INTEL-TAU: A Color Constancy Dataset

Firas Laakom*, Jenni Raitoharju§, Alexandros Iosifidis†, Jarno Nikkanen‡¶ and Moncef Gabbouj*

* Faculty of Information Technology and Communication Sciences Tampere University, Finland
§ Programme for Environmental Information, Finnish Environment Center, Jyväskyl, Finland
† Department of Engineering, Electrical and Computer Engineering Aarhus University, DK-8200 Aarhus, Denmark
‡ INTEL, Insinöörinkatu 41, 33720 Tampere, Finland
¶ XIAOMI, Hermiankatu 6-8H, 33720 Tampere, Finland

Abstract—In this paper, we describe a new large dataset for illumination estimation. This dataset, called INTEL-TAU, contains 7022 images in total, which makes it the largest available high-resolution dataset for illumination estimation research. The variety of scenes captured using three different camera models, i.e., Canon 5DSR, Nikon D810, and Sony IMX135, makes the dataset appropriate for evaluating the camera and scene invariance of the different illumination estimation techniques. Privacy masking is done for sensitive information, e.g., faces. Thus, the dataset is coherent with the new General Data Protection Regulation (GDPR) regulations. Furthermore, the effect of color shading for mobile images can be evaluated with INTEL-TAU, as we provide both corrected and uncorrected versions of the raw data. We provide in this paper evaluation of several color constancy approaches.

Index Terms—Color constancy, illumination estimation

I. INTRODUCTION

The observed color of an object in a scene depends on its spectral reflectance and the spectral composition of the illuminant. As a result, when the scene illuminant changes, the light reflected from the object also changes [1]. Thus, a dissimilar color is perceived. The human visual system has the ability to discount this effect, giving a consistent color representation of the object under various illuminant. This ability to filter out for the color of the light source is called color constancy [1]. Mimicking this ability is a fundamental prerequisite for many computer vision applications. For a robust color-based system, the illumination effects of the light source need to be discounted, so that the colors of presented in the image reflects the intrinsic properties of the objects in the scene. This is important for many higher level image or video applications. Without color constancy, colors would be an unreliable cue for object recognition, detection and tracking. Thus, color constancy research, also called illumination estimation, has been extensively studied and several approaches has been proposed to tackle it [2]–[4].

With the advancement of machine learning in general and deep learning in particular in various computer vision tasks, many machine learning-based approaches has been proposed for color constancy [5]–[12]. However, these approaches usually require a large amount of data for training and evaluation. Acquiring labeled datasets for illumination estimation is a challenging tasks [13], as in order to extract the ground truth illumination of a scene, a ColorChecker chart needs to be included in the scene. Several datasets have been proposed [13]–[17], although more and larger datasets are always needed. In addition, after the introduction of General Data Protection Regulation (GDPR) act [22], data privacy in the collected datasets needs to be addressed and sensitive information needs to be masked out.

In this paper, we describe a new INTEL-TAU dataset for color constancy research. The dataset containing 7022 high-resolution images is by far the largest publicly available high-resolution dataset for training and evaluation of color constancy algorithms. Furthermore, we have masked all recognizable faces, license plates, and other privacy sensitive information. Thus, the dataset is now fully GDPR compliant. A subset of 1558 images of the current dataset was previously published as Intel-TUT dataset [15], but we had to remove the dataset due to its GDPR incompliance. We collected the images in INTEL-TAU dataset using three different camera models: Canon 5DSR, Nikon D810, and Mobile Sony IMX135. The images contain both field and lab scenes and both printout. It has mainly real scenes along with some lab printout with the corresponding white point information. Thus, this dataset is suitable for scene and camera-invariance estimation of color constancy algorithms.
The rest of this paper is organized as follows. First, we review the available color constancy datasets and approaches in Section II. In Section III we describe the proposed dataset and highlight its main novelties. We also propose several protocols for using this dataset for illumination estimation research. In Section IV, we evaluate the performance of the baseline statistic based models. In the end, we conclude the paper in Section V.

