

Color calibration of digital images for agriculture and other applications

S. Sunoj^a, C. Igathinathane^{a,*}, N. Saliendra^b, J. Hendrickson^b, D. Archer^b

^a Department of Agricultural and Biosystems Engineering, North Dakota State University, 1221 Albrecht Boulevard, Fargo, ND 58102, USA

^b Northern Great Plains Research Laboratory, USDA-ARS, 1701 10th Avenue SW, Mandan, ND 58554, USA



ARTICLE INFO

Keywords:
Agriculture
Image processing
Phenocam
Terrestrial cameras
Vegetation mapping

ABSTRACT

Image processing in agriculture relies primarily on correlating color changes within images' region of interest to specific quality attributes (e.g., plant phenology, plant health, crop stress, maturity). Changes in lighting conditions during image acquisition affect image color, even though there is no change in quality, and produces misleading inference when used without calibration. The focus of this study was to develop a method for calibrating images to make them homogeneous to improve phenological comparisons. The method was developed with synthetic images and validated with actual plant images in laboratory and field conditions using a standard ColorChecker (X-Rite) chart. Six different color schemes were tested to determine the effect of patch order, and minimum number of patches required for efficient calibration. A user-coded ImageJ plugin named 'ColorCal' was developed in Fiji package for color calibration that derived and applied a $[3 \times 3]$ color calibration matrix, based on selected color patches and standard values. Modified total error and calibration performance index (CPI) were developed to evaluate calibration performance. Calibration using any 12 color patches taken in any order gave equal performance ($0.14 \leq CPI \leq 0.26$). Calibration performance using only commonly followed neutral color patches (e.g., white, gray) was poor ($0.26 \leq CPI \leq 1.0$). Using red (R), green (G), and blue (B) color patches was recommended as it produced visually similar images, the performance was comparable with 24 color patches ($0.21 \leq CPI \leq 0.24$), and was simple and practical. The developed plugin took ≈ 7 s for calibration (Windows laptop, Intel Core i5, and 8 GB RAM). Determining phenological and other applications using the plugin was more reliable than using the raw images.

1. Introduction

Color of the digital image is an important feature in many image analysis applications; however, the calibration of images by users is still a great challenge. An image's true color is vital in obtaining the physical, chemical, and surface characteristics of the object from the image. Images captured using different cameras at the same time produce color variations due to the different spectral responses of the camera sensor, or due to the different digitizers of the cameras. Furthermore, images obtained using the same camera at different times of the day produce color variations, due to either the ambient temperature or the light irradiance at that time, despite no change in the object characteristics (Andresen et al., 2018). Such variations in color tend to critically impact the object's characteristics evaluation when color is considered as the major criterion in the image analysis. The only way to overcome this issue of color variations is to calibrate the pixels color values in the captured image.

Color calibration is broadly classified into two categories: (i) device-based, and (ii) image-based. Many researchers have explored device-

based color calibration in which camera sensors were developed to automatically adjust to the color values based on the scene lighting (Porikli, 2003; Brown et al., 2005). Image-based color calibration is used to compensate for changes in the image due to illumination conditions. This is the most common method in professional photography to maintain the same color tone in all scenes. It used a standard color calibration chart to linearly map the pixel values obtained in the image to the standard color values of the calibration chart (Kagarlitsky et al., 2009). A relationship was developed between the color values obtained in the image to the standard color values, which can then be applied to image pixels for standardization (calibration). Commercially, there were several calibration charts available (ranging from 6 to 100s of color patches); however, the 24 color patch calibration chart was widely used in many applications (McCamy et al., 1976).

Many fields of study, such as space (Levin and Levin, 2004), medical (Haeghen et al., 2000), underwater biology (Akkaynak et al., 2011), and agriculture (Scharf and Lory, 2002) have had to solve the color variation issues while processing images. In some studies using satellite and aerial images, radiometric calibration was performed by linearly

* Corresponding author.

E-mail address: lgathinathane.Cannayan@ndsu.edu (C. Igathinathane).

relating the known reflectance of ground control points (e.g., road, concrete roof, ocean, sand) to the reflectance of corresponding points in the image (Xu et al., 2012). This simplified calibration was performed for individual channels (red, green, blue, and infrared) to obtain the calibrated reflectance. However, comprehensive information on color calibration of images related to agricultural applications is limited. Brown et al. (2012) used hue, saturation, and intensity (HSI) color space and derived a parameter called color strength, which was used to improve color correction. Akkaynak et al. (2014) developed a color calibration using all 24 color patches in the calibration chart by transforming to XYZ color space, and later back-transforming to RGB color space in underwater images.

Over the last decade, application of digital image processing in agriculture has been well researched. The major application in agricultural research was mainly focused on color-based identification/classification (Beeri and Peled, 2009). The usage of unmanned aerial vehicles and phenocam systems in agriculture has increased the use of digital color images to assess plant phenology and health (van Iersel et al., 2018; O'Connell et al., 2015). Because the images are taken in open atmospheric conditions, the brightness and darkness caused by clouds and the time of the day result in wide color variations. Barbedo (2016) reported lighting variation during image acquisition in agricultural fields is the most common issue in image segmentation and feature extraction and it is inevitable. Thus, color-based characteristics gathered from such digital images to assess crop stress, nutrient deficiency, and crop productivity would lead to inaccuracies due to inconsistent lighting conditions.

Researchers generally use various calibration boards with color patches of known color value in agricultural fields. A few studies on calibration boards used in phenological comparisons include white, gray, and black patches in estimating the leaf emergence in a beech forest ecosystem (Ahrends et al., 2008); gray painted board in estimating spring green-up and autumn senescence in a deciduous forest ecosystem (Richardson et al., 2009); and five color boards, such as red, green, black, gray, and beige in estimating biomass and nitrogen levels in a cornfield (Hunt et al., 2005). However, the color rendering efficiency of these color patches was not investigated, and the calibration accuracy and the methodology followed in those studies remain unknown.

The overall research goal of this study was to produce calibrated images from the images acquired under varying lighting conditions. Such calibrated images will serve as better input images for further analysis and interpretation of various plant characteristics (e.g., plant health, phenology). However, there is no standardized method for image-based color calibration. Therefore, this research was undertaken with the following objectives: (i) develop an open source user-coded ImageJ plugin for color calibration of digital images; (ii) develop calibration performance measures to evaluate the calibrated images; and (iii) determine and recommend a minimum number of color patches for efficient color calibration.

2. Materials and methods

Several processing stages to obtain the calibrated image suitable for further analysis are summarized in Fig. 1. In the development process, a synthetic image (two leaves) printed from FreePNGimg (2018) were used to evaluate the effect of color patch order and to determine the minimum number of color patches for calibration. Reproducibility and simplicity were the reasons for using the synthetic image in the method development process. Actual plant images acquired in a laboratory environment (cut cilantro under controlled lighting), and in field conditions (actual plants in the field under natural lighting) were used to validate the method developed.

Images were taken with a digital single-lens reflex (DSLR) camera (Model: Nikon D5100, Nikon Corp., Japan) with the following camera settings: exposure mode = manual, shutter speed = 1/250 s,

aperture = f/7.1, ISO = 250, focal length = 18 mm along with a wireless shutter release remote for hands-free imaging and avoiding camera shake. To calibrate, it is important to include a standard color calibration chart of known color values in the image. We used the standard 24-patch ColorChecker classic target (X-Rite Inc., Grand Rapids, MI, USA) in this study.

2.1. Image acquisition

2.1.1. Synthetic image from internet

An image of the front and back side of two leaves was selected from the FreePNGimg (2018) stock pictures. A color printout of the leaves image along with the original ColorChecker classic target beside was arranged on a white poster sheet for imaging under different lighting conditions, and these served as the synthetic images. Two benchtop fluorescent lights (65 W, 4300 lm, Model: Designers Edge L-2006, Coleman Cable Inc., Waukegan, IL, USA) were used for lighting, and the lighting conditions were varied by choosing one or two lights at a time, as well as by arbitrarily altering the distance between the object and the lights. The combinations of the number and distance of lights used produced the necessary lighting levels ranging from dark (dim) to bright (well-lit) conditions for image capture.

The camera was placed on a tripod at a height of 900 mm facing down, and three images (replication) were captured at each lighting condition. This arrangement ensured crisp images without disturbing the object position for each image while varying only the lighting condition.

2.1.2. Actual plant images from laboratory

Laboratory images of the actual plants were acquired using bunches of cut cilantro (*Coriandrum sativum* L.). The sample was purchased fresh locally and stored in the refrigerator until used for imaging. Before image acquisition, the cilantro bunches were rolled on a paper towel to remove moisture. Bunches were then arranged in a recyclable paper cup so that the top portion of the bunch represented the canopy of the standing plant viewed from the top. The color calibration chart was placed alongside and the images were captured. The camera settings and lighting conditions were maintained the same as in synthetic image acquisition.

2.1.3. Actual plant images from field

Field images were acquired using experimental plots at the Northern Great Plains Research Laboratory, Mandan, ND (46.81°N, 100.92°W). The camera was fitted on a tripod facing north at a distance of ≈ 2700 mm (≈ 9 ft) from the field plot boundary. The color calibration chart was placed near the plot edge and the images were captured.

In general, the lighting variations in the image are due to the intensity of ambient lighting that hits the camera sensor. We simulated the varying lighting conditions by adjusting camera aperture setting (f-stop). Typically in a professional camera (DSLR), the aperture setting varies from f/3.5 to f/22. This range indicates the size of aperture opening, with f/3.5 represents wide opening (more light to sensor—lighter image), and f/22 represents narrow opening (less light to sensor—darker image).

However, the two extreme ranges were not necessary for our application as it produced an overexposed image at f/3.5, and an underexposed image at f/22. Hence we used the aperture range between f/6.3 and f/16, and kept the rest of the camera settings constant (shutter speed = 1/250 s, ISO = 250, and focal length = 18 mm).

