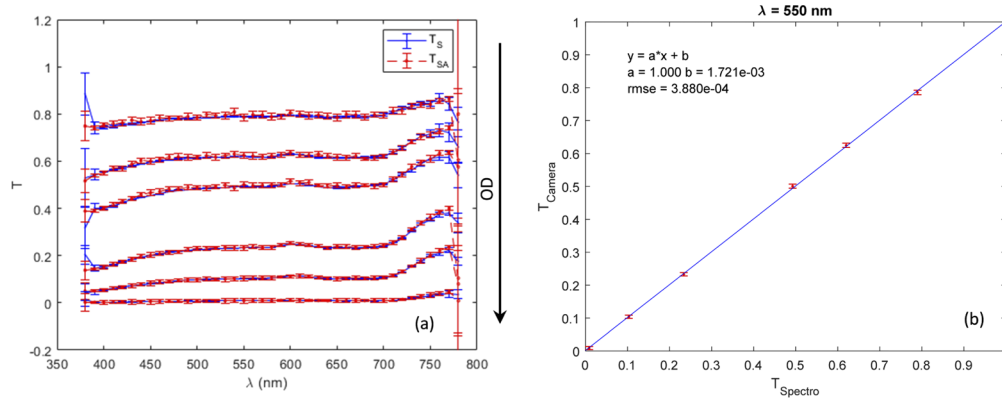
Fig. 4. (a) Comparison of T_S, the transmittance spectra measured with a spectroradiometer, and T_{SA}, the transmittance measured with the camera (spatial average over the image) for a set of KW gelatin neutral density filters with $OD = [0.1,\ 0.2,\ 0.3,\ 0.6,\ 1.0,\ 2.0]$. T_{SA} and T_S overlap within the error bars (coverage factor $k = 2$) for $\lambda = 390$ nm *to* 770 nm for most ND filters; (b) T_{SA} versus T_S fitted with a linear model for $\lambda = 550$ nm.

As an example, Fig. 4(b) presents the results of the linear interpolation at $\lambda = 550$ nm, with a slope $a = 1.000$ and an intercept $b = 1.721 \times 10^{-3}$ for a root mean square error (rmse) of 3.880×10^{-4}.

Fig. 5. The CIE 1931 color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ the relative cumulative weight.

Table 1 presents the CIELAB coordinates results (L^*, a^*, b^*) for the set of neutral density filters, their uncertainties and the median value of the statistical distributions of ΔE^*_{ab}, $\mathrm{med}_{\Delta E^*_{ab}}$, obtained for each sample. One should note that despite the proximity of the (L^*, a^*, b^*) values issued from T_{SA} and T_S, there is not always an overlap within the error bars at $k = 2$. For $OD = 2.0$, despite the relative proximity of the CIELAB coordinates mean values their large uncertainties explain the larger $\mathrm{med}_{\Delta E^*_{ab}} = 3.51$ obtained for this sample for which the T_{SA} and T_S values are close to zero over the detection wavelengths. Figure 6(a) presents the transmittance data for $OD = 0.3$ and shows that T_{SA} and T_S overlap within the error bars at all wavelengths. For $\lambda = 780$ nm the overlap is explained by the large uncertainty on T_{SA}. Figure 6(b) (c) and (d) present the CIELAB coordinates and 95% confidence regions issued from T_S and T_{SA} measurements and

their uncertainties in the (a^*, L^*), (b^*, L^*) and (b^*, a^*) projection planes, respectively. There is no overlap between the 95% confidence regions apart from in the (b^*, a^*) projection plane which points to a systematic error on the color coordinates. However, the median value of ΔE^*_{ab} ($\mathrm{med}_{\Delta E^*_{ab}} = 0.68$) is relatively small and the agreement between the color coordinates issued from T_{SA} and T_S is considered reasonable.