II. PREVIOUSLY PUBLISHED COLOR CONSTANCY DATASETS

One of the most commonly used dataset in color constancy is the ColorChecker dataset [14] introduced by Gehler et al. It is composed of 568 high-resolution RAW images acquired by two camera models: Canon 1D and Canon 5D. Shi and Funt [23] proposed a methodology to reprocess the original images and to calculate the ground truth. The images are demosaiced and available as TIFF images. The location of the color chart and the saturated and clipped pixels are also provided with the database. Later, Finlayson et al. [24] raised a problem with the Shi reprocessed dataset. For this end, Recommended ColorChecker dataset with an updated ground truth was introduced in [13], [25].

Another dataset is the SFU HDR database [26], [27]. It is a set of 105 high dynamic range images, captured using a calibrated camera. 9 images per scene were captured in order to generate the high dynamic range images. For an accurate measure of the global illumination, 4 color charts were used at different locations of the scene.

NUS-8 [17] is one of the largest color constancy datasets. It contains 1736 RAW images. Eight different camera models were used to capture the scenes of this dataset and a total of 210 images were captured by each camera model. Although the dataset is relatively large, a commonly used protocol is to perform tests on each camera separately and report the mean of all the results. As a result, each experiment involves using only 210 images for both training and testing, which is not enough for appropriate training of machine learning and deep learning based approaches.

Banic and Loncaric introduced Cube dataset in [16]. This dataset is composed of 1365 RGB images. All the scenes available in the dataset are outdoor scenes acquired with a Canon EOS 550D camera in Croatia, Slovenia, and Austria. This dataset was later extended into Cube+ dataset [16]. This extension is enriched by an additional 342 images, containing indoor and outdoor scenes. The overall distribution of illuminations in the Cube+ is similar to the ground truth distribution of the NUS-8.

Other hyperspectral datasets [18]–[21] are available for color constancy research. However, these dataset are relatively scarce and thus unsuitable machine learning based solution with the exception of [19] which contains 11000 images. Nonetheless, this dataset is actually composed on video frames as a results most of the images are highly correlated and only 600 are not [14]. Moreover, this dataset has low-resolution images that were subject to correction.

Intel-TUT was proposed in [15]. It contained a subset of 1558 images of the novel INTEL-TAU dataset. Due to the aforementioned problems with the GDPR regulations, it was recently removed. Furthermore, a larger subset of 3420 images was published and used for experiments in [11] and [12], but similar privacy issues were encountered. The privacy masking, which we applied for the current INTEL-TAU dataset, solves all the privacy problems, while we have preserved all the advantages of the previously published subsets and we provide further benefits as described in the next section. Table I presents a comparison of different color constancy datasets.

III. DATASET DESCRIPTION

We introduce a new color constancy dataset, called INTEL-TAU, which

- is the currently clearly largest publicly high-resolution available color constancy dataset with 7022 images with ground truth illumination.
- is available at http://urn.fi/urn:nbn:fi:att:2cd872a6-d18a-4d02-9bc8-7c893b0b51a0.
- provides the training images without the color charts (i.e., there is no need for color chart masking).
- contains images taken with 3 different cameras to allow camera invariance evaluation.
- contains images grouped by the scene type to allow scene invariance evaluation.
- contains mobile images before and after color shading to allow studying the effect.
- is fully GDPR compliant with privacy masking applied on all sensitive information.

INTEL-TAU contains both outdoor and indoor images captured in 17 countries: Finland, India, Malta, Israel, Estonia, USA, Spain, France, Italy, Tenerife, Austria, China, Croatia, Iceland, Belgium, United Arab Emirates, and Latvia. There are 7022 images in total, captured using the three different camera models: Canon 5DSR, Nikon D810, and Mobile Sony IMX135. The dataset has four folders per camera:
field_1_cameras, containing unique field images captured by the camera, field_3_cameras containing the same field images captured by the different camera models, lab_printouts, containing lab printouts, and lab_realscenes consisting of real lab scenes. Table II reports the number of images per category.

