https://doi.org/10.2352/CIC.2023.31.1.30

A general-purpose pipeline for realistic synthetic multispectral image dataset generation

Marco Buzzelli[1]***, Mikhail K. Tchobanou***[2]***, Raimondo Schettini***[1]***, Simone Bianco***[1]
[1] ***Department of Informatics Systems and Communication, University of Milano – Bicocca; Milan 20126, Italy***
[2] ***Moscow Research Center, Huawei Technologies Co. Ltd, Russia***

Abstract

A pipeline for the generation of synthetic dataset of spectral scenes, with corresponding sensor readings, is here proposed. The pipeline is composed of two main parts: Part 1: Image pixel reflectance assignment. Individual pixels from an input sRGB image dataset are replaced with appropriate reflectance spectra from a given non-image reflectance dataset. The resulting dataset of reflectance images is considered the starting point for simulated sensor acquisition. Part 2: Simulated sensor acquisition. Each spectral reflectance image in the dataset is illuminated with an illuminant spectra to produce a radiance image. The resulting dataset of radiance images is then synthetically read from the simulated sensors (camera and ambient multispectral sensor) of the Huawei P50 phone, using the corresponding sensors transmittance information. The capability of generating any large-scale, diverse, and annotated synthetic spectral datasets can facilitate the development of data-driven imaging algorithms, and foster reproducible research.

Introduction

Spectral imaging, which captures rich spectral information at each pixel, has emerged as a powerful technique with applications in various fields, including remote sensing, biomedical imaging, material analysis, digital photography, computational photography and more. However, the availability of high-quality spectral datasets is often limited, hindering the development and evaluation both of spectral and RGB imaging algorithms.

In recent years, the use of synthetic datasets has gained prominence as a valuable resource for training and evaluating computer vision algorithms. Synthetic data provides the flexibility to control various imaging parameters, facilitating algorithm development and benchmarking in controlled settings. However, most existing synthetic datasets focus on conventional RGB images, neglecting the crucial spectral dimension. Spectral datasets, on the other hand, are often limited to patch-based (e.g. “Mondrian-like [1]”) content, which for example limits the applicability of advanced methods for spectral reconstruction that rely on the analysis of scene content.

To address this gap, we present a comprehensive pipeline for generating synthetic datasets specifically tailored for spectral images. Our pipeline aims to provide researchers with a diverse and content-customizable source of spectral images, enabling the advancement of spectral imaging algorithms and applications.

The benefits of synthetic spectral datasets are manifold. Firstly, they offer a controlled environment for algorithm development, allowing researchers to assess the performance of spectral imaging techniques under various conditions, such as different illuminants, surface reflectance properties, and atmospheric effects. This controlled experimentation may help in understanding the limitations and strengths of different algorithms. Furthermore, synthetic datasets offer a practical solution to the scarcity of real-world spectral datasets, which often suffer from limited coverage, restricted access, or prohibitive costs. The availability of large-scale, diverse, and annotated synthetic spectral datasets can bridge this gap, facilitating the development of data-driven algorithms and fostering reproducible research. Our synthetic dataset generation also enables the exploration of new applications and algorithms that leverage spectral information. For instance, spectral image analysis plays a crucial role in precision agriculture, where the identification of crop health, disease detection, and nutrient analysis can be enhanced using spectral data. Additionally, in the field of material analysis, synthetic spectral datasets can aid in material identification, classification, and characterization tasks.

Our pipeline consists of two main components: image pixel reflectance assignment, and simulated sensor acquisition under the chosen light. These components work together to produce high-quality synthetic datasets that mimic real-world image acquisition processes and exhibit realistic visual characteristics. We evaluate the fidelity of our pipeline, using the INTEL-TAU dataset as an example benchmark [2]. To optimize the dataset generation pipeline and reduce computational time, we introduce a lookup-table (LUT) mechanism, through which we achieve a significant reduction in dataset generation time. We quantify the performance improvement through experimentation and highlight the efficiency gained by implementing the LUT mechanism.

Methodology

Our pipeline consists of two main research components: reflectance assignment and simulated sensor acquisition.