2.2. Color calibration process

Images obtained at varying lighting conditions produced variations in the pixel color intensities and resulted in color casts (unwanted tint on the whole or part of the image). The color calibration eliminated such color casts and reduced the variability among the calibrated

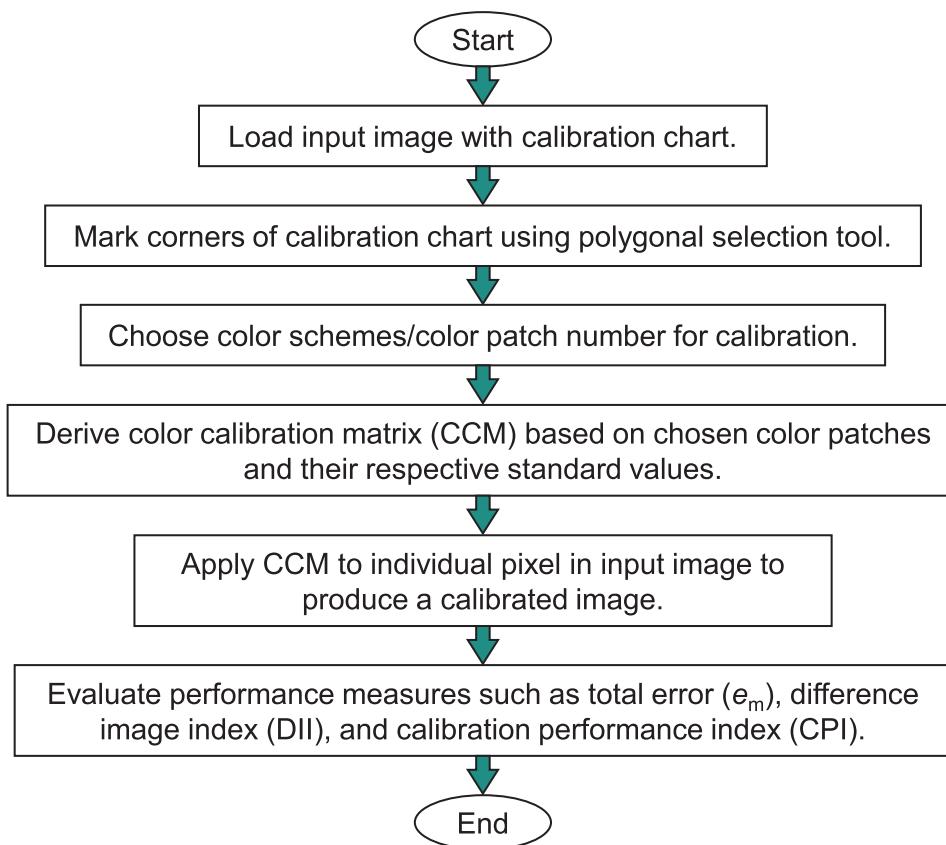


Fig. 1. General process flow of digital image color calibration.

images. The calibrated images represented images acquired at a uniform lighting condition, which are suitable for further color-based analysis.

The color calibration chart included in each input image was used to map the average RGB values of each color patch (of the color chart in the image) to the standard RGB values of the corresponding color patch as supplied by X-Rite (X-Rite, 2018). The calibration essentially means to bring the RGB values of color patches in the input image to the standard RGB values. A color calibration matrix (CCM) was derived based on this mapping and when applied to the entire image produced the calibrated image.

In matrix notation, the process of image color calibration is expressed as:

$$\mathbf{S}_{\text{RGB}} = \mathbf{M}_{\text{CC}} \times \mathbf{I}_{\text{RGB}} \quad (1)$$

where \mathbf{S}_{RGB} is a $3 \times N$ matrix of standard RGB values from N color patches provided by X-Rite, \mathbf{M}_{CC} is the 3×3 CCM to be derived, and \mathbf{I}_{RGB} is a $3 \times N$ matrix of average RGB values from the corresponding N color patches in the input image.

Rearranging Eq. (1) by multiplying the transpose of \mathbf{I}_{RGB} on both sides, and taking inverse determined \mathbf{M}_{CC} as follows:

$$\begin{aligned} [\mathbf{S}_{\text{RGB}} \times \mathbf{I}_{\text{RGB}}^T] &= \mathbf{M}_{\text{CC}} \times [\mathbf{I}_{\text{RGB}} \times \mathbf{I}_{\text{RGB}}^T] \\ \mathbf{M}_{\text{CC}} &= [\mathbf{S}_{\text{RGB}} \times \mathbf{I}_{\text{RGB}}^T] \times [\mathbf{I}_{\text{RGB}} \times \mathbf{I}_{\text{RGB}}^T]^{-1} \end{aligned} \quad (2)$$

where $\mathbf{I}_{\text{RGB}}^T$ is the transpose of \mathbf{I}_{RGB} .

The above equations (Eqs. (1) and (2)) work for any number of color patches (1, 2, 3, ...N) and the resulting \mathbf{M}_{CC} is always a 3×3 matrix. Irrespective of the number of color patches considered, the \mathbf{M}_{CC} need to determine only 9 coefficients (3×3). This leads to underdetermined ($N < 3$) and overdetermined ($N > 3$) systems of equations Eq. (2) and for which the least squares approximation solution approach was followed. Thus, the derived \mathbf{M}_{CC} when multiplied to the individual pixel values (RGB) in the input image produced the calibrated image. Unlike

the simplified radiometric calibration with linear interpolation of individual channels (Xu et al., 2012), this calibration procedure with 3×3 CCM will have interaction among the RGB channels starting from a single color patch.

2.3. Color calibration plugin (ColorCal) development

A user-coded color calibration plugin named ‘ColorCal’ was developed in Fiji (Ver. ImageJ 1.52d, Java 1.6, lifeline version, Rasband (2018)), an open source image processing software. Fiji is a distribution of ImageJ that has core ImageJ applications and extends additional functionality for plugin development (Schindelin et al., 2012). The ColorCal plugin had 925 lines of Java codes was divided into two major sections: (i) semi-automatic color patch selection from the color chart for average RGB values extraction, and (ii) CCM derivation and image calibration (Supplementary material).

2.3.1. Semi-automatic color patch selection

The primary step in a color calibration process extracts the average RGB values of individual color patches of the calibration chart included in the input images. If the camera and calibration chart are placed in a fixed position in a scene, the color values will be extracted by either overlaying a region of interest (ROI) mask on the input image or through the coordinates of ROI of each color patch, which is known and can be hard-coded into the program. However, maintaining an accurate fixed location of calibration chart with respect to the camera is not always feasible.

In our experiments, as the calibration chart was included at different positions in the images, we developed a semi-automatic method for color patches ROI extraction with the least inputs from the users. Using this method, the user selects only the four overall corners of the calibration chart (A, B, C, D; Fig. A.1) in the input image using ImageJ’s polygonal selection tool, and the ROI for each color patch with its

average RGB values (I_{RGB} ; Eq. (2)) were extracted for determining the M_{CC} . The detailed derivation of this developed method is given in [Supplementary material Section A.1](#). It is also possible to automate the color patch selection through feature descriptors (e.g., SURF, SIFT, FAST) and machine learning algorithms; however, that requires additional software to implement the algorithm and will be complicated to integrate in this ImageJ plugin development.

2.3.2. Color calibration matrix derivation in plugin

Matrix manipulations such as transpose and matrix multiplication were available through appropriate methods in the `RealMatrix` class from Apache Commons Mathematics (Ver. 3.6) package ([ASF, 2018](#)). The matrix inverse was calculated using the singular value decomposition method available in Java class `SingularValueDecomposition` and interface `DecompositionSolver`. These methods from the Apache Commons Mathematics were useful in performing the necessary matrix operations of Eq. (2).

The average RGB values of each extracted color patch (I_{RGB}) and the standard RGB values as supplied by the X-Rite (S_{RGB} ; [X-Rite \(2018\)](#)) were stored in respective `RealMatrix` objects (Eq. (1)). These two matrices were solved to obtain the CCM (M_{CC}) as shown in Eq. (2). Each pixel RGB values [1×3] of the input image when multiplied with the derived CCM [3×3] returned the RGB values [1×3] of corresponding output pixel of the calibrated image.

Sometimes, while multiplying CCM with each pixel values of the input image, the resulting calibrated pixel value fell beyond the range of 0–255. Such pixel values produced a false-colored image and were rectified. Reassigning such “out of the range” values to the respective limiting pixel values as 0 and 255, with appropriate plugin code, solved this issue.

2.4. Minimum number of color patches for calibration

Usage of all 24 color patches of the standard color chart for color calibration is expected to produce good calibration. Even though the analysis CPU times were similar (3.27 ± 0.16 s for CCM; and 7.28 ± 0.25 s for overall) among the number of patches ($N = 1$ –24), in other practical applications (e.g., unmanned aerial vehicle images, online color-based sorting system) usage of all 24 color patches for calibration becomes tedious.

Therefore, in the present study, the minimum number of color patches required for color calibration was determined. For this purpose, six different color schemes (Set 1–6; [Fig. 2](#)) were used with the numbers of color patches considered were progressively increased one by one, and their calibration performances were calculated and compared. Choosing a combination of color scheme and the number of color patches results in a unique group of colors available for image calibration.

Out of six color schemes, we considered two color schemes ('Set 1' and 'Set 2'; [Fig. 2](#)) suggested by [Akkaynak et al. \(2014\)](#), and also

developed four new color schemes (Set 3–6; [Fig. 2](#)). The color scheme 'Set 1' was the standard color scheme provided by the X-Rite ([X-Rite, 2018](#)). The color scheme 'Set 2' started with the six neutral colors in the 4th row of the calibration chart ([Fig. A.1 in Supplementary material](#)) followed by the other 18 color patches in the regular order.

