DOI: 10.1111/j.1467-8659.2012.03024.x
EUROGRAPHICS 2012 / P. Cignoni, T. Ertl
(Guest Editors)

Volume 31 (2012), Number 2

Practical Spectral Photography

Ralf Habel[1], Michael Kudenov [2] & Michael Wimmer[1]

[1]Institute of Computer Graphics and Algorithms, Vienna University of Technology, Austria
[2]Optical Detection Lab, University of Arizona, USA

Abstract

We introduce a low-cost and compact spectral imaging camera design based on unmodified *consumer cameras and a custom camera objective. The device can be used in a high-resolution configuration that measures the spectrum of a column of an imaged scene with up to 0.8 nm spectral resolution, rivalling commercial non-imaging spectrometers, and a mid-resolution hyperspectral mode that allows the spectral measurement of a whole image, with up to 5 nm spectral resolution and 120x120 spatial resolution. We develop the necessary calibration methods based on halogen/fluorescent lamps and laser pointers to acquire all necessary information about the optical system. We also derive the mathematical methods to interpret and reconstruct spectra directly from the Bayer array images of a standard RGGB camera. This objective design introduces accurate spectral remote sensing to computational photography, with numerous applications in color theory, colorimetry, vision and rendering, making the acquisition of a spectral image as simple as taking a high-dynamic-range image.*

1. Introduction

Hyperspectral imaging, also called imaging spectroscopy, is a method of obtaining the spectral content of each pixel in a 2D image. This technology divides the image data, pixel by pixel, into very narrow wavelength (color) bands, resulting in a 3D data cube, consisting of two spatial image dimensions and a wavelength dimension (see Figure 2).

The applications of hyperspectral imaging are numerous, as the high-resolution spectral data can be used in many ways, such as identification of chemical compounds or materials, which can be essential in quality control, geology or mining, just to mention a few industrial applications. In graphics, knowledge of a full spectrum instead of only a RGB color triple is essential for fields such as color theory, colorimetry, vision, accurate image-based lighting design, spectral rendering and spectral material models. Most spectral measurement devices can only take spatially integrated measurements, whereas ideally, one would like to know the full spectrum of each individual pixel in an image. So far, this problem can only be solved using very expensive apparati (starting at $ 25,000 for even the most basic devices) that are fabricated for a specific application, for example, airborne remote sensing.

Recently, so-called "snapshot" spectral imaging has been developed based on the principles used in computed tomography. These *computed tomography image spectrometers* (CTIS) [OTY93] use a diffraction grating to split the incoming light into a number of spectral projections onto a single image plane. The spectra are then recovered by solving an underdetermined linear system. CTIS devices were developed for applications where very short acquisition times are important, like biomedical imaging, airborne sensing, and space surveillance [Hor10], in specialized devices. However, to obtain enough light in the diffraction orders, CTIS devices require expensive diffraction gratings, specialized sensors and lens arrangements that increases their size and weight.

Figure 1: *The assembled CTIS camera objective.*

In this paper, we show that by relaxing the requirement of sub-second exposure times, and instead using standard HDR image acquisition, one can perform CTIS spectral imaging using *unmodified* consumer cameras. The other princi-

pal advantages of CTIS are maintained, i.e., avoiding any opto-mechanical components, requiring only a simple optical path, and having all information recorded in one (HDR) image. In addition, we gain the further advantages of low cost and compact design, because we can use inexpensive, compact components. These features allow us to build a functional prototype in the form of a fully portable camera objective that can be cheaply constructed from off-the-shelf photographic components, leading to drastically lower cost than any other spectral imaging device while every component can still be used for its original purposes. Despite the severe compromises made, our prototype offers a spectral resolution of 0.8 nm in slit configuration (masking out all but a single pixel-wide column) and 4.89 nm resolution (54 spectral bands) in a hyperspectral configuration of 120x120 pixels. These resolutions are competitive with non-imaging spectrometers as well as dedicated CTIS devices, but at the expense of exposure times in the range of a few seconds rather than milliseconds, depending on the dynamic range of the scene. Such resolutions cannot be achieved at all with filter-based systems. The availability of a cheap and practical imaging spectrometer offers many interesting opportunities in almost all areas of research related to computer graphics, vision and color theory, as this device strongly simplifies the acquisition of spectra in general settings.