The reflectance assignment component focuses on assigning appropriate reflectance spectra to the initial dataset of RAW and sRGB images. While the reflectance assignment component lays the foundation for generating realistic synthetic datasets, the simulated sensor acquisition component simulates the image formation process that occurs in digital cameras or Ambient light Multispectral Sensors (AMS). Figure 1 provides an overview of the pipeline architecture.

Image Pixel Reflectance Assignment

The reflectance assignment step in our pipeline involves generating a synthetic dataset of spectral reflectance images based on existing image and spectra datasets. This step is crucial in creating a diverse and representative dataset that aligns with the content

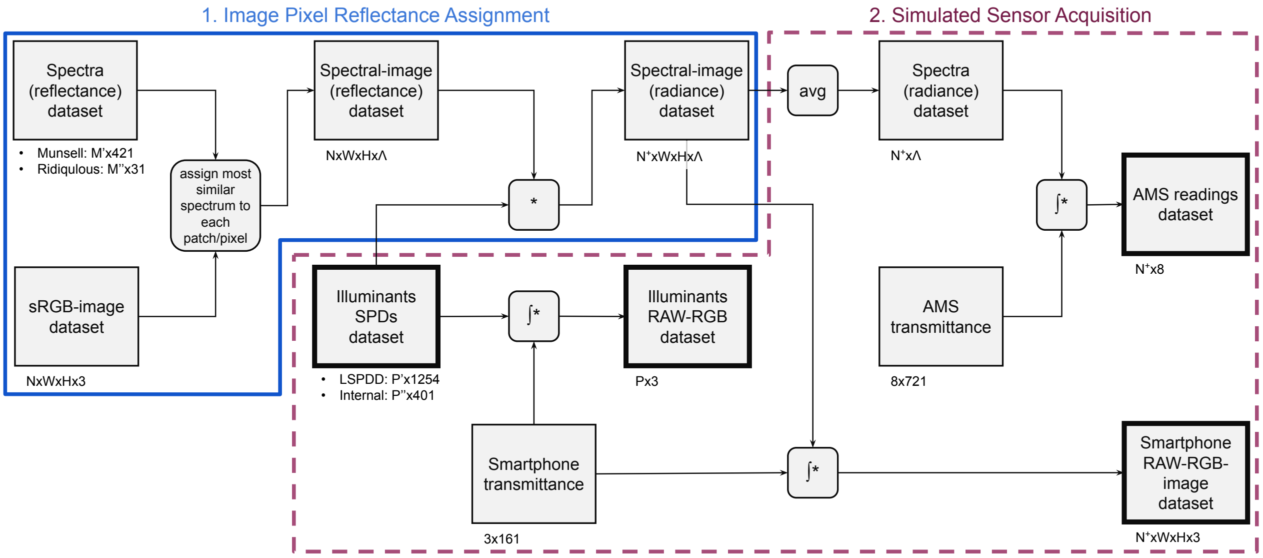
1. Image Pixel Reflectance Assignment
2. Simulated Sensor Acquisition
Spectra (reflectance) dataset
• Munsell: M'x421
• Ridiqulous: M''x31
assign most similar spectrum to each patch/pixel
Spectral-image (reflectance) dataset
NxWxHxΛ
*
Spectral-image (radiance) dataset
N⁺xWxHxΛ
avg
Spectra (radiance) dataset
N⁺xΛ
∫*
AMS readings dataset
N⁺x8
sRGB-image dataset
NxWxHx3
Illuminants SPDs dataset
• LSPDD: P'x1254
• Internal: P''x401
∫*
Illuminants RAW-RGB dataset
Px3
AMS transmittance
8x721
Smartphone transmittance
3x161
∫*
Smartphone RAW-RGB-image dataset
N⁺xWxHx3

Figure 1. Overview of the synthetic dataset generation pipeline. Different cardinalities are reported below each data element, including: the number of reflectance spectra (M), the number of sRGB images (N) with weight W and height H, the number of spectral channels for the output images (Λ), and the number of spectral illuminants (P).

distribution of existing sRGB datasets while incorporating accurate spectral reflectance information. The approach of assigning metameric reflectance spectra to RGB pixels has been presented in the past [3], applied in the context of computer graphics scenes. While this approach has the significant advantage of enabling the modelling of mutual surface and complex materials interaction, it depends on the costly creation of ad-hoc three-dimensional models. In our case, the initial sRGB dataset is selected based on specific requirements, including a wide range of subjects and neutral or known illuminants. In our pipeline, we utilize the INTEL-TAU dataset [2] as the starting point for synthetic dataset generation, as documented later on.