The basis for considering new color schemes was as follows: The color scheme 'Set 3' started with the additive colors (R, G, and B), followed by subtractive colors (cyan: C, magenta: M, and yellow: Y), then followed by the regular color patch order. The combination of different ratios of additive colors (R, G, and B) produced a wide color gamut and has been used in many image processing applications, camera sensors, televisions, and monitors. Similarly, the use of subtractive colors (C, M, and Y) produced a wide range of colors as used in commercial printers to produce the same color printed output as seen in the display monitor. Hence, the first 6 colors of 'Set 3' and 'Set 4' were based on the combination of RGB, and CMY color patches, followed by the regular color patch order of the chart.

The first two color patches of 'Set 5' and 'Set 6' were based on the previous studies on phenocam network ([Ahrends et al., 2008](#); [Richardson et al., 2009](#)), white and gray patches were considered primarily, which was same as first two color patches in 'Set 2'. The effect of RGB and CMY color patches after white and gray color patch was observed in 'Set 5' and 'Set 6', respectively. The calibrated images produced by different color schemes were evaluated by different developed performance measures to determine the suitable color scheme and the minimum number of color patches required for efficient color calibration.

2.5. Calibration performance measurement

To evaluate the performance of the calibration process, the resulting calibrated image quality had to be accessed based on the uniformity in color values among calibrated images. Manual observation cannot discern the amount of color differences among the calibrated images. Therefore, along with the available performance measure (total error; [Akkaynak et al. \(2014\)](#)), we have developed a new measure named 'calibration performance index' (CPI), using a 'difference image index' (DII) by comparing two images, for evaluating the calibration performance.

2.5.1. Modified total error

The calibration process alters the RGB values of color patch in the input image (I_{RGB} ; Eq. (1)) closer to the standard RGB values (S_{RGB} ; Eq. (1)). The total error is a measure of closeness that the calibrated image has attained with respect to the standard reference RGB values of the chart. The modified expression of total error ([Akkaynak et al., 2014](#)) after normalization is given as:

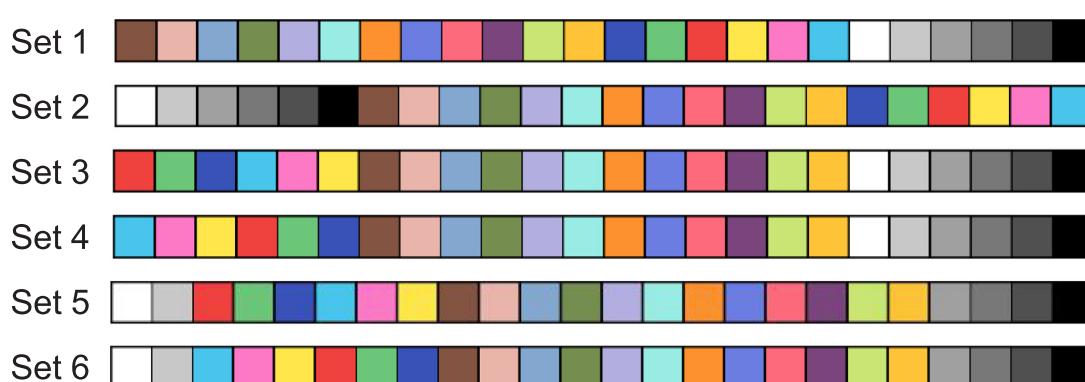


Fig. 2. Order of color schemes tested to determine the minimum number of color patch required for efficient color calibration.

$$e_m = \frac{\sum_{i=1}^{N=24} \sqrt{(R_s - R_i)^2 + (G_s - G_i)^2 + (B_s - B_i)^2}}{N \times (MD_R + MD_G + MD_B)} \quad (3)$$

where e_m is the modified total error (range between 0.0 and 1.0); N is the number of color patches in the ColorChecker and is fixed at 24; R_s , G_s , and B_s are the standard reference values of the individual color patch in ColorChecker (supplied by X-Rite); R_i , G_i , and B_i are the R, G, and B values of corresponding color patch in the calibrated image; and MD_R ($255-8 = 247$), MD_G ($255-52 = 203$), and MD_B ($255-31 = 224$) are the maximum difference of ColorChecker standard reference values obtained from 255 and the minimum value of the 24 patches, for R, G, and B, respectively.

A perfect calibration, where RGB values of all color patches in the calibrated image exactly match the standard RGB values, produced e_m of 0.0, and obtaining e_m of 1.0 is unlikely for actual images as the denominator of Eq. (3) is a constant and values might not represent most natural images. Attaining $e_m = 0.0$ is practically impossible; however, any lower value (close to 0.0) is desirable.

2.5.2. Calibration performance index

The CPI is a developed measure of how well the two input images of varied lighting conditions (dark, and light) are calibrated. CPI was developed to test the calibration performance of different color schemes and the number of color patches selected for calibrating the input images. For this purpose, two synthetic images of extreme practical lighting variation (dark and light) were used as the fixed input, as the results from these included any other intermediate lighting variations.

As images have to be compared for CPI, a method of comparing two images and expressing the difference numerically was developed using a DII. For comparing two images, a 'difference image' was first obtained using the `ImageCalculator()` class of ImageJ, which was the 'bound checked' RGB values (0-255) of pixelwise absolute difference between two images (DI). From this absolute difference image the DII was evaluated as:

$$DII = \frac{\sum_{i=0}^W \sum_{j=0}^H (DI_{R,i,j} + DI_{G,i,j} + DI_{B,i,j})}{W \times H \times 3 \times 255} \quad (4)$$

where DII, the difference image index, is a normalized value ranging between 0 and 1, where 0 indicates no difference between the two images, and 1 indicates the images are totally opposite; i and j are the pixel coordinates in an image; W and H are width and height of the difference image; and DI_R , DI_G , and DI_B are the absolute difference image's R, G, and B values of individual pixels.

For CPI measure, the plugin took 2 input images (dark and light), and generated 4 output images (2 calibrated images with N patches, and 2 reference calibrated images with 24 patches) and evaluated DII between images. The reference calibrated images were obtained using the maximum number of patches ($N = 24$) from Set 1 as they were reported to be the best (Joshi et al., 2005; Ilie and Welch, 2005). The CPI was calculated as a summation of the DII's obtained between images among the calibrated images as:

$$CPI = W_1 \cdot DII_A + W_2 \cdot DII_B + W_3 \cdot DII_C \quad (5)$$

where W_1 , W_2 , and W_3 are the respective weights attached to DII values of DII_A (between calibrated dark and reference calibrated dark images), DII_B (between calibrated light and reference calibrated light images), and DII_C (between calibrated dark and light images), respectively.

From all the combinations of color patch order and number of patches, the maximum value of CPI was evaluated. The maximum CPI value was used to normalize the CPI values to have a range from 0.0 to 1.0. Thus, the best calibration method will have the lowest CPI value and a poor calibration a value towards 1.0.

2.6. Features of the ColorCal plugin

The flowchart of the ColorCal plugin capable of performing (i) calibration, (ii) comparison, (iii) calibration and comparison, and (iv) comparison with reference, based on user inputs through options, is illustrated in Fig. 3. The user loads one or two images based on the desired operation, chooses appropriate option, and selects any of the

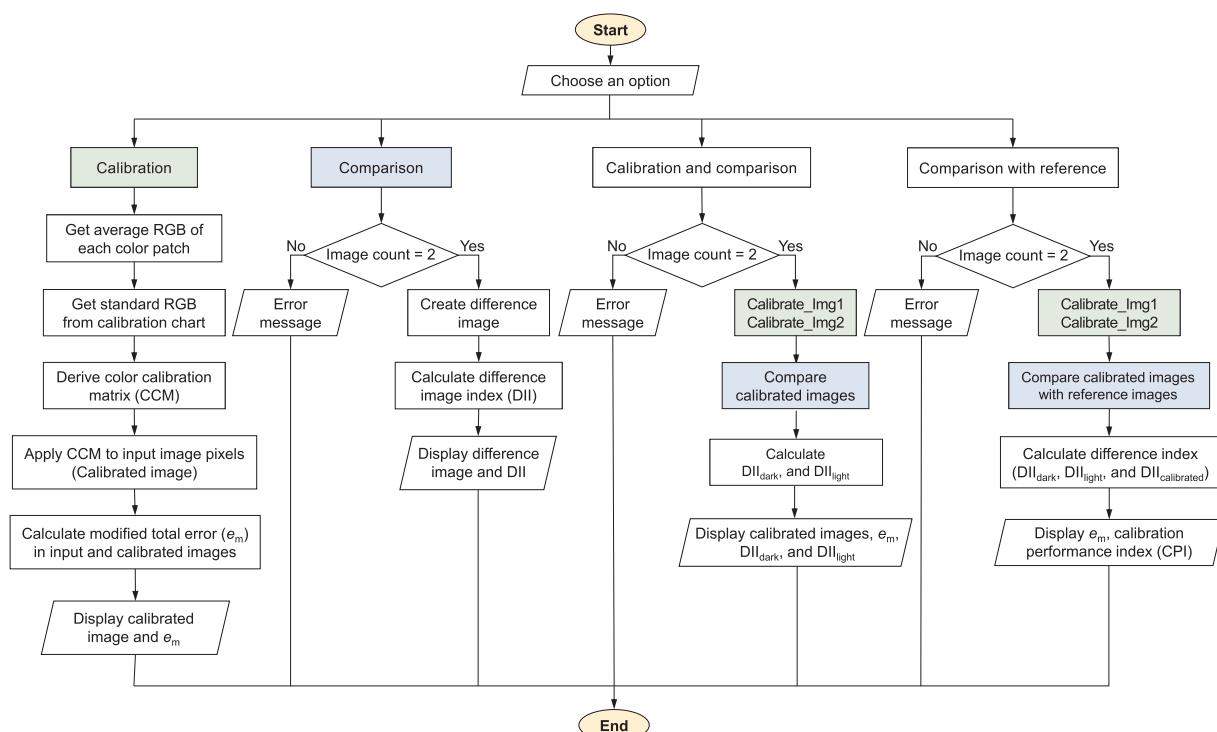


Fig. 3. Flowchart showing various operations performed by the ColorCal plugin from user input options.

predefined ('Set 1'–'Set 6'; Fig. 2) or 'User-defined' color scheme and runs the plugin (Fig. A.2). This user-defined color scheme allows the user to try out different sets, the number of patches, and their order to test or establish efficient image calibration. Further details of the operation and pseudocode of the plugin is provided in the [Supplementary material](#).

