Evaluation of Targets for Color Calibrating Digital Images from an Optical Bright-Field Transmission Microscope

Hong Wei, Michael H.Brill,* Taeyoung Park

Datacolor, Inc., Lawrenceville, NJ

Received 27 August 2014; revised 28 October 2014; accepted 10 November 2014

Abstract: Color calibration of digital microscope images of stained biological slides may use a special slide (called a calibration slide) with known color samples that are usually filters. To evaluate quantitatively how well the calibration has been done by a particular system, another set of different known samples is needed, with transmittance spectra that are typical of stained biological material. A slide with such colors comprises what is called a reference slide. In this article, a method is described for selecting the colors for such a slide. A reference slide was created by this prescription and then used to evaluate two embodiments of a color management system for several microscope and illumination combinations. It was found that a set of 20 reference colors spans the gamut of colors produced by a particular kind of stain. Using the reference slide, Datacolor's ChromaCal™ system performance was quantitatively evaluated. The performance is robust to change of microscope and illumination and does not degrade much when a tungsten light is increased in intensity (hence in correlated color temperature) with only one initial calibration at a single color temperature followed by automatic white balancing of a traditional sort. Col Res Appl, 40, 577–584, 2015; Published Online 24 November 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/col.21932

Key words: color management; calibration; microscopy; white balance

INTRODUCTION

In this article, we will describe our recent work to assess color management systems for digital images from bright-field transmission microscope images. A bright-field transmission microscope is familiar from school biology laboratories: the sample illumination is transmitted white light (e.g., illuminated from below and observed from above), and features in the sample are made visible by absorption of some of the transmitted light in dense areas of the sample.

Digital microscopy is an evolving new technology that combines traditional optical microscopy and the digital camera. The recorded digital images can be stored, processed, analyzed, and distributed much more easily than the traditional film images. To maintain the integrity of these images, varieties of the standard procedures and targets have been implemented. For example, stage micrometers are used to calibrate the scale of the image. However, a missing link is the color integrity of these digital images as the recorded color of the same specimen can vary from one digital microscope system to another. Some pioneering studies have been done to color calibrate digital images from bright-field transmission microscopes using a set of targets with known transmission spectra and color coordinates.[1–4] These colors were selected to match their transmission spectra with those of the specimens under test. A common problem of these color targets is that they are too big to be captured all together in the typical field of view of a digital microscope system and therefore require a tedious and lengthy calibration process. It is a huge technical challenge to manufacture the color targets at such a small scale (usually at 100-μm scale) with preferred transmission characteristics.

Recognizing the need for a solution to these problems, we developed a new color target set for color calibrating optical bright-field transmission microscope images. The target set contains 20 colors arranged as a 5-by-4 matrix. Among the 20 colors, four of them are neutral colors with flat transmission spectra in the visual range. The remaining 16 are highly saturated colors with narrow-band transmission spectra equally spaced between 400

*Correspondence to: Michael H. Brill (e-mail: mbrill@datacolor.com)

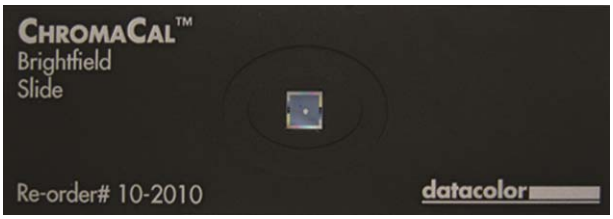

Fig. 1. Calibration slide.

and 700 nm. The sizes of the color targets are engineered so that the entire 5-by-4 matrix can be imaged in the field of view of typical microscope objectives. The color targets are deposited on a 4.5 cm × 4.5 cm glass chip using a state-of-the-art microfabrication technique with process control that ensures high reproducibility. The glass chip is mounted over an opening on a 3 inch × 1 inch metal slide. The assembly (called a calibration slide) is used to transform camera values to tristimulus values by a matrix technique.

How can such a calibration slide be tested once it is developed? A strategy that seems possible is to make a test slide with biologically stained specimens, acquire a color image of the slide, and compute (pixel by pixel) the color difference between the actual *XYZ* values in the stained specimen and the *XYZ* values of the color-calibrated image. However, there is a technical hurdle: the actual *XYZ* values at a pixel are unknown because a pixel-by-pixel measurement of the transmission spectrum is very difficult to reproduce. Furthermore, biological stains are known to change color with time and temperature, and thus, the measured spectrum at a pixel would likely not match the spectrum that determined the camera values later on.

