Colour Calibration for Agricultural Imaging: A Pilot Study

James Bennett[1,2], Jinhuan Liu[2], Daniel Freer[2], Marc Jones[2] and Graham Finlayson[1]

[1]School of Computing Sciences, University of East Anglia, Norwich, UK.
[2]Antobot, Chelmsford, UK.

E-mail: james.r.bennett@uea.ac.uk

Abstract. Cameras are becoming central to agricultural automation, where they are used to measure and interpret an unstructured biological environment. Applications such as monitoring plant health and predicting yield depend on consistent image formation, yet images often contain variation. A key question is whether the observed variation is caused by genuine biological diversity or introduced by the camera pipeline. We present a pilot study of the colour variation of ripe strawberries imaged in the field. Images are captured with fruit still on the plant, deriving ripe fruit annotations from in-field observations of what was subsequently picked. We compare the chromaticity variation of ripe strawberries in images rendered with the camera's default processing pipeline and in images rendered after calibration from a colour checker placed in each image. The calibration reduces the observed variation suggesting that the cause is not solely biological and that in-field calibration can improve the consistency of colour measurements in agricultural imaging.

1 Introduction

Camera-based vision systems are increasingly central to agricultural robotics [1] and used in many tasks including, disease detection [2] and fruit counting for yield estimation [3, 4]. Colour is the primary measurement signal in many agricultural applications [5], and nowhere more so than in strawberry ripeness estimation [6, 7]. This work is motivated by deployment on Antobot's agricultural imaging platform, which uses embedded cameras to visually assess strawberries in commercial polytunnel environments (Figure 1). A key question arises: is the colour variation we observe in images of ripe fruit a property of the fruit itself, or is it introduced by the capture environment and/or camera pipeline? Regarding the former, illumination in the polytunnel environment is uncontrolled and varies with solar angle, cloud cover, and canopy shading. And, regarding pipelines, it is known that while cameras attempt to account for the colour of the light and exposure, it is also known that their success in doing so is only partial.

The image signal processor (ISP) in a digital camera applies a sequence of operations to convert raw sensor data into a rendered output image. These stages typically include black level correction, demosaicing, white balance, colour space conversion, and gamma encoding [8]. Several of these stages are scene dependent. White balance in particular estimates the colour of the prevailing illuminant and applies a compensating gain to each colour channel, a problem that is well-studied but remains imperfect in unconstrained conditions [9]. These operations are not invertible from the rendered output and the individual contributions of each pipeline

Figure 1: Antobot's robot operating in a polytunnel of tabletop strawberries.

stage cannot be disentangled [10]. Most consumer and embedded camera platforms offer no access to the underlying pipeline. The Raspberry Pi ISP (PiSP) is unusual in this respect: its configurable, documented architecture makes it possible to estimate calibration parameters from raw sensor data and inject them directly into the rendering pipeline, a property we make use of in this paper.

Illumination variability is a frequently cited source of error in agricultural imaging [6, 11], yet the typical response in agricultural robotics is workaround rather than calibration - converting colour space, augmenting training data with colour perturbations, or applying learned illumination normalisation. Rigorous chart-based colour calibration is used in plant phenotyping [12] and UAV remote sensing [13], but is rarely applied to ground-based robots. The problem is structurally well understood in adjacent scientific imaging domains, including underwater ecology [10] and biomarker quantification from smartphone images [14] with demonstrable in-field efficacy. However, to our knowledge, it has not been addressed in agricultural robotics.

In this paper, we present a pilot study of colour variation in ripe strawberries imaged in the field. A 24-patch Classic ColourChecker is placed in each scene which is used to estimate white balance colour gains and a colour correction matrix per image. Calibrated parameters are injected into the PiSP back-end, so the result is rendered by the same pipeline the camera uses in deployment. We find that calibration reduces the variance of ripe berry chromaticity centroids by 48% on images rendered through the PiSP pipeline. This suggests that a substantial fraction of the observed colour variation in ripe strawberries is instrumental rather than biological, and is recoverable by straightforward in-scene calibration - with direct implications for how colour-based ripeness systems should be designed.

2 Image acquisition

Images were captured using Antobot's Insight In-a-Box system, which consists of a Raspberry Pi 5 paired with two Raspberry Pi HQ cameras using the Sony IMX477 sensor. The system was operated from a wireless tablet interface; pressing a button triggers simultaneous capture of a JPEG, a DNG file containing the raw Bayer data, and a metadata file recording the parameters computed by the camera pipeline for that frame - including white balance colour gains, black levels, and the colour correction matrix applied by the ISP. The default `imx477.json` tuning file was used throughout, with the exception of exposure parameters tuned for the field conditions.

Figure 2: Acquisition experimental setup (left) and uncalibrated pipeline output of initial image (middle) then after ripe fruit harvested (right).

