DOI: 10.1111/cgf.15280

COMPUTER GRAPHICS *forum*
Volume 0 (2025), number 0, e15280

Constrained Spectral Uplifting for HDR Environment Maps

L. Tódová and A. Wilkie

Charles University, Prague, Czech Republic
todova@gmail.com, wilkie@cgg.mff.cuni.cz

Abstract

Spectral representation of assets is an important precondition for achieving physical realism in rendering. However, defining assets by their spectral distribution is complicated and tedious. Therefore, it has become general practice to create RGB assets and convert them into their spectral counterparts prior to rendering. This process is called spectral uplifting. *While a multitude of techniques focusing on reflectance uplifting exist, the current state of the art of uplifting emission for image-based lighting consists of simply scaling reflectance uplifts. Although this is usable insofar as the obtained overall scene appearance is not unrealistic, the generated emission spectra are only metamers of the original illumination. This, in turn, can cause deviations from the expected appearance even if the rest of the scene corresponds to real-world data. In a recent publication, we proposed a method capable of uplifting HDR environment maps based on spectral measurements of light sources similar to those present in the maps. To identify the illuminants, we employ an extensive set of emission measurements, and we combine the results with an existing reflectance uplifting method. In addition, we address the problem of environment map capture for the purposes of a spectral rendering pipeline, for which we propose a novel solution. We further extend this work with a detailed evaluation of the method, both in terms of improved colour error and performance.*

Keywords: HDR, image-based lighting, rendering, spectral uplifting

CCS Concepts: • Computing methodologies → Rendering

1. Introduction

The last few years have seen a significant improvement in the field of spectral rendering due to its ability to simulate light transport in a physically correct manner. As opposed to tristimulus rendering, where the colours are represented as an RGB value, spectral renderers model colour stimuli as they occur in nature, which is a distribution of wavelengths. This allows for a physically correct simulation of light transport and results in more realistic calculations of reflection, transmission, absorption and so forth. An additional benefit is its capability to simulate natural phenomena, such as fluorescence or phosphorescence.

However, directly creating spectral assets, such as textures or environment maps, is, in most cases, rather difficult. Either their real-life counterparts need to be precisely measured with a spectrometer (or, in case of environment maps, captured with a hyperspectral camera), which is a tedious and complicated process, or they can be modelled in the spectral space, which is very unintuitive for artists. For the purposes of VFX workflows, the preferred pipeline is to model or capture assets in the RGB space, and then convert them into the spectral domain. This conversion process is called *spectral uplifting*.

The relationship between the spectral and the RGB domain is not bijective, as multiple different spectra, called *metamers*, attain the same value when converted to RGB. Therefore, the process of spectral uplifting is not straightforward and a number of techniques have been proposed. The selection of a suitable technique depends on the scene at hand—for example, when uplifting reflectances of objects usually found in nature (such as wood or vegetation), methods producing smooth and simple curves are usually preferred in order to obtain results in accordance with real-life counterparts. On the other hand, the dyes of fabrics tend to have more complex spectra and would therefore benefit from a distinct uplifting approach.

However, most existing spectral uplifting techniques focus mainly on reflectance uplifting, that is, spectral curves that only have values between 0 and 1. For the purposes of emission uplifting, the current general approach is the downscaling of input RGB values into low dynamic range RGB (i.e., the RGB components are bounded by 0 and 1), and then utilising one of the existing reflectance uplifting methods. The resulting spectral curve is then scaled back into the high dynamic range. However, due to the typically more complex and spiky nature of emission spectra, the spectral power distributions obtained this way rarely resemble

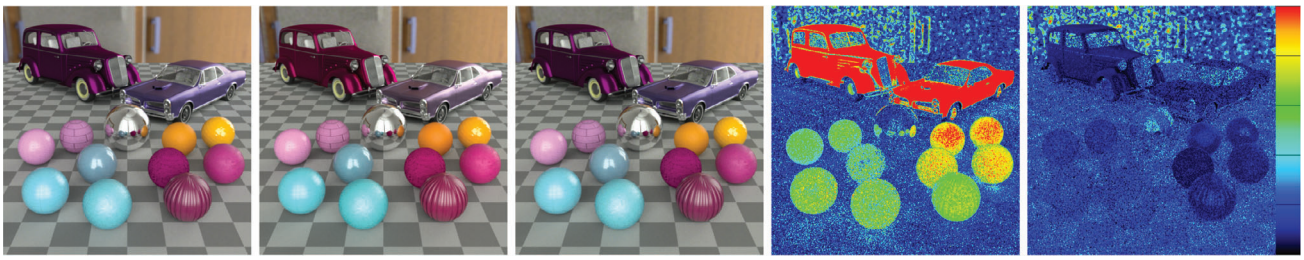

Figure 1: *Different approaches to rendering image-based lighting.* **Ground Truth** *uses a spectral HDR environment map as an input, that is, no uplifting is performed.* **Scaled Reflectance Uplift** *employs the current state-of-the-art environment map uplifting, while* **Our Uplift** *uses the technique proposed in this paper. Both uplifts take the RGB counterpart of the ground truth spectral map as input. The error images are relative to CIE Delta E 2000 = 5, and the reflectance measurements of the objects in the scene are from the Pantone Atlas. Note that the environment map was obtained as a spectral render rather than a capture for increased precision for comparison purposes.*

their real-life counterparts. This, in turn, may cause metameric artefacts in the final render.

This is a problem in the VFX industry, where it is common to freely mix plate footage and its digital counterpart. These switches pronounce even slight deviations between the two scenes, making the discrepancies visible to the human eye. To prevent this, specific light sources present in the scene are therefore usually uplifted manually, by measuring their corresponding spectral power distributions with a spectrometer. The problematic case is that of image-based lighting, that is, HDR environment maps, where such an option is not easily possible.

In this paper, we present a technique capable of uplifting HDR environment maps for image-based lighting in a manner that simulates real-world behaviour. Our approach builds upon the observation that the emissive properties of pixels in HDR environment maps are mainly due to one dominant light source (or a small set of strong light sources). This, in turn, implies that the spectral power distribution of all pixels must be influenced by the emission spectra of these light sources. With our method, we identify these sources and utilise their spectral power distributions to constrain a universal reflectance uplifting technique in order to obtain plausible uplifts for HDR environment maps.

In the VFX industry, currently, the generally used approach to camera calibration for environment map capture is to use a colour chart. However, such captures do not contain enough information for the proper identification of the present light sources. Therefore, we furthermore propose a novel way in which to perform camera calibration. In addition to maximising the accuracy of our technique, we explain how the process allows the capture to retain valuable spectral information.

2. Background

2.1. Image-based lighting

In order to achieve a realistic scene appearance, especially in outdoor or more complex settings, the VFX industry relies on the use of image-based lighting via HDR environment maps. During path tracing, each pixel of the map is treated as a separate parallax-free emitter located at infinity. Although environment maps mainly serve as a tool that aids realism, their incorrect capture and colour calibration can have negative impact on the final render by causing an undesired tint. As different cameras can have distinct spectral response curves, it is easy to obtain such erroneous captures without prior calibration.

Generally, camera calibration is performed with the help of a colour target, such as the Macbeth Colour Checker [MMD*76], which is a chart with colourful patches that have known reflectance properties. Every patch is assigned a *target RGB value*. First, captures are performed with the colour target in the scene. The captured RGB values of the patches are expected to deviate from the target values, and based on the differences, a colour correction profile is determined. This profile is then applied to the actual environment map footage. The RGB values of the patches are selected so that they provide a wide coverage of the RGB space, thus ensuring better consistency of colour appearance regardless of the camera sensitivity curves.

Although this is a sufficient approach for use in RGB renderers, we note that the proposed calibration is, only a mean of colour correction in order to simulate human perception and therefore achieve a more realistic appearance. It does not provide any meaningful information about the colour properties of the original scene—for example, if the captured patches have a yellow tint, there is no way of identifying whether it is due to the spectral response curves of the camera or, for example, due to a yellowish fluorescent light source. Such captures therefore render the identification of spectral properties of individual objects impossible.

An important tool that addresses the problem of environment map capture for the spectral rendering pipeline, called PhysLight, was introduced by Langlands and Fascione [LF20]. It has the capability of bringing photometric units and other physical parameters to the digital pipeline. However, it also does not attempt to identify the spectral properties of present light sources.

Currently, the only way to achieve such calibration is to use an emissive colour target, such as the camSPECS measurement device [Eng]. It works by emitting a controlled spectrum of light,

which the camera captures and uses to calibrate its colour response. However, the use of such targets is of limited practicality both due to their high cost and restricted accessibility. Their handling may be cumbersome for an average user, due to both their complex calibration and limited portability. Additionally, they are dependent on a stable, controlled environment, which is not always available.

We additionally note that even though most of the objects in an HDR environment map capture are not by themselves emissive, their emissive properties in the final image are due to the dominant light source(s) being reflected off a non-emissive surface. Specifically, the final spectral power distribution of a pixel can be computed as a per-wavelength multiplication of the reflectance spectrum $R(\lambda)$ of the object and the incoming emission $E(\lambda)$ that reaches the object. We utilise this observation in this work.

2.2. Spectral uplifting

Since spectral rendering as a research area is comparatively new, there is not a very large variety of spectral uplifting techniques, and many of those that exist are limited. Initially, research was not even focused on the final curve shape, but rather on satisfying the most important constraints of a proper spectral uplifting technique—specifically, a negligible round-trip error between the original RGB and the RGB the uplifted curve evaluates to, and, in case of reflectance uplifting, a proper boundary on the curve's values (i.e., it has to be bounded by 0 and 1). Successful evaluation of the techniques for the whole gamut (e.g., sRGB) also posed a problem. Therefore, although the initially proposed methods, such as by MacAdam [Mac35], Meng et al. [MSHD15] or Otsu et al. [OYH18] significantly advanced the field of spectral uplifting, they lack in the fundamental aspects.