To avoid the problems of single-pixel spectrum measurement of an unstable material, we manually prepared a glass slide with attached squares of filters selected from a database of Rosco® gels. Each of the selected filters has a color typical of a common biological stain or stain mixture. Because of their large size and stability in time and temperature, it is possible to obtain reproducible camera values from the slide and to use the tabulated values of the transmission spectra (together with the microscope light spectrum) to compute *XYZ* values. At this point, we can perform a valid test of the calibration slide using filters that are not like the filters in the calibration slide but typical of the stains encountered in practical microscopy.

We used this slide (called a reference slide) to test the color calibration incurred by use of the calibration slide. Because the filter array of the reference slide is large, it must be moved by hand to measure the separate filter patches, a time-consuming procedure we do not recommend for on-line microscopy. However, performed off-line in a laboratory, the test gives a way to test color calibration procedure. The research question is whether, although transmission spectra of most of the calibration slide filters do not match those of the specimens under test (such as biological stains), a simulation and an experiment will show sufficient calibration accuracy to meet the basic need of color integrity.

It should be noted that this article is an exploratory feasibility study, and it is expected that future works will explore such questions as completeness of the color gamut on a slide and robustness to repetition of measurement.

BACKGROUND

As in any digital imaging pipeline, almost every step along the digital microscope imaging pipeline imparts color. Some key components that would affect the color are the light source of the microscope, the type of objective, the type of optical filters, the type of camera, data manipulation in the camera software/firmware, and finally the display monitor. As a result, the color of the specimen displayed on the monitor is different from the color of the specimen that would be seen through the eyepiece. Furthermore, as color performance is typically uncontrolled in the image-capturing pipeline, the color will be inconsistent from session to session and from one microscope system to another.

A proper color management solution with a color standard is necessary to maintain the color integrity of the captured images. Generally, such a solution would include an image calibration with standard color targets and the monitor calibration. The image calibration typically would happen after the raw specimen image is taken. The monitor calibration can be done any time before the calibrated image is displayed.

A Color Management System

One such color management solution designed for microscope imaging is the Datacolor ChromaCal® product, which includes a calibration slide for image calibration and a colorimeter for monitor calibration. The monitor calibration is not part of this study and will be discussed elsewhere. The calibration slide (shown in Fig. 1) contains filter sets that constitute calibration colors.

If we zoom in to the center of the slide (Fig. 2), we can see two color arrays with different size. Both have the same 20 colors. The first four columns are narrowband interference filters with their peaks ranging from deep blue to red. The last column comprises neutral

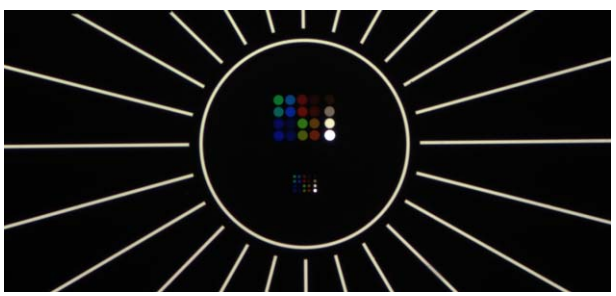

Fig. 2. Zoomed-in view of calibration slide.