The scheme of this first part of the dataset generation pipeline is detailed in Figure 2, assuming the availability of a RAW-image dataset instead of an already sRGB image dataset.

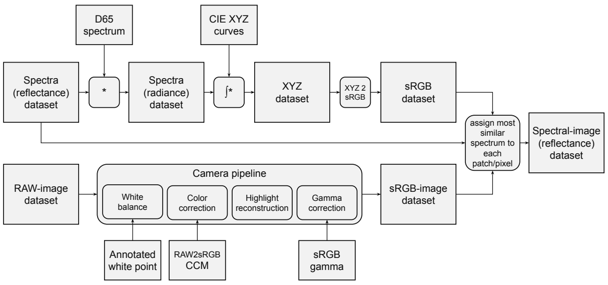

Figure 2. Detailed pipeline of part 1 of synthetic dataset generation: image pixel reflectance assignment.

For the actual reflectance data, we rely on existing spectra datasets that provide accurate measurements of spectral reflectance across different materials and surfaces. These datasets serve as a reliable source of ground truth reflectance information, enabling the creation of realistic spectral images.

During the reflectance assignment process, we determine the most similar sRGB triplet for each pixel in the synthetic images. This involves finding the reflectance spectrum that best matches the desired sRGB values. By comparing the spectral characteristics of the available reflectance spectra, we identify the spectrum that produces the closest resemblance to the target sRGB triplet.

Once the reflectance spectrum is identified, we assign it to the corresponding pixel in the synthetic image. This step ensures that each pixel in the generated dataset has an associated spectral reflectance value that aligns with the desired sRGB color. By performing this assignment for every pixel, we create a dataset with pixel-level accuracy in terms of spectral reflectance information.

It is important to note that once the synthetic dataset is generated, the original sRGB reference becomes irrelevant. The newly created spectral image dataset is considered the ground truth for subsequent steps and experiments. This allows researchers to work with a dataset that provides accurate spectral information and facilitates the evaluation and comparison of spectral imaging algorithms.

Lookup-Table (LUT) Computation

To optimize the time required for reflectance assignment, we employ a lookup-table (LUT) mechanism. The LUT is precomputed once for a given reflectance spectra dataset and similarity function. It covers all possible sRGB triplets, allowing for efficient and fast retrieval of the most similar reflectance spectrum during the dataset generation process.

The LUT computation involves the following steps:

1. Preprocessing: The reflectance spectra dataset is preprocessed to remove noise and irrelevant information. Additionally, any necessary normalization or transformation is applied to ensure consistency.
2. Similarity Function Selection: A similarity function is chosen to measure the similarity between sRGB triplets and re-

flectance spectra. The choice of similarity function depends on specific requirements and research goals.

3. LUT Construction: For each sRGB triplet, the LUT is constructed by finding the most similar reflectance spectra based on the selected similarity function. This process involves comparing the sRGB triplet with all reflectance spectra in the dataset and storing the corresponding indices of the most similar spectra in the LUT.

Once the LUT is computed, it can be utilized during the reflectance assignment phase to significantly reduce computational time.

Simulated Sensor Acquisition

The simulated sensor acquisition component replicates the image formation process that occurs in digital cameras and Ambient-light Multispectral Sensors. This process involves integrating surface reflectance information, illuminant sources, and specific transmittance channels to compute the observations captured by the imaging sensors.

The simplified equation for image formation is as follows:

$$O^{(\rho)} = \int_{\lambda} R(\lambda) \cdot I(\lambda) \cdot S^{(\rho)}(\lambda) d\lambda \qquad (1)$$

Here, $O^{(\rho)}$ represents the observation at channel ρ, obtained by integrating the product of surface reflectance $R(\lambda)$, illuminant source $I(\lambda)$, and specific transmittance channel $S^{(\rho)}(\lambda)$ over the wavelength λ. In the case of an AMS, the spatial information related to radiance data ($R(\lambda) \cdot I(\lambda)$) is spatially averaged according to specifications before applying the specific transmittance channel and integrating the result.