The first widely used technique was proposed by Smits [Smi99]. Although it is also prone to minor round-trip errors and does not necessarily satisfy the [0,1] boundary constraint, it is simple and efficient for use during the spectral rendering process. Additional focus of the technique was on the physical plausibility of the uplifted spectra.

To counter the deficiencies of the method by Smits [Smi99] and to address the blockiness of the resulting spectra, Jakob and Hanika [JH19] proposed a method that can be considered state of the art for general reflectance uplifting. In their work, they present a low-dimensional parametric model for spectral representation, which stores spectra with only 3 floating-point coefficients. The simplicity of the model allows them to precompute a set of RGB to spectra mappings, which they store in a 3D table. The uplifting itself is then performed by a lookup in the table, and, in case the desired RGB value is not present, the spectra of its closest entries are interpolated. This approach both achieves smooth reflectance curve shapes similar to those found in nature and satisfies the fundamental requirements of a correct uplifting technique. As this technique was originally proposed for standard reflectances within the sRGB gamut only, the work by Tódová et al. [TWF22] extends it to support Adobe Wide Gamut RGB. To simplify the lookup, they use an evenly-spaced RGB cube as the uplift model instead of the proposed 3D table.

As the method by Jakob and Hanika [JH19] finally solved the issue of satisfying the fundamental uplifting constraints, it allowed research in the area of spectral uplifting to focus on improving the physical realism of the uplifts based on specific material properties. Their model is, for example, used as a basis for the work presented by Jung et al. [JWH*19] in order to uplift a wider colour gamut by adding fluorescent components. The work by Tódová et al. [TWF21] also takes inspiration from the pre-computed uplift model and proposes the novel idea of constrained reflectance uplifting, which allows the user to preserve specific spectral shapes during the uplifting process. This aids in preserving desired metameric artefacts when matching plate footage to its digital counterpart in the VFX industry. This method is later extended for support of Adobe Wide Gamut RGB [TWF22]. Van de Ruit and Eisemann [VDRE23] also address the problem of metamerism during uplifting. They propose a method that selects the most plausible metamer by allowing the user to define texture appearance under multiple illuminants. Belcour et al. [BBG23], on the other hand, propose a technique that uplifts to a family of metamers rather than a single spectrum, allowing the user to select the preferred uplift.

Although recent years have seen significant advancements in the area of spectral uplifting, most of the proposed techniques focus on reflectances only. The main reason for that is the lack of emissive objects in a scene in comparison to reflective surfaces. Additionally, in the VFX industry, the main light sources of the scenes are always going to be uplifted manually, by performing measurements on the set. This is because of their significant influence on the scene appearance and therefore a need for high precision.

Currently, general emission uplifting in common spectral renderers is performed by downscaling the input HDR values into the low dynamic range and utilising a known reflectance uplifting method (for example, the Mitsuba renderer uses the technique by Jakob and Hanika [JSR*22], while the Manuka renderer uses a modified version of the method by Smits et al. [FHL*18]). The resulting reflectances are then scaled back into the high dynamic range and are therefore considered emissive. In terms of the round-trip error, the results of this approach are satisfactory, however, the final spectra are only synthetic and do not correspond to real-life data. In contrast to reflectance uplifts, where this does not cause a problem, as the shapes of the synthetic spectra resemble real-life reflectances, emission spectra tend to attain much more complex and spikier shapes. Especially in the case of indoor light sources, such as LED or fluorescent lamps, the emission uplifts significantly differ from the actual measurements. This may, in turn, cause metameric artefacts in the final render. While these can be avoided by manually uplifting the light sources from measurements, this is not an option for the case of image-based lighting, where they remain prevalent.

Both our previous work [TW24] and this extension address this issue by constraining the uplifting process of input HDR environment maps with the illumination that was present during their capture.

3. Environment Map Uplifting

Our proposed uplifting method consists of two parts that can be considered separate processes. The first part involves an analysis

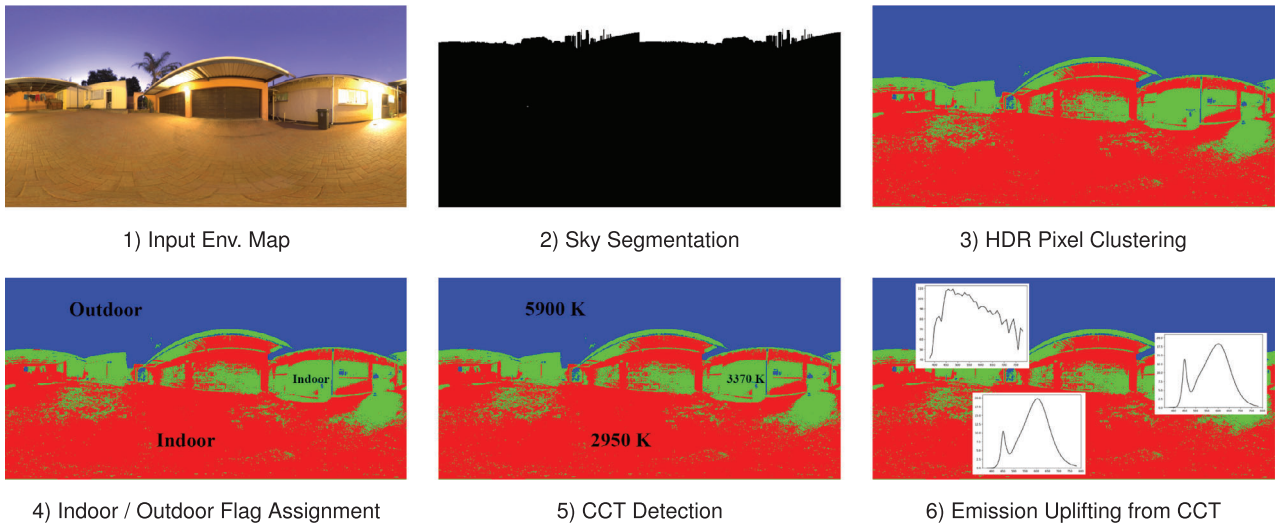

Figure 2: *Individual steps of our proposed light source detection process shown on an input environment map.*

of the input HDR environment map in order to identify its main light sources and their emission spectra. The second part consists of per-pixel uplifting of the environment map constrained by the identified illuminants.

3.1. Light source identification

In this section, we describe our proposed solution to light source identification in input HDR environment maps. Our implementation consists of multiple parts—first, we perform detection of the most luminous light sources in the map and identify which pixels they affect. Second, we determine whether the light sources are indoor or outdoor and establish their correlated colour temperatures. We use this information alongside a database of measured spectra to determine their spectral power distribution.

We provide an overview of this process applied to an input environment map in Figure 2. In the following, we describe the individual steps of the process in more detail.

3.1.1. *HDR pixel clustering*

At first glance, light source identification in an input RGB environment map image appears to be an image segmentation problem requiring a robust trained neural network. However, our analysis of the HDR pixels of existing environment maps gives rise to a more elegant solution.

Every neutral light source has a specific correlated colour temperature—for example, indoor tungsten sources tend to have warm temperatures around 2000 K, while daylight illumination typically reaches the temperature of 6500 K. This results in a colour cast that is transferred to the objects illuminated by the light source. Specifically, in HDR environment maps, pixels are tinted towards yellow or blue depending on which light source they are affected by the most. While the colour difference between pixels in the lower dynamic range (i.e., the pixels' luminosity is lower than the average in the image) is too small to be able to distinguish any type of colour cast, in the higher dynamic range, even slight shifts in correlated colour temperature represent significant changes in the RGB colour distance.

In Figure 3, we present a set of HDR environment maps and a visualisation of their pixels in the 3D RGB space. Although the pixels with lower luminosity values tend to create a seemingly random cluster in the low dynamic range, in the high dynamic range, we can observe rather distinguishable clusters forming around certain axes. We call these clusters the *light source clusters*, since the highest luminosity values of each of them belong to light sources of a similar temperature. Upon closer inspection, we further observe that the rest of the pixels in each cluster belong to areas that are primarily illuminated by light sources of this temperature.

In order to identify the visible clusters, we implement a modified version of the K-Means clustering algorithm. For each input point, in addition to storing its RGB coordinates, we also compute its spherical coordinates - specifically the angles θ and ϕ. The means (centroids) also carry this information. When assigning a point to its nearest mean, as opposed to the least squared Euclidean distance that is used as a distance metric in the standard K-Means algorithm, we utilise the least squared distance between the θ and ϕ angles. By omitting the radial distance r from the calculation, we force the algorithm to create clusters based on the points' relative angular position from the origin rather than their Euclidean distance from each other. This, in turn, forces the points to cluster around specific axes.

As they have no informative value for the algorithm and their processing only hinders performance, we omit the pixels in the low dynamic range from the clustering. We assign them to their corresponding clusters only after its termination, by utilising the same distance metric as in our modified version of the K-Means algorithm.