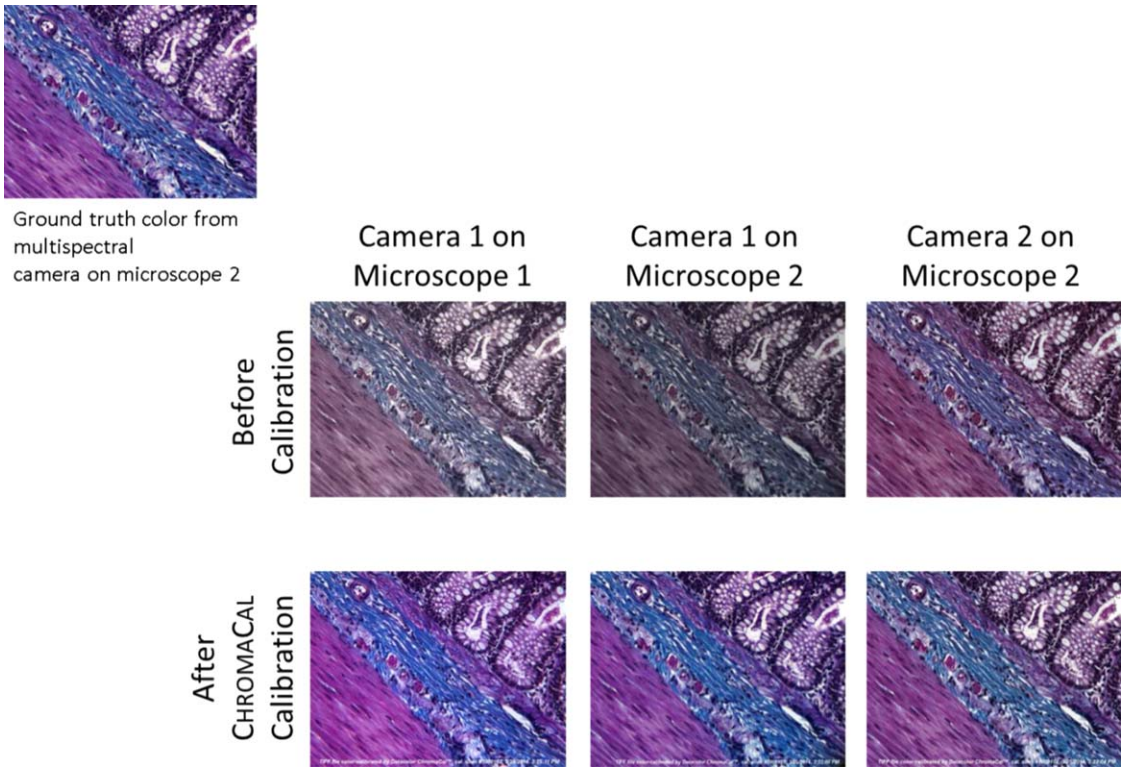

Fig. 3. Example of color calibration.

density filters with transmission from almost zero to almost 100%. The large array at the top has a color patch diameter of 30 μm, and thus, the entire matrix can be captured with the magnification up to 40×. The small array at the bottom has a diameter of 9 μm for magnification up to 100×. With the additional image of the calibration slide captured during a normal microscope session, the specimen images can be calibrated by our image calibration software.

Figure 3 shows an example of the usage of the ChromaCal image calibration. The images are captured from a human muscle slide (product number HD 5-1T) ordered from Triarch. Based on the vendor's description, the region as shown in the images is smooth muscle that was stained with Trichrom stains. The slide was imaged by two microscopes and two cameras to demonstrate the color inconsistency from different systems. Microscope 1 as indicated in Fig. 3 is a Motic AE31 inverted microscope with a Motic LWD PH20× objective and an integrated tungsten light source. This microscope has a condenser with fixed numerical aperture of 0.30. Microscope 2 as indicated in Fig. 3 is a Zeiss Axio Scope A1 upright microscope with a Zeiss N-Achroplan 20× objective. The numerical aperture of the condenser is set at 0.45 to match the objective. A Zeiss 100 W tungsten illuminator is used as the light source. Camera 1 is a Moticam 3 color camera. Camera 2 is a Qimaging QClick Monochrome camera attached with a Qimaging three-band tunable filter (product number 01-RGB-HM-NS, discontinued) for color formation. The color image of the top-most image is rendered by a multispectral camera on Microscope 2 and serves as the ground truth. The multispectral camera is composed of a Qimaging QClick monochrome camera, a liquid crystal tunable filter (LCTF) with 7 nm bandwidth from PerkinElmer, and a relay lens assembly from Channel Systems that projects the images from microscope onto the QClick camera through the LCTF. This multispectral imaging system was calibrated to achieve 1 ΔE_{00} of color accuracy with respect to a NIST-certified Ocean Optics USB4000 spectrometer.

Unsurprisingly, without color management, a mix and match of two cameras and two microscopes gives us totally different colors (see middle row of images). In Fig. 3, the bottom row shows the images after calibration. All the calibrated colors appear to be much closer to the ground truth and are also more consistent between different systems.

Reference Color Slide as a Quantitative Test

The above claims "closer to the ground truth" and "more consistent between different systems" need quantitative support, and such support can come from the use of a reference color slide, as discussed above in the "Introduction" section. Meeting this need is a goal of the Medical Image Working Group (MIWG) of the International Color Consortium (ICC). We internally developed a test methodology including a reference slide with multiple test colors. The goals were to evaluate the ChromaCal performance and to provide the community a tool to test other systems that may arise. Such ideas have been

