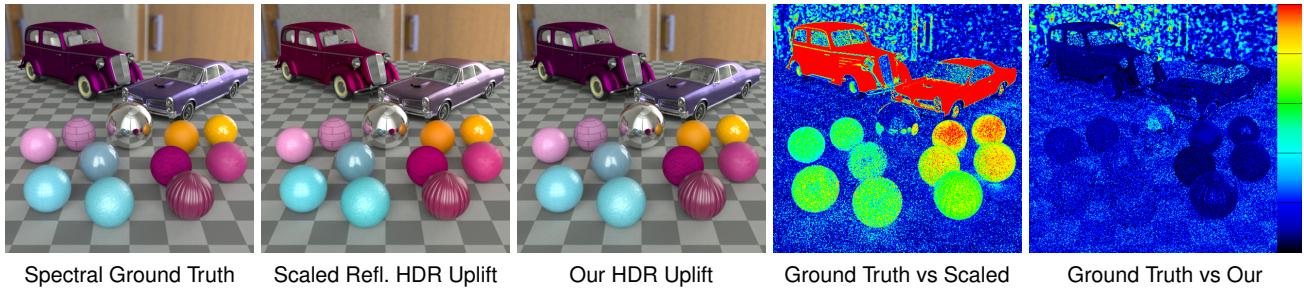


# Constrained Spectral Uplifting for HDR Environment Maps

L. Tódová<sup>1</sup>  and A. Wilkie<sup>1</sup> 

<sup>1</sup>Charles University, Czech Republic



**Figure 1:** Different approaches to rendering image-based lighting. **Ground Truth** uses a spectral HDR environment map as an input, i.e. 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.

## 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.

We propose 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.

## 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 as a distribution of wavelengths. This allows for physically correct simulation of light transport and results in more realistic calculations of reflection, transmission, absorption, etc. 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 to 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 of the existing spectral uplifting techniques focus mainly on reflectance uplifting, i.e. 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 their real-life counterparts. This, in turn, may cause metamerism 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, i.e. HDR environment maps, where such 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 e.g. 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.

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 wider colour gamut by adding fluorescent components. The work by Tódová et al. [TWF21] also takes inspiration from the precomputed 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 metamerism 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 Eismann [VE23] 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 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, which may in turn cause metamerism 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.