The final aspect left to determine is the number of clusters. As the standard K-Means algorithm takes the desired number of clusters as a parameter on input (denoted k), the general approach is to iteratively run the algorithm for an increasing k, and design a cluster

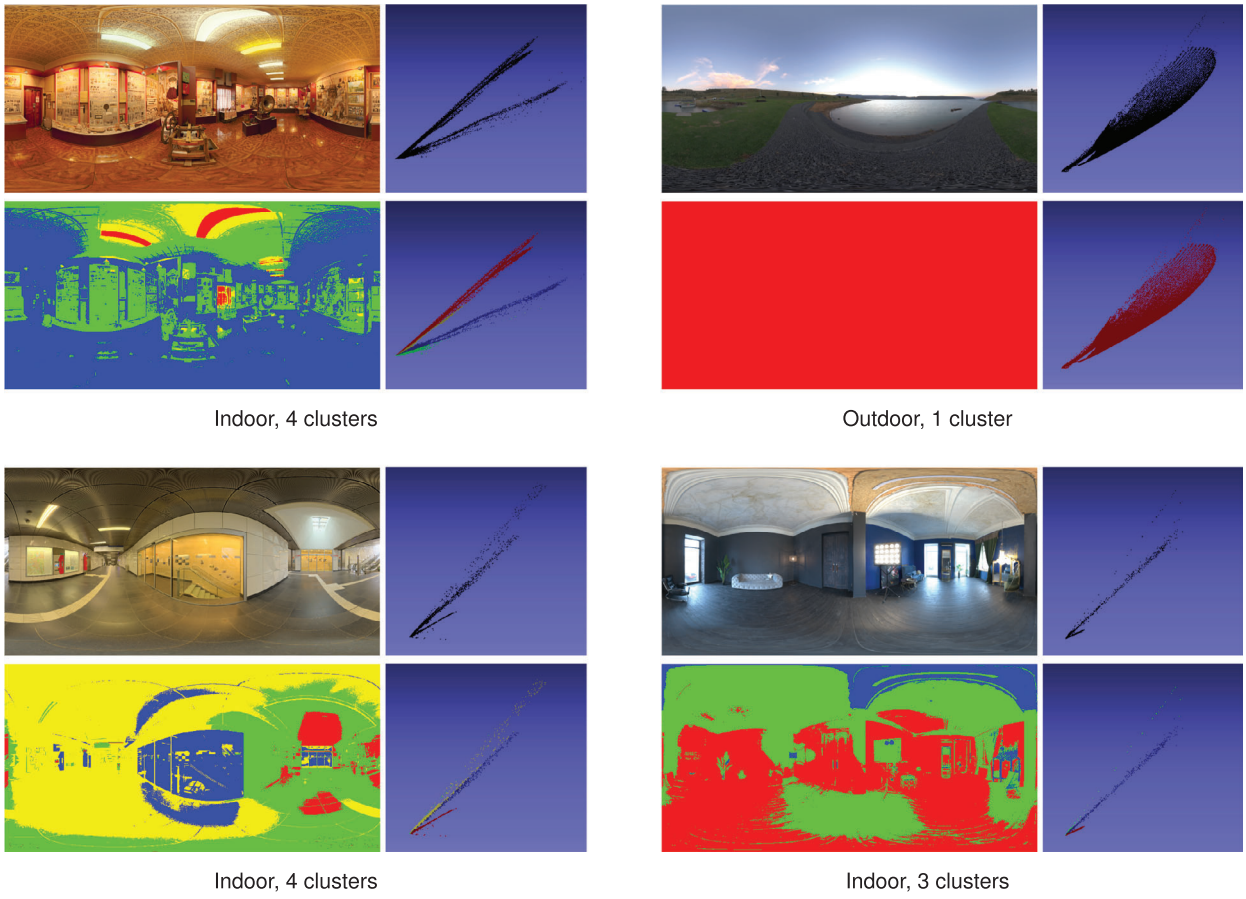

Figure 3: *Our modified version of the K-Means algorithm applied to 4 input HDR environment maps. For every environment map, we visualise: its pixels in the RGB space; its colour-coded clustered version (one colour belongs to one cluster); a colour-coded visualisation of its pixels. Note that pixels affected by distinct illuminants of similar temperatures fall into the same cluster.*

sufficiency metric based on heuristics that identifies the most desirable clustering. We also implement this approach, and design a set of heuristic rules based on the most common properties of environment maps—for example, we do not allow more than 1 outdoor cluster, we discard clusterings that result in clusters with similar colour temperatures, we do not allow more than 6 clusters overall (as it generally results in overclustering, which hinders performance) and so forth.

Such a heuristic approach is bound to have its deficiencies. In our situation, these are most pronounced for corner cases, such as indoor settings with a lot of small lamps or a night sky capture covered with bright stars. While we believe that some of the shortcomings could be resolved using a distinct approach, such as employing a neural network, we find the current algorithm sufficient for our purposes, especially since its design was not the main goal of the paper. Additionally, we note that even the most robust methods proposed for the problem of environment map segmentation are prone to failure on certain inputs.

We present the results of our proposed clustering on a couple of example environment maps in Figure 3. We note that the assumption of cluster exclusivity (that is, one pixel belongs to one cluster), is not entirely correct, as in real life, objects are often illuminated by more than one light source. However, this assumption does not have a significant effect on the final result. This is because the effect of the environment map on the rest of the scene is mainly determined by the spectral uplifts of the pixels with the highest luminosity, whose emission is generally a result of exclusively one light source (as they either belong to the light source or are extremely close to it).

3.1.2. *CCT detection*

The general idea behind determining the emission spectrum of a cluster's light source is to first determine its correlated colour temperature (or *CCT*), and then utilise existing light source measurements to identify the one that attains this temperature.

We explored multiple approaches to most accurately determine the CCT of a cluster, such as picking the most luminous pixel, using linear regression to fit a line through the cluster and so forth. Our final and most accurate solution works as follows:

First, the clusters are divided into *indoor* and *outdoor* clusters depending on whether the cluster contains any portion of the sky. The division is performed with a simple sky segmentation algorithm,

specifically the one proposed by Shen and Wang [SW13]. While this algorithm sometimes fails to recognise the specific edges (see e.g., Figure 2), we employ it solely for the purpose of determining the presence of sky, for which it performs sufficiently. Note that this approach manages to classify even, for example, outdoor streetlamps as indoor clusters (as their illumination rarely affects the sky pixels), which is the desired behaviour.

To determine the CCT of the indoor clusters, we utilise the results of the K-Means algorithm. Using McCamy's approximation [McC92], we convert the RGB values of the clusters' means to their corresponding correlated colour temperatures, which are then established to be the final CCTs.

Unfortunately, we found this approach to perform inadequatly for specific corner cases in outdoor settings—particularly for sunset and sunrise environment maps captured when the sun is just below the horizon. In such cases, the mean CCT was not able to accurately balance out the extremely saturated red or yellow tints attained at the horizon with the cool colour cast of the rest of the capture. We address these cases by applying heuristics and averaging the CCT of the cluster mean with the CCT of the pixel with the highest luminosity in the cluster, which we found to result in a much closer final estimation. We avoid the need to detect these corner cases by applying this heuristic to all outdoor clusters, since for daylight illumination, the CCT of the pixel with highest luminosity is practically identical to the CCT of the mean.

We note that we do not perform sky segmentation only to address a specific corner case, and that the information about cluster type is also utilised for the purposes of uplifting emission from CCT as explained later in Subsubsection 3.1.4.

3.1.3. *Environment map calibration*

The CCT detection process heavily relies on the light sources to attain colour values identical to their real-life properties. However, as previously mentioned in Subsection 2.1, the current environment map capture and calibration process only aids with simulating human visual perception. And while there exist methods that are capable of preserving this information (i.e., using an emissive colour target such as camSPECS [Eng]), they are too cumbersome and costly to use in practice.

We therefore propose a novel capture process for the purposes of a spectral rendering pipeline—specifically, we propose to additionally calibrate the camera according to a light source with a known spectral power distribution that is, in the same manner as the colour target, manually placed and captured in the scene. For our work, we use a multi-LED light source that simulates the spectrum of the D65 distribution (specifically, we use Waveform Lighting's ABSOLUTE SERIES LED D65 Module). While the ACES colour space, which is generally used for production workflows in motion pictures, has a different white point, we chose to use D65 as it is the white point of most current RGB spaces. We, however, note that our proposed technique would work in an identical manner for any chosen white point.

The capture therefore contains both a colour target and a light source. The calibration process itself is then twofold—first, the standard colour target calibration is performed. This results in the change of colour for the whole image, including the light source. As we know the goal RGB value of the light source (in our case, it is $RGB = (1, 1, 1)$), we can obtain a white balance matrix that transforms the image into a colour space, where we can properly distinguish the colour cast. To do so, we use the von Kries chromatic adaptation method [Fai20], with the source white point being the current RGB value of the light source and the destination white point being the goal RGB of the light source (in our case, $RGB = (1, 1, 1)$). We note that the selection of the white balance matrix was only due to personal preference, and that utilising a different one (e.g., the CIECAM02 matrix) would not yield qualitatively different results.

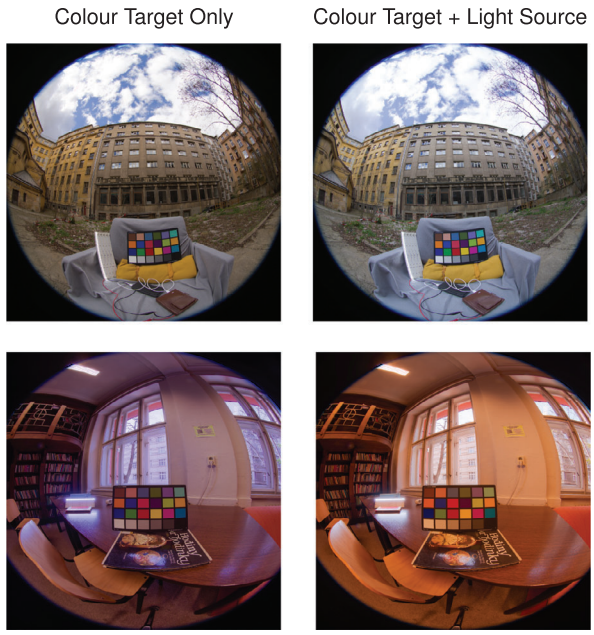

Figure 4: *Our proposed calibration process shown on examples of environment maps.* **Left**. *The capture is calibrated with the Macbeth Colour Chart only.* **Right**. *The capture is additionally white balanced according to the light source (i.e., the light strip present in the images). In the calibrated captures, the pixels of the light source therefore attain the RGB value of (1,1,1) (in sRGB).*

We present examples of environment maps captured in the proposed manner in Figure 4, both before and after the light source calibration. We note that while this process is essential for the method to work perfectly, it does not render our technique useless for environment maps calibrated using a colour target only. Especially for clear daylight illumination, we do not expect significant changes of the image's colours. However, in such cases, we must warn against the possibility of incorrectly estimated emission, which may in turn result in a decrease of realism of the final uplift.