Table 1. CIELAB coordinates for the KW gelatin neutral density filters as measured by the spectroradiometer, T_S, and the spatially averaged images acquired by the camera, T_{SA}. 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 $\mathrm{med}_{\Delta E^*_{ab}}$ of these distributions are reported. The uncertainties are presented with a coverage factor $k = 2$.

| OD	Transmittance	L^*	a^*	b^*	$\mathrm{med}_{\Delta E^*_{ab}}$
0.1	T_S	91.15 ± 0.05	-0.46 ± 0.16	1.46 ± 0.14	0.41
	T_{SA}	91.41 ± 0.16	-0.63 ± 0.49	1.40 ± 0.46	
0.2	T_S	82.83 ± 0.05	-0.55 ± 0.16	2.06 ± 0.14	0.46
	T_{SA}	83.09 ± 0.17	-0.62 ± 0.48	1.84 ± 0.45	
0.3	T_S	75.53 ± 0.06	-0.93 ± 0.19	3.81 ± 0.16	0.68
	T_{SA}	75.85 ± 0.17	-0.61 ± 0.50	3.34 ± 0.50	
0.6	T_S	55.39 ± 0.10	-1.04 ± 0.30	5.78 ± 0.26	0.50
	T_{SA}	55.61 ± 0.24	-1.13 ± 0.69	5.92 ± 0.70	
1.0	T_S	38.37 ± 0.17	-0.84 ± 0.50	6.47 ± 0.45	0.75
	T_{SA}	38.52 ± 0.39	-0.75 ± 1.13	6.39 ± 1.19	
2.0	T_S	7.87 ± 0.86	-0.46 ± 2.55	5.77 ± 2.11	3.51
	T_{SA}	8.06 ± 1.95	-0.53 ± 5.72	5.59 ± 5.48	

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, s_R^2, of sample #32 is used for all KW color filters but for sample #47 to compute the total type A variances, u_A^2, on the measured transmittances. Figure 7 shows that for all samples T_{SA} and T_S curves overlap within the error bars at all wavelengths. Again, for $\lambda = 780$ nm the overlap is explained by the large uncertainty on T_{SA}.

The CIELAB coordinates, their uncertainty and the median value, $\mathrm{med}_{\Delta E^*_{ab}}$, of the statistical distributions of ΔE^*_{ab} obtained for each KW color filters are presented in Table 2. Again, despite the proximity of the (L^*, a^*, b^*) values issued from T_{SA} and T_S, there is not always an overlap within the error bars at $k = 2$.

As an illustration of the uncertainty analysis for the KW color filters, Fig. 8(a) presents the T_S and T_{SA} spectra with their uncertainty for filter #32 (magenta) and shows that the largest discrepancies indeed occur at $\lambda = 380$ nm and 780 nm. The 95% confidence regions around the CIELAB coordinates derived from T_S and T_{SA} and their uncertainties presented in Fig. 8(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, $\mathrm{med}_{\Delta E^*_{ab}}$ ranges from 0.74 to 2.31 and the agreement between the color coordinates issued from T_{SA} and T_S is considered reasonable.

