Reprinted from
Journal of Imaging Science and Technology® 64(5): 050411-1–050411-8, 2020.

A Flying Gray Ball Multi-illuminant Image Dataset for Color Research

Hoda Aghaei and Brian Funt▲
Simon Fraser University, Vancouver, B.C., Canada
E-mail: haghaei@sfu.ca

Abstract. *For research in the field of illumination estimation and color constancy, there is a need for ground-truth measurement of the illumination color at many locations within multi-illuminant scenes. A practical approach to obtaining such ground-truth illumination data is presented here. The proposed method involves using a drone to carry a gray ball of known percent surface spectral reflectance throughout a scene while photographing it frequently during the flight using a calibrated camera. The captured images are then post-processed. In the post-processing step, machine vision techniques are used to detect the gray ball within each frame. The camera RGB of light reflected from the gray ball provides a measure of the illumination color at that location. In total, the dataset contains 30 scenes with 100 illumination measurements on average per scene. The dataset is available for download free of charge.*

[DOI: 10.2352/J.ImagingSci.Technol.2020.64.5.050411]

1. INTRODUCTION

A technique is proposed for measuring the chromaticity of incident illumination at many locations within both indoor and outdoor real-world scenes with the goal of creating a useful database for research on illumination estimation and color constancy. The strategy employs a drone to carry a calibrated gray ball throughout a scene while it is simultaneously being photographed by a calibrated camera at a fixed location. Using the proposed technique, a database of 30 scenes containing 100 to 150 images per scene with accompanying ground-truth illumination measurements is constructed and made publicly available.

Figure 1 shows the complexity of the illumination in some of the typical scenes from the database. For example, the image in the second column, second row (from the top) includes both direct sunlight and light filtered through tree leaves creating strong shadows. The image in column one, row two includes both indoor light and light entering from a window. Column two, row six shows an example of multiple indoor lights. Clearly, a single measurement of scene illumination using a gray card or a Macbeth ColorChecker will not suffice as a measure of scene illumination in these cases.

Among many color constancy approaches such as the methods by Buchsbaum [1], Land [2], Van De Weijer and Gevers [3], Finlayson and Trezzi [4], and Barron and Tsai [5], the most common approach involves two steps. First, illumination chromaticity (which is assumed to be constant throughout the scene) is estimated. Second, the estimated illumination is used to adjust the image colors to appear as similar as possible to the case where the colors would be under some standard chosen "white" light. However, illumination chromaticity is rarely constant throughout a scene. Recently, methods have been proposed for the illumination estimation for multi-illuminant scenes such as the techniques by Bleier et al. [6], Zhu and Funt [7], Gao et al. [8], Bianco et al. [9], and Beigpour et al. [10]. To evaluate the performance of these methods, the ground-truth illumination chromaticity at every (or at least many) image location is required over a significant number of scenes. This article describes the construction and use of such a dataset. This dataset overcomes the problem of the need to have many calibration targets placed throughout a scene by using a drone to carry a single target to many locations.

This work aims to overcome these difficulties by using a new technique that involves moving the illumination measurement target within the scene with the help of a drone. This idea facilitates the process of taking multiple pictures to capture the illumination chromaticity at multiple positions within each scene. The resultant dataset can be used to develop illumination estimation methods for real-world, multi-illuminated scenes.

2. RELATED WORK

There are many large image datasets, such as COCO, which are available for use in a variety of image processing tasks. However, few image datasets satisfy all the necessary criteria required for color constancy and illumination estimation projects. One such factor is that images should be in RAW format, which is a requirement before any gamma adjustment or white balancing methods are applied. Although some datasets, such as RAISE, may provide images in RAW format, they do not provide any information about scene illumination.

The most important and challenging feature required of an image dataset for color constancy and illumination estimation is that it should provide the ground-truth value of illumination chromaticity for each surface in the scene, preferably at every pixel. As this is a particularly challenging task, simplifying assumptions are usually made about scenes.

▲ IS&T Member.

Received Mar. 31, 2020; accepted for publication Aug. 20, 2020; published online Sep. 30, 2020. Associate Editor: Marius Pedersen.
1062-3701/2020/64(5)/050411/8/$25.00