2. Related Work

2.1. Spectral Imaging

The three-dimensional hyperspectral signal can be measured using various methods, which mainly differ in the way they separate incident light into its spectral components. This can be done using dispersive elements, diffractive media or exchangeable/tunable bandpass filters [HBG*00]. To acquire the three-dimensional signal, almost all approaches mask out one of these dimensions, and use the time dimension to scan along the masked dimensions. For example, filter-based systems scan along the spectral dimension, while whisk broom-based spectrometers scan one pixel at a time. In general, these approaches require dedicated high-cost devices due to the necessity of accurate and dynamic opto-mechanical constructions and arrangements. Hybrid approaches are also possible. For example, using mosaicking for the spatial dimension and a linear variable interference filter for the spectral dimension simultaneously. While this approach works with commodity cameras [SN02], it requires a large number of image samples of the same panorama and provides only low spectral resolution. Du et al. [DTCL09] propose a quasi-hyperspectral approach by using a standard prism/slit configuration in a slit array, which does not allow a compact construction and has a lower resolution than our approach. Furthermore the result is discontinuous along the spatial dimension due to the masking required to interlace spatial and spectral data in the recorded image. Another approach, which is limited to measuring reflectance instead of directly measuring arbitrary spectra (e.g., light sources), is to use controlled lighting [PLGN07, CYBE10, KF09], creating the spectral separation by applying spectrally different light sources. These methods rely on the low dimensionality of material reflectance [Mal86]. They have only a limited spectral resolution and accuracy and require a very controlled dynamic lighting environment. Mohan et al. [MRT08] use an optical configuration similar to CTIS, consisting of an imaging lens, diffraction grating and re-imaging lens, to be able to modulate the spectrum with a mask to achieve color spectrum control of cameras and projectors.

2.2. Computed Tomography Imaging Spectrometry

Computed tomography imaging spectrometry is unique when compared to these approaches because it does not require temporal scanning. Instead, the signal is diffracted, distributing the spectral elements *spatially* by projecting the 3D volume at various angles on the image sensor. The full 3D volume can be reconstructed by inverting the projection. Due to its resulting "snapshot" quality, CTIS have found many applications in different areas as it overcomes problems such as long scanning times of dynamic scenes (causing motion artifacts), and fragility of optical devices with moving parts. CTIS have been used in biomedical imaging [JWF*07], microscopy [GKHT10], space surveillance [HOWJD03] and even defense applications [DDD95], just to name a few examples. In contrast to our device, this previous work requires expensive custom-built devices.

2.3. The CTIS Principle

At the heart of any CTIS sits a diffraction grating, which consists of dense grooves with sizes roughly on the same scale as the wavelength of the electromagnetic radiation it diffracts. Due to interference effects, light is diffracted in different directions dependent on its wavelength, creating higher order modes of diffractions with decreasing intensity (see [Hec01] or any optical physics book, Figure 12 shows an example of the zero mode and 4 first-order diffraction modes). In a CTIS, an image is formed by using the diffraction grating to divert the light path depending on the wave length. The formed image is a superposition of the responses to the different wavelengths, and can be interpreted as parallel projections of the data cube in different directions (see Figure 2). The data cube consists of voxels that are defined by the image resolution of the acquisition and the grating density. The size of the zero mode defines the spatial axis resolution, and the angle of parallel projection sets the wavelength resolution. To obtain a spectral image, the full data cube needs to be reconstructed from the projections (hence computed tomography) formed by the grating and in our case, the zero mode.