Based on measurements by the reference spectroradiometer, the gamut of the color filters composing the color phantom is represented in the CIELAB space along with their color appearance (Fig. 9(a)). Figure 9(b) is a boxplot of the corresponding ΔE^*_{ab} values computed from T_S, T_{SA} and their uncertainties. A boxplot displays the distribution of data using a five-numbers
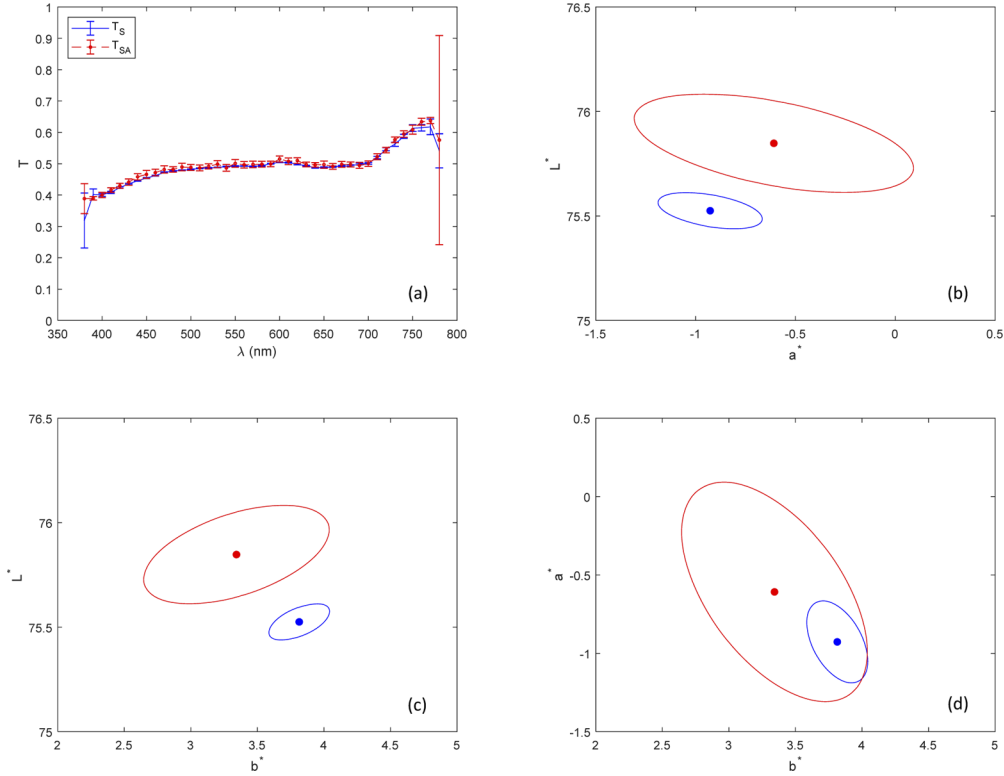
Fig. 6. (a) Comparison of T_S, the transmittance spectra measured with a spectroradiometer, and T_{SA}, the transmittance measured with the camera (spatial average over the image) for KW gelatin neutral density filter $OD = 0.3$ (error bars with a coverage factor $k = 2$); (b) (c) and (d) CIELAB coordinates and 95% confidence regions from T_S (blue) and T_{SA} (red) in the (a^*, L^*), (b^*, L^*) and (b^*, a^*) projection planes, respectively.

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, T_S, and the spatially averaged images acquired by the camera, T_{SA}. 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, $\mathrm{med}_{\Delta E^*_{ab}}$, of these distributions are reported. The uncertainties are presented with a coverage factor $k = 2$.