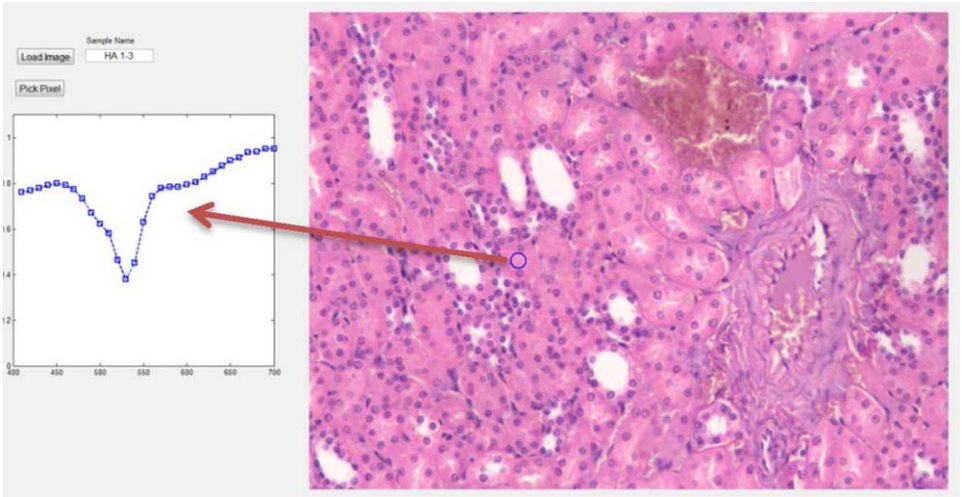

Fig. 4. One of the 22 H&E stained cells, and the transmission spectrum from one of its pixels.

intensively pursued and discussed in the microscopy community. Dr. Yukako Yagi, Massachusetts General Hospital, has helped to develop the concept of a reference slide, the ICC MIWG has adopted it, and the current study advances the concept further.

There are several challenges to achieving our goals. First of all, there are no commercially available test color charts that are designed for microscope images. People have used test colors that were designed for traditional scanning but are not optimized for digital microscopes. Second, we do not have enough knowledge of the colors of the chemical stains. We have not found any publications that provide device-independent color information (such as CIELAB color coordinates) of the stains, particularly of the stains as applied to biological material. (The Biological Stain Commission has paid some attention to colors of raw stains, but not to colors of stained biological specimens.[5]) We would have to know the color gamut (and, actually, the spectral gamut) of the stained biological specimens to find the best test colors. Finally, there are no standard color management solutions for digital pathology, and thus our only guideline is the white paper that ICC MIWG is drafting.[6]

METHOD OF REFERENCE SLIDE CREATION

We used four steps to develop the reference slide:

a. Collect transmission spectra of biological samples with a selected stain or class of stains by multispectral imaging.
b. Compute the CIELAB color coordinates of the sample with the stain(s) using a preferred illuminant, such as CIE D50.
c. Group all the found colors together in the color space and delimit from them the color gamut of the stain(s).
d. Identify test color filters that match or are close to the stain colors.

The basic concept is to build a separate reference slide for each individual stain that will represent all the colors that stain can possibly provide and comprise this slide from a few test colors that are representative of all the possible colors that could emerge by using this stain on biological material. Currently, our activity has been limited to the hematoxylin and eosin (H&E) stain as it is a popular stain. However, the same concept can be applied to other stains.

We need to discuss the color gamut of the H&E stain. One may wonder how a single stain can generate a 3D gamut of color. First of all, the H&E stain is actually two stains mixed together (H&E), and they are taken up differently by different biological materials. Second, the colors of these individual stains are also affected by the biological material that takes them up.

Explanation is also due about the method of determining whether a set of reference colors "covers" the gamut of H&E sampled colors. Our method is to start with a ΔE_{00} threshold grid and to determine the occupation number of image pixels within each box in the grid. The colors from a set of about 200 Rosco gels were selected so as to be in boxes of nonzero occupation number, and then gels that were close to each other in that space were eliminated until 20 colors remained. Now, we are in a position to explain the details of the slide generation.

Step 1: Collection of Transmission Spectra with a Multispectral Camera. First, we constructed a multispectral imaging system on a microscope. The multispectral camera had the following properties: 31 bands, 400–700 nm, 7 nm bandwidth; 12 bits of dynamic range; and pixel format of 1392 × 1040 (1,447,689 transmission spectra). The system was calibrated against a NIST-certified spectrometer to have a color accuracy close to 1 ΔE_{00}.

With this system, we acquired the pixels from 22 H&E stained slides, for a total of 31,848,960 spectra. An example of such a spectrum, from a human kidney slide, is shown in Fig. 4.