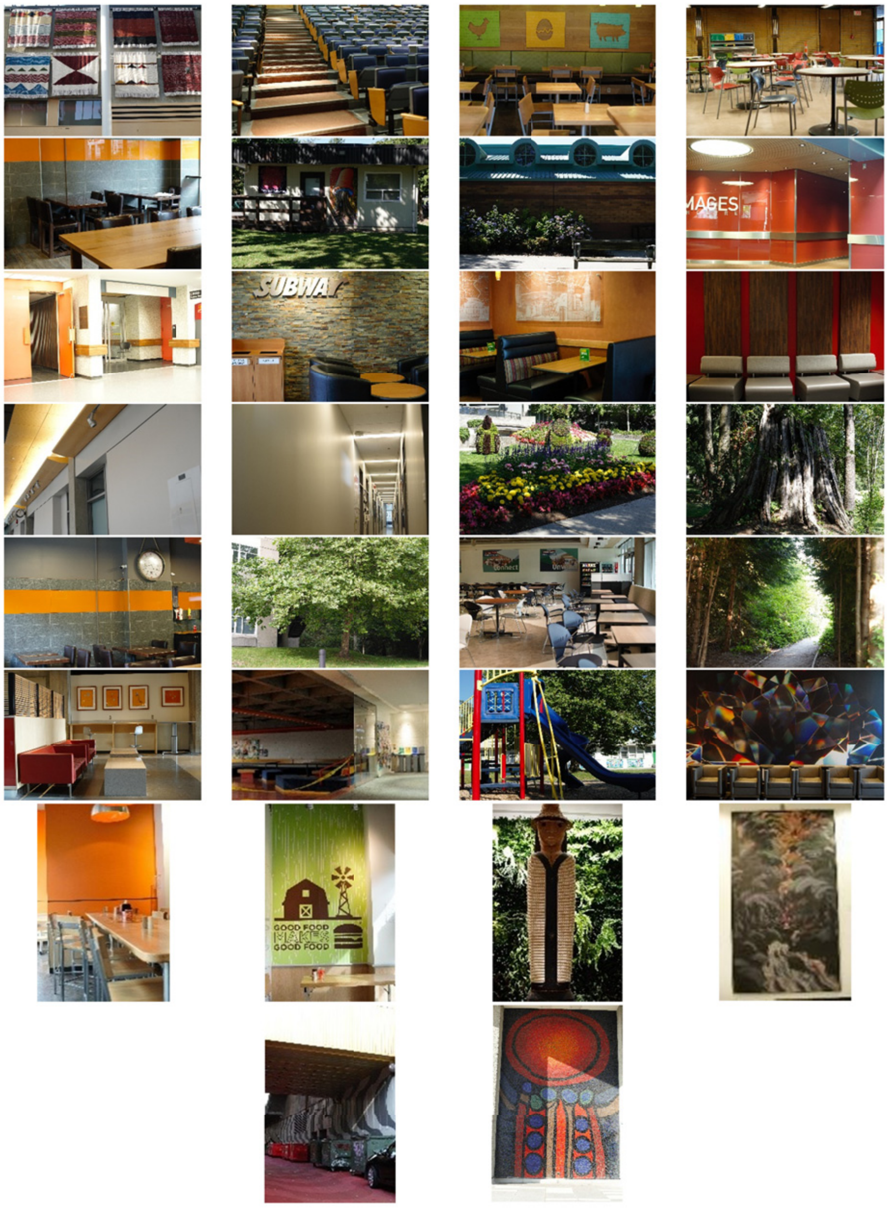

Figure 1. Examples of the multi-illuminant scenes in the database showing direct sunlight, shadow, and mixtures of indoor and outdoor illumination.

These assumptions affect the approaches and instruments that are used to measure illumination chromaticity.

The most common assumption is that illumination chromaticity is uniform across each scene. Therefore, only one measurement of illumination chromaticity would suffice for the entire scene. This measurement could be obtained by estimating the RGB of a standard calibrated gray card. An ideal gray surface reflectance ($(S(\lambda)_{\text{visible}} = 0.5)$ has $R = G = B$ under ideal white light ($(E(\lambda)_{\text{visible}} = \text{constant}$), and thus any deviation from this represents the illumination chromaticity. There are several image datasets available based on the uniform-illumination assumption, including those by Gehler et al. [11], Shi and Funt [12], Hemrit et al. [13], Cheng et al. [14], Banić and Lončarić [15], and Ciurea and Funt [16].

The first version of the Color Checker dataset was released by Gehler et al. [11]. It consists of 568 images provided in both RAW and nonlinear gamma adjusted format. A Macbeth color chart was placed in each scene to measure illumination chromaticity. This dataset has been widely used in color constancy research due to its high-quality images and a large number of uncorrelated images. In 2011, Shi and Funt [12] provided a reprocessed version in linear 16-bit TIFF format derived directly from the original RAW image dataset. However, there was some ambiguity in the instructions provided with the dataset, which led to inconsistency in how researchers interpreted the black level for a subset of images. To overcome these problems, new ground-truth data have been recommended by Hemrit et al. [13] as the new standard.

The NUS dataset by Cheng et al. [14] contains 1,736 images from eight different cameras with a color checker in each image. As the images are taken by using different cameras, this is helpful for carrying out analysis independent of the cameras. Information is provided in Matlab metadata file format (*.mat*) for each image, including the coordinates of the color checker, masks for the color patches on the color checker, and the black level of each camera.

The Cube image dataset by Banić and Lončarić [15] consists of 1,365 outdoor images taken by using a Canon EOS 550D camera. For this dataset, a measurement object named SpyderCube is placed at the lower corner of each scene. There is also an extension to the dataset, which adds 342 images including indoor and nighttime outdoor images, named the Cube+ image dataset [17]. The SpyderCube calibration object has the advantage that it includes faces at different angles, and the faces contain different shades of gray including 18% gray, 96% white, and 4% black, and a hole that is nearly absolute black. These different shades allow users to adjust and correct the exposure and the black level of the images. Additionally, in scenes with multiple illuminants, it is useful to have two faces as a measure of the illumination.