3.1.4. *Emission uplifting*

To determine the correct emission spectrum for an input correlated colour temperature, we implement different methods depending on the type of the light source. In the following, we review the light

source categories that we support, along with their viable input CCTs, and the approaches to their uplifting:

Daylight (over 5000 K). In addition to the standard daylight CIE illuminants (D50, D55, D65 and D75) whose temperatures range from roughly 5000 K to 7500 K, Judd et al. [JMW*64] defined a method that computes the spectral power distribution of a D-series illuminant (i.e., emission of natural daylight) for colour temperatures also outside of this range. Our uplifting therefore utilises the linear interpolation of D-series CIE illuminants for input CCTs of values from 5000 K to 7500 K, and performs the D-series illuminant computation by Judd et al. [JMW*64] for inputs over 7500 K.

Sunset/Sunrise (2200 K–5000 K). To obtain sky emission spectra at sunset or sunrise, we utilise the Prague Sky Model presented by Wilkie et al. [WVBR*21] and its implementation proposed in the later extension by Vévoda et al. [VBRKW22]. We locked 4 of the model's 5 parameters to values that simulate sunset the closest (specifically, we set *ground albedo* to 0, *solar azimuth* to 90 degrees, *camera view* to side-facing fisheye and *ground level visibility* to the maximum value of 131.8 kilometres). We then performed multiple renders with varying levels of *solar elevation* ranging from 0 to 20 degrees, as higher elevation values can already be considered daylight. For each of these renders, we obtained the emission spectrum of its most luminous pixel, which resulted in a list of spectra with temperatures ranging from 2200 K to 5000 K stored in an ascending order. The final spectrum is obtained by a linear interpolation of two neighbouring spectra from this list in terms of CCT. Since any two neighbouring spectra are close to each other in terms of curve shape, this approach does not cause any distortions or colour artefacts.

LED Light Source (1700 K–30,000 K). In their paper, Kokka et al. [KPB*18] present a large database of over 1500 measured LED light sources, from which they compute 5 representative distributions with colour temperatures ranging from 2700 K to 6500 K that are currently the CIE LED standard distributions. As the temperatures of LEDs can span a significantly wider range, we pick additional 15 distributions from the provided measurements, ranging from 1700 K to 30,000 K. To ensure proper coverage of the 2700 K–6500 K range and to avoid any interpolation artefacts, we additionally add 4 more spectra, summing up to the overall list of 19 distributions. The uplifting itself is then also performed as an interpolation of the two closest neighbouring spectra in terms of CCT. We note that LEDs with temperatures over 20,000 K are not suitable for regular lighting purposes—however, we offer this possibility for specific corner cases.

Fluorescent Light Source (2900 K–6500 K). Similarly to sunset and LED spectra, we obtain a set of fluorescent emission distributions and perform their interpolation according to the input CCT. So far, we only support the CIE standard fluorescent illuminants, which are the F1–F6 lamps from the CIE Illuminant series F. Since other available fluorescent lamp measurements have distinct spectral shapes, performing their interpolation with any of the F1–F6 measurements is undesirable, as it would provide unrealistic results.

Incandescent Light Source (1000 K–3000 K) The spectral power distribution of the emission M of an incandescent object heated to a temperature T is given by the Planck's law, which follows:

$$M(\lambda, T) = \frac{c_1}{\lambda^5} \frac{1}{exp\left(\frac{c_2}{\lambda T}\right) - 1}$$

where c_1 is the first radiation constant and c_2 is the second radiation constant. We utilise this equation to compute the final incandescent emission spectrum for an input temperature T.

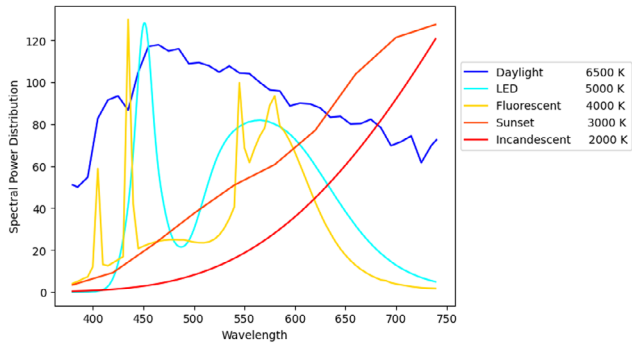

Figure 5: *Spectral shapes of different light source types obtained with our HDR environment map uplifting system. Note the complexity of the shapes of the LED and fluorescent light sources in comparison to, for example, daylight or an incandescent light source—when used to illuminate an object with a smooth reflectance curve, it is evident that the former are prone to cause visible metameric artefacts.*

While the resulting emission spectra obviously differ depending on the input temperature, their general shapes remain the same throughout every light source category. We visualise the emission shapes that our system is capable of obtaining in Figure 5.

To distinguish the light source type of each cluster, our system takes advantage of the *indoor* and *outdoor* flags obtained during CCT detection (see Subsubsection 3.1.2). Depending on their CCT, outdoor clusters use either the *daylight* ($CCT \geq 5000\ K$) or the *sunset* ($2200\ K \leq CCT \leq 5000\ K$) uplifting method. Note that our decision to uplift to daylight illumination as opposed to direct sun illumination is motivated by the typical composition of outdoor settings, where daylight sky is the most dominant. Additionally, in the visible range, the D-series illuminants are a reasonable approximation of solar radiance [Kur84], and therefore utilising sun illumination measured during daylight would provide similar results.

Unfortunately, the input environment map captures do not contain enough information for us to be able to accurately determine the type of indoor light sources. Therefore, all indoor light sources are treated as LED unless otherwise specified by the user. We do not see this as a severe shortcoming of our approach, as it is reasonable to ask users what basic types of illumination are present in a capture.

We note that while some of our proposed uplifting scenarios perform linear interpolation of two spectra based on their input CCT, this relationship is not linear. Therefore, the final spectrum will not be evaluated exactly to the desired temperature. However, in practice, this does not pose a problem. As the resulting CCT cannot be outside of the range of the neighbouring spectra, and as the

temperatures of the distributions are sampled densely, the error is minimal to cause any visible colour distortions.

3.2. Uplifting from light source

The process of uplifting a specific input HDR pixel is based on the simple fact that its emissive properties are due to the combination of the object's original reflectance and the incoming light, specifically:

$$P(\lambda) = E(\lambda) \cdot R(\lambda)$$

where $P(\lambda)$ is the final SPD of the pixel, $R(\lambda)$ is the reflectance of the object the pixel belongs to and $E(\lambda)$ is the incoming emission that hits the object in the original scene. We simplify this observation by making the assumption that $E(\lambda)$ is only a result of one specific light source (i.e., the cluster's light source). As previously explained in Subsubsection 3.1.1, with regard to the realism of the final uplifts, this assumption is acceptable. Therefore, to obtain the final uplift $P(\lambda)$, the only variable that remains to be determined is the object's reflectance $R(\lambda)$.

To do so, we utilise a reflectance-uplifting method. We specifically take advantage of the wide gamut sigmoid uplift cube implemented as part of the work by Tódová et al. [TWF22], which is technically the uplift model proposed by Jakob and Hanika [JH19] extended for support of Adobe Wide Gamut RGB. We select this model due to its favourable results in the area of general reflectance uplifting, as it creates smooth and simple curves similar to those of real-life objects (e.g., wood, vegetation, soil, etc.), which is what typical environment maps tend to consist of.

As previously mentioned in Subsection 2.2, the uplift cube consists of evenly-spaced lattice points that contain a mapping from their coordinate (or RGB value) to a coefficient representation of a reflectance spectrum. This spectrum then evaluates to the RGB value of the lattice point.

However, every RGB value in the cube is already stored *with respect to a white point*. In practice, this means that the reconstructed curve is first multiplied by the illuminant of this white point and only then is the final spectral power distribution converted to its RGB counterpart. For this purpose, both the model proposed by Jakob and Hanika [JH19] and its Adobe Wide Gamut RGB extension [TWF22] use the white point of the D65 daylight illuminant. To determine the reflectance of the objects in our scene, though, we need to know their properties with respect to the white point of the distinguished light source as opposed to D65.

While it is possible to create a new model in a similar manner as proposed in the previous works (i.e., with the process of *cube fitting*), we found this approach to be too costly to perform on the fly. Instead, we make use of the already existing D65 cube and create the new model by reconstructing the reflectances of all of the D65 cube's points and determining their RGB values under the cluster's identified illuminant.

As the conversion of a spectrum to RGB is not a linear operation, the resulting structure does not necessarily have to be an evenly-spaced cube. As shown in Figure 6, this transforms both scales and skews the model. Additionally, as the wide gamut cube also contains empty voxels that do not have viable reflectance mappings,

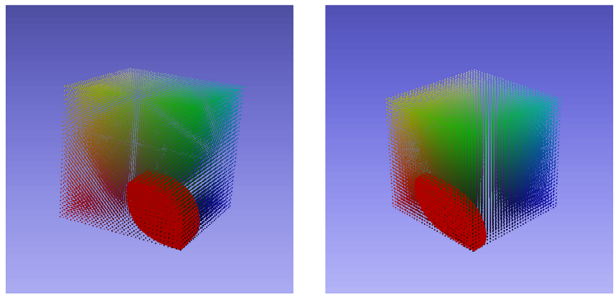

Figure 6: *Comparison of the wide gamut reflectance uplift cube to our reflectance uplift model (visualised in red) constrained by a LED spectrum with a temperature of roughly 5000 K (left) and a sunset spectrum with a temperature of roughly 2700 K (right). Note that the scale of the models is proportional to the scale with which the illuminants are stored, and does not affect the final uplifting results.*

the model loses its original cube shape and rather attains a blob-like shape of the Adobe Wide Gamut RGB colour space. Therefore, as opposed to the regular cube, performing a constant lookup in the structure based on the input RGB value does not guarantee that we will obtain the correct voxel corners for trilinear interpolation. We address this by still utilising constant lookup and, if its results are unsatisfactory, by moving into neighbouring voxels depending on the neighbours' coordinates. We note that, due to the skewing of the model, the trilinear interpolation does not necessarily have to be performed within a voxel, but rather between the closest 8 surrounding points. This penalises the lookup process in terms of execution time. However, we find the performance overhead to be worth the improvements in colour accuracy (see Subsection 4.3, specifically 4.3.1).