Steps 2 and 3: Compute and Plot CIELAB D50/2 Color Coordinates of Each Acquired Spectrum. Then, a total of 31 million color points were plotted in the three-

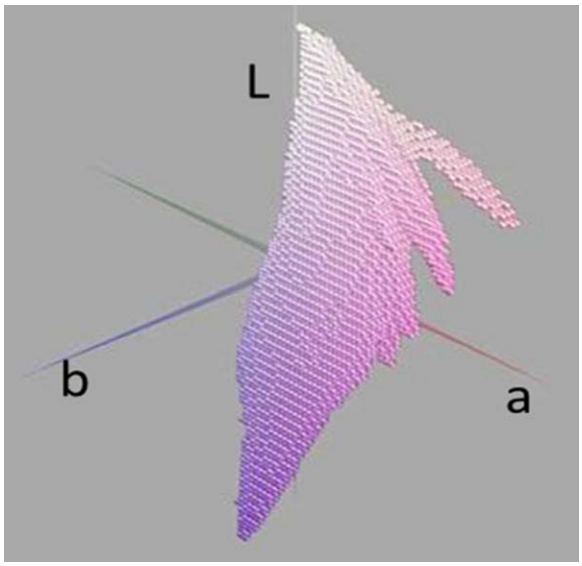

Fig. 5. H&E stain color gamut in CIELAB 50/2° space.

dimensional CIELAB space as shown in Fig. 5. We call the 3D region spanned by these points "the color gamut of the H&E stain." Ideally, the color gamut represents all the colors that H&E stains could possibly provide. The color of each point in the figure is the real color based on its CIELAB values. The blue/purple/red colors as shown conform to our knowledge that H&E stained materials have blue, red, and purple colors. This H&E color gamut that we defined serves as the reference point for finding the proper set of test colors.

Step 4: Identify Test Color Filters That Match or Are Close to the Stain Colors. There are 20 test colors selected in the above way. Figure 6 shows a photograph of the finished slide: a typical glass slide, with 3.8 mm by 3.8 mm Rosco filters applied, and overlaid by a protective cover glass.

These 20 color filters are selected from 231 Roscolux filters to which we have access. The selection rule is that the color of the filter should be located in or no more than three units (ΔE_{00}) away from the H&E color gamut as defined in Fig. 5. A total of 31 filters were selected based on this rule. The number is further reduced to 20 by eliminating some that are very close to each other. The transmissions of the 20 selected Roscolux filters, as shown in Fig. 7, were measured by an Ocean Optics

Fig. 6. Photo of the reference slide.

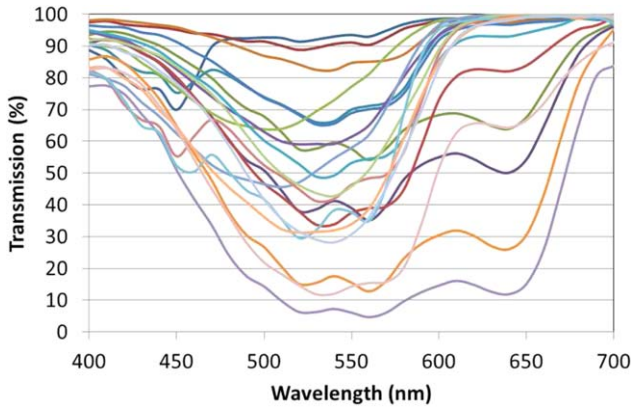

Fig. 7. The transmission spectra of the 20 color filters on the H&E reference slide. A bare glass slide was used as the reference.

USB4000 spectrometer with a laboratory-grade tungsten light source. A bare glass slide was used as the reference. It is worth noting that the number of filters selected is more or less arbitrary as the purpose of this step is to find some colors that are in or close to the H&E color gamut as defined in Fig. 5. It is never a claim that the H&E reference slide would precisely represent the H&E color gamut.

In Fig. 8, the colors are shown as coordinate-plane projections CIELAB, coplotted with projections of the H&E stain gamut. It can be seen that the distribution of reference slide colors is fairly uniform in color space and also covers the gamut of the 31 million plotted colors.

The reference slide we have developed seems to meet all the requirements that ICC recommended for a reference slide to be used to test color performance of microscope digital images.[5] Generally, ICC posted five requirements for such reference slide. First of all, the color should be representative of the color of stains, and this is suggested by Fig. 8. Second, ICC requires that the color should be measurable, and Figs. 7 and 8 show such measurements, and such measurement should be repeatable. It essentially requires that the size of the color patch needs to be big enough to be measured. The third requirement is that the color should be stable. It should not fade significantly over time. The Roscolux filters have this property. ICC also requires that the slide should be serialized for easy identification. Finally, it should be easy to scan the slide on a digital microscope or a whole slide imaging system. As we have shown here, our H&E reference slide meets these requirements.

COLOR MANAGEMENT SYSTEM TEST WITH THE REFERENCE SLIDE

With this reference slide following strictly the ICC recommendations, ChromaCal's performance was evaluated.