Data selection and filtering

Reflectance data

The considered dataset of (non-image) reflectance spectral data are summarized in Table 1.

Table 1. Summary statistics of considered reflectance datasets.

Name	# spectra	Range (nm)	# bands	step (nm)
Munsell [4]	1269	380-800	421	1
Ridiqulous [5]	43M / 7M / 114K	400-700	31	10

The chromaticity of the reflectance spectra are visualized in Figure 3 under D65 reference illuminant. The Ridiqulous dataset is composed from various sources of multispectral and hyperspectral data (full list at [5]). Three versions of the dataset are publicly available, with spectra clustered at various levels to reduce redundancy.

The Ridiqulous dataset was found to be an adequate source of information for synthetic dataset generation, specifically as the 7M and the 100K versions which reduce redundancy and allow for agile data management. However, some of the involved spectra are characterized by saturated and/or highly-quantized bands. We therefore defined an automated procedure, illustrated in Figure 4, aimed at automatically removing such elements:

- We characterize each spectrum based on $v1$: the maximum number of repeated consecutive elements in the spectrum.

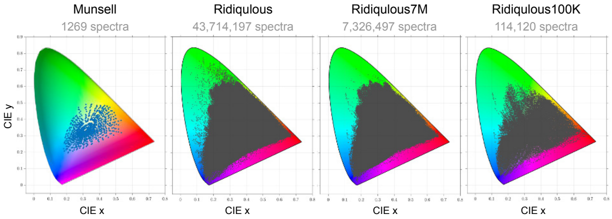

Figure 3. Chromaticity (under D65) of the illuminants of considered spectra reflectance datasets.

- We eliminate all spectra where $v1 \geq 3$ (out of 31 bands/samples).

After filtering, Ridiqulous100K had 365/114120 spectra rejected (0.32%), and Ridiqulous7M had 33017/7326497 spectra rejected (0.45%).

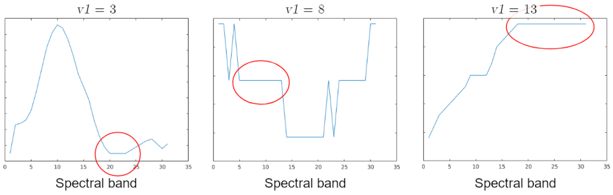

Figure 4. Examples of the characterization of the Ridiqulous spectra based on $v1$: the maximum number of repeated consecutive elements in the spectrum. Samples with $v1 \geq 3$ are removed from the dataset.

Illuminant data

The considered datasets of spectral illuminant data are summarized in Table 2.

Table 2. Summary statistics of considered illuminants datasets.

Name	# spectra	Wavelength range (nm)	# bands	step (nm)
LSPDD [6]	307	273-899.5	1254	0.5
Internal	62	380-780	401	1

The LSPDD dataset offers a wide set of illuminants, with fine spectral resolution, and a wide variety of classes. Due to limitations in the corresponding license, however, we focus our experimental setup on using an internal source of illuminant spectral data. Figure 5 offers a detailed view of the 62 illuminants in the Internal illuminant dataset, as grouped by class information (fluorescent, LED, CIE-D series, CIE-A series).

Transmittance data and AMS characterization

The considered sensors for synthetic dataset generation, are summarized in Table 3.

Table 3. Summary statistics of considered sensors.

Type	# spectra	Range (nm)	# bands	step (nm)
Camera	3	400-720	161	2
AMS	8 out of 10	380-1100	721	1

The camera is that of a Huawei P50 smartphone. The AMS is characterized by 10 spectra (channels), of which only the first

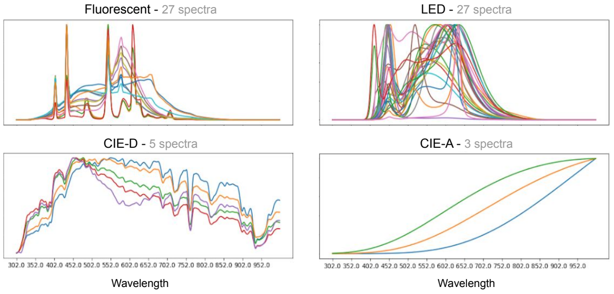

Figure 5. *Illuminant spectra of our internal illuminant dataset, grouped by class.*

8, contained in the visible spectrum, are considered within this research project.