When capturing the images, we avoided strong mixed illumination. Instead, we targeted the framing so that one illumination is dominating in the scene. To define the ground truth, there is one ground truth raw Bayer image associated with each raw Bayer image in the database. The ground truth image has a X-Rite ColorChecker Passport chart positioned in such way that it reflects the main illumination in the scene. The actual database image does not have the chart in it, except for a handful of images in which it was intentionally put there as image contents. The same ground truth image can be associated with multiple database images if the illumination is the same in those images. We calculated the ground truth white point from the grey patches #20 − #23, omitting the brightest grey patch #19 and darkest grey patch #24, and additional saturated patches if any. Noise was reduced by 9 × 9 averaging kernel before recording the color component values inside the center area of the grey patch. We manually checked the annotation was for each image.

Only the database images are made publicly available along with the ground truth illumination. The ground truth images, i.e., images with the color chart, are not published in this version of the dataset. Thus, no color chart masking needs to be done before evaluating color constancy approaches using INTEL-TAU dataset. The white point is then stored as [R/G, B/G] coordinate. The spectral responses of the different camera models and the spectral power distributions of the lab light sources are also provided.

The associated .ccm was not calculated based on the ground truth image, but selected from a pre-calculated set of CCMs according to the estimate of the illumination (daylight, indoor fluorescent, indoor tungsten-halogen). Consequently, the .ccm should not be treated as a very accurate color conversion matrix, but just as convenience for illustration purposes, and also as the means to guide the color shading correction that was done on the Sony IMX135 images. Figure 2 presents an example ground truth and database image pair as an illustration, not actual raw Bayer content.

Figure 3 presents the actual raw images of an example ground truth and database image pair as a reminder to the reader that the database has raw Bayer images. Different camera and images characteristics are presented in Table III and Table IV.

Following the GDPR regulations, we applied privacy mask-

---

### Table II

| Camera Model   | field_1_camera | field_3_camera | lab_printouts | lab_realscenes |
|----------------|----------------|----------------|---------------|----------------|
| Canon 5DSR     | 1645           | 144            | 300           | 20             |
| Nikon D810     | 2329           | 144            | 300           | 20             |
| Sony IMX135    | 1656           | 144            | 300           | 20             |

---

### Table III

| Camera Model   | Image Width | Light Shielded Pixels at Left | Image Height | Light Shielded Pixels at Top |
|----------------|-------------|------------------------------|--------------|------------------------------|
| Canon 5DSR     | 8896        | 160                          | 5920         | 64                           |
| Nikon D810     | 7380        | 0                            | 4928         | 0                            |
| Sony IMX135    | 3264        | 0                            | 2448         | 0                            |

| Camera Model   | Raw Data Bit Depth | Data Pedestal/Black Level | Saturation Point |
|----------------|--------------------|--------------------------|------------------|
| Canon 5DSR     | 14                 | 2047                     | 15380            |
| Nikon D810     | 14                 | 601                      | 16383            |
| Sony IMX135    | 10                 | 64                       | 1023             |

(*) The raw frames are stored as uint16 value per each pixel
(/**): Note that the saturation point is not necessarily 2 − \(\text{satuation point} \) − 1
(****): Some of the Sony IMX135 images are upside down

---

### Experimental protocols

We propose the following two evaluation protocols for using INTEL-TAU as a benchmark:

- Using all scenes acquired by one camera for the training, one for the validation and one for the testing in three experiments:
  1. Images acquired by Canon as a training set, images acquired by Nikon for validation, and Sony images for testing (training: 2109 images, validation: 2793 images, testing: 2120 images).
  2. Images acquired by Nikon for training, images acquired by Sony for validation, and Canon images for testing, (training: 2793 images, validation: 2120 images, testing: 2109 images).
  3. Images acquired by Sony for training, images acquired by Canon for validation, and Nikon images for testing.