First of all, the H&E reference slide is imaged by the same multispectral camera that was used to construct the H&E color gamut. Spatially averaging the monochrome image of each wavelength band gives the transmission

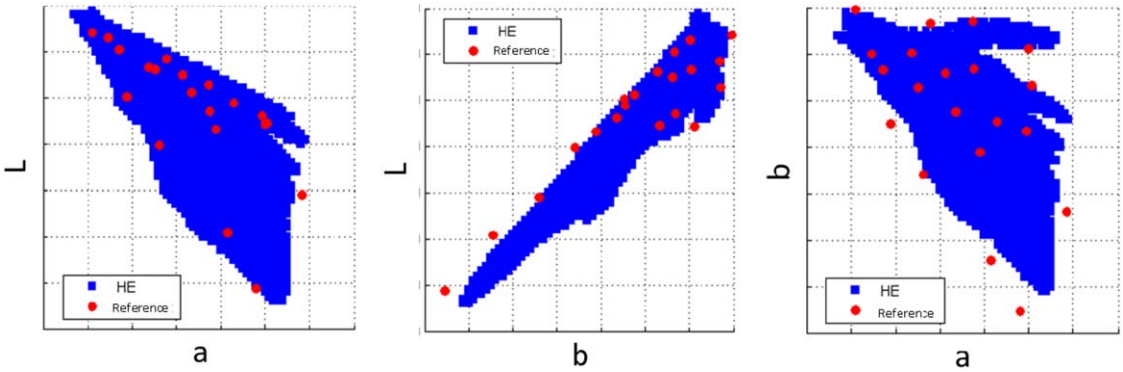

Fig. 8. Color gamut of H&E stain coplotted with 20 reference colors in CIELAB coordinate-plane projections.

spectrum of each test color patch. From the transmission spectra, the CIELAB values of all 20 test colors can be calculated.

The same H&E stain reference slide is then imaged by a regular RGB camera. Averaging each of the 20 raw RGB images gives the *RGB* values of each color. To show what happens in the absence of color management, *RGB* values are converted by camera default to the CIELAB values for each color. The color difference between the true color values and the rendered color values from camera *RGB* values is a metric of the color accuracy of the microscope system. We call the color difference between the true color and the rendered color without any color calibration as Raw_ΔE.

We also process the raw RGB images through ChromaCal image calibration software to calculate the CIELAB values of all 20 colors after color calibration. To do that, an additional image of the ChromaCal calibration slide needs to be taken before loading it and the test color images into the software. One needs to take the image of the calibration slide image at the same settings as that when the test color images were taken. We call the color difference between the true color and ChromaCal calibrated color as Cal_ΔE. The comparison of the Raw_ΔE and the Cal_ΔE can tell us how much of the color accuracy improvement can be gained with ChromaCal image calibration.

Figure 9 shows one example of the color accuracy improvement as a result of ChromaCal image calibration. The microscope system we tested is a Zeiss Axio Scope A1 with a Lumenera Infinity 2-1R camera (traditional Bayer style RGB color camera). The light source is a tungsten halogen with a correlated color temperature of 3000 K. The blue bars are the color accuracies of the 20 test colors from such system without any color calibration. The red bars are for 20 colors after ChromaCal calibration. As can be seen, ChromaCal image calibration improved the color accuracies of all the test colors. Overall, the mean ΔE_{00} of the 20 colors is reduced from 11.5 to 2.9 after calibration, a 75% improvement. The maximum ΔE_{00} of the 20 colors reduced from 14.5 to 9.5, a 35% improvement. The worst color is dark blue #16. It is typical to have poor color performance on a dark color.

As an exploratory study, we added in a Hoya 80A daylight filter to the optical path of the microscope system of the previous example. The 80A filter absorbs almost 80% of the long-wavelength energy and less than 20% of the short-wavelength energy. As a result, the light source color temperature was increased from 3000 to 4830 K, which is very close to the 5000 K that is the color temperature of the D50 illuminant that ChromaCal used as the destination illuminant. This addition improved the postcalibration accuracy (see Fig. 10). The color accuracy is improved in almost all 20 colors. The mean ΔE_{00} improved to 80% (from 2.87 to 2.05), whereas the max ΔE_{00} improved to 67% (from 9.44 to 4.75).

Figure 11 shows additional evidence for the correlation between the light source color temperature and color accuracy after ChromaCal calibration. We compared the color accuracy of a system with a white LED source versus a tungsten source with a daylight filter. These two light sources have very similar color temperatures. One is at 4830 K and the other one is at 5269 K. As in Fig. 11, the color accuracy of the systems with these two light sources is also very similar.