Annotating ripeness from images alone risks introducing the very bias this paper seeks to measure: a data labeller working from a monitor has no ground truth for ripeness, and the camera's uncalibrated colour response may distort their judgement. Our approach instead derives annotations from the harvesting decisions of an experienced picker.

The camera system was mounted on a tripod facing the side of a strawberry row (Figure 2 left), with a 24-patch Classic ColourChecker placed in the scene and positioned to receive the same illumination as the fruit. An initial image was captured containing all visible berries (Figure 2 middle). The picker then harvested the ripe fruit, and a second image was captured from the same position (Figure 2 right). Segmentation masks for ripe berries were drawn on the initial image by identifying berries present in the first image but absent in the second - that is, those judged ripe by the picker. Data collection took place on two days in September and October 2025, towards the end of the growing season. The first day was characterised by strong direct sunlight with intermittent cloud cover; the second was overcast and more consistent. The subset of data used in this work comprises 48 images, 112 annotated ripe berries, resulting in 413,635 ripe pixels in total, split roughly equally between the two days.

3 Camera pipeline and calibration

The PiSP consists of a front-end that computes per-frame scene statistics, a suite of image processing algorithms that consume those statistics and determine rendering parameters, and a back-end that applies those parameters - performing black level correction, demosaicing, white balance, colour correction, and gamma encoding - to produce the output image. The estimated parameters are applied and the result encoded as a JPEG; once rendered, the individual contributions of each stage cannot be recovered from the output alone. The DNG files saved alongside each JPEG capture the raw Bayer data before any of these operations, and the accompanying metadata records the parameters the pipeline computed for that frame. Together these give us everything needed to re-render the image with different parameters.

3.1 White balance and colour correction

White balance gains are estimated directly from the raw Bayer data rather than from the JPEG that has the camera's estimated parameters baked in. In Python, the 24-patch Classic ColourChecker in each scene is detected using the `colour-checker-detection` library, aided by a bounding-box annotation, improving reliability against cluttered backgrounds. To estimate

the white balance gains, we use the six neutral patches of the ColourChecker and select the brightest that is not saturated. In the raw Bayer image, the R, G, and B channels of this patch are sampled and averaged in a small neighbourhood around the patch centre. The white balance gains are then

$$R_{\text{gain}} = \frac{G_{\text{avg}}}{R}, \quad B_{\text{gain}} = \frac{G_{\text{avg}}}{B}, \tag{1}$$

where G_{avg} is the mean of the two green sub-channels, normalising the neutral patch so that all three channels are equal.

The colour correction matrix (CCM) is fitted in camera-native linear RGB - the same space in which the Raspberry Pi ISP hardware applies its CCM - so the fitted matrix can be injected directly into the pipeline. Each DNG is rendered into camera-native space using `rawpy` with the patch-derived white balance gains applied at render time. Both the measured patch values and the 24-patch linear sRGB reference are normalised by the white-patch measurement, mapping white to $[1, 1, 1]$ in both. For each image, the CCM is then found by least squares,

$$\hat{M} = \arg\min_{M} \|AM - B\|, \tag{2}$$

where A is the 24×3 matrix of normalised measured patch values and B is the corresponding matrix of normalised linear sRGB reference values.

3.2 Re-rendering via the PiSP back-end

Rather than rendering calibrated images in Python, we re-render using the actual PiSP back-end hardware on the Raspberry Pi 5. This ensures the output is faithful to what the camera would produce in deployment, including the hardware gamma LUT and the specific fixed-point arithmetic of the CCM stage. The back-end is driven by a lightweight C++ executable which streams the raw Bayer data into the pipeline via a V4L2 device interface and injects the calibration parameters directly into the hardware ISP. Black level correction, demosaicing, colour correction, and gamma are applied in sequence by the hardware to produce the rendered output. All parameters other than the white balance gains and CCM are taken from the default `imx477.json` tuning file. Critically, the baseline images are also re-rendered through the same path, using the white balance gains and CCM recorded in the capture metadata rather than the patch-estimated values. The comparison is therefore direct and any systematic rendering artefact affects both renditions equally.

4 Results and discussion

We quantify colour variation using chromaticity, which captures the colour of a pixel independently of its brightness. Rather than normalised rg chromaticity, we use the opponent chromaticity representation [15]

$$rg = \frac{R - G}{R + G + B}, \quad by = \frac{R + G - 2B}{R + G + B}, \tag{3}$$

which is a more regular mapping of chromaticities and better separates colour variation from residual luminance differences.

Each annotated berry contributes a cloud of pixels in chromaticity space. Within-berry variation is mostly biological - surface texture, specular highlights, shadowing across the fruit - and is not something calibration can or should be expected to remove. We are instead interested in shot-to-shot consistency: whether two berries of equal ripeness, imaged under different conditions, appear at the same location in chromaticity space. We therefore compute the chromaticity

centroid of each berry and treat the distribution of those 112 centroids as our primary measure of variation.