### 3. Environment Map Uplifting

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

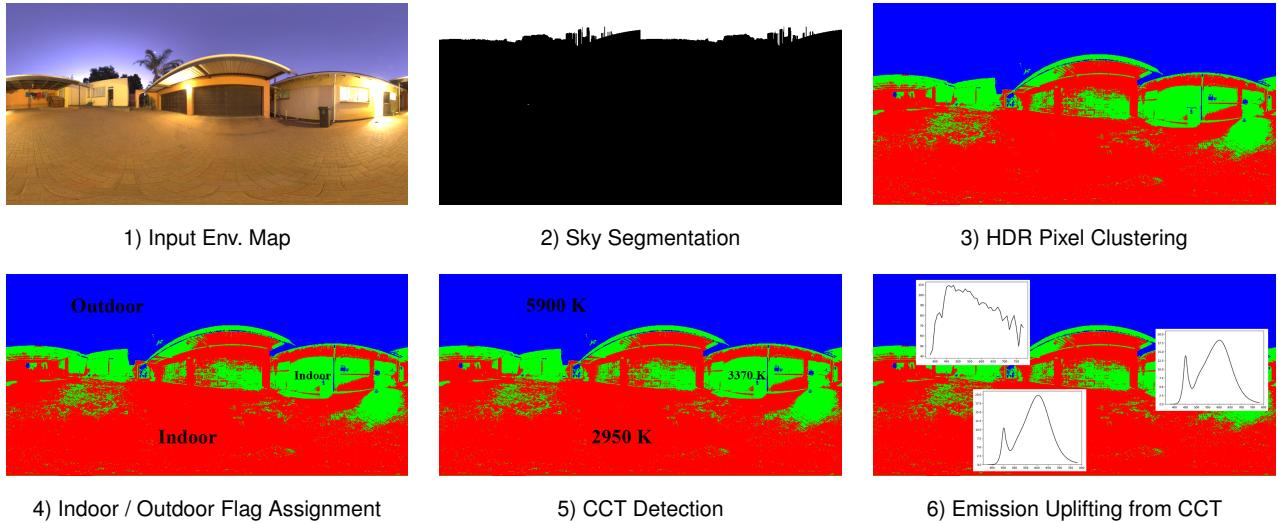
#### 3.1. Light Source Identification

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 color 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 each step 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.



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

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  $\theta$  and  $\phi$ . 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  $\theta$  and  $\phi$  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.

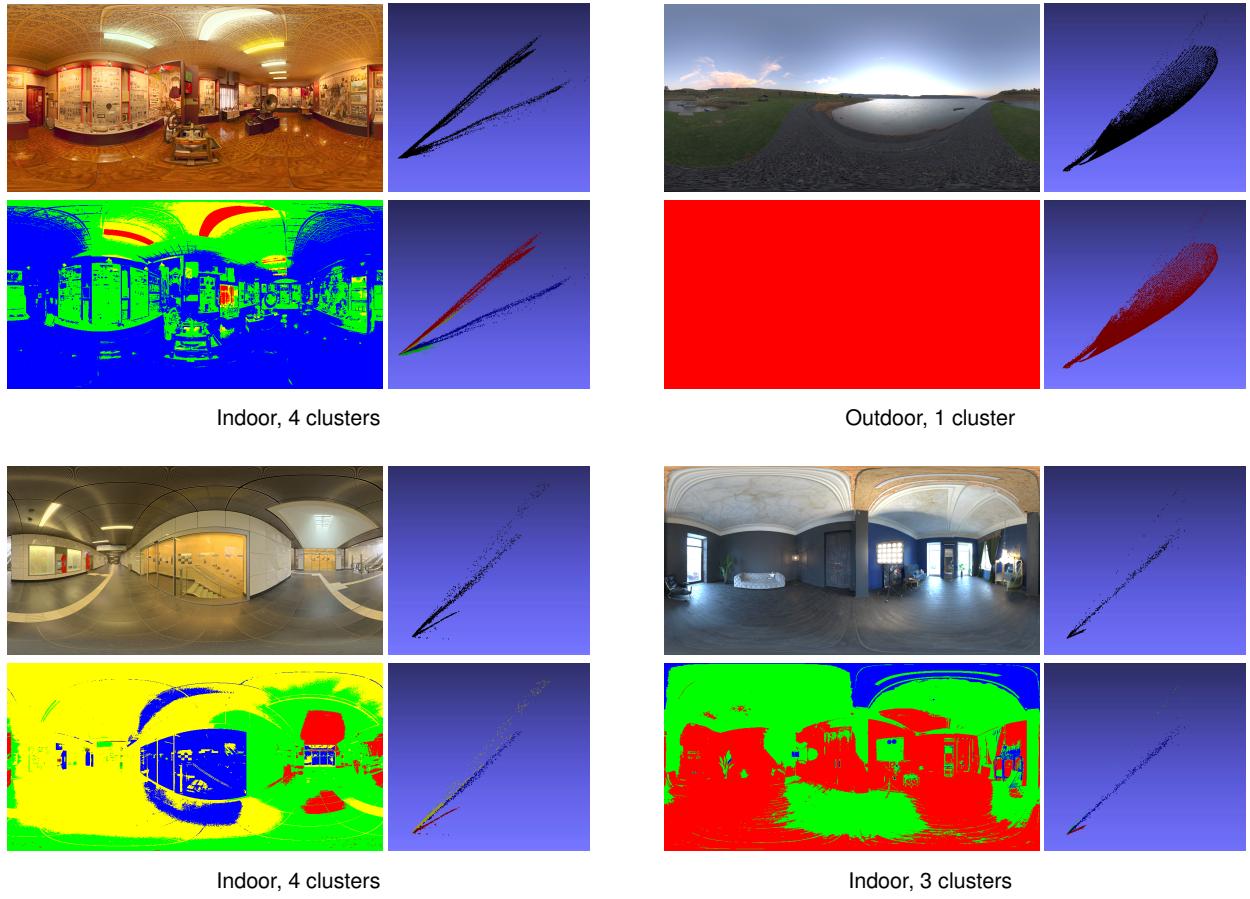
We present the results of our proposed clustering on a couple of example environment maps in Figure 3. We note that the assumption of cluster exclusivity, i.e. one pixel belongs to one cluster, is not entirely correct, as in real life, objects are often illuminated by more than one light source. However, 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, etc. 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. Note that this approach manages to classify even e.g. outdoor streetlamps as indoor clusters (as their illumination rarely affects