For the Simon Fraser University (SFU) gray ball image dataset by Ciurea and Funt [16], a gray ball attached to a stick so that it is held in the camera's field of view is used to measure the chromaticity of illumination. This is one of the largest image color constancy datasets with approximately 11,000 images, but there is some correlation between the images. The NTSC RGB (not sRGB) values of the ball are provided. However, the values are recorded after post-processing; so they are not linear RAW values and are effectively uncalibrated. Cusano et al. [18] provided an image dataset, which includes 46 conditions that differ in terms of illuminant direction, chromaticity, and intensity. Their dataset contains images of 68 samples of raw food (various types of meat, fish, cereals, and fruit) under each of the 46 lighting conditions for a total of $68 \times 46 = 3128$ images. The Burghouts and Geusebroek [19] dataset is based on a combination of four different camera viewing angles, five illumination angles, and two different illumination chromaticities for each material. The dataset contains images of 250 rough textures and also linear mixtures of 12 materials under these different conditions for a total of more than 27,500 images. The Ojala et al. [20] dataset contains images of 319 macrotextures and microtextures under illuminants of three different correlated color temperatures (CCTs) at nine different camera angles and six different camera resolutions. In total, the image dataset contains 51,678 images. Although all these image datasets include many different illumination conditions, they are constrained in that the illumination across each individual image is effectively constant.

The uniform-illumination assumption does not hold for most real-world scenes. The challenge is how to efficiently measure the illumination throughout a scene. Some approaches to this problem include those by Bleier et al. [6], Bell et al. [21], Hao et al. [22], Gijsenij et al. [23], and Nascimento et al. [24].

The image dataset by Bell et al. [21] includes more than 5,000 indoor images that have been annotated using crowdsourcing. The annotations provide ground-truth estimates of surface reflectance properties based on human judgment. During crowdsourcing, some sample points are selected across the images. Following this, users are asked to give their opinion about a pair of sample points as to which point has a darker reflectance value. The dataset is quite large, and it provides beneficial information about surfaces in scenes. However, as the authors point out, the annotation ignores the effect of colored illumination on what users perceived as the color of the surface. Another limitation of the work is that the users only give their judgments on the reflectance's intensity and not its chromaticity. This particularly raises issues when there is potentially a comparison between two points with distinctly different colors. In this case, it is a difficult task to obtain an accurate opinion about the comparison of the brightness of the points.

Another image dataset by Bleier et al. [6] was generated under laboratory conditions, and it contains 408 images. There are four different scenes, mostly composed of objects with diffuse surface materials, with 17 illumination conditions for each scene. The images are captured under different exposures and shutter times in RAW format. Following this, the images are converted to linear images by *DCRAW* without white balancing. To measure the ground truth, the entire scene is painted gray and rephotographed. Although it is a creative method, it is not a practical approach for real-world scenes, and the images in the dataset are of limited variety.

Hao et al. [22] used ray-traced computer graphics to render full spectrum (physically correct as possible) photorealistic images. Since images are synthesized, the ground truth is available on every pixel. The dataset provides 900 stereo pairs. The renderings are limited to indoor scenes.

Gijsenij et al. [23] placed several gray balls in each scene to measure the illumination chromaticity at several scene locations. Images were captured by using a Sigma SD10 camera with a Foveon X3 sensor having a spatial resolution of 384×256. The camera's white balancing mode was set to the overcast setting for the entire dataset. The images are captured in RAW format and converted to linear sRGB for the purpose of experimentation. The dataset consists of 59 images under laboratory settings and 9 outdoor images. Therefore, the number of real-world images is still quite limited.

Nascimento et al. [24] used a similar approach but captured hyperspectral rural and urban scenes with gray spheres embedded in each. Their goal was to analyze spatial variations in local illumination. The spectral information of the local illumination in each scene is extracted from which a total of 1,904 chromaticity coordinates and CCTs are derived. In each scene, there are multiple spheres; from each sphere, 17 sample points are uniformly selected.

As can be seen from these projects, it is clear that capturing the ground-truth value for illumination chromaticity in real-world scenes is a challenging task. Hence,

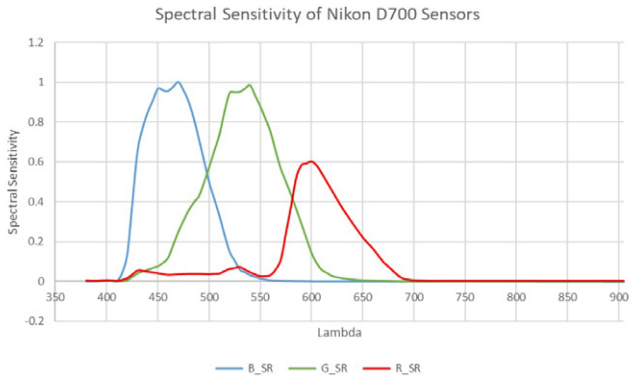

Figure 2. Spectral sensitivity functions of the Nikon D700 camera used to capture all images as generously measured by Image Engineering GmbH & Co.KG using their camSPECS device.