Motivated by these two results, we collected data at other color temperatures. In Fig. 12, the mean ΔE_{00} and max ΔE of the 20 test colors on our HE reference slide

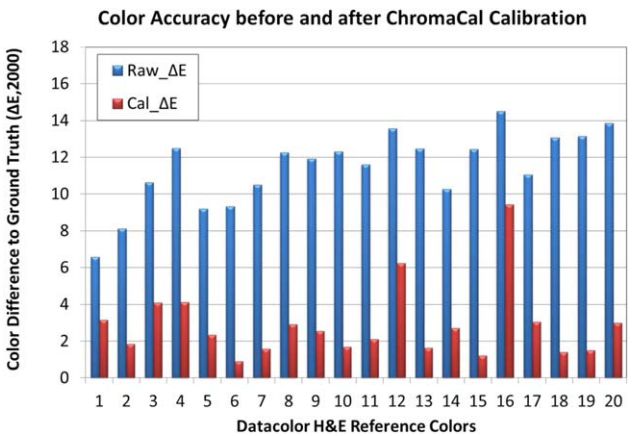

Fig. 9. Color accuracy before and after calibration, as assessed by reference color slide.

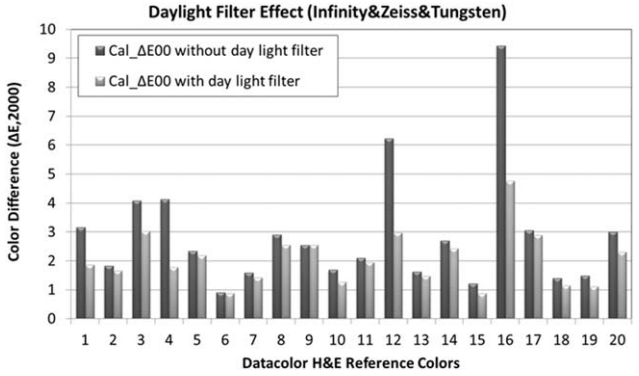

Fig. 10. Effect of imposing a daylight filter on the calibration. The notation "Infinity&Zeiss&Tungsten" means that we used a Luminera Infinity camera, a Zeiss Axio scope A1, and a tungsten-halogen light with correlated color temperature of 3000 K.

after ChromaCal calibration is plotted against light source color temperature. Sources with low color temperature between 2000 and 3000 K are associated with typical tungsten light sources without any filters. The last two points are the re-plot of the result we have shown with daylight filter + tungsten and LED. It is clear to us that light source color temperature plays an important role in the color accuracy after ChromaCal image calibration.

TEST OF AUTOMATIC WHITE BALANCE

Up to now, all the results shown here were generated with the ChromaCal slide image and the reference slide images being taken under exactly the same microscope and camera settings. However, from our beta test, it was realized that most microscopists tend to change settings all the time. The most frequently changed and the most unrepeatable setting is light source intensity, for example, one needs higher light intensity as the magnification is increased. As we know, for the tungsten light source, a different intensity produces a different color. Higher intensities are always bluer than lower intensities. As a result, if the ChromaCal slide image and the reference slide or the specimen slide are taken under different light intensities (hence different color temperatures), the color

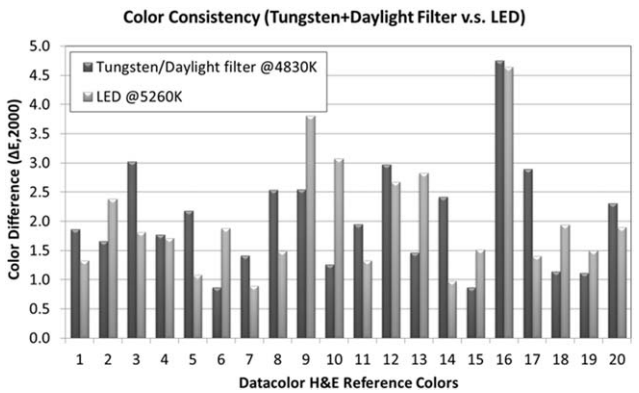

Fig. 11. Color consistency, tungsten/daylight filter, and LED.