Image dataset

Existing image datasets with spectral reflectance information are available, but they are limited in size and content distribution (e.g. NUS (2014) [7], Stanford HS (2013) [8], Stanford MS (2008) [9], CAVE (2008) [10], MS Image db (2004) [11]). We are therefore generating a synthetic dataset of spectral reflectance images according to these criteria:

- The image content distribution is guided by existing sRGB datasets.
- The actual reflectance data comes from existing spectra datasets.

For the choice of initial RAW (and consequently sRGB) dataset, we set the following requirements:

- Wide range of subjects
- Neutral or known illuminants

To this extent, we exploit images from AWB (Automatic White Balance) datasets, which are provided with a known illuminant that can be neutralized before application of arbitrary illuminants. We rely on our previous analyses [12, 13], synthesized and illustrated in Figure 6. The INTEL-TAU dataset [2] is characterized by high cardinality and a fair distribution of image content, and it is therefore selected as the starting point for synthetic dataset generation.

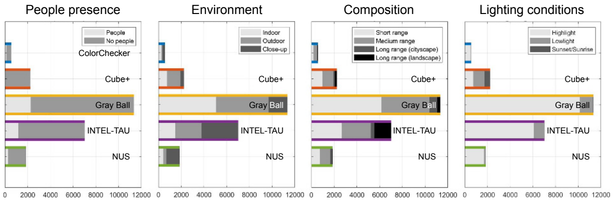

Figure 6. *Cardinality and class distribution of popular AWB datasets.*

Experimental validation

Fidelity validation for reflectance assignment

We validate the fidelity of part 1 of the dataset generation pipeline, by comparing the "sRGB image dataset" with the "Spectral-image (radiance) dataset", as indicated in Figure 2. We perform these comparisons by measuring distances in several color spaces: RGB as it is directly accessible and consequently characterized by low computational requirements, CIELab and ProLab [14] as they are designed to be perceptually uniform. We then evaluate perceptually the results using both ΔE_{76} and ProLab Euclidean distances. Additional experiments might be performed in the future using other distances, including variations of CIELab's ΔE_{76} such as ΔE_{00}. These validations allow us to assess the accuracy of the reflectance assignment process and make informed decisions regarding the choice of reflectance datasets, spectral sampling, and the space for assigning the closest spectrum.

Table 4. Fidelity validation of part 1 of our synthetic dataset generation pipeline (reflectance assignment).

		Optimization	**Evaluation**	
Sampling	**Reflectance db**	**Space**	ΔE_{76}	**proLab eu.d.**
1nm	Munsell	RGB	7.148	5.010
1nm	Munsell	ProLab	6.685	4.699
1nm	Munsell	CIELab (ΔE_{76})	5.328	5.653
1nm	Ridiqulous100K	RGB	1.162	0.625
1nm	Ridiqulous100K	ProLab	1.144	0.535
1nm	Ridiqulous100K	CIELab (ΔE_{76})	0.895	0.725
10nm	Munsell	RGB	7.155	4.997
10nm	Munsell	ProLab	6.702	4.677
10nm	Munsell	CIELab (ΔE_{76})	5.335	5.648
10nm	Ridiqulous100K	RGB	1.160	0.624
10nm	Ridiqulous100K	ProLab	1.144	0.534
10nm	Ridiqulous100K	CIELab (ΔE_{76})	0.895	0.725
10nm	Ridiqulous7M	ProLab	0.227	0.114
10nm	Ridiqulous7M	CIELab (ΔE_{76})	0.180	0.149

Results are reported in Table 4. The following observations can be derived from the results.
Sampling:

- The advantage of having a finer sampling (1 nm vs 10 nm) is not evident at this step
  - Recall that Ridiqulous comes with a 10nm step, and Munsell with a 1 nm step. Sprague interpolation [15] was used to augment the spectral resolution of Ridiqulous.