Fig. 2. An example ground truth and database image pair (illustration, not actual raw Bayer content)

Fig. 3. An example ground truth and database image pair (actual raw image)
TABLE IV
CHARACTERISTICS OF THE CAMERA MODELS USED IN INTEL-V3

| Camera Model        | Resolution         | Focal length     | Aperture size | Pixel size | Raw data bit depth |
|---------------------|---------------------|------------------|---------------|------------|-------------------|
| Canon EOS 5DSR      | 52Mpix (8896H × 5920V) | EF 24-105/4L @ 28mm (*) | F8.0(***)     | 4.14µm     | 14bpp             |
| Nikon D810          | 36Mpix (7380H × 4928V) | AF-S 24-70/2.8G @ 28mm (*) | F8.0(***)     | 4.88µm     | 14bpp             |
| Sony IMX135         | 8Mpix (3264H × 2448V)  | 30.4mm (actual 4.12mm) | F2.4          | 1.12µm     | 10bpp             |

(*)-28mm was the closest to the mobile device focal length that was easy to set consistently based on the markings on the objectives
(**): Smaller aperture was used in order to reduce the depth-of-field difference between the DSLRs and the mobile modules

We provide results for the static methods Grey-World [28], White-Patch [29], Spatial domain [17], Shades-of-Grey [31], and Weighted Grey-Edge [32]. We report the mean of the worst 25% of the recovery angular error (RAE) between the ground truth white point and the estimated illuminant defined as follows

\[
\text{RAE}(\rho^{gt}, \rho^{Est}) = \cos^{-1}\left(\frac{\rho^{gt} \cdot \rho^{Est}}{||\rho^{gt}|| \cdot ||\rho^{Est}||}\right),
\]

where \(\rho^{gt}\) is the ground truth illumination and \(\rho^{Est}\) is the estimated illumination.

Table V and Table VI report the results on the test sets of the first and second protocols defined in previous Section, respectively. We note high error rates for all unsupervised methods.

V. CONCLUSION

In this paper, a new large dataset INTEL-TAU dataset is presented. This dataset is suitable for color constancy research. The diversity of scenes and camera models makes the new database appropriate for evaluating the camera and scene invariance of different illumination estimation techniques. Privacy masking is done for sensitive information, e.g., faces. Thus, it is coherent with the new GDPR regulations. Furthermore, the effect of color shading for mobile images can be evaluated with INTEL-TAU, as it provides both corrected and uncorrected versions of the raw mobile data.

In the future, more approaches color constancy approaches, specially learning based approaches needs to be evaluated on this dataset.

ACKNOWLEDGMENT

This work was supported by the NSF-Business Finland Center for Visual and Decision Informatics project (CVDI) project AMALIA, Dno 3333/31/2018, sponsored by Intel...
Finland. We thank Harish Essaky Sankaran, Uygar Tuna, and Lauri Suomela for assistance in capturing the new images.