Figure 3 shows the chromaticity distributions for all ripe berry pixels from the PiSP renders, with berry centroids overlaid as white circles. The spread of centroids is characterised by the 1σ and 2σ confidence ellipses shown for the centroid distribution. The left panel shows the camera-metadata render (our *fair* re-rendition of the actual camera output); the right panel shows the patch-fitted calibrated render. The tightening of the centroid distribution is visually apparent: the 2σ ellipse shrinks substantially and the centroids cluster more tightly around a common locus. The limits of the rg axis in the chromaticity space are $[-1, 1]$ and for the by axis are $[-2, 1]$, so Figure 3 is zoomed in to show more detail.

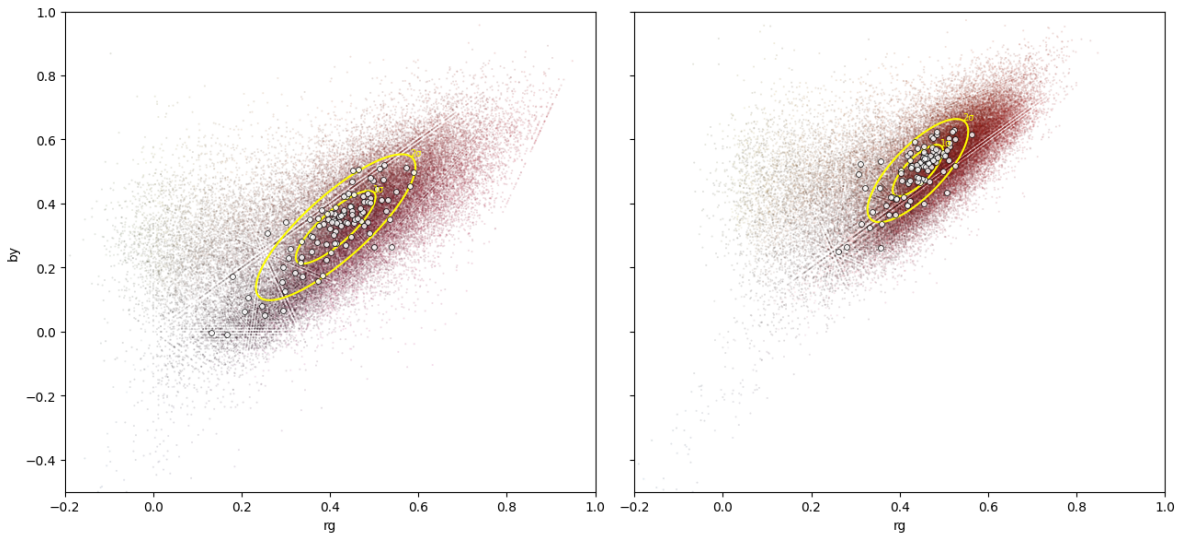

Figure 3: Chromaticity scatter plot showing uncalibrated (left) and calibrated (right) ripe pixel cloud, per berry centroid (white circles) and 1σ and 2σ confidence ellipses for the centroids.

Quantitatively, the berry centroid 2σ ellipse area is reduced from 0.072 to 0.037, a reduction of **48%**. The 2σ ellipse for the total pixel cloud variance is reduced in area from 0.28 to 0.18, a reduction of 36%. The larger reduction in centroid variance than pixel cloud variance is expected: calibration corrects for shot-to-shot illuminant variation, which shifts the centroid of each berry, while within-berry variation is largely unaffected.

Figure 4 shows example ripe berries rendered with camera-metadata parameters (left group) and patch-fitted calibration (right group). The greater colour consistency of the right group is apparent, particularly in hue: the camera-metadata renders range from dark magenta to bright orange-red depending on the illuminant conditions at capture, while the calibrated renders are more uniform in appearance.

The calibration is implemented entirely within the operating space of the PiSP hardware - white balance gains in raw Bayer space, CCM in camera-native linear RGB - and the parameters are injected directly into the PiSP back-end without modification to the rendering architecture. Alternative rendering approaches yield larger reductions in chromaticity variance, but rigorous comparison is complicated by differences in the rendering chain and evaluation space. Here we prioritise faithfulness to the deployed pipeline, so the result reported reflects the gain achievable without modification to the camera's rendering architecture.

The principal limitation of the current work is that the ColourChecker must be placed in

Figure 4: Example ripe berries rendered with camera estimated parameters (left) and calibrated parameters (right).

the scene by hand at each imaging position. This is manageable for a pilot study but is not compatible with autonomous operation at scale. A realistic path towards deployment exists, however. Polytunnel strawberry production already makes use of fiducial markers for row identification and navigation, and coloured reference patches could be incorporated into this existing infrastructure. Even a single grey patch of known reflectance visible in frame would support per-frame white balance updates, with a full CCM estimated periodically as the illuminant changes. The frequency of recalibration need not be per-frame since white balance gains in the PiSP are updated gradually over multiple frames. An investigation into the periodicity of recalibration would be a valuable future contribution.