Although this model achieves proper round-trip when uplifting under the identified illuminant, it is still bounded and therefore unsuitable for all HDR pixels. We address this in the same manner as the current state-of-the-art emission uplifting, that is, with *scaling*. The scaling factor is determined in the same manner as the current emission scaling in Mitsuba 3, which follows:

$$scale = 2 \cdot max(rgb.r, rgb.g, rgb.b)$$

This ensures unproblematic uplifts of all values as opposed to just using the maximum component and attempting to uplift at the gamut boundary. The final uplift is then obtained as:

$$P(\lambda) = scale \cdot E(\lambda) \cdot R(\lambda)$$

where $R(\lambda)$ is the reflectance obtained from the model created for the white point of $E(\lambda)$. This approach ensures the proper scaling of the illuminant and therefore prevents the loss of spectral shape information.

Lastly, we address the problem of out-of-gamut input pixels. As the emission spectrum of the light source is only an estimation, it may happen that some of the pixels in the image fall outside of the gamut of the uplift model regardless of its scale. This is usually the case of dark pixels close to $RGB = (0, 0, 0)$, which suffer from the skewing. As such pixels are barely affected by the light

source anyway, we address these cases by employing the standard scaled emission uplifting with the D65 cube. For the rare cases of HDR out-of-gamut pixels, we also support gamut mapping. Similarly to the wide gamut cube, the process maps to the closest point within the model.

Currently, the uplift model for every cluster is created prior to the actual uplifting process and is stored in memory. Each pixel then contains a pointer to the uplift model (or cluster) it belongs to. As we do not allow more than 6 clusters per environment map (see Subsubsection 3.1.1), this does not significantly affect memory requirements. The uplifting itself is then performed per pixel.

4. Results

The evaluation of our technique consists of three parts. First, we prove that our method is technically sound by assessing the round-trip error for a diverse and extensive set of environment maps. Second, we evaluate how our technique improves the current state of the art by comparing a set of ground truth hyperspectral environment maps to both our and the state-of-the-art uplifts of their RGB counterparts. Last, we address the performance overhead of our method in comparison to the current state of the art, both in terms of execution time and memory requirements.

Due to our lack of access to commercial spectral renderers, we present our uplifting system as a standalone implementation. The input is an HDR environment map, and the output an uplift file that contains mappings from individual pixels to their corresponding spectra.

In order to see the performance of our method in terms of the final colour deviations in practice, we incorporated this uplift file for environment map uplifting in the Mitsuba 3 renderer. Unfortunately, the spectral variant of Mitsuba 3 has a hard-wired 3 coefficient representation for every spectrum, which is incompatible with the shape of our uplifted spectra. To fully integrate our solution into the renderer, we would have had to redesign its low-level data structures, which was deemed to be outside the scope of this work. Therefore, we use our implementation in the Mitsuba 3 renderer only for visualisation purposes (e.g., to create Figures 1 and 8). All colour accuracy and performance tests are carried out in our standalone implementation.

For evaluation purposes, our uplifting system also supports the current state of the art, that is, the *scaled reflectance uplifting*. We specifically use a variant of the technique used in the Mitsuba 3 renderer, however, we can expect the scaling of any other type of smooth reflectance spectra to perform similarly.

4.1. Uplift accuracy

To evaluate the round-trip error, we compile an extensive set of HDR environment maps. We make sure to encapsulate a variety of different scene settings—both indoors and outdoors; with one or multiple distinct illuminants; with varying light source temperatures; captured at different times of the day. For each environment map in the set, we perform our proposed uplift and obtain its RGB counterpart. We then compare the uplifted RGB image to the original input image by means of the CIE Delta E 2000 metric. To cover all proposed uplifting scenarios, we treat the light sources in indoor environment maps as LED, fluorescent and incandescent respectively and, if applicable, perform a different uplift for each light source type.

On average, we obtain an error of only $\Delta E = 4.85 \times 10^{-12}$, while the maximum Delta E achieved is $\Delta E = 6.53 \times 10^{-12}$. Generally, colour differences of $\Delta E <= 1$ are considered to not be perceivable by the human eye [MT11], which our method definitely satisfies. The pixels that are not uplifted with our model (i.e., the pixels that fall out of its gamut and are therefore uplifted with the D65 cube) take up only 0.0035% of the input and none of them are in the high dynamic range. We can therefore claim that the distinct approach to their uplifting does not reduce the realism of the final uplift.

Of the 166 performed uplifts (counting in the multiple uplifts of indoor environment maps), we found 2 to have pixels out of gamut. On average, the out-of-gamut pixels amounted to 0.000016% of the maps. In practice, this is barely visible to the human eye—however, when matching plate footage to its digital counterpart in the VFX industry, even a couple of erroneous pixels can cause noticeable artefacts. In some cases, we measured errors as high as $\Delta E = 25$, which is worrisome in terms of the final appearance. We attribute these errors to the insufficient estimation of the light sources in the scene. We specifically found the biggest deficiency of our pipeline to lie in the CCT detection process, and we strongly suggest focusing on its improvement as future work. However, as the focus of this paper was on spectral uplifting rather than white balance and clustering, we consider the overall accuracy of our method to be satisfactory.

4.2. Comparison to the state of the art

In order to determine how our technique improves the current state-of-the-art, we perform comparisons of the following three spectral environment maps: the *ground truth*, that is, an environment map captured with a hyperspectral camera; the *scaled reflectance uplift*, that is, a spectral environment map acquired by uplifting the RGB counterpart of the ground truth with the scaled reflectance uplifting technique; and *our uplift*, which is obtained by applying our technique.

We determine the level of similarity between two spectral environment maps by their effect on the appearance of multiple reflectances when treated as a light source. For each pixel and its two uplifts treated as emission, we iterate over multiple colour atlases of reflectance measurements and compute the RGB of the reflectances under both uplifts. To determine their difference, we once again use the CIE Delta E 2000 metric. The atlases used are the Munsell Book of Colour (1598 samples), the Pantone Colour Matching System (1853 samples) and the Macbeth Colour Chart (24 samples). We note that the proposed test is performed in an isolated environment, that is, in a setting with only a single illuminant and a single Lambertian surface. No rendering or path tracing is performed, and the final colour is obtained by a simple multiplication of the present emission and reflectance spectrum and a subsequent conversion into RGB. Therefore, the results in terms of absolute error measurements do not convey the exact colour error that would be present if both techniques were to be, in turn, applied for the

purposes of environment map uplifting in a state-of-the-art spectral renderer. The final absolute colour error can either be greatly accentuated or diminished, depending on the different types of light interactions with distinct materials during path tracing. Therefore, the more meaningful measure obtained during this test is rather the *relative* colour difference, that is, the Delta E 2000 ratio, between the two techniques. This, on average, remains the same even if applied to more complex scene settings.

While our proposed test is robust insofar as it tests our method's performance on the whole of the Adobe Wide Gamut RGB gamut through the colour atlases, its main drawback is the lack of ground truth data, that is, hyperspectral environment maps. As we do not possess means to perform hyperspectral captures, we rely on existing resources. Specifically, we opted for 4 large datasets from real-world settings (by Morimoto et al. [MLNS24], Chakrabarti and Zickler [CZ11], Kider et al. [KJKN*14] and the HSSOD dataset [IOZ*18]), as they, to our best knowledge, most closely align with our requirements. However, similarly to other available data, they also lack in three regards.

First, the data is usually inadequate in certain aspects which are crucial for image-based lighting—for example, the measured values are bounded ([MLNS24], [CZ11]) or the captures are not a 360-degree view but rather a closeup of a small object ([IOZ*18], [CZ11]). Therefore, for most of the datasets, our technique requires custom adjustments (i.e., scaling of values or forcing certain types of uplifting due to the unreliability of the sky detection process). Second, there are very few indoor captures available, and those that exist do not contain information about the type of present illumination. Third, the capture process consists of a number of steps, none of which are flawless [PHG19]. For example, the results of both colour calibration and geometric calibration may be hindered by imperfections in the target (e.g., a white patch is only 99% white). System stabilisation is susceptible both to human error and to limitations of the used technology, and even environmental conditions, such as wind or dust particles, may introduce noise when present during the capture. Each of these factors can decrease the accuracy of the final measured spectra and, as a result, lead to their deviations from what happens in nature. As our goal is to assess accuracy, we find this to be the most worrisome aspect.

Therefore, in addition to the existing captures, we create our own spectral environment map dataset in terms of renders. Specifically, we use the ART renderer [Wil18] to obtain 20 distinct outdoor environment maps (using the sky model by Wilkie et al. [WVBR*21] and Vévoda et al. [VBRKW22]), and the Mitsuba 3 renderer to obtain 60 indoor renders—20 using LED lighting, 20 with fluorescent light sources, and 20 with incandescent illumination. In order to cover all possible scenarios, the temperature of the illuminants in individual renders is evenly sampled across the supported CCT range.

As both used renderers are physically based and all the materials of the scenes are defined with physically-plausible spectra, the process of light transport is almost identical to what would happen in nature. However, in contrast to the existing datasets, all inaccuracies caused by performing manual hyperspectral measurements are eliminated. This makes the results of the renders even closer to real world than captures. Additional benefit of this dataset is our ability to cover a wide range of cases and, as we possess knowledge of the scene's original illuminant, our capability to analyse the accuracy of individual steps of the proposed pipeline.