| Filter	Transmittance	L^*	a^*	b^*	$\mathrm{med}_{\Delta E^*_{ab}}$
12	T_S	88.27 ± 0.03	2.92 ± 0.11	134.72 ± 2.85	2.23
	T_{SA}	88.22 ± 0.12	2.52 ± 0.40	136.13 ± 5.07	
25	T_S	43.81 ± 0.05	80.74 ± 0.18	74.74 ± 3.32	2.31
	T_{SA}	44.11 ± 0.13	80.48 ± 0.30	75.23 ± 5.60	
32	T_S	42.47 ± 0.05	98.33 ± 0.19	-61.85 ± 0.22	1.21
	T_{SA}	42.74 ± 0.14	97.93 ± 0.30	-60.76 ± 0.48	
47	T_S	17.03 ± 0.09	70.94 ± 0.38	-87.37 ± 0.20	0.74
	T_{SA}	16.66 ± 0.08	71.57 ± 0.54	-87.35 ± 0.34	
58	T_S	57.02 ± 0.06	-89.90 ± 0.43	67.52 ± 1.68	1.45
	T_{SA}	56.79 ± 0.13	-90.49 ± 1.06	67.50 ± 2.85	
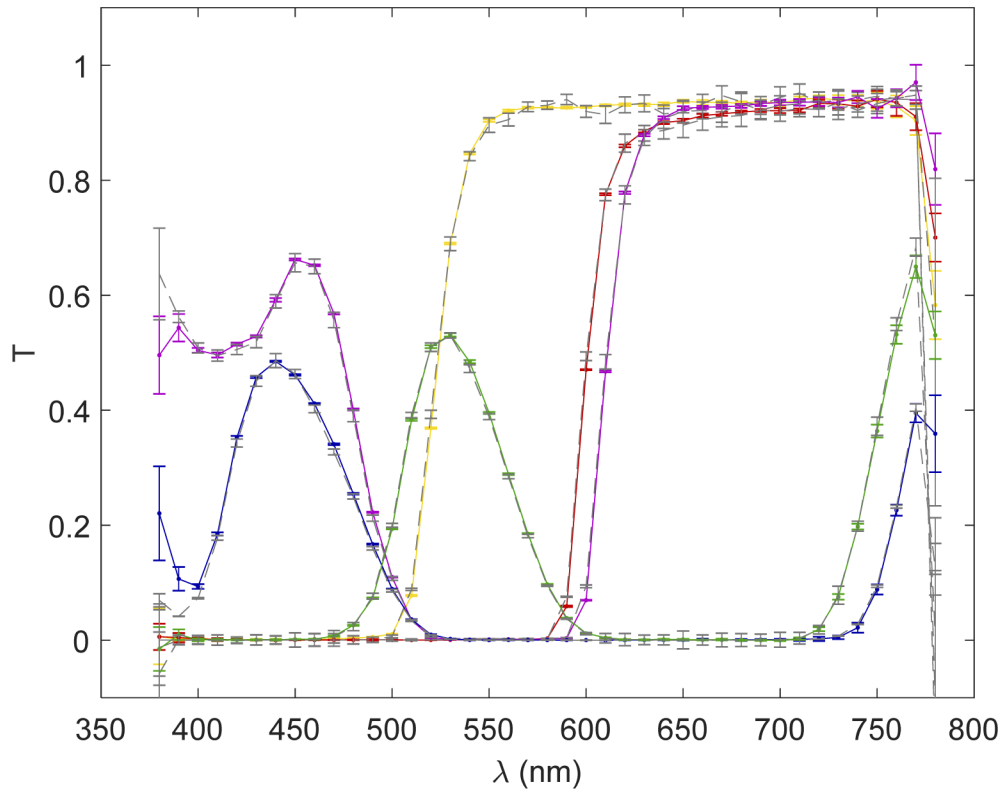
Fig. 7. Transmittance spectra of five KW color gelatin filters (#12: yellow; #25: red: #32: magenta; #47: deep blue: #58: green) measured by the spectroradiometer (T_S: color, plain) and the camera (T_{SA}: grey, dash). T_{SA} and T_S overlap within the error bars (coverage factor $k = 2$) for $\lambda =$ 390 to 770 nm.

summary. The upper limit of the box is the first quartile Q_1, the bottom is the third quartile Q_3 with the median values Q_2 set between these two limits. The "minimum" and "maximum" values are defined as $Q_1 - 1.5 \times IQR$ and $Q_3 + 1.5 \times IQR$, respectively, with the interquartile $IQR = Q_1 - Q_3$. They respectively correspond to the minimum and maximum values excluding any outliers which are represented as red crosses in the boxplot. Median value of ΔE_{ab}^* 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 T_S and T_{SA} measurements. The phantom will be used provide traceability of the measurements by the hyperspectral microscope.