REFERENCES

[1] Ebner, Marc, “Color constancy,” in John Wiley & Sons, 2007.
[2] A. Gjisjenj, T. Gevers, and J. van de Weijer, “Computational color constancy: Survey and experiments,” IEEE Transactions on Image Processing, 2011.
[3] J. T. Barron and Y.-T. Tsai, “Fast fourier color constancy,” in CVPR, 2017.
[4] A. Chakrabarti, K. Hirakawa, and T. Zickler, “Color constancy with spatio-spectral statistics,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012.
[5] D. Cheng, B. Price, S. Cohen, and M. S. Brown, “Effective learning-based illuminant estimation using simple features,” in IEEE Conference on Computer Vision and Pattern Recognition, 2015.
[6] S. Bianco, C. Cusano, and R. Schettini, “Color constancy using CNNs,” in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015.
[7] Y. Hu, B. Wang, and S. Lin, “FC4: Fully convolutional color constancy with confidence-weighted pooling,” in IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[8] W. Shi, C. C. Loy, and X. Tang, “Deep specialized network for illuminant estimation,” in European Conference on Computer Vision. Springer, 2016.
[9] J. T. Barron, “Convolutional color constancy,” 2015 IEEE International Conference on Computer Vision, 2015.
[10] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, “Automatic color constancy algorithm selection and combination,” Pattern Recognition, 2010.
[11] F. Laakom, N. Passalis, J. Raitoharju, J. Nikkanen, A. Tefas, A. Iosifidis, and M. Gabbouj. Bag of color features for color constancy, 2019.
[12] F. Laakom, J. Raitoharju, A. Iosifidis, J. Nikkanen, and M. Gabbouj. Color constancy convolutional autoencoder, 2019.
[13] G. Hemritt, G. Finlayson, A. Gjisjenj, P. Gehler, S. Bianco, B. Funt, M. Drew, and L. Shi. “Rehabilitating the colorchecker dataset for illuminant estimation,” in Color and Imaging Conference, 2018.
[14] P.Y. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp. “Bayesian color constancy revisited,” in IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[15] C. Aytekin, J. Nikkanen, and M. Gabbouj. “A Data Set for Camera-Independent Color Constancy,” IEEE Transactions on Image Processing, 2018.
[16] B. Nikola and L. Sven, “Unsupervised learning for color constancy,” in arXiv preprint arXiv:1712.00436, 2017.
[17] 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” in Journal of the Optical Society of America. A, Optics, image science, and vision, 2014.
[18] Kobus Barnard, Lindsay Martin, Brian Funt, and Adam Coath. A data set for color research. Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, 27(3):147–151, 2002.
[19] Florian Ciurea and Brian Funt. A large image database for color constancy research. In Color and Imaging Conference, volume 2003, pages 160–164. Society for Imaging Science and Technology, 2003.
[20] Sérgio MC Nascimento, Flávio P Ferreira, and David H Foster. Statistics of spatial cone-excitation ratios in natural scenes. JOSA A, 19(8):1484–1490, 2002.
[21] Michael J Vrhel, Ron Gershon, and Lawrence S Iwan. Measurement and analysis of object reflectance spectra. Color Research & Application, 19(1):4–9, 1994.
[22] The EU general data protection regulation (GDPR), “The EU general data protection regulation (GDPR),” A Practical Guide, 1st Ed., Cham: Springer International Publishing, 2017.
[23] Lilong Shi and Brian Funt. Re-processed version of the gehler color constancy dataset of 568 images, 2010.

Graham D Finlayson, Ghiaia Hemritt, Arjan Gijsjenj, and Peter Gehler. A curious problem with using the colour checker dataset for illuminant estimation. In Color and Imaging Conference, 2017.

S. D. Hordley and G. D. Finlayson. “Re-evaluating colour constancy algorithms,” in International Conference on Pattern Recognition, 2004.

B. Funt and L. Shi. The effect of exposure on maxrgb color constancy. In Human Vision and Electronic Imaging XV, 2010.

B. Funt and L. Shi. The rehabilitation of maxrgb. In Color and imaging conference, 2010.

J. Cepeda-Negrete and R.E. Sanchez-Yanez. Gray-world assumption on perceptual color spaces. In Image and Video Technology, pages 493–504, 2014.

Al. Rizzi, C. Gatta, and D. Marini. Color correction between gray world and white patch. In The International Society for Optical Engineering, 2002.

Joost Van De Weijer, Theo Gevers, and Arjan Gijsjenj. Edge-based color constancy. IEEE Transactions on image processing, 16(9):2207–2214, 2007.

G. Finlayson and E. Trezzi. Shades of gray and colour constancy. In Color Imaging Conference, pages 37–41, 2004.

Arjan Gijsjenj, Theo Gevers, and Joost Van De Weijer. Physics-based edge evaluation for improved color constancy. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pages 581–588. IEEE, 2009.