This work is a pilot study and several questions remain open. The dataset spans two days at the end of the growing season; a more complete study would cover a wider range of conditions, growing stages, and camera configurations. The primary finding is that a substantial fraction of the observed chromaticity variation in ripe strawberries is instrumental rather than biological. Improving the consistency of the colour signal in this way is likely to reduce the burden on downstream ripeness classifiers, though quantifying that improvement directly remains an important open question. More broadly, the case for calibration is arguably stronger wherever the diagnostic colour signal is more subtle: in canopy health assessment, for instance, the colour shift from chlorophyll-rich green towards yellow-green indicating nutrient stress or disease onset may be entirely masked by the instrumental variation we observe here.

5 Conclusion

We have presented a pilot study of colour variation in ripe strawberries imaged in a commercial polytunnel environment, asking whether the variation observed is biological or instrumental. Using a ColourChecker placed in the scene to estimate white balance gains and a colour correction matrix, and re-rendering images through the actual Raspberry Pi ISP back-end hardware, we find that calibration substantially reduces the shot-to-shot chromaticity variation of ripe berries. This demonstrates that a meaningful fraction of the observed variation is introduced by the camera pipeline rather than the fruit, and that it is recoverable by straightforward in-scene calibration implemented within the deployed rendering architecture.

Acknowledgements
We would like to thank our picker, Tsakani, and the wider *Wilkin & Sons Limited* team for accommodating our data collection.

References

[1] Duckett T, Pearson S, Blackmore S, Grieve B, Chen WH, Cielniak G, et al.. Agricultural Robotics: The Future of Robotic Agriculture; 2018. arXiv:1806.06762 [cs.RO].

[2] Mahlein AK. Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. Plant disease. 2016;100(2):241-51.

[3] He L, Fang W, Zhao G, Wu Z, Fu L, Li R, et al. Fruit yield prediction and estimation in orchards: A state-of-the-art comprehensive review for both direct and indirect methods. Computers and Electronics in Agriculture. 2022;195:106812.

[4] Koirala A, Walsh KB, Wang Z, McCarthy C. Deep learning – Method overview and review of use for fruit detection and yield estimation. Computers and Electronics in Agriculture. 2019;162:219-34.

[5] Abdalla A, Cen H, Abdel-Rahman E, Wan L, He Y. Color Calibration of Proximal Sensing RGB Images of Oilseed Rape Canopy via Deep Learning Combined with K-Means Algorithm. Remote Sensing. 2019;11(24):3001.

[6] Kirk R, Cielniak G, Mangan M. L*a*b*Fruits: A Rapid and Robust Outdoor Fruit Detection System Combining Bio-Inspired Features with One-Stage Deep Learning Networks. Sensors. 2020;20(1):275.

[7] Mohamed I, Williams D, Stevens R, Dudley R. Strawberry ripeness calibrated 2D colour lookup table for field-deployable computer vision. IOP Conference Series: Earth and Environmental Science. 2019 May;275(1):012003.

[8] Karaimer HC, Brown MS. A Software Platform for Manipulating the Camera Imaging Pipeline. In: Leibe B, Matas J, Sebe N, Welling M, editors. Computer Vision – ECCV 2016. Cham: Springer International Publishing; 2016. p. 429-44.

[9] Finlayson GD. Colour and illumination in computer vision. Interface Focus. 2018;8(4):20180008.

[10] Akkaynak D, Brown MS. Underwater imaging without color distortions requires RAW capture; 2026. arXiv:2603.20823 [eess.IV].

[11] Barbedo JGA. A review on the main challenges in automatic plant disease identification based on visible range images. Biosystems Engineering. 2016;144:52-60.

[12] Berry JC, Fahlgren N, Pokorny AA, Bart RS, Veley KM. An automated, high-throughput method for standardizing image color profiles to improve image-based plant phenotyping. PeerJ. 2018;6:e5727.

[13] Sunoj S, Igathinathane C, Saliendra N, Hendrickson J, Archer D. Color calibration of digital images for agriculture and other applications. ISPRS Journal of Photogrammetry and Remote Sensing. 2018;146:221-34.

[14] Finlayson G, Nixon-Hill M. Computer-implemented method and system for image correction for a biomarker test; 2026. Assignee: Vital Signs Solutions Ltd. Filed 2022-06-29. US Patent US12530754B2. Available from: https://patents.google.com/patent/US12530754B2.

[15] Berens J. Image Indexing using Compressed Colour Histograms. Norwich, UK: University of East Anglia; 2002.