Our proposed spectrometer camera objective can operate in two modes: using a slit aperture and using a square aperture. In the first case, only a single column of the scene is

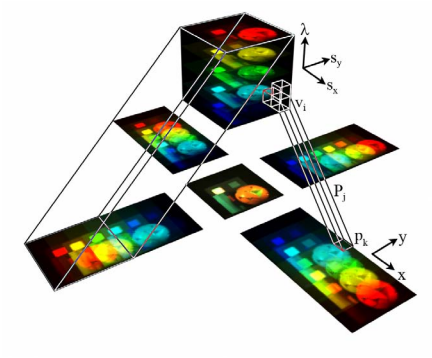

Figure 2: *Diffractions interpreted as parallel projections P_j of data cube voxels v_i to an image plane pixel p_k.*

imaged, and the spectrum can be recovered directly from the first mode (see Section 5). In the second case, a whole image of the scene is formed, which requires reconstruction of the data cube by inverting the projection (see Section 6).

3. Overview

The main idea of our approach is to exploit the CTIS principle to build an inexpensive imaging spectrometer using a consumer camera, avoiding the extreme cost of specialized devices. This is enabled by four main contributions:

- Use of HDR acquisition instead of relying on expensive diffraction gratings, specialized sensors, and large optical arrangements to collect enough light in the projections.
- Use of the Bayer filter array present in a consumer camera instead of an expensive full-spectrum monochromatic sensor.
- Showing which standard optical components are required and how they can be combined in an objective to create the optical path required by CTIS.
- Simple calibration methods to achieve very high accuracy, exploiting every single element in the Bayer filter array.

In Section 4, we describe the physical and optical construction of the proposed CTIS device and our functional prototype. By using two different apertures, we can measure spectral data in slit mode, which is covered in Section 5, and in hyperspectral mode, shown in Section 6. For both modes, we discuss the necessary calibration steps, the mathematics required for reconstruction, and demonstrate results.

Calibration has two parts: spatial and spectral. The *spatial calibration* localizes the positions of individual wavelengths on the image grid (in slit mode), or determines the coefficients of the projection of the data cube. In both cases, the calibration will allow a precise mapping from the projected image to wavelengths in the spectrum. The *spectral calibration* calculates the spectral intensity response of the optical system and the sensor pixels. The pixels are arranged in a RGGB Bayer color filter array, so each pixel responds more in the red, green or blue range of the spectrum. This means that similar to standard imaging, we need to apply demosaicing algorithms to recover the full spectra, described in Section 6.4.

4. The CTIS Camera Objective

4.1. Optical Design

The principal optical path of a CTIS imaging system is displayed in Figure 3. An imaging lens images the scene onto a slit or square aperture. The divergent light that passes through the aperture is collimated (i.e., made parallel) by the collimating lens, after which the light passes through the transmissive diffractive medium. Both the non-diffracted as well as the diffracted light get re-imaged to the camera sensor through the re-imaging lens. It is essential that light is collimated before diffraction because divergent diffracted light would show the imaged scene from a different perspective, which creates an inconsistency of the zero mode and higher order diffraction modes.

In photographic terms, we can intuitively understand this design as consisting of two cameras: the first camera takes a picture of a scene, but instead of imaging it to a sensor, the image is formed at the aperture plane. The second camera, using a re-imaging lens with a macro lens attached, records the image in the aperture, i.e., the image created by the first camera. It is this intuitive description of the optical path that leads to a design involving standard photographic components for the prototype.

4.2. Camera Objective Construction

Starting from this intuition, construction becomes a matter of positioning the optical components within one optical path. To achieve an appropriately long optical path, we use a 75mm and 40mm diameter PVC drain piping system in our prototype to mount the components and to provide the possibility to easily focus and position the optical components via slip-on pipe connection sleeves.