As the plugin output, the user can have either the CCM or the calibrated image or both as output by choosing appropriate output options. All the inputs choices will output an appropriate calibration performance measure in the textual form.

If the user chooses the 'Calibration' option, then the plugin takes one input image and calibrates it with the color scheme and outputs e_m performance measure (Eq. (3)) and displays the calibrated image if the output option was chosen. With 'Comparison' option, the plugin takes two input images, creates the difference image and outputs the DII performance measure (Eq. (4)). The comparison can be performed with any two input images, either with raw or calibrated. When both 'Calibration' and 'Comparison' options are chosen, the plugin takes two input images, calibrates both images with the specified color scheme, and produces the DII between the two calibrated images (Fig. 3).

A smaller DII (darker difference image) indicates better calibration performance; however, this will be misleading, especially when two poorly calibrated but similar images will have the smallest DII. To address this issue, the 'Comparison with reference' option (Fig. 3) was included in the plugin that took two input images, produced the calibrated images with the specified color patches (N), also generated the reference calibrated images ($N = 24$ color patches), compared between the calibrated and reference calibrated images, and produced the CPI (Eq. (5)). Thus, the CPI is an overall performance measure which does

not suffer the misleading comparison produced by DII.

2.7. Plugin validation using actual plant images

The ColorCal plugin was developed using the synthetic image and the calibration process was evaluated to determine the best performing color scheme and the efficient number of color patches. The best color scheme and the minimum color patches, selected based on performance measures, were validated with the actual plant images from laboratory and field conditions. The intensity variation among the calibrated images was also studied to verify the image calibration process.

3. Results and discussion

3.1. Plugin outputs and interpretations

The input synthetic images obtained from dark and light conditions, their corresponding calibrated images from the developed plugin; and the difference image between dark and light input images, and dark and light calibrated images are presented in Fig. 4. The shown calibrated images were obtained by choosing 'Set 1' option from the plugin front panel, which selected all 24 color patches for calibration (Fig. 3B). Previous studies (Joshi et al., 2005; Ilie and Welch, 2005) suggested that using 24 color patches in the calibration chart results in the best calibrated image, and is followed in professional photography as well. An example of CCMs derived from dark and light input images, obtained through display CCM option in the front panel (Fig. 3), shows the dominance of the principal diagonal elements and its difference between the dark and light image CCMs:

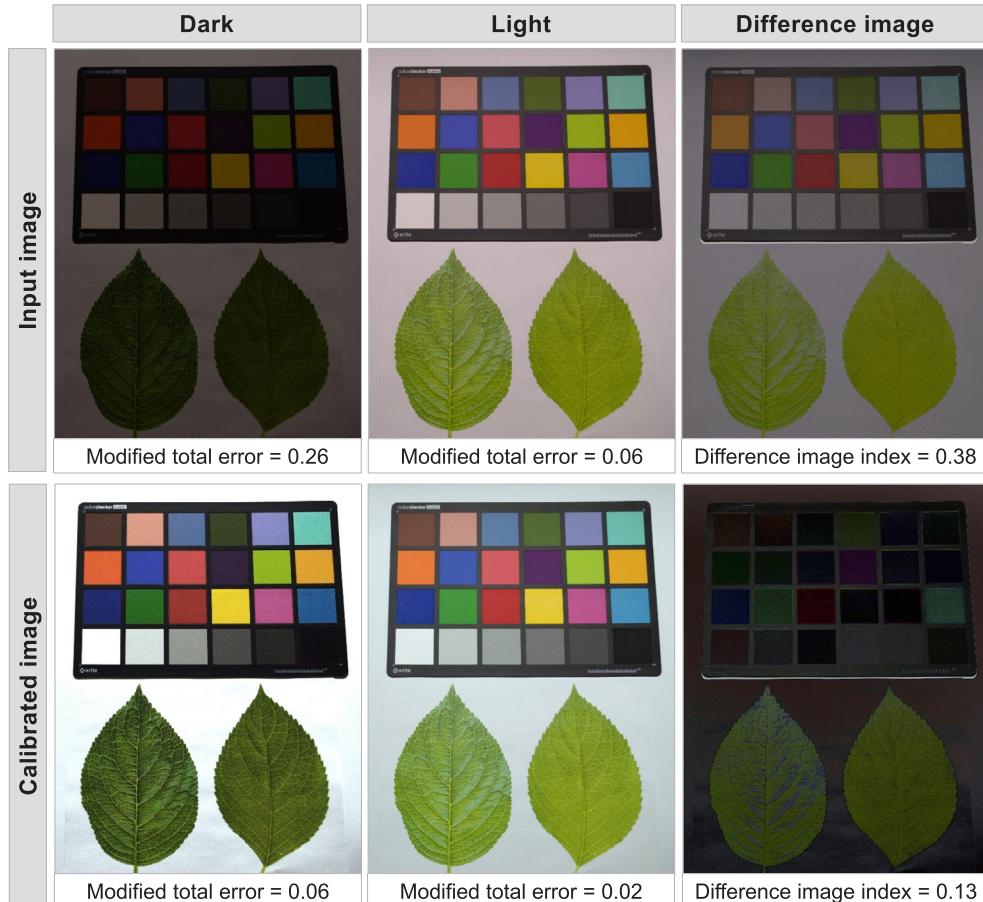


Fig. 4. Synthetic input images acquired in a laboratory at dark and light conditions, respective calibrated images (using 'Set 1'), and the difference image between two input images and two calibrated images.

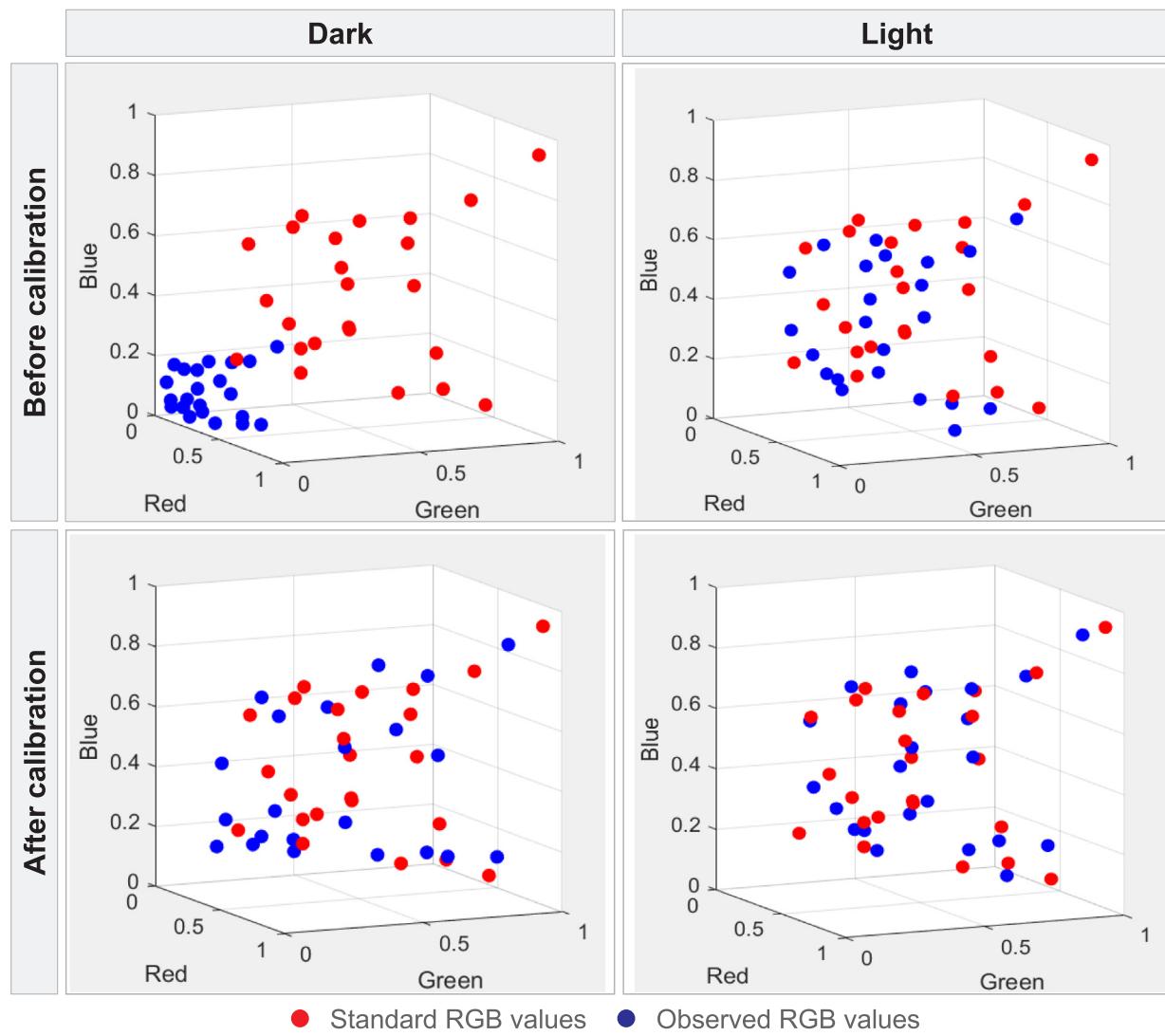


Fig. 5. Scatter plot of normalized RGB values of 24 color patches in the image before and after calibration in dark and light conditions.

$$\mathbf{M}_{CC_dark} = \begin{bmatrix} 1.06 & 0.09 & 0.02 \\ 0.07 & 1.16 & -0.09 \\ 0.09 & 0.08 & 1.02 \end{bmatrix}; \quad \mathbf{M}_{CC_light} = \begin{bmatrix} 0.87 & -0.07 & -0.23 \\ -0.03 & 0.79 & -0.14 \\ -0.06 & -0.13 & 0.83 \end{bmatrix}$$

As seen in Fig. 4, leaf color in the dark input image almost looked black due to the dark lighting and caused e_m of 0.26; however, the calibration process brought the leaf color to green and reduced the e_m to 0.06. Similarly, the e_m was reduced from 0.06 to 0.02 in the light condition. The difference images shown in the third column were obtained by choosing the ‘Comparison’ option after loading the two input images (dark and light), and the two corresponding calibrated images, respectively.