The evaluation of our method is therefore performed on both a set of existing datasets and our own set of 80 rendered environment maps. Specifically, to test the accuracy of outdoor uplifting, we use 160 spectral captures obtained from the following datasets: 60 captures from the HSSOD dataset [IOZ*18], 50 daylight captures by Chakrabarti and Zickler [CZ11], 20 sky measurements by Kider et al. [KJKN*14], 10 hyperspectral environment maps by Morimoto et al. [MLNS24], and 20 of our own custom renders from the ART renderer. The indoor uplifting is tested on 27 hyperspectral captures performed by Chakrabarti and Zickler [CZ11], and on a set of 60 renders from the Mitsuba 3 renderer.

4.2.1. *Outdoor uplifting*

First, we assess the accuracy of the outdoor uplifting. As mentioned previously in Subsection 3.2, the current state-of-the-art technique utilises the scaled results of the D65 cube for the purposes of HDR uplifting. So if the illumination of the scene is determined to be natural daylight, the estimated spectral power distribution of the light source is bound to be extremely close to D65 in terms of spectral shape. Therefore, our goal is to obtain results very similar to the state of the art for natural daylight, with our technique slightly outperforming in cases where warmer or cooler temperatures are present, and significantly outperforming for extremely warm sunset or sunrise settings (due to visibly distinct spectral shapes of the illuminants).

We provide the results of our measurements in Table 1a. These show that both techniques perform very similarly in terms of the final spectral shape. The observed differences are too small to clearly determine whether the error is caused by one technique being superior, or rather by inaccuracies in the initial ground truth measurements. We attribute them to a combination of the two. However, a noteworthy observation can be made when examining the error for the dataset by Kider et al. [KJKN*14], which consists of precise sky measurements that are not susceptible to capture errors. Here, our technique outperforms the current state of the art, which implies that its results are closer to real-world natural light measurements. On the other hand, our suboptimal performance for the HSSOD database [IOZ*18] can be attributed to the calibration of its images. Although they were captured during daylight, they contain warmer, sepia-like tones, which cause our system to incorrectly estimate the final illumination as sunset.

An even more significant observation is the colour accuracy for the *Renders* dataset. As it contains multiple images with sunset settings, we expected our technique to outperform the current state of the art by a notable margin. However, the actual results are only slightly in our favour. Upon closer examination, we attribute this solely to the failure of the CCT detection process, the implementation of which is rather simple and does not account for every possible environment map scenario (see Section 3.1.2). As the goal of this paper was not the design of a CCT detection algorithm, but rather to prove that our pipeline is technically correct, we therefore additionally evaluate how an incorrect CCT estimation affects the colour accuracy of the overall result. To do so, we force the CCT detection

Table 1: *Accuracy of our proposed* ***(a) outdoor*** *and* ***(b) indoor*** *uplifting method and the current state-of-the-art uplifting (i.e., the* scaled reflectance *uplifting) when compared to real-world hyperspectral measurements from distinct datasets.*

(a) Outdoor uplifting

Dataset	Images	Scaled refl. ΔE	Our ΔE
HSSOD	60	0.81	0.90
Chakrabarti and Zickler	50	1.33	1.31
Morimoto et al.	10	0.62	0.57
Kider et al.	20	0.23	0.17
Renders	20	1.19	1.14

(b) Indoor uplifting

Dataset	Images	Scaled refl. ΔE	Our ΔE
Chakrabarti and Zickler	27	1.25	0.99
Renders LED	20	1.16	0.41
Renders fluorescence	20	0.92	0.34
Renders incandescence	20	0.97	0.48

Note: The provided error measurements are the average CIE Delta E 2000 errors caused by illuminating a wide set of reflectances by the ground truth and the two uplifts respectively and comparing the results. Also note that small perceived error differences may not be caused by one technique outperforming the other, but rather by inaccuracies in ground truth captures. Additionally, we emphasise that the error perceived for the *Renders* dataset is mainly caused by incorrect CCT detection and that the differences are much smaller when CCT is established beforehand.

Table 2: *The effects of correct CCT estimation on the colour accuracy of our uplifting method, evaluated for images from the outdoor* Renders *dataset.*

		Avg ΔE		$\Delta E \geq 1$	
Dataset	Images	Scaled refl.	Our	Scaled refl.	Our
Daylight	7	0.56	0.49	12%	10%
Sunset	8	0.70	0.32	26%	4%
Sunset ($\leq 2k$)	5	0.87	0.21	29%	2%

Note: To better distinguish between different types of uplifting and CCT categories, the images are divided into multiple smaller datasets. Also note that daylight images with $CCT \geq 10k$ were omitted from the tests, as we do not find them representative of typical lighting conditions.

process to output the correct values and we repeat the tests for the *Renders* dataset.

We provide the results in Table 2. In addition to the average Delta E, it also contains information about the percentage of samples that attain the error over $\Delta E >= 1$. We find this metric especially meaningful, as $\Delta E = 1$ is the maximum error perceivable by a standard observer. It is clear that once the CCT is correctly detected, our technique behaves exactly as expected.

Lastly, we address the maximum error attained during our tests. While we can observe noticeable improvements (i.e., decreasing the error from $\Delta E = 35$ to $\Delta E = 5$ for particular sunset cases, or even from $\Delta E = 34$ to $\Delta E = 15$ for daylight settings), we note that the specific values are misleading. This is mainly because the maximum errors are generally achieved for emitters with the highest luminosity, which greatly depends on the capture process and on the scaling of the emission spectra. Therefore, we once again emphasise the importance of relative colour differences as opposed to the absolute error values.

Table 3: *Percentage of samples from individual indoor datasets where $\Delta E > 1$, measured for tests performed as described in Subsection 4.2.*

Dataset	Images	Scaled refl.	Our
Chakrabarti and Zickler	27	48%	35%
Renders LED	20	49%	9%
Renders fluorescence	20	57%	11%
Renders incandescence	20	33%	3%

Note: We emphasise the error ratio between our and the scaled reflectance technique, as the absolute Delta E colour errors are mostly dependent on the luminosity range of the input data.

In Figure 7, we present two outdoor examples from the performed tests, along with the achieved spectral shapes and the final colour appearance. While these are mainly for visualisation purposes, we note the spectral shapes and the considerable deviation of the current state of the art from both the ground truth and our technique.

In conclusion, our proposed outdoor uplifting technique improves the current state of the art and can result in significant improvements for warm sunset and sunrise settings. Its only serious drawback is the accuracy of the CCT detection process, which does not diminish the validity of the proof of concept.

4.2.2. *Indoor uplifting*

In order to test the performance of indoor uplifting on real-life hyperspectral measurements, we utilised the indoor dataset of 27 captures by Chakrabarti and Zickler [CZ11]. As the type of illumination was not provided, we examined the performed measurements and eventually estimated the light sources to be the closest to narrowband fluorescent illuminants, that is, the CIE F10-F12 lamps. To better resemble the original illumination, we additionally performed slight smoothing of the illuminants. We attribute the need for this to either imprecise hyperspectral captures or our inaccurate estimate.

Furthermore, we also perform the tests on our own 3 sets of renders that cover every supported indoor illuminant type and its respective temperature range.

We provide the results of our tests in Table 1b. In all cases, it is clear that our method significantly outperforms the current state of the art. The average error is effectively reduced to at least half in all of the rendered cases, even if the individual steps of the pipeline do not perform exactly as desired. Even the hyperspectral dataset with imprecisely estimated illumination exhibits noticeable improvement. Additionally, in Table 3, we present the percentage of samples that obtained an error of $\Delta E > 1$. For all the provided datasets, our technique once again significantly improves the current state of the art.

We emphasise that all of the results obtained for indoor uplifting were achieved with the current version of our system. This means

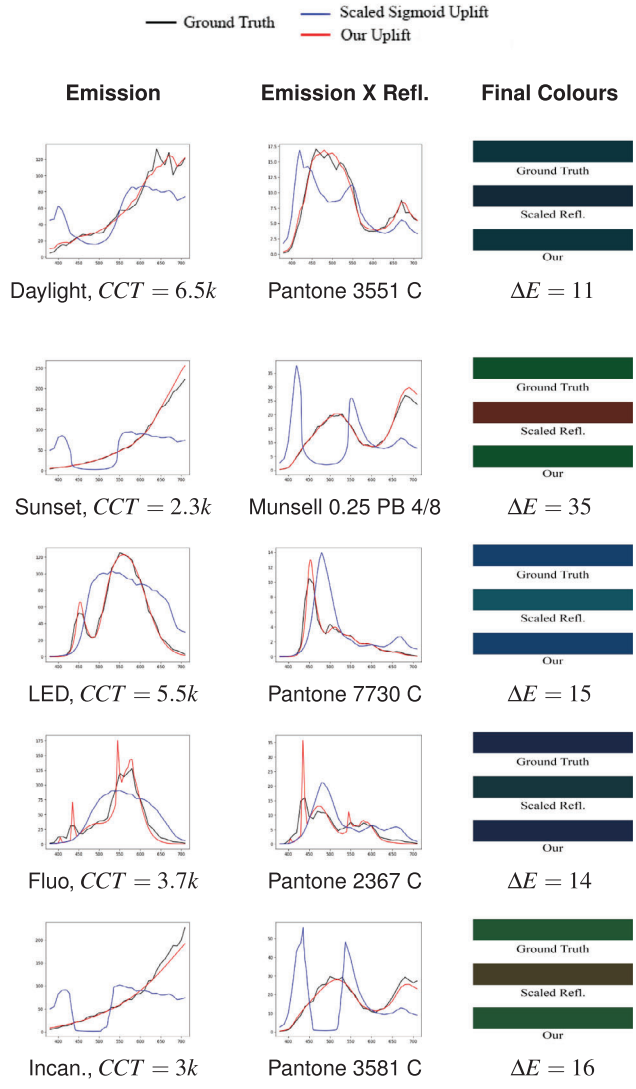

Figure 7: *Examples of ground truth hyperspectral measurements* ***(black)*** *and their two uplifts (obtained with the current state of the art* ***(blue)*** *and our technique* ***(red)****) when used to illuminate specific reflectances from available colour atlases. The first column labelled* Emission *represents the raw spectral power distribution of the emitter, while the second column labelled* Emission x Refl. *represents the final SPD when the emission spectrum is used to illuminate the disclosed reflectance patch.* Final colours *visualises the actual RGB values the spectra in the second column evaluate to. The information regarding the type of applied uplifting, the CCT and the illuminated reflectance patch is available for each image. The provided Delta E measurement is the CIE Delta 2000 error between the ground truth and the state-of-the-art uplift. The Delta E between the ground truth and our uplift remains below 1 for all displayed cases. Note that the colours are visualised in sRGB and their brightness is slightly increased for the reader. Also note that the provided colour error information is only our measured value, and would significantly differ with distinct luminosity of the illuminant.*

that we did not force the CCT detection process to output a specific value, nor was any other step altered in any way. However, while improving the CCT detection could additionally aid our technique, we note that this would not be to such an extent as for the outdoor uplifting presented in Table 2. This is because the CCT estimates are already quite accurate for indoor settings, as opposed to outdoor scenes, which are prone to significant deviations. As our CCT detection process relies on a roughly even distribution of RGB values of pixels to estimate the present colour tint, we attribute these failures mainly to the outdoor environment maps not satisfying this requirement (e.g., by containing only a blue sky region and a green grass field).