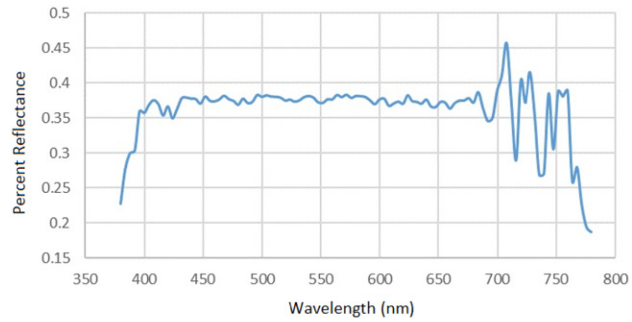

Figure 3. The percent surface spectral reflectance of the painted surface of the gray ball.

the gray-ball-on-a-drone method is proposed as a possible solution. A camera at a fixed location and orientation is used to photograph the gray ball on the drone as it is flown through the scene. The images are then processed using machine vision techniques to locate the gray ball in each frame, record its RGB, adjust for the percent surface spectral reflectance of the ball, and thereby obtain the chromaticity of the light incident on the ball at each location within the scene.

3. METHOD AND EXPERIMENT

3.1 *Hardware Setup*

The camera used to capture images is a Nikon D700 DSLR with a Nikon 50 mm 1:1.4G lens. The spectral properties of the camera's sensor using this lens are shown in Figure 2. The shutter is activated by a remote control in "serial shot" mode. The exposure is automatic, and then the exposure value is recorded separately. Images are saved in RAW format, which is done before any white balancing or gamma correction is applied. The images are therefore linear and require demosaicing.

A light plastic ball of diameter 13 cm is painted uniformly with a high-volume, low-pressure sprayer using Munsell N5 Neutral Matt Gray Vinyl Latex Emulsion Paint. The spectral reflectance of the ball is measured under a standard daylight simulator, D65. The measurement is subsequently divided by the spectrum of the D65 simulator. The final result is shown in Figure 3. The spectrum is measured using the Photo Research SpectraScan PR-650 spectroradiometer. As shown in Fig. 3, the spectral reflectance of the gray paint is effectively flat across the visible wavelength range.

A drone is used to carry the gray ball around the scene and specifically near its surfaces. Figure 4 shows the drone and the gray ball attached to it. The drone used is CineBee 4K Whoop (model F006661RX). Because of its short battery life combined with the additional weight of the ball, each flight lasts only 10 minutes. The propellers are guarded, which allows the drone to be flown next to surfaces without the risk of damage. The drone is flown by a pilot with the basic certificate required under Canadian regulations for drone flight safety.

3.2 *Image Processing*

After recording the image data, the ground-truth chromaticity data are extracted as described in the following sections.

3.3 *Black Level Correction*

Many cameras have a black level greater than zero. The black level of the Nikon D700 camera is measured by taking several exposure shots with the lens cap on and averaging the pixel values of the images. The black level for the camera is effectively zero at 0.8 out of 4096.

3.4 *Ball and Drone Detection*

To measure the chromaticity of the gray ball in an image, it needs to be accurately located, and given the large number of images, an automatic method is needed. Simple background subtraction was tested, but it failed too often. As a result, the *Mask R-CNN* [25] method is used. This requires a large training and validation image dataset. A synthetic dataset of 800 images was created by inserting drone images on top of background images randomly chosen from the COCO dataset. The images of all 30 scenes without a gray ball in them were also included in the training set as background images.

The *Mask R-CNN* model is trained for 300 epochs; 100 steps are present in each iteration, and the batch size is 1. The training requires approximately 3 hours on NVIDIA GeForce GTX 1080. Finally, the trained model is applied to the actual dataset. Figure 5 shows a sample of the results in which the bounding box correctly includes the entire ball and drone.

The ball still needs to be isolated within the bounding box. The Hough circle transform [26] applied to the image area inside the bounding box reliably detects the ball as shown in Fig. 5. To do so, a range of possible values for the radius of the circle is searched to find the circle with the highest confidence value.

4. RESULTS

4.1 *Dataset Properties*

The dataset contains images of 30 scenes. These include both indoor and outdoor scenes captured during the day

Figure 4. (Left) The CineBee 4K Whoop drone used to carry the gray ball; (right) the gray ball being flown on top of the drone.