Reflectance dataset:

- Ridiqulous100K has a clear advantage w.r.t. to Munsell:
  - ~1 unit in ΔE_{76} vs. ~5 in ΔE_{76}.
- Ridiqulous 7M has an even larger advantage than Ridiqulous100K:
  - ~0.2 in ΔE_{76} vs. ~1 in ΔE_{76}.
  - Limited experiments due to time requirements.

Space for assigning the closest spectrum:

- As expected, optimizing for ΔE_{76} will produce the best solution according to ΔE_{76}, and optimizing for proLab will produce the best solution according to proLab
- Interestingly, when evaluating with proLab, the second best solution is optimizing for RGB, not ΔE_{76}.

Table 5. Fidelity validation for simulated sensor acquisition, depending on the RAW-to-XYZ correction matrix optimized on different illuminants combinations (rows), as tested on different illuminant conditions (columns, with illuminant cardinality in parentheses). Using a linear 3×3 matrix (top), or 10×3 polynomial expansion. The main table reports $\Delta E76$ on the XYZ computed from radiance spectra vs. the XYZ computed from the camera RAW-RGB readings. The last column reports the intrinsic optimization error.

3 × 3		Test illuminants						Matrix optim.
		A (3)	D65 (1)	E (1)	FL (15)	LED (1)	ALL (21)	ΔE_{76}
Train illuminants	A (3)	1.159	3.810	3.870	3.397	1.187	3.014	1.541
	D65 (1)	3.170	1.405	1.603	2.492	2.137	2.478	1.311
	E (1)	2.777	1.346	1.508	2.500	1.911	2.409	1.393
	FL (15)	4.251	2.319	2.447	1.719	3.498	2.228	2.395
	LED-RGB (1)	2.331	7.896	7.052	5.501	0.526	4.999	0.705
	ALL (21)	2.407	1.370	1.377	2.028	2.122	2.024	3.073

10 × 3		Test illuminants						Matrix optim.
		A (3)	D65 (1)	E (1)	FL (15)	LED (1)	ALL (21)	ΔE_{76}
Train illuminants	A (3)	1.422	7.417	6.892	4.955	1.639	4.502	1.346
	D65 (1)	2.325	1.784	1.830	4.329	1.665	3.675	1.358
	E (1)	2.213	1.932	1.903	4.888	1.974	4.084	1.444
	FL (15)	4.560	2.782	2.818	1.736	3.897	2.344	2.319
	LED-RGB (1)	2.214	11.248	9.475	6.000	0.584	5.617	0.698
	ALL (21)	3.474	2.297	2.036	2.049	4.045	2.359	2.919

An example of sRGB images rendered after the image pixel reflectance assignment step is provided in Figure 7, visually highlighting the difference in using the Munsell or the Ridiqulous7M as the spectra reflectance dataset.

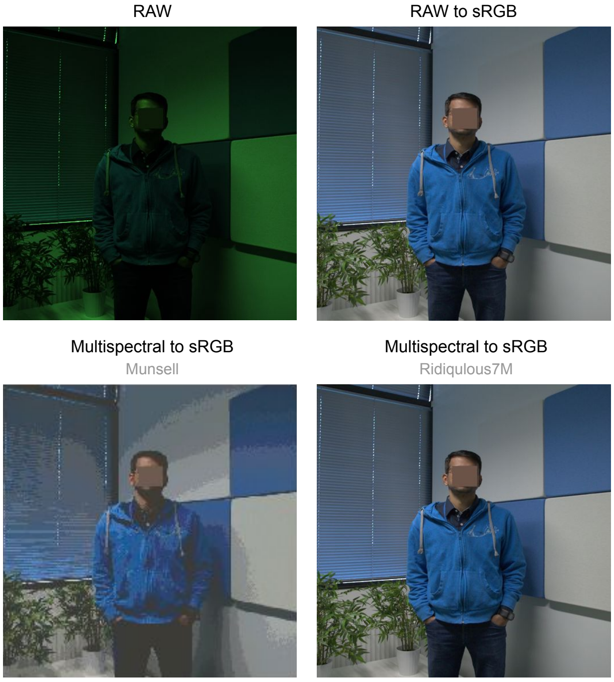

Figure 7. *Example image in its original RAW format, in sRGB representation obtained directly from the RAW, and in sRGB obtained after multispectral reflectance assignment (using Munsell or Ridiqulous as the spectra reflectance dataset).*

Code optimization

After optimization, the time to generate a synthetic dataset is ∼0.2 seconds per image using the Ridiqulous100K spectra, which translates to **∼20 minutes** for 5000 images @ 500×400 pixels (0.46% of the original time). The LUT computation step itself (to be performed once per spectra dataset / similarity function) takes ∼45 minutes for the Ridiqulous 100K spectra, and ∼48 hours for the Ridiqulous7M spectra.