Let us follow the light path from the sensor (right-hand side in Figure 3) and discuss the components of the prototype (Figure 4): The re-imaging lens is a standard camera lens that mounts directly to the camera. Attached to the re-imaging lens is a gel holder which contains a diffraction gel. The gel holder's lens hood provides the means to mount the 75 mm PVC tube in front of the gel holder. Inside the PVC tube, the collimating lens is positioned directly in front of the diffraction gel. The imaging lens is mounted using a slip-on pipe connection sleeve at the other end to allow for easy focusing of the collimation through moving the slip-on sleeve. The aperture behind the imaging lens is mounted to it using 40 mm PVC tubes and a cut-through camera lens cap, again via a slip-on connection sleeve for easy positioning and exchange of the aperture. The apertures themselves,

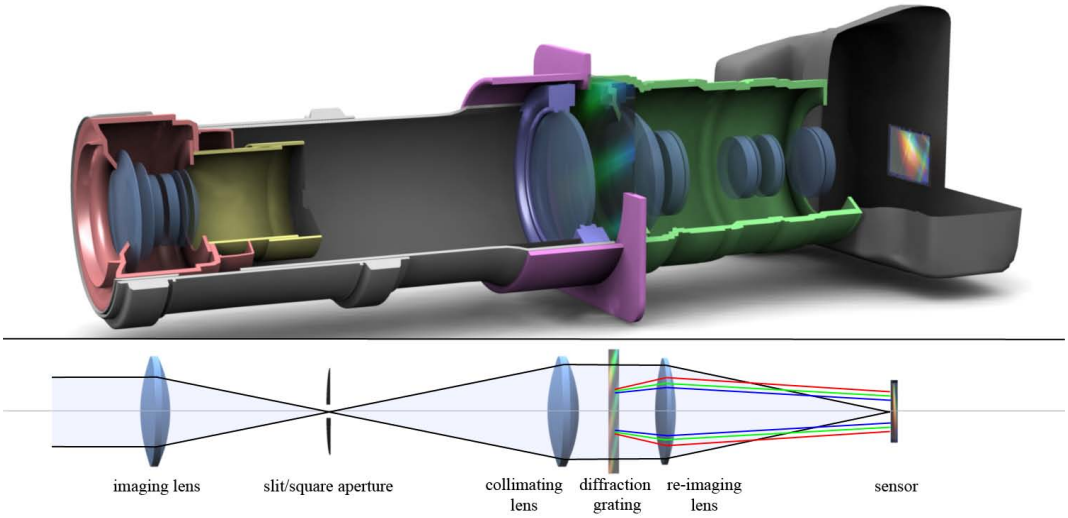

Figure 3: *A schematic cut-away of the prototype camera objective (top) consisting of a 50mm/1:1.8 camera lens as the imaging lens (red), a +10 Diopter macro lens (blue) as the collimating lens and a 17-40mm/1:4 camera lens as the re-imaging lens (green). The aperture is mounted to the imaging lens using a 40 mm tube and slip-on sleeve (yellow). The imaging lens is mounted in a 75mm slip-on sleeve and tube (grey) which is fixed to the gel holder (purple). The gel holder places the diffraction gel between the collimating and re-imaging lens. The principal optical path is shown on the bottom.*

both slit and square, are formed with razor blades to ensure straight edges. To suppress stray light, both the 75 mm tube and the 40 mm aperture tube are lined with light absorbing black paper. The used diffraction gels are cheaply available as sheets [Edm11] or come in the form of novelty glasses. The diffraction gels are holographically produced phase gratings and therefore very accurate and produce very little stray light. Several linear gratings and a cross grating are available. It is the existence of this gel that makes a simple construction possible because glass-based gratings are usually highly fragile and are not to be touched, whereas we can handle the diffraction gel like any other filter gel.