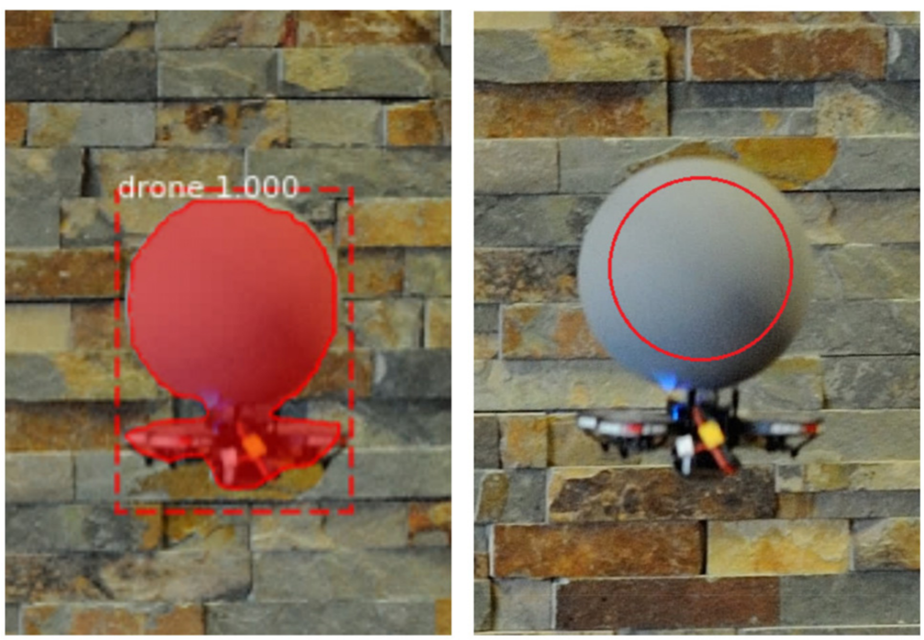

Figure 5. (Left) Automatically detected bounding box surrounding ball and drone; (right) circular area detected on the gray ball from which its average chromaticity is determined.

in summer (July and August) in Vancouver. For each scene, there are 150 images on average. The final version of the image dataset has a size approximately 60 GB. The images are freely available at https://www2.cs.sfu.ca/research/groups/Vision/data2/DRONE-Dataset/. They are in three formats: NEF (Nikon RAW format), linear TIFF, and nonlinear PNG. The scenes are located at the SFU Burnaby and Downtown campuses and at Confederation Park. There are a variety of lighting conditions including direct sunlight, daylight in shadow, daylight in cloudy weather, artificial indoor lights, and reflections from colorful surfaces.

The RAW images in the dataset are first demosaiced using DCRAW. Measurements of illumination chromaticities for each scene are provided in two ways. The first format is a NumPy dictionary. NumPy [27] is a library for the Python programming language for processing multi-dimensional arrays. For each image, the data include the coordinates of the detected gray ball, its radius, and the RAW RGB values

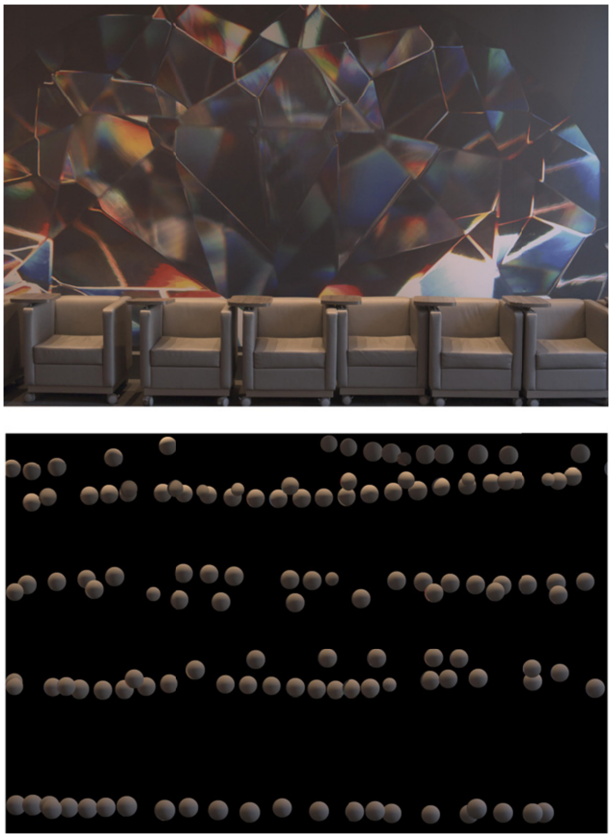

Figure 6. (Top) Indoor scene with chairs and a mural painted on the wall; (bottom) the gray ball as it appeared at the different locations in the scene during its flight.

from the gray ball in the camera's native RGB color space. In all cases, the location of the gray ball is confirmed by visual inspection and adjusted manually when necessary.

The second format is an image, an "illumination map" within which the illumination chromaticity information is embedded. The illumination map for a scene consists of a black ($R = G = B = 0$) background with the images of the detected gray balls placed on it. To account for the fact that the gray ball is not a perfect gray, each RGB in the map is divided by the computed RGB of the gray ball under ideal white light. This value can be calculated by multiplying the camera sensitivity for each channel by the spectral reflectance of the gray ball, which are both already measured.

A 15×15 uniform averaging filter is applied to the images of scenes prior to cropping out the gray ball to average out the effects of the ball's slight surface roughness. Finally, the RGB values in the illumination map are converted to rg-chromaticity values by normalizing each pixel with its $R + G + B$ value. The illumination map is stored as an 8-bit PNG image. An example is shown in Figure 6.

4.2 *Illumination Chromaticity Distribution*

The goal of this project is to obtain the true chromaticity of the incident illumination at many locations in images of typical scenes. The question naturally arises as to how much variation in the chromaticity actually occurs. This is evaluated in terms of the angular difference. Each rg-chromaticity 2-tuple is converted to an rgb 3-tuple using $r + g + b = 1$. The angle between the given gray ball's illumination chromaticity and the chromaticity of the gray ball under ideal white illumination is then computed for each ball location in the scene.