Similarly to our evaluation of the outdoor uplifting, while we observe some significant improvements in the maximum error (e.g., decreasing from $\Delta E = 41$ to $\Delta E = 4$ for specific renders with the incandescent light source), we do not focus on the absolute measurements but rather on the uplifted spectral shapes. We present an example for each of the illumination types in Figure 7, where it can be observed that our method follows the ground truth spectral shape rather closely, while the scaled reflectance uplifting significantly differs.

Although the proposed tests are the most reliable measure for colour accuracy, we remind that they are still performed in an isolated environment. Therefore, while we claim that our technique definitely outperforms the current state of the art, we additionally conclude this section with a couple of examples of the uplifts applied to an actual rendered scene. This is simply for visualisation purposes, that is, to illustrate how the techniques perform in practice when incorporated into the whole path-tracing pipeline. We present the final renders in Figure 8.

4.3. Performance

Lastly, we summarise the performance of our proposed method in terms of both the execution time and memory usage. All the tests in this section were executed on an Intel Core i7-11800H 2.3GHz (8 cores, 16 threads) processor and 32GB DDR4 3200MHz RAM.

4.3.1. *Execution time*

To evaluate the execution time, we use our standalone implementation. This is the preferred method of testing, as it eliminates any potential disruptions caused by other processes. Therefore, even if our technique were to be fully integrated within a state-of-the-art spectral renderer, we would still refer to the standalone implementation for accurate performance measurements.

Similarly to our colour accuracy tests described in Subsection 4.1, we use a dataset of 166 HDR environment maps of the same resolution that cover a wide variety of distinct settings and illumination conditions. For each of these maps, we perform the following: their loading; the initialisation of the uplifting system for the given map (including all the steps of the pipeline, that is, sky detection, CCT detection, etc.); and finally the per-pixel uplifting. The execution time of each step is recorded. In order to compare the performance of our technique, we additionally run the same test on the scaled reflectance uplifting method that is supported by our system. The

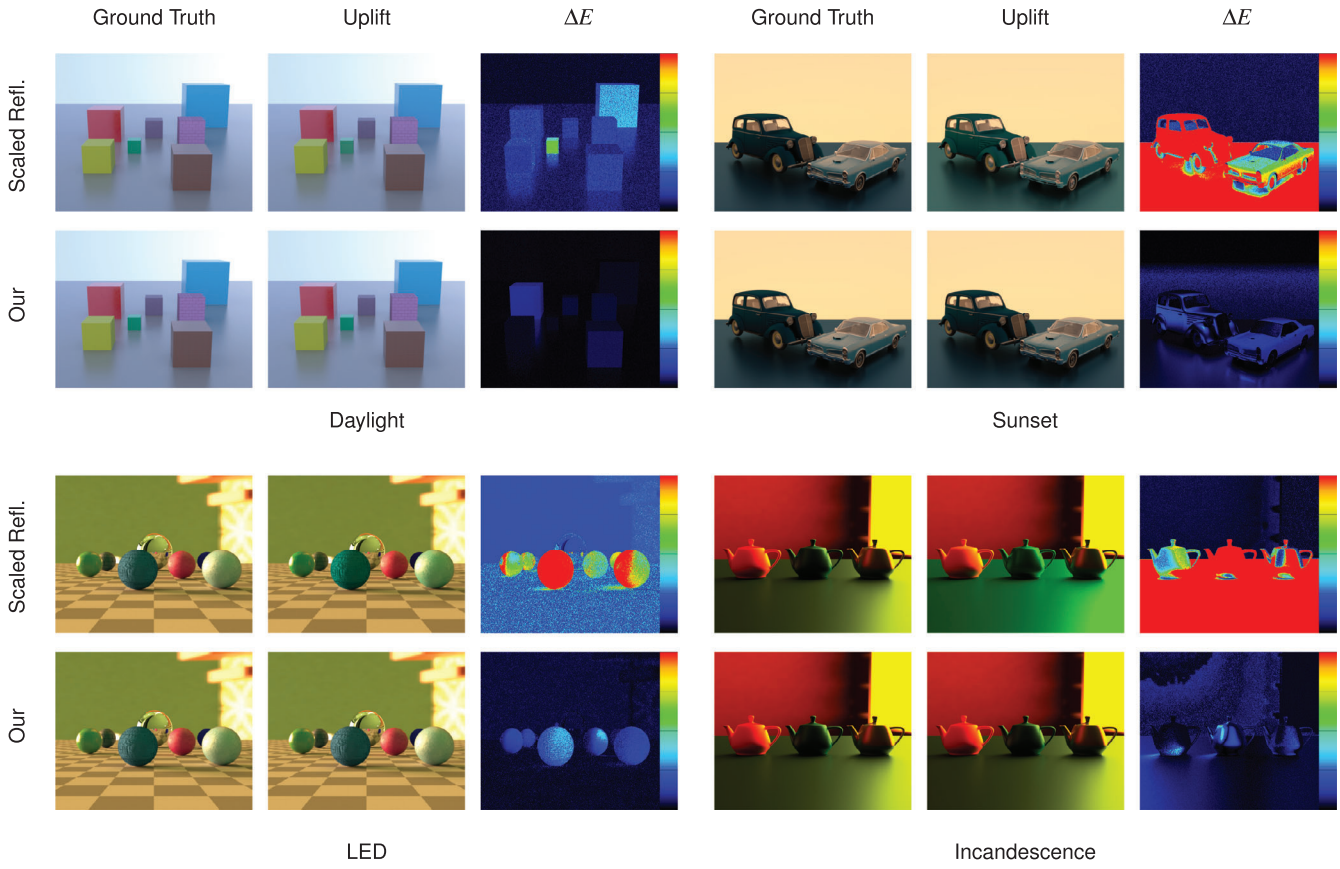

Figure 8: *Examples of both our technique and the state-of-the-art (or the* scaled reflectance*) uplifting when used for image-based lighting in the Mitsuba 3 renderer. Both techniques are compared to the* ground truth *in terms of the CIE Delta E 2000 error. All error images are relative to* $\Delta E = 5$*. All used environment maps are from our own rendered datasets created for the colour accuracy tests in Subsection 4.2. Note that 4 of the 5 available uplifting categories are present, with the* Fluorescent *uplifting visualised in Figure 1.*

Table 4: *Average execution time of the individual steps of the uplifting pipeline, both for our and the current state of the art (i.e., the* scaled reflectance*) technique.*

Step	Scaled refl. *(s)*	Our *(s)*
Image loading	0.1	0.1
Sky detection	N/A	2.8
Clustering	N/A	3.0 [1-13]
Per cluster *(up to 6x)*:		
CCT Detection	N/A	0.001 *(x6)*
Cube initialization	0.85	0.85 *(x6)*
Per-pixel Uplifting	174	540

Note: The clustering process is the only one with variable execution time, with range specified in brackets.

method is identical to the one used for testing in Subsections 4.1 and 4.2.

We present our result summary in Table 4. We specifically divide the individual steps into *initialisation* of the uplifting system (which includes image loading, sky detection, clustering, and both the CCT detection and cube initialisation) and the actual *uplifting* of the input environment map.

We first address the *initialisation* process. As our technique contains a multi-step pipeline in contrast to only a single cube loading of the current state of the art, it is clear that its execution time is worse. Specifically, on average, our initialisation takes 5.45 s as opposed to 0.95 s of the scaled reflectance technique. While the best-case scenario only worsens the current state of the art by a factor of 3, worst-case initialisation cases take up to 16 s.

However, although the initialisation time is significantly worse for our technique, we do not see this as a fundamental problem. In terms of the whole rendering pipeline, this process needs to only be executed once, and we find the performance overhead to be more than acceptable for the improved colour accuracy. Nevertheless, we see possible future work both for the sky detection and the clustering process, the performance of which was not specifically addressed during our research. The clustering process especially is heavily penalised by suboptimal heuristics.

The more worrisome results are for the actual *per-pixel uplifting*. On average, our current implementation takes 3 times longer

than the scaled reflectance technique. While some of this can be attributed to the out-of-gamut values and the overhead caused by multiple clusters, the majority of the performance penalty is caused by the lookup in the new illuminant-based uplift models. While in the D65 cube, a constant lookup is sufficient to find the correct voxel, the new models are scaled and skewed and a lookup within them can also require checking of neighbouring voxels (see Section 3.2). On average, our current implementation requires 6 lookups for each RGB value.