Figure 4: *The disassembled objective showing the imaging lens (left). The collimating lens is mounted inside the PVC tube with the slip-on pipe connection sleeve (middle). The gel holder containing the diffraction grating is directly mounted to the re-imaging lens on the camera (right).*

4.3. Choosing the Optical Components

We only use standard camera lenses without any modifications (in particular, without disassembling them) for the prototype, so any optical component, including the camera itself, can still be used for their original purpose. Because camera lenses are accurate in their imaging properties, we can ensure high quality and avoid artifacts such as chromatic aberrations. Chances are great that a photographic enthusiast already has a set of camera lenses that can be combined to an imaging spectrometer. Any combination of lenses and aperture sizes are permissible that fulfil the following conditions:

- The first diffraction modes must be fully imaged.
- The first modes must be far enough away from the zero mode to avoid the influence of glare in the first order.
- In each color channel, the first and second modes must not overlap.

Depending on the grating, two lenses with short focal lengths (20-50mm) for the imaging lenses and one with a long focal length (80-150mm) as the collimating lens is required, though any camera lens can act as the imaging lens.

In our case, we chose as imaging lens a Canon EF 50mm/1:1.8 due to its light weight, large field of view and large aperture. Collimation is done by an achromatic macro lens with +10 Diopter (100 mm focal length), because it is light in weight and does not limit the effective collimating aperture. As can be seen in Figure 3, the focal lengths of the lenses determine the respective distances between compo-

nents, as the slit/square aperture sits in the focal point of both lenses. The re-imaging lens is a Canon EF 17-40mm/1:4 which, due to its large size, increases the overall length of the imaging spectrometer objective, but the varying focal length proved to be valuable since we can trivially change the area of projection on the sensor to accommodate for the different situations in slit and square aperture configurations. However, this choice also comes at the cost of a small re-imaging aperture so that the light efficiency is not optimal. All lenses are used at their maximum aperture setting. The camera in use is a Canon EOS 5D Mark I.

The size of the slit aperture should ideally project to exactly 2 Bayer pixels (~0.1 mm in our prototype) in its image – any slit larger than that will reduce the effective resolution but also increases the intensity of the diffractions. The size of the square aperture can be set to fulfil the previously mentioned conditions if only a limited set of lenses is available. In our case, the square aperture has a size of 5mm $\times$ 5mm.

4.4. HDR Image Acquisition

Apart from the physical design, operating mode of the camera is also important when creating a CTIS. In particular, we need to take the limitations of light sensitivity and dynamic range of the used sensor into account. We therefore record images using standard high-dynamic-range (HDR) image capturing [DM97]. Using a HDR approach solves several problems at once: First and foremost, HDR imaging largely extends the dynamic range, so we can image strong emissive spectra (such as from bright light sources) and far weaker reflective spectra (such as from surfaces) even in the same image. Furthermore, the extended dynamic range is important to properly expose both the zero mode and the diffractions, which have only about 10% of the zero mode intensity with the diffraction grating used our prototype. (Note that this is in contrast to expensive blazed diffraction gratings used in other CTIS devices: they have much more intensity in the diffractions, overexposing the zero-mode, but also have diffraction patterns which allow omitting the zero-mode from the reconstruction, thus avoiding the need for HDR.) Finally, there are regions at the boundaries of the visible spectrum where the response of the RGGB sensor is weak, and longer exposures can increase the accuracy in those regions.

To avoid any processing of the recorded data, we directly read out the image in raw format because we need to take into account the relative positions of the pixels of the Bayer filter. This also keeps the 12-bit accuracy recorded by the camera. However, instead of recovering the intensity response of the whole system using an optimization procedure as proposed by Debevec and Malik [DM97], we use the spectral calibration procedure based on a halogen lamp, described in Section 5, and apply it to each color channel separately.

4.5. Camera and Imaging Spectrometer Objective Parameters

Due to the nature of the optical path, we only get low light intensities onto the sensor and we need to balance the camera and lens parameters according to the lighting situation.