The difference images served as an illustration of the intensity difference between the images; however, the difference images worked only when both images were the same with matching location, which was obtained without changing the camera position. It can be seen that the darker the difference image, the greater was the similarity between input images, and vice versa. The difference image between the two input images was bright (DII = 0.38) because of a wide difference in input lighting conditions, but the difference image between the two calibrated images was dark (DII = 0.13) because the difference was less between the calibrated images. This demonstrates the process of calibration in bringing the input images towards a common intensity.

The calibration process can be better visualized with the 3D scatter plot shown in Fig. 5. Each red marker denotes the normalized RGB

values of each color patch in the calibration chart (standard values), and the blue marker denotes the normalized average RGB values of each color patch in the input image. Before calibration, the RGB values in dark condition were closer to 0, while the RGB values of color patches in light condition were well scattered similar to the standard RGB values. After calibration, the RGB values of color patches in dark condition were closer to the standard values, and in light condition already distributed input values got adjusted and most of the calibrated RGB values overlapped with the standard values.

3.2. Selection of the minimum number of color patches

The variations in e_m (Eq. (3)) for calibrated images obtained from dark and light input images are shown in Fig. 6. There was no difference in e_m beyond 13 color patches in the calibrated light image, while the calibrated dark image required >18 color patches. This was also observed in the 3D scatter plot (Fig. 5), where the light image had well distributed normalized RGB values of each color patches before calibration and the e_m (0.06) was less than dark image (0.26) as they were well separated before calibration. Previous studies (Alsam and Finlayson, 2008; Akkaynak et al., 2014) also followed a similar pattern for color schemes ‘Set 1’ and ‘Set 2’, where the magnitude of e_m was lesser in the light calibrated image than the dark calibrated image.

The color scheme ‘Set 2’ required more color patches (> 9 patches)

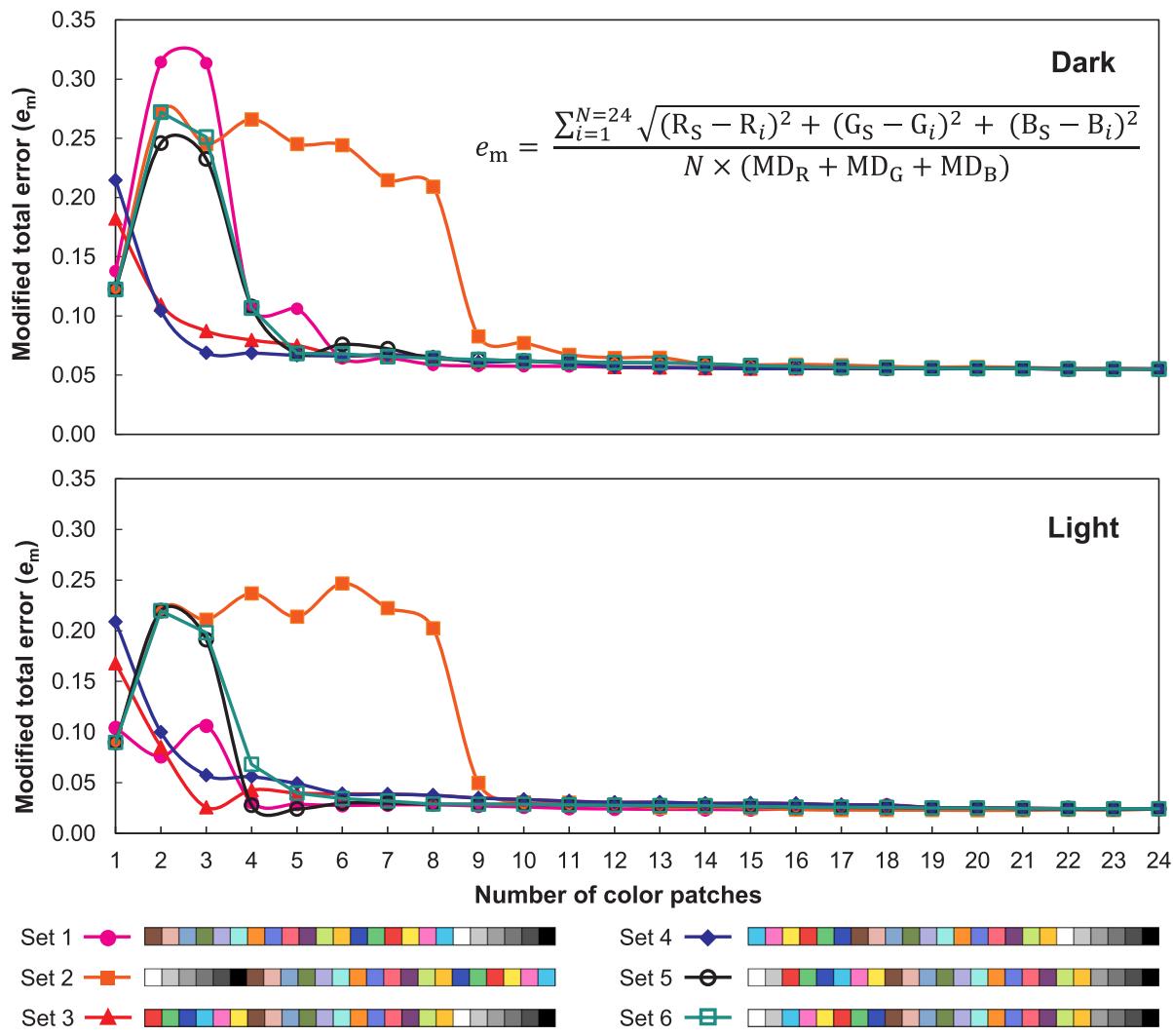


Fig. 6. Variations in modified total error in calibrated images by different color schemes in dark and light conditions (data in Section A.3).

for producing the calibrated image with minimum error, because the first 6 colors (neutral shades) were achromatic, and did not contribute efficiently to the color calibration (Fig. 6). Other than 'Set 2', all other color schemes (Set 1, 3–6) attained the minimum value after the first few color patches (≤ 6). The color schemes 'Set 3' and 'Set 4' were evidently seen as efficient with only 3 color patches producing the minimum error in both dark and light calibrated images compared to other color schemes. In the dark calibrated image, 'Set 4' produced the minimum error with 3 color patches, while in the light calibrated image 'Set 3' produced the minimum error with 3 color patches. However, the calibrated image produced by 'Set 4' using the first 3 color patch produced a greenish color cast and was not suitable compared to the calibrated image by 'Set 3'. Therefore, color scheme 'Set 3' was found efficient in rendering the colors of the image based on e_m .

3.2.1. Comparison of dark and light calibrated images

The CPI was also used in the selection of color scheme and number of patches to be considered in calibration. We found that images calibrated using the color scheme 'Set 1' with all 24 color patches produced the minimum e_m (Fig. 6) and considered as the reference image for CPI calculation in both dark, and light conditions. Additionally, weight factors as used in Eq. (5) is important in deciding the contribution of differences. Even though the difference between the two calibrated images was kept to a minimum, the similarity between the calibrated image (with N color patches) and the reference image (with 24 color

patches) needs to be considered. Therefore, more weights were considered for W_1 and W_2 , and slightly less weight for W_3 . The weights were fixed, after preliminary trials, as 0.35 for each W_1 and W_2 , and 0.30 for W_3 in Eq. (5) to evaluate CPI for each color patch.

The plot of CPI values and number of color patches of different color schemes is shown in Fig. 7. The CPI showed a pattern similar to the e_m performance measure (Fig. 6). The CPI values from all color schemes converged and attained a low CPI when half of the color patches (12) were used. Overall, below 9 color patches, all color schemes showed distinct variations in the CPI values. It was interesting to note that the order of color patches did not have any effect on calibration, i.e., using R, G, B or B, G, R produced identical calibrated images.

The color schemes 'Set 1', 'Set 2', 'Set 5', and 'Set 6' showed lower CPI with a single color patch, but the calibrated images had increased color cast (dark skin and white). This was also observed when only 2 color patches were included for calibration. By including the 3rd color patch, other than the schemes 'Set 3' and 'Set 4', the calibration with the rest of the schemes resulted in a color cast that raised the CPI values (Fig. 7). The grayish color cast with color schemes 'Set 2', 'Set 5', and 'Set 6' was because of the achromatic color patches (white and gray), and with color scheme 'Set 1' was because of skin tones in the starting patches that were inefficient in overall image calibration.