Figure 10 presents an example of a tissue image (H&E stained, cancer adjacent cervix tissue, CNS801 B10, US Biomax, Inc., Derwood, MD, USA) obtained by the HIMS. The transmittance values are measured from $\lambda =$ 380 nm to 780 nm at the pixel level and the CIEXYZ coordinates are computed for each pixel. The sRGB coordinates are computed from the CIEXYZ coordinates for each pixel [15].
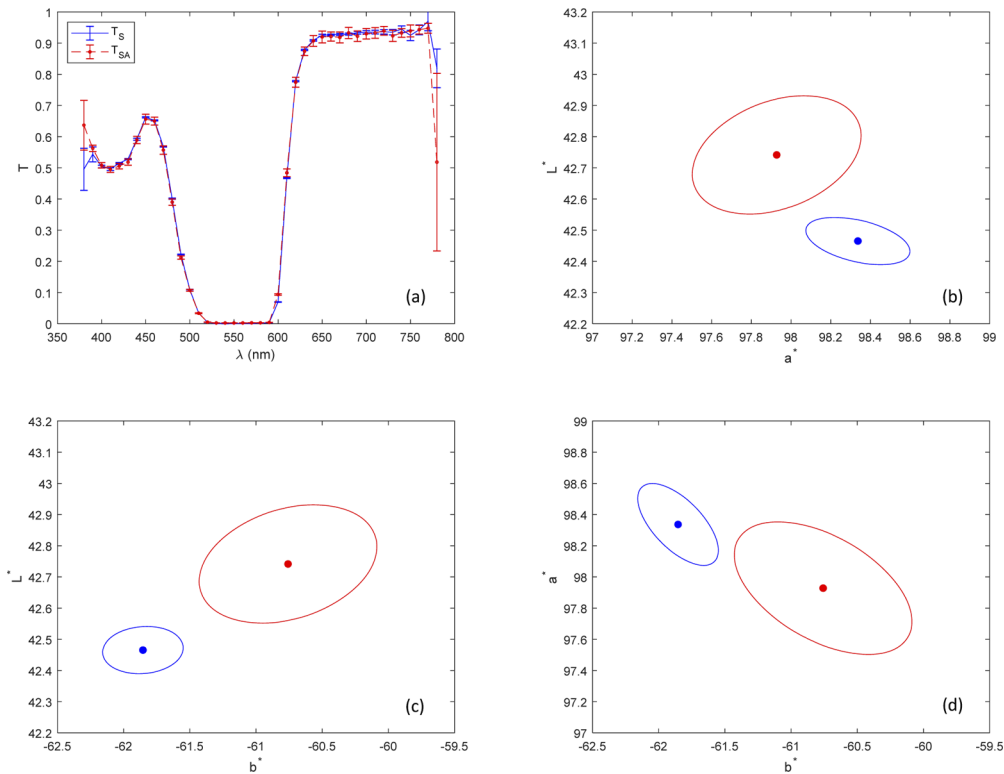
Fig. 8. (a) Comparison of T_S, the transmittance spectra measured with a spectroradiometer, and T_{SA}, the transmittance measured with the camera (spatial average over the image) for KW color gelatin filter #32 (magenta) (error bars with a coverage factor $k = 2$); (b) (c) and (d) CIELAB coordinates and 95% confidence regions from T_S (blue) and T_{SA} (red) in the (a^*, L^*), (b^*, L^*) and (b^*, a^*) projection planes, respectively.
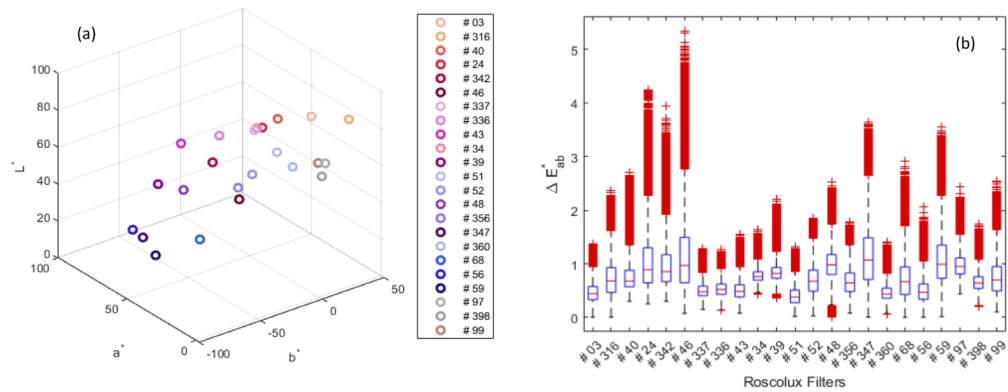
Fig. 9. (a) CIE LAB representation of the 23 patches composing the color filter phantom using measurements by the reference spectroradiometer; (b) Corresponding ΔE^*_{ab} issued from T_S (spectroradiometer) and T_{SA} (camera) measurements.
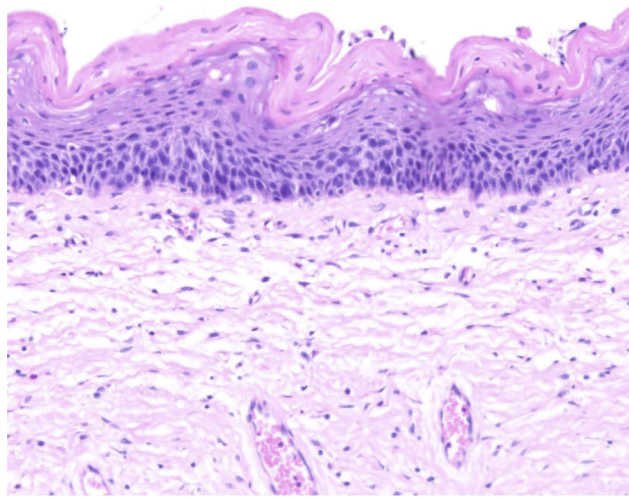
Fig. 10. Image of a H&E stained tissue (H&E stained, cancer adjacent cervix tissue, CNS801 B10) obtained by the HIMS.