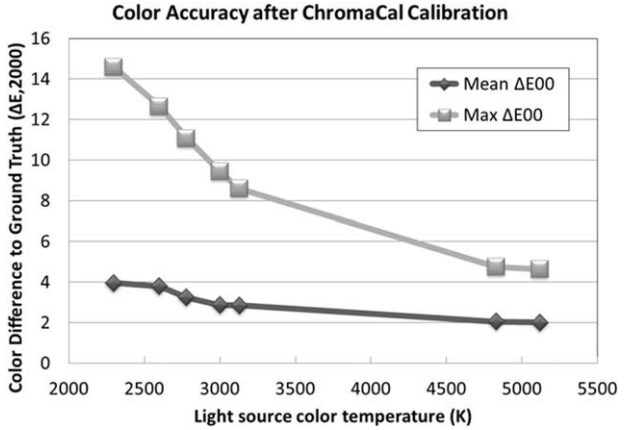

Fig. 12. Mean and max ΔE as a function of color temperature for automatic white balance used to compensate the measurements to a color temperature of about 5000 K prior to ChromaCal calibration.

accuracy will be damaged. To ease the use of the ChromaCal software, we introduced a process that automatically white balances both the ChromaCal slide image and the specimen image, allowing users to take images with different light intensities and different color temperatures without retaking the ChromaCal slide every time the light intensity is changed.

Figure 13 is a demonstration of the automatic white balance (AWB) that we implemented in ChromaCal software. This is the color accuracy of the 20 test colors on the reference slide after ChromaCal calibration. For this test, the test color image and ChromaCal slide image being taken under the same light-source color temperature and the AWB in ChromaCal software is bypassed. However, if we took the test color image under a different color temperature light, the color accuracy will be greatly degraded if the AWB is not used. If we include AWB in the ChromaCal calibration, the color accuracy is largely recovered.

CONCLUSION

We developed a reference slide representative of H&E stained biological specimens, in the form of 20 large

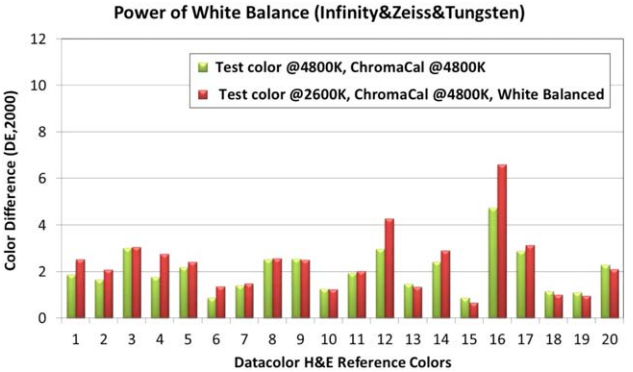

Fig. 13. The performance of calibration under 2600 K with automatic white balance to 4800 K when compared with that of calibration under 4800 K directly. The AWB from a different calibration temperature is nearly as good as a direct calibration.

color patches selected from an extensive catalog of such patches. The reference slide meets the requirements of the ICC MIWG as a test for microscope color management systems. We also evaluated the Datacolor ChromaCal system as an example of use of the new reference slide. Finally, the AWB was found to dramatically improve a match with a color acquired under a different color temperature than the calibration color temperature.

Further improvement of the reference slide is possible in the future, by incorporating more samples with H&E stains and selecting other reference colors that enhance the complicity with the true H&E color gamut. Repeat measurements would also reinforce the statistical robustness of the method described here. Finally, the method might profitably be applied to other stains.

1. Murakami Y, Gunji H, Kimura F, Yamaguchi M, Yamashita Y, Saito A, Abe T, Sakamoto M, Bautista PA, Yagi Y. Color correction in whole slide digital pathology. In: Proceedings of the 20th Color and Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications. IS&T, Springfield VA, 2012. p 253–258.
2. Baxi V, Keay T. Omnyx/GE Healthcare Color Calibration Procedure. ICC Medical Imaging Working Group Meeting, November 2013. Available from: http://www.color.org/groups/medical/MIWG-Vancouver2013.pdf. Accessed Nov 18, 2014
3. Varga VS. Calibration based on IT8.7/1. ICC Medical Imaging Working Group Meeting, November 2013. (See pdf of Ref. 2.)
4. Hulsken B. Calibrating the Philips Slide Scanner. ICC Medical Imaging Working Group Meeting, November 2013. (See pdf of Ref. 2.)
5. Horobin RW, Kiernan JA, editors. Conn's Biological Stains, 10th Edition. Biological Stain Commission; Taylor & Francis, New York, 2002.
6. Revie WC, Jackson P, Cheng WC, Shires M, Yagi Y, McDowell D, Treanor D. International Color Consortium. ICC draft white paper: Digital microscope test materials and test methods, document version 3. Available from: http://www.color.org/groups/medical/Digital_microscope_test_materials_and_test_methods-v3.pdf. Accessed Nov 18, 2014.