The variations of CPI values for all color schemes with 3, 6, 12, and 24 color patches, indicated in Fig. 7, are presented in Table 1. In all image samples, the CPI values from 'Set 3' and 'Set 4' with 3 color

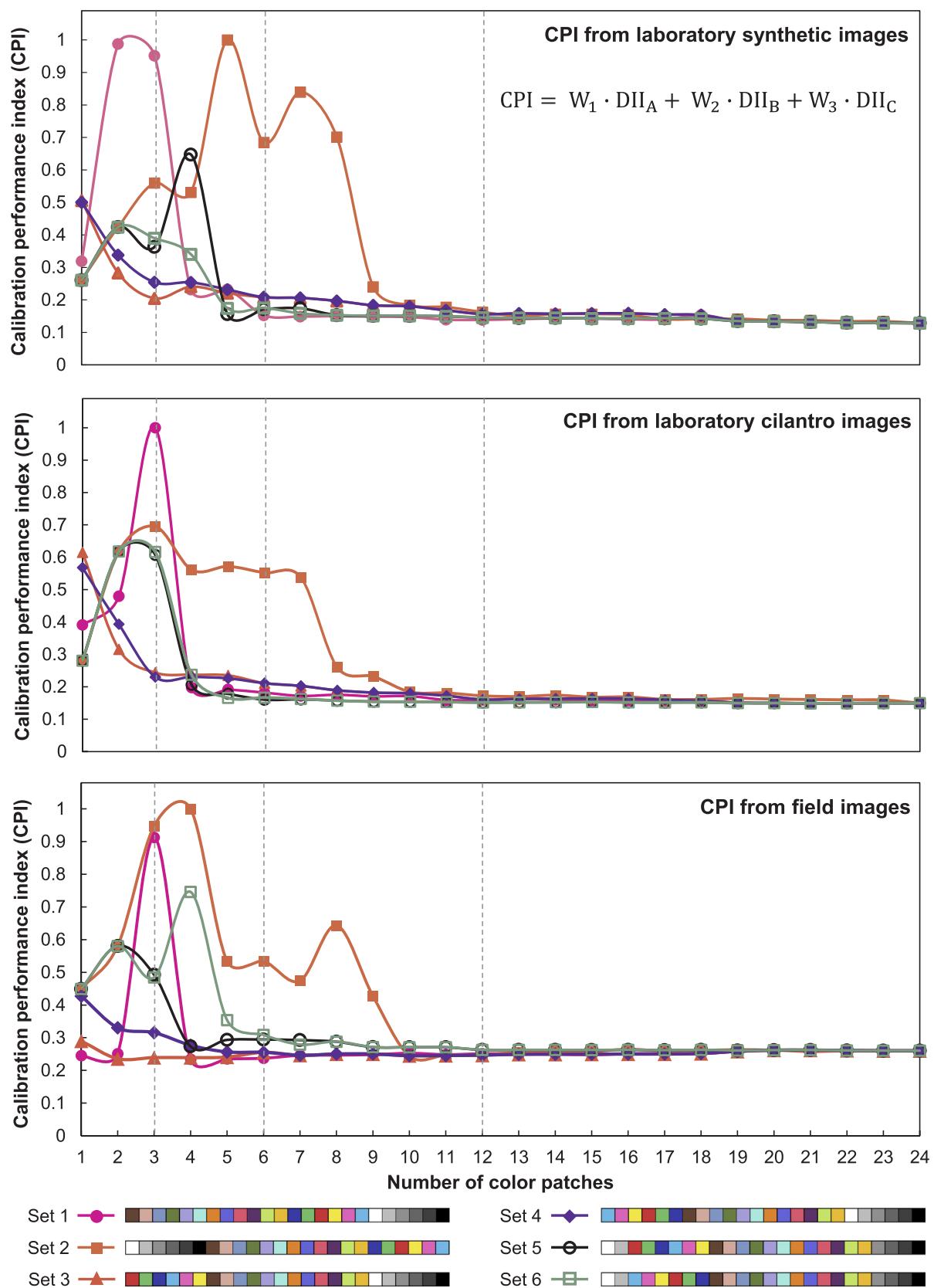


Fig. 7. Variations in calibration performance index (CPI) in calibrated images with the number of color patches by different color schemes (data in Section A.4).

patches were the lowest (Set 3: 0.205–0.244; Set 4: 0.230–0.315) compared to other color schemes. The calibrated images produced from 'Set 3' and 'Set 4' with 3 color patches were visually similar. The next lower CPI values were obtained with 6 color patches in all color

schemes except 'Set 2'. Including additional color patches further reduced the CPI values; however, considering practical applications and simplicity reasons for using color patches in a field setting, using a minimum number of color patches is desirable. Therefore, including 3

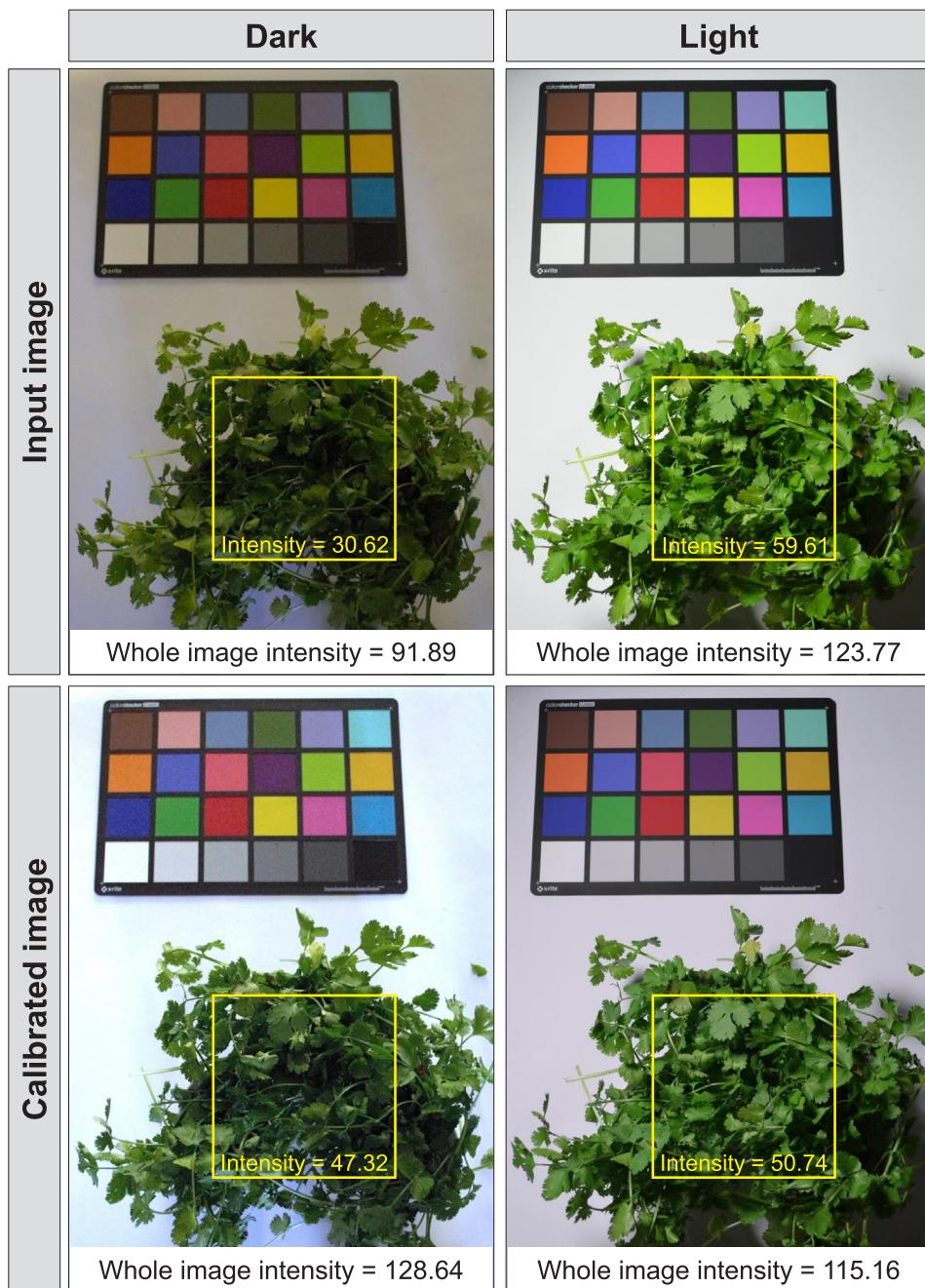


Fig. 8. Cilantro cut plant images acquired in the laboratory in dark and light conditions as inputs and their respective calibrated image obtained using only the three R, G, and B color patches (calibration performance index, CPI = 0.244).

color patches—R, G, and B were recommended for an efficient color calibration as it produced calibrated images that were visually similar and had low CPI values.

3.3. Color calibration plugin validation

3.3.1. Laboratory conditions—Cilantro cut plant

We validated the color calibration method with 3 color patches (RGB) using cut cilantro plant samples in the laboratory under varying lighting conditions (Fig. 8). We observed that with actual samples having three-dimensional shapes, shadows occurred in artificial or natural lighting. However, the shadows have not interfered in extracting of an average color value from the region of interest (ROI) and in the process of image calibration.

The calibration performance of the cilantro images in dark and light conditions had a low CPI value of 0.244 (Table 1), which was good. The DII (Eq. (4)) between dark calibrated image and its dark reference image was 0.12 (DII_{dark}), and the difference index between light calibrated image, and its light reference image was 0.13 (DII_{light}). However, the dark and light calibrated images produced slightly higher difference index of 0.52 ($DII_{calibrated}$). This is expected when two images obtained under widely varying lighting conditions are calibrated and compared.

The calibration process raised the whole image intensity value from 92 to 129 in dark condition (ROI intensity from 31 to 47) and lowered from 124 to 115 in light condition (ROI intensity from 60 to 51). The intensity value of the whole dark calibrated image was higher than light calibrated image because the white background used had become saturated in the calibration process and increased the whole image

Table 1

Calibration performance index (CPI) variations for different color schemes with selected number of color patches.

Image samples	Color scheme ^a	Number of color patches			
		3	6	12	24
Synthetic images	Set 1	0.952	0.152	0.138	0.128
	Set 2	0.560	0.685	0.162	0.128
	Set 3 ^b	0.205	0.208	0.156	0.128
	Set 4 ^c	0.254	0.208	0.156	0.128
	Set 5	0.363	0.171	0.146	0.128
	Set 6	0.390	0.175	0.146	0.128
Cilantro images	Set 1	1.000	0.182	0.159	0.150
	Set 2	0.695	0.552	0.172	0.150
	Set 3 ^b	0.244	0.211	0.161	0.150
	Set 4 ^c	0.230	0.211	0.161	0.150
	Set 5	0.608	0.160	0.152	0.150
	Set 6	0.617	0.167	0.152	0.150
Field images	Set 1	0.913	0.237	0.251	0.260
	Set 2	0.948	0.534	0.247	0.260
	Set 3 ^b	0.239	0.255	0.247	0.260
	Set 4 ^c	0.315	0.255	0.247	0.260
	Set 5	0.492	0.295	0.263	0.260
	Set 6	0.485	0.308	0.263	0.260

These selected CPI values were obtained from the Section A.4 and are also indicated in Fig. 7.