The average angle in degrees, along with the corresponding standard deviation, is computed for each scene. Based on the results shown in Figure 7, the significant standard deviation in the angles found for most scenes indicates that they contain a significant variation in illumination chromaticity.

The next issue is whether or not the scene illuminations represent a reasonable sample of a range of illumination chromaticities that are likely to be found in real-world scenes. Plots of the measured illumination chromaticities are shown in Figures 8–10.

Each illumination chromaticity found in a scene is represented by a dot on the plot. The first chromaticity plot (in camera rg-chromaticity space) corresponds to the scene "Uncle Fatih's Pizza" at SFU. The chromaticities are primarily located along the Planckian locus and cover a range of color temperatures. Figs. 8 and 9 separately plot the chromaticities found for indoor and outdoor scenes.

4.3 *Optimal Single-Illuminant Estimate*

Another measure of the variation in illumination across the scene is provided by analyzing the optimal single-illuminant estimate of the illumination. This involves solving for the illuminant chromaticity that minimizes the angular error across all the ground-truth illumination measurements. Hao et al. [22] describe the Oracle method for solving this optimization problem. The Oracle method is based on complete information about the ground-truth illumination; so it is not an illumination estimation method. However, it does tell us what the minimum error any possible single-illuminant method can generate. It also provides a target for potential multi-illuminant methods to attain.

5. DISCUSSION

5.1 *Applying the Dataset*

The drone flight path for each scene is specifically planned so that each distinct surface is sampled by the drone flying near it at least once. Due to the limitations of the drone's battery life and the large size of the RAW images exhausting the camera's memory capacity, the number of captured frames for each scene is usually limited to 150. However, this is expected to be sufficient for making comparisons among illumination estimation methods. Note that the camera's position and orientation remain constant during the drone's flight. An image of the scene without the drone is also saved.

Using this dataset, an illumination estimation algorithm can be evaluated by applying its algorithm to the drone-free image; its estimate of the illumination can be compared to the ground-truth measurements at the corresponding image locations. This method has the possible disadvantage that the gray ball only flies near the surfaces but is not actually on the surfaces. This is unlikely to make a significant difference,

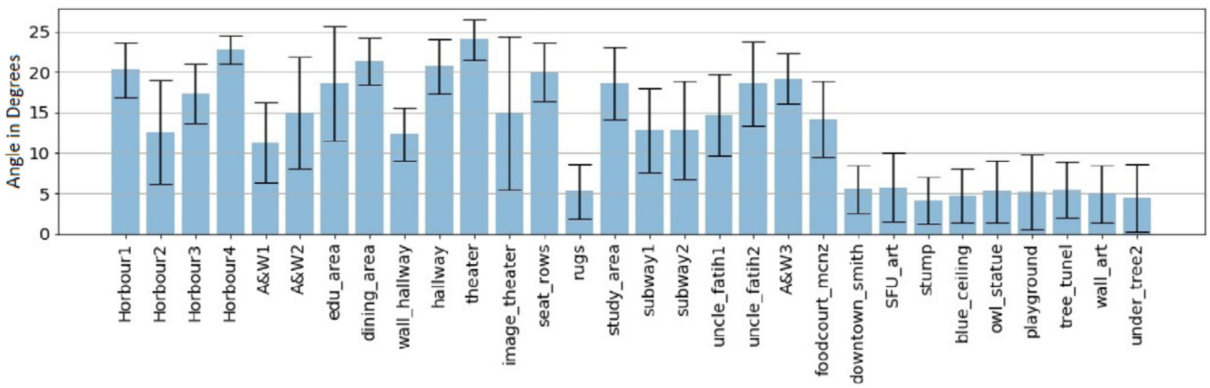

Figure 7. A plot of the mean angle in degrees for each scene between the chromaticity of the gray ball as measured at multiple locations as it is flown throughout the scene and the ball's chromaticity under ideal white illumination. The bars indicate 1 standard deviation from the mean.

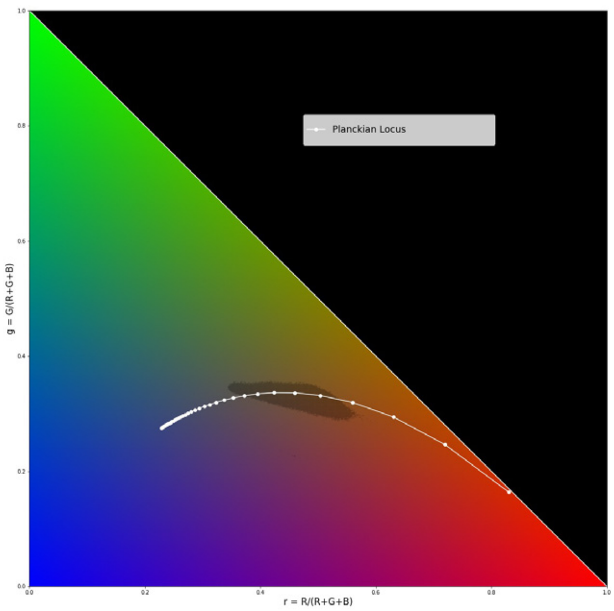

Figure 8. Plot in camera rg-chromaticity space for the indoor scene "Uncle Fatih's Pizza" with the grayish area composed of many separate dots, each one representing the chromaticity of the flying gray ball measured at a different location in the scene. The Planckian locus is provided for comparison.