4. Conclusions

Using a HIMS is an effective means of measuring the spectral transmittance of biological tissue slides at the pixel level, which is essential for determining the color truth for evaluating the color performance of WSI systems. WSI systems are medical devices. As a reference instrument, the accuracy and uncertainty of a HIMS need to be quantitatively assessed with an objective method. In this study, a bench test method is described and demonstrated to assess a sample HIMS. A color target consisting of 23 miniature optical filters is crafted to be compatible with standard glass slides to test a HIMS, WSI, or microscope device for traceability. The spectral transmittance of each miniature optical filter is measured with a spectroradiometer as the truth. The errors of the HIMS under test are estimated by computing the Euclidian distance in the CIELAB space, i.e., ΔE_{ab}^*, with a Monte Carlo simulation. The results show that the median value of the sample HIMS are less than 0.75 ΔE_{ab}^* for neutral patches and less than 1.0 ΔE_{ab}^* for chromatic ones. These results are adequate for assessing commercial WSI systems, where the nominal errors are usually greater than 20 ΔE_{ab}^*. Thus, the color performance testing data of a WSI device measured by the HIMS under test are considered validated for regulatory purposes.

Other than HIMS, the described test method can also be adapted to assess the other spectrum-sensing optical instruments such as colorimeters, photometers, and densitometers. Using alternative metrics such as CIEDE2000 to estimate color errors is part of the future work in addition to crafting color targets for the other staining methods.

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**(2), 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 Res. Appl. **43**(2), 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," J. Digit Imaging **28**(1), 41–52 (2015).
5. 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).
6. A. R. Robertson, "The CIE 1977 color-difference formulae," Color Res. Appl. **2**(1), 7–11 (1977).
7. P. Shrestha and B. Hulsken, "Color accuracy and reproducibility in whole slide imaging scanners," J. Med. Imag. **1**(2), 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. B. N. Taylor and C. E. Kuyatt, "Guidelines for evaluating and expressing the uncertainty of NIST measurement results," NIST Technical Report 1297 (1994).
12. "Guide to the Expression of Uncertainty in Measurement (GUM)–Supplement 1: Numerical Methods for the Propagation of Distributions," International Organization for Standardization (2004).
13. J. Gardner and R. Frenkel, "Correlation coefficients for tristimulus response value uncertainties," Metrologica **36**(5), 477–480 (1999).
14. G. Wübbeler, J. Campos Acosta, and C. Elster, "Evaluation of uncertainties for CIELAB color coordinates," Color Res. Appl. **42**(5), 564–570 (2017).
15. M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes, "Proposal for a standard default color space for the internet—srgb," in *Color and imaging conference* (Society for Imaging Science and Technology, 1996), pp. 238–245.