While the goal of this paper was to propose a proof of concept as opposed to a fully functional method, we still see this as a drawback of our technique. In the future, we propose focusing on improving this aspect by optimising the data structures. We specifically suggest researching the idea of creating new cubes by interpolating data from the skewed models.

4.3.2. *Memory usage*

In terms of memory requirements, the main difference between our technique and the current state of the art lies in the number of stored uplift models. While the scaled reflectance uplifting method stores only one cube of size 32^3, our new system allows for a maximum of 6 models (i.e., one for each recognised cluster). Additionally, the cube uplift structure does not require the storage of RGB values of individual lattice points, as they can be computed from the voxels' indices. However, this is not an option for our skewed models, which therefore additionally store 3 floating point values for every point.

This specifically means increasing the memory requirements from 0.4MB to a maximum of 4.8MB in the worst-case scenario. While our method also uses some additional data structures - such as an array of pointers for pixels to their respective clusters, or the clusters' CCT and illumination type information, we do not consider them of significance for two reasons: first, a lot of this data can be stored in optimised structures in the future; and second, even the current state of the art stores additional information as opposed to only the 3 sigmoid coefficients.

Overall, while our method has higher memory requirements than the scaled reflectance technique, we do not find this to be a limiting factor when used in a spectral renderer. With all the other data that needs to be loaded prior to path tracing, we consider the memory overhead negligible.

5. Conclusion

We presented the first method capable of uplifting HDR environment maps based on real-world emission data. Instead of creating synthetic spectra, we rely on estimating the real-world properties of light sources present at environment map capture, which we then use to constrain the uplifting process. The realistic data provided by our method aids in the prevention of colour deviations when using image-based lighting in a spectral renderer. This is important, especially in cases where even slight colour artefacts between the render and its real-life counterpart matter, such as in the VFX industry.

The results of our method are satisfactory in terms of the round-trip error and show a noteworthy improvement of the current state of the art. This is especially pronounced in the case of indoor lighting conditions, where the current techniques exhibit significant drawbacks.

However, we recognise some remaining deficiencies in certain parts of our proposed uplifting pipeline. Specifically, as the precision of these steps was not the main focus of this work (which is to illustrate that such a pipeline can be made to work in the first place), both the process of clustering and CCT detection still generate insufficiently accurate results for certain inputs. As future work, we propose utilising more robust approaches to increase the precision of these aspects of our method. Additionally, we intend to focus on the data structures that are used to store our new uplift models, as their optimisation could significantly improve the performance of the method. Specifically, we suggest creating cubes by interpolating the data already stored in our models.

As a byproduct of our work, we also presented a novel approach to camera calibration for the purposes of environment map capture that uses both a colour target and a light source. In addition to it aiding our uplifting process, we believe it to be a useful tool in the area of spectral asset creation, since calibrating with only a colour target does not retain sufficient information about the light source colour temperature in the captured scene. An additional benefit of this approach is that it is more cost-effective and practical in comparison to existing emissive calibration targets.

Acknowledgements

We acknowledge funding by the Grantová Agentura České Republiky with grant number GAČR-22-22875S, the Horizon 2020 Framework Programme of the European Union (grant number 956585, the PRIME ITN), and by Charles University under internal grant SVV-260699. The proposed environment map captures presented in this work were made by Roman Frátrik as a part of a student project at Charles University. We additionally thank Petr Vévoda for his help with the Prague Sky Model.

Conflicts of Interest

None of the authors have conflict of interest that would affect their objectivity in conducting the presented research.

References

[BBG23] Belcour L., Barla P., Guennebaud G.: One-to-many spectral upsampling of reflectances and transmittances. *Computer Graphics Forum 42*, 3 (2023), e14886.

[CZ11] Chakrabarti A., Zickler T.: Statistics of real-world hyperspectral images. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Los Alamitos, CA, USA, 2011), IEEE, pp. 193–200.

[Eng] Engineering Image: camSPECS. https://www.image-engineering.de/products/equipment/measurement-devices/588-camspecs-express. Accessed: 2024-12-11.

[Fai20] Fairchild M. D.: Von kries 2020: Evolution of degree of chromatic adaptation. In *Color and Imaging Conference*

(Bellingham, WA, 2020), vol. *28*, Society for Imaging Science and Technology, pp. 252–257.

[FHL*18] Fascione L., Hanika J., Leone M., Droske M., Schwarzhaupt J., Davidovič T., Weidlich A., Meng J.: Manuka: A batch-shading architecture for spectral path tracing in movie production. *ACM Transactions on Graphics (TOG) 37*, 3 (2018), 1–18.

[IOZ*18] Imamoglu N., Oishi Y., Zhang X., Ding G., Fang Y., Kouyama T., Nakamura R.: Hyperspectral image dataset for benchmarking on salient object detection. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)* (Piscataway, NJ, 2018), IEEE, pp. 1–3.

[JH19] Jakob W., Hanika J.: A low-dimensional function space for efficient spectral upsampling. In *Computer Graphics Forum* (2019), vol. *38*, Wiley Online Library, pp. 147–155.

[JMW*64] Judd D. B., MacAdam D. L., Wyszecki G., Budde H., Condit H., Henderson S., Simonds J.: Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America 54*, 8 (1964), 1031–1040.

[JSR*22] Jakob W., Speierer S., Roussel N., Nimier-David M., Vicini D., Zeltner T., Nicolet B., Crespo M., Leroy V., Zhang Z.: Mitsuba 3 renderer, 2022. https://mitsuba-renderer.org.

[JWH*19] Jung A., Wilkie A., Hanika J., Jakob W., Dachsbacher C.: Wide gamut spectral upsampling with fluorescence. *Computer Graphics Forum 38*, 4 (2019), 87–96.

[KJKN*14] Kider Jr J. T., Knowlton D., Newlin J., Li Y. K., Greenberg D. P.: A framework for the experimental comparison of solar and skydome illumination. *ACM Transactions on Graphics (TOG) 33*, 6 (2014), 1–12.

[KPB*18] Kokka A., Poikonen T., Blattner P., Jost S., Ferrero A., Pulli T., Ngo M., Thorseth A., Gerloff T., Dekker P., et al.: Development of white led illuminants for colorimetry and recommendation of white led reference spectrum for photometry. *Metrologia 55*, 4 (2018), 526.

[Kur84] Kurucz R. L.: Solar flux atlas from 296 to 1300 nm. *National Solar Observatory Atlas 1* (1984).

[LF20] Langlands A., Fascione L.: Physlight: An end-to-end pipeline for scene-referred lighting. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Talks* (New York, NY, 2020), Association for Computing Machinery, pp. 1–2.

[Mac35] MacAdam D. L.: The theory of the maximum visual efficiency of colored materials. *Journal of the Optical Society of America 25*, 8 (1935), 249–252.

[McC92] McCamy C. S.: Correlated color temperature as an explicit function of chromaticity coordinates. *Color Research & Application 17*, 2 (1992), 142–144.

[MLNS24] Morimoto T., Linhares J. M., Nascimento S. M., Smithson H. E.: How many surfaces can you distinguish by color? Real environmental lighting increases discriminability of surface colors. *Optics Express 32*, 20 (2024), 34246–34253.

[MMD*76] McCamy C. S., Marcus H., Davidson J. G., et al.: A color-rendition chart. *Journal of Applied Photographic Engineering 2*, 3 (1976), 95–99.

[MSHD15] Meng J., Simon F., Hanika J., Dachsbacher C.: Physically meaningful rendering using tristimulus colours. *Computer Graphics Forum 34*, 4 (2015), 31–40.

[MT11] Mokrzycki W., Tatol M.: Colour differenceΔ EA survey. *Mach. Graph. Vis 20*, 4 (2011), 383–411.

[OYH18] Otsu H., Yamamoto M., Hachisuka T.: Reproducing spectral reflectances from tristimulus colours. *Computer Graphics Forum 37*, 6 (2018), 370–381.

[PHG19] Pillay R., Hardeberg J. Y., George S.: Hyperspectral imaging of art: Acquisition and calibration workflows. *Journal of the American Institute for Conservation 58*, 1-2 (2019), 3–15.

[Smi99] Smits B.: An rgb-to-spectrum conversion for reflectances. *Journal of Graphics Tools 4*, 4 (1999), 11–22.

[SW13] Shen Y., Wang Q.: Sky region detection in a single image for autonomous ground robot navigation. *International Journal of Advanced Robotic Systems 10*, 10 (2013), 362.

[TW24] Tódová L., Wilkie A.: Constrained spectral uplifting for HDR environment maps. In *Eurographics Symposium on Rendering*, E. Haines and E. Garces (Eds.) The Eurographics Association (2024). https://doi.org/10.2312/sr.20241153

[TWF21] Tódová L., Wilkie A., Fascione L.: Moment-based constrained spectral uplifting. In *Eurographics Symposium on Rendering (EGSR) (DL)* (Göttingen, Germany, 2021), Eurographics Association, pp. 215–224.

[TWF22] Tódová L., Wilkie A., Fascione L.: Wide gamut moment-based constrained spectral uplifting. *Computer Graphics Forum 41*, 6 (2022), 258–272.

[VBRKW22] Vévoda P., Bashford-Rogers T., Kolářová M., Wilkie A.: A wide spectral range sky radiance model. *Computer Graphics Forum 41*, 7 (2022), 291–298.

[VDRE23] Van De Ruit M., Eisemann E.: Metameric: Spectral uplifting via controllable color constraints. In *ACM SIGGRAPH 2023 Conference Proceedings* (New York, NY, 2023), Association for Computing Machinery, pp. 1–10.

[Wil18] Wilkie A.: The Advanced Rendering Toolkit, 2018. http://cgg.mff.cuni.cz/ART.

[WVBR*21] Wilkie A., Vevoda P., Bashford-Rogers T., Hošek L., Iser T., Kolářová M., Rittig T., Křivánek J.: A fitted radiance and attenuation model for realistic atmospheres. *ACM Transactions on Graphics (TOG) 40*, 4 (2021), 1–14.