**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.

the sky pixels), which is the desired behaviour. The division is performed with a simple sky segmentation algorithm, specifically the one proposed by Shen and Wang [SW13].

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 temperatures.

Unfortunately, we found this approach to fail 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

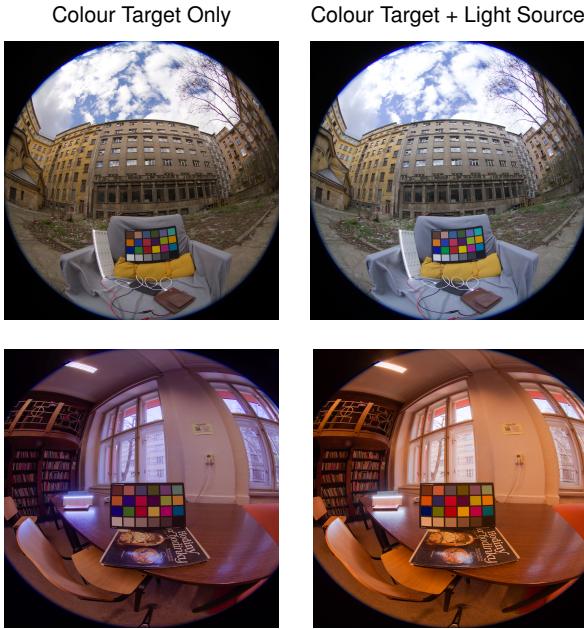
CCT of the pixel with highest luminosity is practically identical to the CCT of the mean.

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 therefore does not guarantee to preserve this information.

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



**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).

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.

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

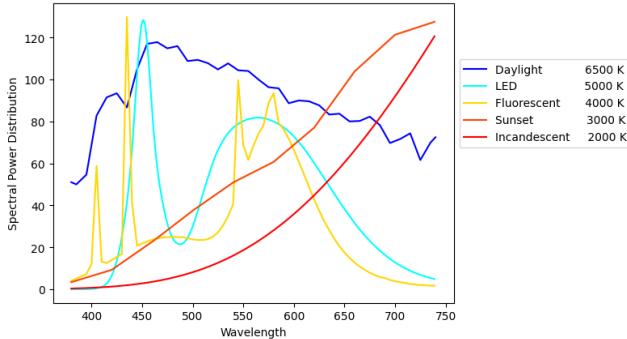
To determine the correct emission spectrum for an input correlated colour temperature, we implement different methods depending on the type of the light source. In the following, we review the light source categories that we support along with their viable input CCTs, and the approaches to their uplifting:

**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 to 5000 K)** To obtain sky emission spectra at sunset or sunrise, we utilise the Prague Sky Model presented by Wilkie et al. [WVB\*21] and its implementation proposed in the later extension by Vévoda et al. [VBKW22]. 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 kilometers). 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 to 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 to 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 to 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.



**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 e.g. 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 metameristic artefacts.

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 to 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:

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

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  $T$ .

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 the 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 evaluate 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.