Fidelity validation for simulated sensor acquisition

The fidelity and correctness of the simulated sensor acquisition part can be verified for the camera sensor by testing the difference between:

- The XYZ computed from radiance spectra.
- The XYZ computed from the camera RAW-RGB readings.

In order to obtain XYZ coordinates from the camera RAW-RGB readings, the validation involves the creation of a RAW-to-XYZ color correction matrix (CCM), optimized using reflectance spectra datasets. We evaluate the optimization process and matrix usage under different illuminants and expansion techniques, providing insights into the performance and generalization capabilities of the matrix.

The procedure for this validation is visualized in Figure 8. This validation first requires the creation of a RAW-to-XYZ color correction matrix (indicated as “Smartphone CCM” in figure). Such matrix is optimized as a Moore-Penrose pseudoinverse [16] starting from a dataset of reflectance spectra (e.g. Munsell or Ridiqulous):

$$\begin{aligned} 3\times 3 &: \text{only RGB} \\ 10\times 3 &: \text{polynomial } (R, G, B, R^2, G^2, B^2, RG, RB, GB, 1) \end{aligned}$$

According to preliminary experiments in RAW-to-XYZ matrix generation, the more simple Munsell dataset allows for better generalization than any version of the Ridiqulous dataset. For this

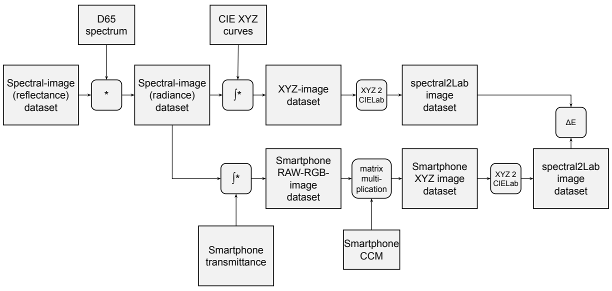
D65 spectrum
CIE XYZ curves
Spectral-image (reflectance) dataset
*
Spectral-image (radiance) dataset
∫*
XYZ-image dataset
XYZ 2 CIELab
spectral2Lab image dataset
ΔE
∫*
Smartphone RAW-RGB-image dataset
matrix multi-plication
Smartphone XYZ image dataset
XYZ 2 CIELab
spectral2Lab image dataset
Smartphone transmittance
Smartphone CCM

Figure 8. *Fidelity validation procedure for the simulated camera acquisition described in part 2 of the dataset generation pipeline.*

reason, in the following evaluation we will use the Munsell dataset specifically for matrix generation. We use either A (3), D65 (1), E (1), FL (15), LED-RGB (1), or all of the above illuminants (21), with or without polynomial expansion. We evaluate two steps:

- The matrix optimization itself (optimization error).
- The matrix usage for simulated camera reading within the dataset generation pipeline.

Results are presented in Table 5 under the following conditions:

- We test on Ridiqulous7Mf (filtered), using either illuminant.
- We use a sample of 10 representative images.

In the current experimental setup, polynomial expansion produces worse (or equivalent) performance w.r.t. a simpler 3×3 matrix. Specifically, polynomial expansion has a positive effect mainly when the images are illuminated by class A illuminants. An improvement is also observed on LED-RGB-illuminated images, when the matrix is optimized using D65.

Conclusions

We presented a pipeline for generating synthetic datasets for spectral imaging research, starting from any sRGB image collection. Our pipeline simulates the image formation process and produces high-quality synthetic datasets for evaluating spectral and RGB imaging algorithms. By employing a precomputed look-up table mechanism, we optimized the reflectance assignment stage, significantly reducing dataset generation time. This allows for efficient exploration of different datasets and similarity functions.