The focal length of the re-imaging lens directly defines the spectral resolution (for a fixed diffraction grating), as it determines the area of projection on the sensor. In case of the hyperspectral mode, together with the size of the square aperture, it also defines the spatial resolution. One of the main concerns in acquiring the signals is to avoid noise, though we can partly counteract noise by averaging several spectra after the spectral reconstruction. To avoid strong noise in the first place, we use the ISO value giving the lowest noise, which we experimentally determined to be ISO 50 in the case of the Canon 5D Mark I. Here again we can trade noise against light intensity by choosing a higher ISO value. We found that signals recorded with an ISO value of under 200 still deliver accurate spectra. The HDR reconstruction process also reduces noise [DM97], as only well-exposed intensities are used to build the HDR image.

5. Slit Calibration and Imaging

In this section, we describe the operation of the CTIS device in slit mode, measuring one column of the scene. We first describe the spatial and spectral calibration procedures, and then how to measure an actual spectrum. Note that the spectral calibration is carried out only once for a given set of components, and is reused for the hyperspectral mode. In both cases, we assume the intensity response function of the sensor has already been taken into account through the HDR acquisition process (Section 4.4). For the slit configuration, we use a linear 600 lines/mm grating, which produces 2 usable first order diffractions.

5.1. Spatial (Wavelength) Calibration

The spatial calibration corrects for distortions in the optical paths, and for misalignments between the slit and diffraction grating. Given an image with pixels (x,y), we want to determine the function $f(x,y) = (\lambda, s)$ that maps an image pixel to a tuple of wavelength and position s along the slit axis. To recover this function, we use a light source with a number of well defined spectral peaks to obtain data points for λ. We mask out parts of the light source (see Figure 5) to obtain a structured light, which gives data points in s-direction. We then fit bicubic polynomials to the given data points to undistort the image [Kil81].

In practice, we use a standard fluorescent lamp, which covers the whole slit in a homogeneous light emission. The fluorescent spectrum has three strong and consistent peaks in red (611.3 nm), green (546.1 nm) and blue (435.7 nm). This spectrum is ubiquitous in our environment, be it fluorescent

lamps, energy saving bulbs or TFT monitors, and therefore easy to come by. The raw output after the HDR composition can be seen in Figure 6. In this image, we mark all the flanks, which are the positions where the intensity drops to zero due to the masked parts, of one of the diffractions directly on the Bayer filter array in order to account for the different positions of the Bayer pixels. In our prototype, the imaged slit has a height of 860 Bayer pixels.

For each flank f_i, we first mark its position in the zero order image to obtain the slit position s_i (as the image y-coordinate), and for each of the three spectral peaks to obtain three triples $(x_{ij}, y_{ij}, \lambda_j), j \in r,g,b$. For N flanks, this gives a total of $3N$ calibration samples $(x_{ij}, y_{ij}, s_i, \lambda_j)$, which we use to fit two bicubic polynomials $s(x,y)$ and $\lambda(x,y)$, giving the complete mapping from image space to slit-axis and wavelength space. This general bicubic fitting approach has the advantage that both lens distortions and misalignments of the diffraction grating and slit aperture orientations can be corrected in a single step. The theoretical spectral resolution that can be reached in the slit configuration can be calculated as $\frac{x_{ir}-x_{ib}}{\lambda_r-\lambda_b}$ for some i.

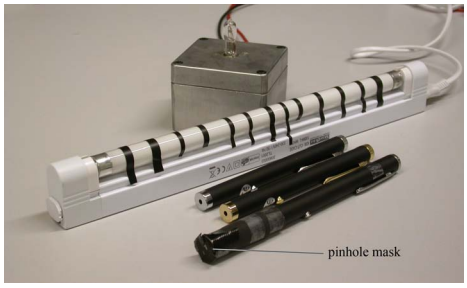

Figure 5: *The calibration equipment consisting of a calibration halogen lamp (top), a striped fluorescent lamp (middle) and red, green and blue laser pointers and a pinhole mask (bottom).*