### 3.1.5. Cluster Count

The final aspect of the light source identification process we have not yet addressed is the number of clusters our method returns. As the standard K-Means algorithm takes the desired cluster count as a parameter on input (denoted  $k$ ), the general approach is to iteratively run the algorithm for an increasing  $k$ , and design a 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), etc.

We note that our proposed clustering method has its deficiencies that 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 its shortcomings could be resolved using a distinct approach, such as employing a neural network, we note that even the most robust methods proposed for the problem of environment map segmentation are prone to failure on certain inputs. Additionally, while the incorrect clustering of the input environment map reduces the realism of the final uplift, it does not render our method useless.

## 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:

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

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 Tólová 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.

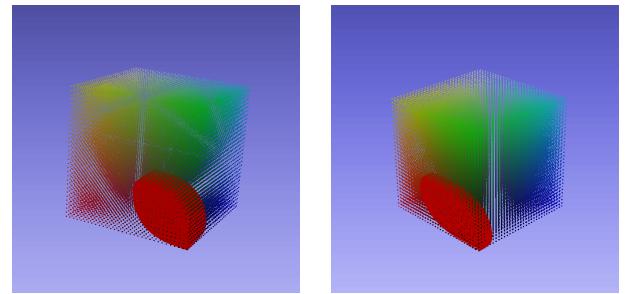
While it is possible to create a new model in a similar manner as proposed in the previous works (i.e. with the process of *cube fitting*), we found this approach to be too costly to perform on the fly. Instead, we make use of the already existing D65 cube and create the new model by reconstructing the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

As the conversion of 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 transform both scales and skews the model. Additionally, as the wide gamut cube also contains empty voxels that do not have viable reflectance mappings, 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 constant lookup in the structure based on the input RGB value does not guarantee us to 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 8 surrounding points. While our proposed lookup does not guarantee a constant complexity on every query, we find the average-case complexity to be similar to that of the regular cube.

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, i.e. with *scaling*. The scaling factor is determined in the same manner as the current emission scaling in Mitsuba 3, which follows:

$$\text{scale} = 2 * \max(\text{rgb.r}, \text{rgb.g}, \text{rgb.b})$$

This ensures unproblematic uplifts of all values as opposed to



**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.

just using the maximum component and attempting to uplift at the gamut boundary. The final uplift is then obtained as:

$$P(\lambda) = \text{scale} * E(\lambda) * R(\lambda)$$

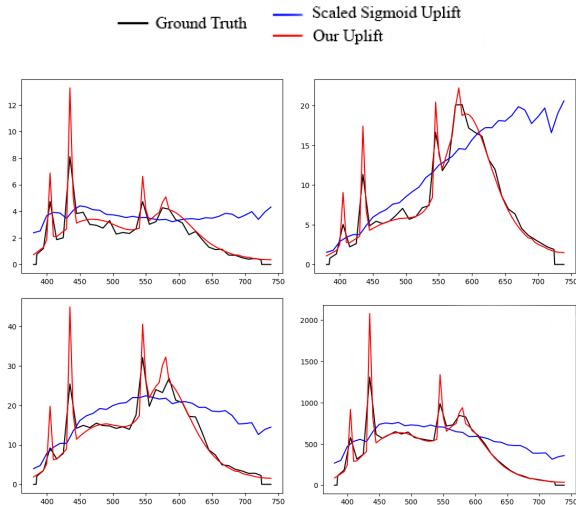
where  $R(\lambda)$  is the reflectance obtained from the model created for the white point of  $E(\lambda)$ . This approach ensures the proper scaling of the illuminant and therefore prevents the loss of spectral shape information.

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 3.1.5), this does not significantly affect memory requirements. The uplifting itself is then performed per pixel.