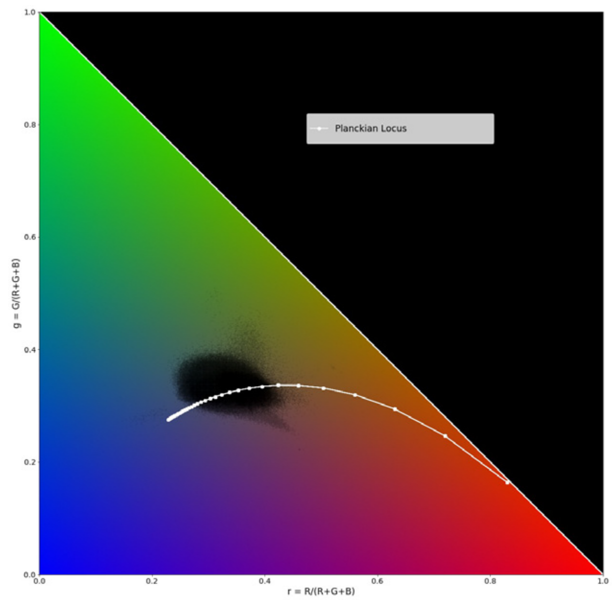

Figure 9. Plot in camera rg-chromaticity space for measurements of the flying gray ball from all indoor scenes combined.

but in an extreme case, the ball could be fully lit while the surface directly behind it is in shadow (not the ball's shadow). However, there are no such cases in the current dataset.

An alternative is to predict the chromaticity of the gray ball itself in each of the separate images. So as long as the algorithm does not explicitly depend on finding the ball or specifically detecting gray areas, this method has the advantage that the ground-truth measurement is precisely at the same three-dimensional location as the ball in the image.

5.2 *Future Work*

Although the chromaticity plots in Figs. 7–9 show that the dataset contains a wide variety of illumination chromaticities, expanding the dataset further to include many more scenes would be desirable. This would make it more useful for training machine learning methods. The range of illumination chromaticities could also be increased by recording scenes at sunrise/sunset, during different seasons, and from different parts of the world.

6. CONCLUSION

In most cases, illumination varies in its chromaticity throughout a scene due to effects such as direct sunlight versus shadow, the filtering of light by plant leaves, indoor light versus daylight through a window, lights of different CCTs within a room, and direct light versus light reflected off colored surfaces. Investigations focused on illumination estimation and color constancy have been hampered by a lack of measured scene illumination data. This project demonstrates that using a drone to carry a gray ball calibration target around a scene while simultaneously photographing it from a fixed location is an effective method for collecting information about the chromaticity of the incident

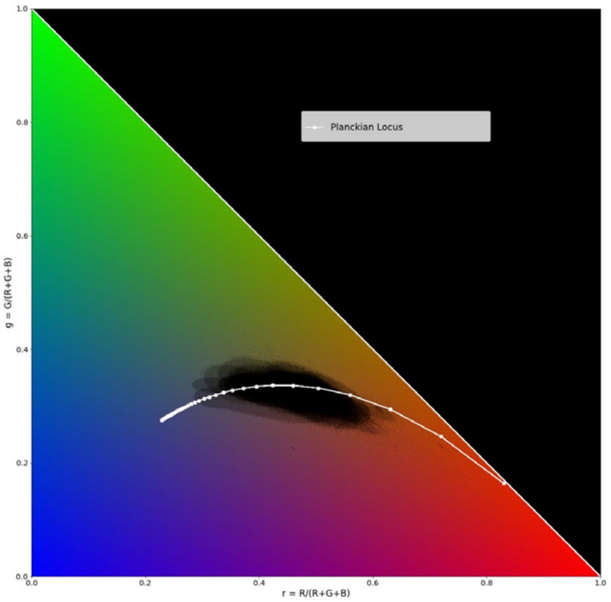

Figure 10. Plot in camera rg-chromaticity space for measurements of the flying gray ball from all outdoor scenes combined.

illumination at many different locations. A database of 30 real-world scenes with illumination measurements at 100 or more locations in each is provided for download. Analysis of the data shows that the database includes many different illumination chromaticities and that there is a significant variation in illumination across most of the scenes.

ACKNOWLEDGMENTS

The authors thank the proprietors of Subway, Uncle Fatih's Pizza, and A&W Canada (all at locations adjoining the SFU Burnaby Campus) for permission to fly the drone on their premises and Image Engineering GmbH & Co. KG for providing the measurements of the camera's spectral sensitivity functions. They also gratefully acknowledge funding by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

[1] G. Buchsbaum, "A spatial processor model for object colour perception," J. Franklin Inst. **310**, 1–26 (1980).