^a Refer to Fig. 2 for details on the color schemes; however, the schemes with lowest CPI are reproduced for ready reference.

^b Set 3: 

^c Set 4: 

intensity. However, comparing the plant canopy and the selected ROI, the color looks more uniform between the dark and light calibrated images than the input images. Therefore, comparison and analysis using the calibrated images would be definitely better than using the raw input images, which had a wider color difference.

3.3.2. Field conditions—Alfalfa and spring wheat

The acquired field images consisted of two crops: alfalfa in the late bud stage that looked green (left), and spring wheat in a mature stage (right) that looked brown (Fig. 9). These images represented a typical field condition for studying the image calibration at two growth stages of field crops. Varying lighting conditions were captured by adjusting the aperture setting from f/6.3 to f/16 that showed a gradual darkening of the image indicated by decreasing image intensities representing the typical lighting variations in field conditions.

In general, in color image processing, a simple ‘white balance’ method is widely followed (Lam, 2005; Richardson et al., 2009), which makes the image white regions look white and correspondingly updating all image pixels. This method calculated a factor that brings an object/section/color-patch of the image from discolored white to perfect white ($R = G = B = 255$). The factor was uniformly applied to all image pixels to produce the white balanced image. This process simply altered all the pixel values only based on white, which is similar to the satellite and aerial images radiometric simplified calibration using linear interpolation of known ground control points reflectance (Xu et al., 2012). The processed images, however, still suffered similar intensity variations as the input images (Fig. 9 - White balance). It is interesting to observe that the white patch (bottom-left patch of the color chart) can be distinctly seen as perfect white in all white balanced images.

The calibrated images using the developed plugin with R, G, and B color patches showed an overall homogeneity across the images (3×3 calibration; Fig. 9), which demonstrated the good calibration performance of the ColorCal plugin. The DII between the calibrated image and their respective reference image ranged between 0.053 (f/16) and

0.280 (f/6.3), with an average of 0.154 and standard deviation of 0.073. The overall CPI evaluated by assuming the image from f/6.3 as light, and f/16 as dark conditions was 0.239 (Table 1). In the field, the CPI value for 3 color patches (0.239) was better than 24 color patches (0.260) indicating that for field images 3 color patches (RGB) were sufficient. Further analyses with other input images may be required to establish this trend. The plugin took ≈ 7 s (CPU time) per input image of resolution 0.68 mm/pixel to produce the respective calibrated image in a laptop PC (Windows 10, 2.20 GHz Intel Core i5-5200U processor, and 8 GB RAM).

All the calibrated images appear to fall toward the darker intensity, but this did not impact the comparison and analysis of the images for any practical applications because they tend to be relatively uniform independent of lighting conditions. However, the actual variation in the crop phenology (growth stages) were actually captured by the raw images, while the images calibrated with ColorCal plugin had eliminated the lighting variation and standardized the visual tone or overall intensity.

The intensity variation plot among these images helped to visualize the effect of white balance and calibration (3 color patches) on the input images from field conditions (Fig. 10) and allowed us to perform further analysis. The steeper negative slopes for the input and white balance images indicated a decrease in intensity values from light to dark.

Ideally, the goal of the calibration process is to make the intensity variations flat (zero slope), while maintaining a desirable intensity. A hypothetical flat line with a lower intensity value (e.g., < 30) indicates an underexposed image, and that with a higher intensity value (e.g., > 160) indicates an overexposed calibrated image, and these are not desirable. The slopes of the fitted linear models for input and white balanced images were -11.7 and -9.8, respectively. The slope for calibrated images was only -2.42, indicating less variation in intensity among calibrated images.

When the left (alfalfa) and the right (spring wheat) image portions were considered separately (Inset; Fig. 10), we found similar variations in slopes of linear regressions (Models; Fig. 10). Slopes for alfalfa and spring wheat ranged between -11.96 and -12.04 for input images, -9.72 and -10.14 for white balanced images, and -2.31 and -2.69 for calibrated images. Although the calibrated images were slightly darker than input and white balanced images, this indicated a more homogeneous “color feel” (intensity). The brightness of calibrated images can be improved through further image processing operations to increase the intensity uniformly, without changing the color value (hue) of the pixels. This forms a potential aspect of future research.

3.4. Suggestions for future work

Future work will consider studying the calibration performances with higher order CCM by including the square terms and interaction terms. Also, chromatic adaptation techniques, such as white balance, gray balance, or RGB equalization before or after color calibration to improve the brightness of the calibrated images, and comparing the results using the developed performance measures should be considered. The color calibration should be explored in other color spaces, such as $L^*a^*b^*$, HSI, and XYZ because these color spaces cover a wide gamut of colors and might produce better results. The performance of the developed method should be compared with the calibration methods offered by commercial software (e.g., Adobe Lightroom, Aperture). Finally, the calibration performance should be tested in images with a wide range of natural variations (brightness and color) in the phenocam network systems.

4. Conclusions

The artificial lighting and camera aperture (f-stop) settings produced a wide range of lighting variation in the laboratory environment

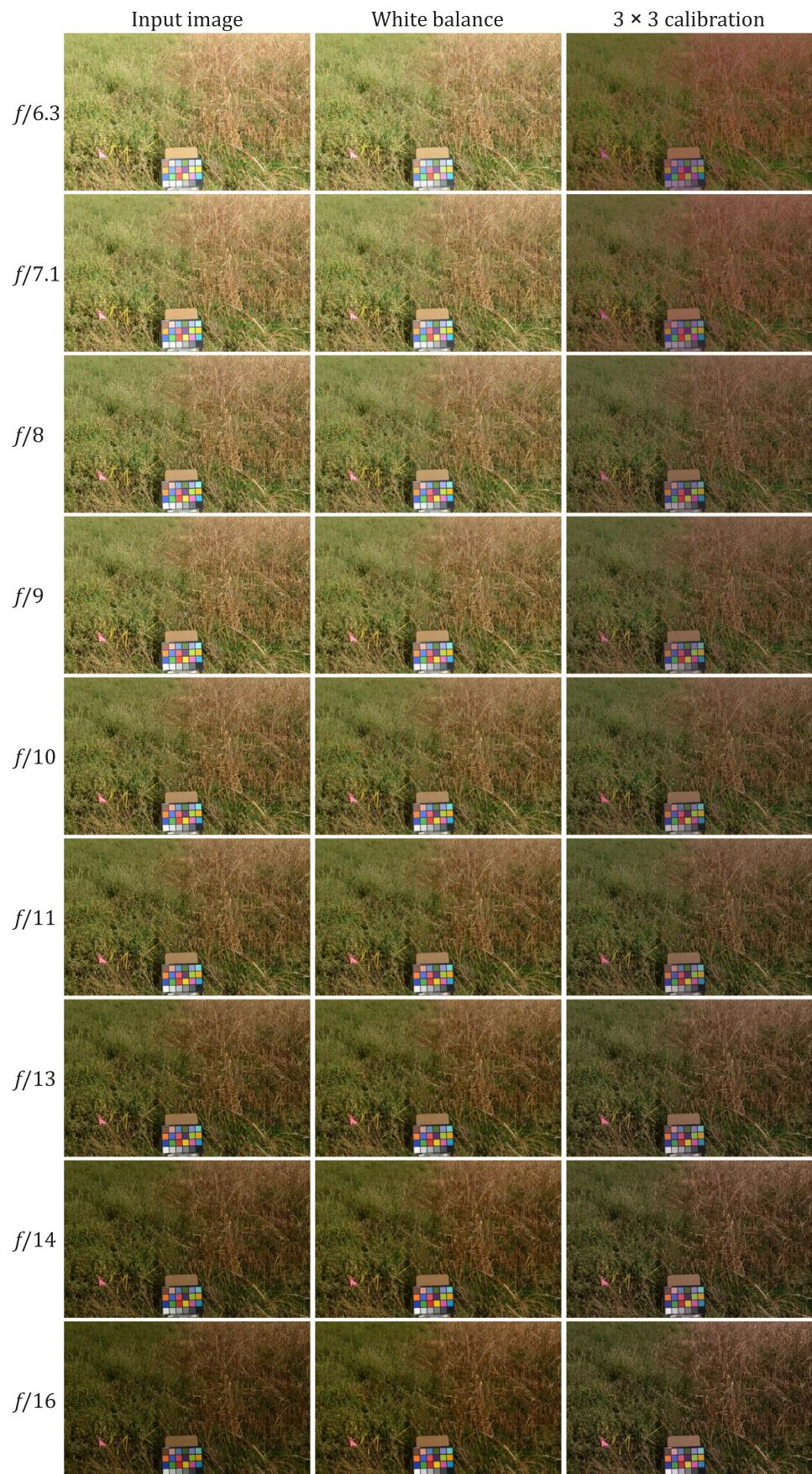


Fig. 9. Comparison of input field images of alfalfa and spring wheat, and their respective white balanced and calibrated images obtained using the three R, G, and B color patches.