#### 4. Results

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 test the accuracy of our method, 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



**Figure 7:** Visualisation of the spectral curves obtained from renders presented in Figure 1 for 4 distinct pixels - specifically pixels belonging to a metal handle (top left), a wooden wardrobe (top right), a cardboard box (bottom left) and a light source (bottom right). The Ground Truth plot represents the spectral curve obtained from the original spectral render of the environment map, while the plots labelled Scaled Sigmoid Uplift and Our Uplift represent the results of applying the respective uplifting techniques to the HDR RGB value the ground truth curve evaluates to.

is incompatible with the shape of our uplifted spectra. To make our solution perform on a level that is on par with existing Mitsuba 3 rendering, we would have had to redesign low level data structures used by the renderer, which was deemed to be outside the scope of this work. Therefore, while our uplift files are used in a technically correct way that does not decrease the accuracy of our method, our proof of concept implementation comes with a substantial performance penalty. As a result, we do not provide any execution time measurements, as they would not provide any meaningful information to the reader.

As we do not possess means for performing spectral captures of environment maps, an additional complication when evaluating the correctness of our method is the lack of ground truth information. So far, the only way for us to obtain spectral environment maps is by performing spectral renders and storing their results prior to their conversion into RGB. We show the results of this process in Figure 1, where we employ the Mitsuba 3 renderer to obtain the ground truth, which we then uplift with our system and the system used by Mitsuba 3 respectively. We show that the render obtained when using our uplifting method is significantly closer to the ground truth. In Figure 7, we additionally present a comparison of emission spectra of specific pixels of the ground truth and the uplifted environment maps respectively, which arrives at a similar conclusion. However, we note that the renders are not identical to real-life captures, and that apart from the provided examples, we cannot provide a conclusive measure of the physical realism of our technique.

The evaluation of our technique therefore consists of two parts: First, we prove the correctness of our method by assessing the round-trip error for an extensive set of environment maps. Then, to demonstrate the added value of our technique, we compare the resulting curves of our uplifts to the state-of-the-art uplifts and show their effect on the appearance of multiple reflectances when treated as a light source.

#### 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  $\delta E = 4.85e - 12$ , while the maximum Delta E achieved is  $\delta E = 6.53e - 12$ . Generally, colour differences of  $\delta E \leq 1$  are considered to not be perceivable by the human eye, 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 by the human eye - however, when matching plate footage with 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 prove that the results of our uplifting process deviate from the current state of the art (or the *scaled reflectance uplifting*), we compare the two methods in terms of their effect on scene appearance. As the scaled reflectance uplifting method, we specifically use a variation of the technique used in the Mitsuba 3 rendered, however, we can expect the scaling of any other type of smooth reflectance spectra to perform similarly.

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).

On average, we obtain an error of  $\delta E = 1.478$ . Although this is not significant in terms of human perception, as previously mentioned, digital counterparts of plate footage are sensitive to even slight colour deviations, making  $\delta E = 1.478$  too high. Additionally, for uplifts affected by more complex curves, such as e.g. the emission of fluorescent light sources, the average error goes up to  $\delta E = 1.73$ . Uplifts of environment maps with the LED light sources achieve an even higher error of  $\delta E = 1.88$ .

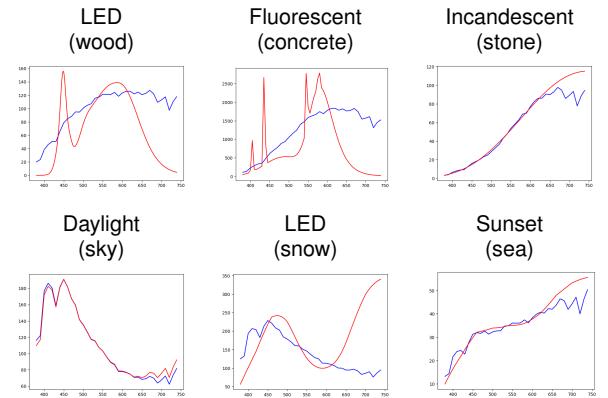
Another even more meaningful result is the *maximum error* between our uplifts and the scaled reflectance uplifts. For every combination of an atlas, an environment map and a light source type, we found the maximum error to attain values over  $\delta E > 20$ , and, for specific complex reflectances in the Pantone Atlas, even over  $\delta E > 40$ . We note that most of the high error values were achieved in the high dynamic range, for pixels that belonged to the light source. We also attribute some of the error to the deficiencies of our implementation of the scaled reflectance uplifting technique, which consists of scaling the reflectances obtained from the wide gamut sigmoid uplift cube implemented by Tólová et al. [TWF22]. As the cube is rather sparsely sampled (specifically, every dimension is evenly divided into 32 parts, summing up to  $32^3$  lattice points), even uplifting in the low dynamic range leads to some inaccuracies. These are even more pronounced in the high dynamic range, capable of achieving a round-trip error of over  $\delta E > 2$ . However, this is not high enough to be a significant influence on the maximum error achieved between the two techniques.