[2] E. H. Land, "The retinex theory of color vision," Sci. Am. **237**, 108–129 (1977).

[3] J. Van De Weijer and T. Gevers, "Color constancy based on the grey-edge hypothesis," *IEEE Int'l. Conf. on Image Processing 2005* (IEEE, Piscataway, NJ, 2005), Vol. 2, pp. II–722.

[4] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," *Proc. IS&T/SID CIC12: Twelfth Color Imaging Conference* (IS&T, Springfield, VA, 2004), pp. 37–41.

[5] J. T. Barron and Y.-T. Tsai, "Fast Fourier color constancy," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 2017), pp. 886–894.

[6] M. Bleier, C. Riess, S. Beigpour, E. Eibenberger, E. Angelopoulou, T. Tröger, and A. Kaup, "Color constancy and non-uniform illumination: Can existing algorithms work?," *2011 IEEE Int'l. Conf. on Computer Vision Workshops (ICCV Workshops)* (IEEE, Piscataway, NJ, 2011), pp. 774–781.

[7] L. Zhu and B. Funt, "Colorizing color images," Electron. Imaging **2018**, 1–6 (2018).

[8] S.-B. Gao, Y.-Z. Ren, M. Zhang, and Y.-J. Li, "Combining bottom-up and top-down visual mechanisms for color constancy under varying illumination," IEEE Trans. Image Process. **28**, 4387–4400 (2019).

[9] S. Bianco, C. Cusano, and R. Schettini, "Single and multiple illuminant estimation using convolutional neural networks," IEEE Trans. Image Process. **26**, 4347–4362 (2017).

[10] S. Beigpour, C. Riess, J. Van De Weijer, and E. Angelopoulou, "Multi-illuminant estimation with conditional random fields," IEEE Trans. Image Process. **23**, 83–96 (2013).

[11] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, "Bayesian color constancy revisited," *2008 IEEE Conf. on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 2008), pp. 1–8.

[12] L. Shi and B. Funt, "Re-processed version of the Gehler color constancy dataset of 568 images", http://www.cs.sfu.ca/~color/data/, 2000.

[13] G. Hemrit, G. D. Finlayson, A. Gijsenij, P. V. Gehler, S. Bianco, M. Drew, B. Funt, and L. Shi, "Providing a single ground-truth for illuminant estimation for the colorchecker dataset," IEEE Trans. Pattern Anal. Mach. Intell. **42**, 1286–1287 (2019).

[14] D. Cheng, D. K. Prasad, and M. S. Brown, "Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution," J. Opt. Soc. Am. A **31**, 1049–1058 (2014).

[15] N. Banić and S. Lončarić, "Unsupervised learning for color constancy", Preprint arXiv:1712.00436, 2017.

[16] F. Ciurea and B. Funt, "A large image database for color constancy research," *Proc. IS&T/SID CIC11: Eleventh Color Imaging Conf.* (IS&T, Springfield, VA, 2003), pp. 160–164.

[17] Color constancy the cube+ dataset. https://ipg.fer.hr/ipg/resources/color_constancy.

[18] C. Cusano, P. Napoletano, and R. Schettini, "Local angular patterns for color texture classification," *Int'l. Conf. on Image Analysis and Processing* (Springer, Genova, Italy, 2015), pp. 111–118.

[19] G. J. Burghouts and J.-M. Geusebroek, "Material-specific adaptation of color invariant features," Pattern Recognit. Lett. **30**, 306–313 (2009).

[20] T. Ojala, T. Maenpaa, M. Pietikainen, J. Viertola, J. Kyllonen, and S. Huovinen, "Outex - new framework for empirical evaluation of texture analysis algorithms," *Object Recognition Supported by User Interaction for Service Robots, Quebec City, Canada* (IEEE, Piscataway, NJ, 2002), Vol. 1, pp. 701–706, Print ISBN: 0-7695-1695-X.

[21] S. Bell, K. Bala, and N. Snavely, "Intrinsic images in the wild," ACM Transactions on Graphics (TOG) **33**, 159 (2014).

[22] X. Hao, B. Funt, and H. Jiang, "Evaluating colour constancy on the new mist dataset of multi-illuminant scenes," *Proc. IS&T CIC27: Twenty-seventh Color and Imaging Conference* (IS&T, Springfield, VA, 2019), Vol. 2019, pp. 108–113.

[23] A. Gijsenij, R. Lu, and T. Gevers, "Color constancy for multiple light sources," IEEE Trans. Image Process. **21**, 697–707 (2011).

[24] S. M. C. Nascimento, K. Amano, and D. H. Foster, "Spatial distributions of local illumination color in natural scenes," Vis. Res. **120**, 39–44 (2016).

[25] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *Proc. IEEE Int'l. Conf. on Computer Vision, Venice, Italy* (IEEE, Piscataway, NJ, 2017), pp. 2961–2969.

[26] H. K. Yuen, J. Princen, J. Illingworth, and J. Kittler, "Comparative study of Hough transform methods for circle finding," Image Vis. Comput. **8**, 71–77 (1990).

[27] T. E. Oliphant, *A Guide to NumPy* (Trelgol Publishing, USA, 2006), Vol. 1.