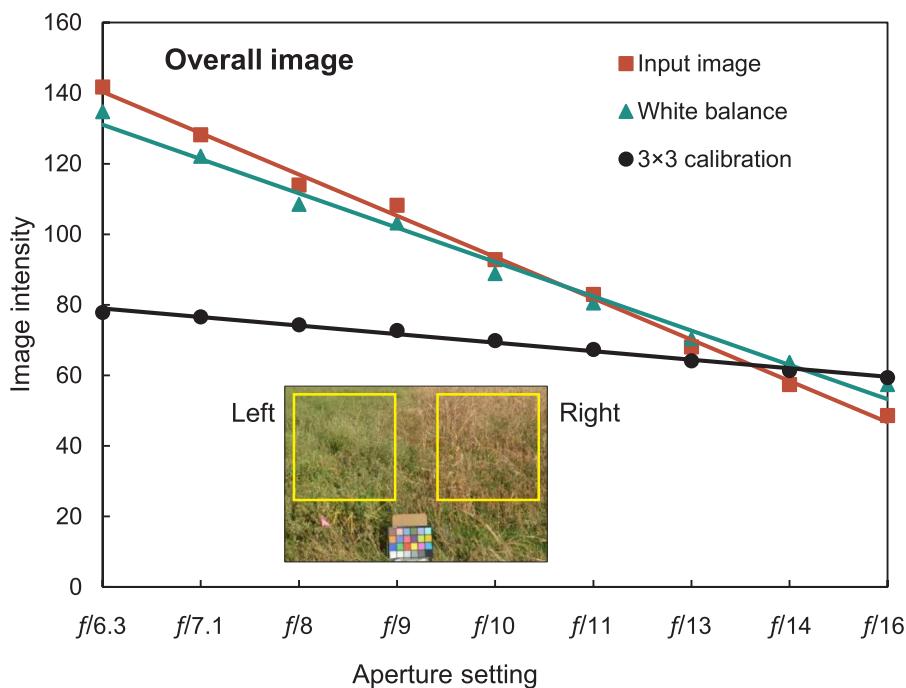


Fig. 10. Intensity variation of input field images of alfalfa and spring wheat at different aperture settings, white balanced image, and the calibrated image obtained using the three R, G, and B color patches and the fitted linear models (data in Section A.5).

and field setting, respectively, suitable for digital color image calibration study. The user-developed plugin performed well in all conditions (laboratory and field conditions) and took ≈ 7 s/image to produce calibrated images by deriving a color calibration matrix (CCM). The developed calibration performance index (CPI) was useful in observing the effect of the number of color patches and deciding the minimum number of color patches for efficient color calibration. The CPI values converged when only 12 color patches (out of 24) were used. On selected color schemes, 3–6 color patches gave a comparable performance to the 24 patches. We found that the order of selected color patches in a color scheme did not affect the calibration performance. The calibration using only the 3 patches namely red, green, and blue (R, G, and B; $[3 \times 3]$ CCM) produced visually homogenous intensity images with low CPI (0.21–0.24), while the commonly followed neutral color white as well as gray patches produced less performance with high CPI (0.26–1.0). Considering practical application and simplicity, 3 color patches: R, G, and B are recommended for an efficient color image calibration. Future work should consider calibrating using CCM derived from higher order terms (square and interaction terms), and testing with real-time field images with natural lighting variations.

Declarations of interest & disclaimer

No known conflict of interests. The North Dakota State University (NDSU) and the United States Department of Agriculture (USDA) are equal opportunity providers and employers.

Acknowledgement

This work was supported in part by the USDA Agricultural Research Service, Grant No.: FAR0028541, and the USDA National Institute of Food and Agriculture, Hatch Project: ND01481, Accession number: 1014700.

Appendix A. Supplementary material

Supplementary data related to this article describing the semi-automatic color patch selection, developed plugin pseudocode, and some results data plugin pseudocode, and some results data can be found, in the online version, at <https://doi.org/10.1016/j.isprsjprs.2018.09.015>.

References

- Ahrends, H.E., Brügger, R., Stöckli, R., Schenk, J., Michna, P., Jeanneret, F., Wanner, H., Eugster, W., 2008. Quantitative phenological observations of a mixed beech forest in northern Switzerland with digital photography. *J. Geophys. Res. Biogeosci.* 113, G04004. <https://doi.org/10.1029/2007jg000650>.
- Akkaynak, D., Chan, E., Allen, J.J., Hanlon, R.T., 2011. Using spectrometry and photography to study color underwater. In: OCEANS. IEEE, pp. 1–8. <https://doi.org/10.23919/oceans.2011.610693>.
- Akkaynak, D., Treibitz, T., Xiao, B., Gürkan, U.A., Allen, J.J., Demirci, U., Hanlon, R.T., 2014. Use of commercial off-the-shelf digital cameras for scientific data acquisition and scene-specific color calibration. *J. Opt. Soc. Am. A* 31, 312–321. <https://doi.org/10.1364/josaa.31.000312>.
- Alsam, A., Finlayson, G., 2008. Integer programming for optimal reduction of calibration

- targets. *Color Res. Appl.* 33, 212–220. <https://doi.org/10.1002/col.20402>.
- Andresen, C., Tweedie, C., Lougheed, V., 2018. Climate and nutrient effects on arctic wetland plant phenology observed from phenocams. *Rem. Sens. Environ.* 205, 46–55. <https://doi.org/10.1016/j.rse.2017.11.013>.
- ASF, 2018. Apache Commons Math. Apache Commons™. The Apache Software Foundation (ASF), Wakefield, MA, USA. Available at: <http://commons.apache.org/proper/commons-math/download_math.cgi> (accessed August 15, 2018).
- Barbedo, J.G.A., 2016. A review on the main challenges in automatic plant disease identification based on visible range images. *Biosyst. Eng.* 144, 52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>.
- Beeri, O., Peled, A., 2009. Geographical model for precise agriculture monitoring with real-time remote sensing. *ISPRS J. Photogramm. Remote Sens.* 64, 47–54. <https://doi.org/10.1016/j.isprsjprs.2008.07.007>.
- Brown, L., Datta, A., Pankanti, S., 2012. Exploiting color strength to improve color correction. In: 2012 IEEE International Symposium on Multimedia. IEEE, pp. 179–182. <https://doi.org/10.1109/ism.2012.4>.
- Brown, M., Majumder, A., Yang, R., 2005. Camera-based calibration techniques for seamless multiprojector displays. *IEEE Trans. Visual Comput. Graphics* 11, 193–206. <https://doi.org/10.1109/tvcg.2005.27>.
- FreePNGimg, 2018. Green leaf PNG. Available at: <http://www.freepngimg.com/png_2293-green-leaf-png> (accessed August 15, 2018).
- Haeghen, Y.V., Naeyaert, J.M.A.D., Lemahieu, I., Philips, W., 2000. An imaging system with calibrated color image acquisition for use in dermatology. *IEEE Trans. Med. Imaging* 19, 722–730. <https://doi.org/10.1109/42.875195>.
- Hunt, E.R., Cavigelli, M., Daughtry, C.S., McMurtrey, J.E., Walther, C.L., 2005. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* 6, 359–378. <https://doi.org/10.1007/s11119-005-2324-5>.
- van Iersel, W., Straatsma, M., Addink, E., Middelkoop, H., 2018. Monitoring height and greenness of non-woody floodplain vegetation with UAV time series. *ISPRS J. Photogramm. Remote Sens.* 141, 112–123. <https://doi.org/10.1016/j.isprsjprs.2018.04.011>.
- Ilie, A., Welch, G., 2005. Ensuring color consistency across multiple cameras. In: Tenth IEEE International Conference on Computer Vision. IEEE, pp. 1268–1275. <https://doi.org/10.1109/iccv.2005.8>.
- Joshi, N., Wilburn, B., Vaish, V., Levoy, M.L., Horowitz, M., 2005. Automatic color calibration for large camera arrays. [Department of Computer Science and Engineering], University of California, San Diego.
- Kagarlitsky, S., Moses, Y., Hel-Or, Y., 2009. Piecewise-consistent color mappings of images acquired under various conditions. In: 2009 IEEE 12th International Conference on Computer Vision. IEEE, pp. 2311–2318. <https://doi.org/10.1109/iccv.2009.545943>.
- Lam, E.Y., 2005. Combining gray world and retinex theory for automatic white balance in digital photography. In: Proceedings of the Ninth International Symposium on Consumer Electronics. IEEE, pp. 134–139. <https://doi.org/10.1109/isce.2005.150235>.
- Levin, R.L., Levin, G.V., 2004. Solving the color calibration problem of martian lander images. In: Instruments, Methods, and Missions for Astrobiology VII, SPIE, pp. 158–171. <https://doi.org/10.1117/12.504359>.
- McCamy, C.S., Marcus, H., Davidson, J., 1976. A color-rendition chart. *J. Appl. Photog. Eng.* 2, 95–99.
- O'Connell, J., Bradter, U., Benton, T.G., 2015. Wide-area mapping of small-scale features in agricultural landscapes using airborne remote sensing. *ISPRS J. Photogramm. Remote Sens.* 109, 165–177. <https://doi.org/10.1016/j.isprsjprs.2015.09.007>.
- Porikli, F., 2003. Inter-camera color calibration by correlation model function. In: 2003 International Conference on Image Processing. IEEE. <https://doi.org/10.1109/icip.2003.124663> II–133.
- Rasband, W.S., 2018. ImageJ. US National Institutes of Health, Bethesda, MD, USA. <<https://imagej.nih.gov/ij/>> (accessed August 15, 2018).
- Richardson, A.D., Braswell, B.H., Hollinger, D.Y., Jenkins, J.P., Ollinger, S.V., 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecol. Appl.* 19, 1417–1428. <https://doi.org/10.1890/08-2022.1>.
- Scharf, P.C., Lory, J.A., 2002. Calibrating corn color from aerial photographs to predict sidedress nitrogen need. *Agron. J.* 94, 397–404. <https://doi.org/10.2134/agronj2002.3970>.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al., 2012. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. <https://doi.org/10.1038/nmeth.2019>.
- X-Rite, 2018. Colorimetric values for ColorChecker family of targets. Available at: <http://xritephoto.com/ph_product_overview.aspx?ID=824&Action=Support&SupportID=5159> (accessed August 15, 2018).
- Xu, Q., Hou, Z., Tokola, T., 2012. Relative radiometric correction of multi-temporal ALOS AVNIR-2 data for the estimation of forest attributes. *ISPRS J. Photogramm. Remote Sens.* 68, 69–78. <https://doi.org/10.1016/j.isprsjprs.2011.12.008>.