In Figure 8, we visualise the uplifted curves of some of the pixels of the input environment maps for both our and the scaled reflectance uplifting technique. Although it is clear that their differences in terms of shape are, in some cases, significant and that utilising them as illuminants would cause colour artefacts, this alone does not prove that our method is superior in terms of realism to the scaled reflectance uplifting. However, we claim that it achieves results closer to their real-life counterparts due to us utilising real world data as opposed to synthetic data. Comparisons similar to the one shown in Figure 7 further support our claim.

## 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 the environment map capture, which we then use to constrain the uplifting process. The realistic data provided by our method aids 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.

While the results of our method are satisfactory in terms of both



**Figure 8:** Comparison of emission curves achieved with our proposed method (red) to the current state of the art, i.e. scaled reflectance uplifting (blue). Every curve was obtained for a different pixel in a different environment map. In addition to the detected light source type, we also provide information about the type of surface the uplifted pixel belongs to. Note that the scaled reflectance uplifts are also multiplied by the D65 illuminant, as that is the white point of the uplift model. Also note the similarity of the uplifts in case of a daylight setting - this behaviour is expected as the identified illuminant of the environment map is almost identical to the D65 illuminant.

the round-trip accuracy and the shapes of the emission spectra, 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. In order to properly evaluate the accuracy of our method in terms of the uplifts' similarity to their real-life counterparts, we additionally propose obtaining ground truth data with spectral environment map captures.

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.