The simulated sensor acquisition stage accurately models the image formation process using reflectance spectra, illuminant sources, and specific transmittance channels. Leveraging the INTEL-TAU dataset, we ensured fidelity and accuracy in the synthetic datasets. To validate the fidelity of simulated sensor acquisition, we compared computed XYZ values from radiance spectra with those from camera RAW-RGB readings, ensuring accuracy and reliability of the synthetic datasets.

In conclusion, our pipeline provides a valuable resource for spectral imaging research, enabling the evaluation and improvement of spectral and RGB imaging algorithms across various domains. Future work includes optimizing the reflectance assignment process, exploring alternative similarity functions, and expanding the pipeline to incorporate additional sensor models or spectral imaging modalities. We anticipate that this work will inspire new avenues of exploration and innovation, leading to further progress in the field. Furthermore, the benefit of using the generated dataset, in comparison with existing datasets, will be demonstrated in multispectral applications in future works.

References

[1] Yi-Tun Lin and Graham D Finlayson. An investigation on worst-case spectral reconstruction from rgb images via radiance mondrian world assumption. *Color Research & Application*, 48(2):230–242, 2023.

[2] Firas Laakom, Jenni Raitoharju, Jarno Nikkanen, Alexandros Iosifidis, and Moncef Gabbouj. INTEL TAU: A color constancy dataset. *IEEE Access*, 9:39560–39567, 2021.

[3] Xiangpeng Hao and Brian Funt. A multi-illuminant synthetic image test set. *Color Research & Application*, 45(6):1055–1066, 2020.

[4] Munsell. Download spectra sets, 1976. http://cs.joensuu.fi/pages/mhk/ColorDB/color/database/download.htm (Accessed on 1st April 2022).

[5] Qiu Jueqin. Spectral reflectance database from hyperspectral images - RidiQulous, 2016. https://ridiqulous.com/spectral-reflectance-database/ (Accessed on 1st April 2022).

[6] Johanne Roby and Martin Aubé. LSPDD — Light Spectral Power Distribution Database, 2015. https://lspdd.org/app/en/lamps (Accessed on 28 February 2022).

[7] Rang MH Nguyen, Dilip K Prasad, and Michael S Brown. Training-based spectral reconstruction from a single rgb image. In *European Conference on Computer Vision*, pages 186–201. Springer, 2014.

[8] Torbjørn Skauli and Joyce Farrell. A collection of hyperspectral images for imaging systems research. In *Digital Photography IX*, volume 8660, page 86600C. International Society for Optics and Photonics, 2013.

[9] Manu Parmar, Francisco Imai, Sung Ho Park, and Joyce Farrell. A database of high dynamic range visible and near-infrared multispectral images. In *Digital photography iv*, volume 6817, page 68170N. International Society for Optics and Photonics, 2008.

[10] Fumihito Yasuma, Tomoo Mitsunaga, Daisuke Iso, and Shree K Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing*, 19(9):2241–2253, 2010.

[11] Steven Hordley, Graham Finalyson, and Peter Morovic. A multispectral image database and its application to image rendering across illumination. In *International Conference on Image and Graphics*, pages 394–397. IEEE, 2004.

[12] Simone Bianco, Marco Buzzelli, Gianluigi Ciocca, Raimondo Schettini, Mikhail Tchobanou, and Simone Zini. Analysis of biases in automatic white balance datasets. In *Proceedings of the International Colour Association (AIC) Conference 2021. Milan, Italy. AIC*, pages 233–238, 2021.

[13] Marco Buzzelli, Simone Zini, Simone Bianco, Gianluigi Ciocca, Raimondo Schettini, and Mikhail K Tchobanou. Analysis of biases in automatic white balance datasets and methods. *Color Research & Application*, 48(1):40–62, 2023.

[14] Ivan A Konovalenko, Anna A Smagina, Dmitry P Nikolaev, and Petr P Nikolaev. Prolab: A perceptually uniform projective color coordinate system. *IEEE Access*, 9:133023–133042, 2021.

[15] Stephen Westland. *Interpolation of Spectral Data*, pages 1–3. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[16] Roger Penrose. A generalized inverse for matrices. In *Mathematical proceedings of the Cambridge philosophical society*, volume 51, pages 406–413. Cambridge University Press, 1955.