## 6. 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.

## References

- [BBG23] BELCOUR, LAURENT, BARLA, PASCAL, and GUENNEBAUD, GAËL. “One-to-Many Spectral Upsampling of Reflectances and Transmittances”. *Computer Graphics Forum*. Vol. 42. 4. Wiley Online Library. 2023, e14886 3.
- [Fai20] FAIRCHILD, MARK D. “Von Kries 2020: Evolution of degree of chromatic adaptation”. *Color and Imaging Conference*. Vol. 28. Society for Imaging Science and Technology. 2020, 252–257 6.
- [FHL\*18] FASCIONE, LUCA, HANIKA, JOHANNES, LEONE, MARK, et al. “Manuka: A batch-shading architecture for spectral path tracing in movie production”. *ACM Transactions on Graphics (TOG)* 37.3 (2018), 1–18 3.
- [JH19] JAKOB, WENZEL and HANIKA, JOHANNES. “A Low-Dimensional Function Space for Efficient Spectral Upsampling”. *Computer Graphics Forum*. Vol. 38. 2. Wiley Online Library. 2019, 147–155 3, 8.
- [JMW\*64] JUDD, DEANE B, MACADAM, DAVID L, WYSZECKI, GÜNTHER, et al. “Spectral distribution of typical daylight as a function of correlated color temperature”. *Josa* 54.8 (1964), 1031–1040 6.
- [JSR\*22] JAKOB, WENZEL, SPEIERER, SÉBASTIEN, ROUSSEL, NICOLAS, et al. *Mitsuba 3 renderer*. Version 3.0.1. <https://mitsubarenderer.org>. 2022 3.
- [JWH\*19] JUNG, ALISA, WILKIE, ALEXANDER, HANIKA, JOHANNES, et al. “Wide gamut spectral upsampling with fluorescence”. *Computer Graphics Forum*. Vol. 38. 4. Wiley Online Library. 2019, 87–96 3.
- [KPB\*18] KOKKA, ALEXANDER, POIKONEN, TUOMAS, BLATTNER, PETER, et al. “Development of white LED illuminants for colorimetry and recommendation of white LED reference spectrum for photometry”. *Metrologia* 55.4 (2018), 526 6.
- [Kur84] KURUCZ, ROBERT L. “Solar Flux Atlas from 296 to 1300 nm”. *National Solar Observatory Atlas* 1 (1984) 7.
- [LF20] LANGLANDS, ANDERS and FASCIONE, LUCA. “Physlight: An end-to-end pipeline for scene-referred lighting”. *Special Interest Group on Computer Graphics and Interactive Techniques Conference Talks*. 2020, 1–2 2.
- [Mac35] MACADAM, DAVID L. “The theory of the maximum visual efficiency of colored materials”. *JOSA* 25.8 (1935), 249–252 3.
- [McC92] MCCAMY, CALVIN S. “Correlated color temperature as an explicit function of chromaticity coordinates”. *Color Research & Application* 17.2 (1992), 142–144 5.
- [MMD\*76] MCCAMY, CALVIN S, MARCUS, HAROLD, DAVIDSON, JAMES G, et al. “A color-rendition chart”. *J. App. Photog. Eng* 2.3 (1976), 95–99 2.
- [MSHD15] MENG, JOHANNES, SIMON, FLORIAN, HANIKA, JOHANNES, and DACHSBACHER, CARSTEN. “Physically meaningful rendering using tristimulus colours”. *Computer Graphics Forum*. Vol. 34. 4. Wiley Online Library. 2015, 31–40 3.
- [OYH18] OTSU, HISANARI, YAMAMOTO, MASAFUMI, and HACHISUKA, TOSHIYA. “Reproducing spectral reflectances from tristimulus colours”. *Computer Graphics Forum*. Vol. 37. 6. Wiley Online Library. 2018, 370–381 3.
- [Smi99] SMITS, BRIAN. “An RGB-to-spectrum conversion for reflectances”. *Journal of Graphics Tools* 4.4 (1999), 11–22 3.
- [SW13] SHEN, YEHU and WANG, QICONG. “Sky region detection in a single image for autonomous ground robot navigation”. *International Journal of Advanced Robotic Systems* 10.10 (2013), 362 5.
- [TWF21] TÓDOVÁ, LUCIA, WILKIE, ALEXANDER, and FASCIONE, LUCA. “Moment-based Constrained Spectral Uplifting.” *EGSR (DL)*. 2021, 215–224 3.
- [TWF22] TÓDOVÁ, LUCIA, WILKIE, ALEXANDER, and FASCIONE, LUCA. “Wide Gamut Moment-based Constrained Spectral Uplifting”. *Computer Graphics Forum*. Vol. 41. 6. Wiley Online Library. 2022, 258–272 3, 7, 8, 10.
- [VBKW22] VÉVODA, PETR, BASHFORD-ROGERS, THOMAS, KOLÁŘOVÁ, MONIKA, and WILKIE, ALEXANDER. “A wide spectral range sky radiance model”. *Computer Graphics Forum*. Vol. 41. 7. Wiley Online Library. 2022, 291–298 6.
- [VE23] VAN DE RUIT, MARK and EISEMANN, ELMAR. “Metameric: Spectral Uplifting via Controllable Color Constraints”. *ACM SIGGRAPH 2023 Conference Proceedings*. 2023, 1–10 3.
- [WVB\*21] WILKIE, ALEXANDER, VEVODA, PETR, BASHFORD-ROGERS, THOMAS, et al. “A fitted radiance and attenuation model for realistic atmospheres”. *ACM Transactions on Graphics (TOG)* 40.4 (2021), 1–14 6.