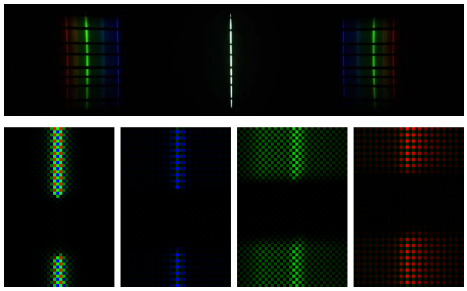

Figure 6: *Raw output after the HDR composition of the masked fluorescent tube (top) and close-ups of diffractions of the main peaks from a masked section in Bayer array pixels (bottom).*

To extract the spectral responses from an input image, we can either directly interpolate the responses of single slit positions using the given polynomials, or we can undistort the complete diffraction of the slit directly on the Bayer array to arrive at a two-dimensional signal that contains the undistorted diffractions of every slit position.

5.2. Spectral Response Calibration

The purpose of the spectral response calibration is to determine the spectral sensitivities of each color channel. The spectral response can be determined by photographing a light source that has a spectrum that is continuous and known with high accuracy. For this purpose, we use a dedicated calibration halogen lamp. Note that this is the only step in the prototype construction that requires the application of previously calibrated equipment. We average over a few responses and then simply divide through the known spectrum of the halogen lamp. By considering red, green and blue Bayer pixels separately, we arrive at the spectral response functions $S_{r,g,b}(\lambda)$ of each channel. We assume that the high-frequency structures of the response functions are noise and filter slightly with a 99% Fourier low-pass filter. The results of this calibration step are displayed in Figure 7, showing the spectral response of the complete optical system.

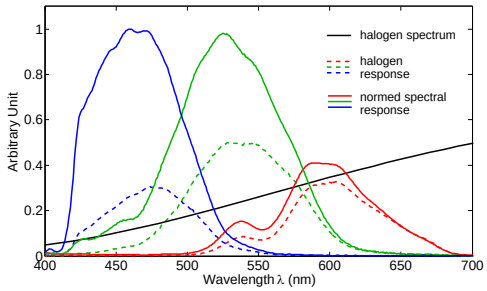

Figure 7: *The response to the halogen spectrum and the normed spectral response $S_{r,g,b}(\lambda)$ of our optical system.*

5.3. Measuring Slit Spectra and Demosaicing

To reconstruct the true spectrum from a measurement, in addition to applying the spatial calibration $f(x,y)$, we need to take into account the different sensitivities of the different color channels in the Bayer filter. For regular image processing, this is usually called demosaicing. Since we have determined the accurate spectral response functions for each color channel, any of the channels could be used alone to reconstruct a spectrum. However, the spectral response of each channel covers only part of the visible spectrum and the positions of the Bayer pixels are interlaced, so we need to combine the measurements of adjacent Bayer pixels with interpolated values. In order to suppress noise, we combine the

channels with weights $w_{r,g,b}(\lambda)$ corresponding to the amplitude of their spectral response:

$$w_{r,g,b}(\lambda) = \frac{S_{r,g,b}(\lambda)}{\sum_{c\,=\,r,g,b} S_c(\lambda)}. \quad (1)$$

A spectrum $M(\lambda)$ of a response measurements $m_{r,g,b}(\lambda)$ is then reconstructed with

$$M(\lambda) = \sum_{c\,=\,r,g,b} \frac{m_c(\lambda)}{S_c(\lambda)} w_c(\lambda) = \frac{\sum_{c\,=\,r,g,b} m_c(\lambda)}{\sum_{c\,=\,r,g,b} S_c(\lambda)} \quad (2)$$

with interpolated values where necessary. To actually measure spectra, we simply take an HDR image of a scene and carry out the above calculations. While emissive spectra only need short exposures (< 250 ms), reflective spectra require exposures of up to 8 seconds at an ISO value of 50 for a high-quality signal. But this is not different from standard HDR image acquisition or dedicated spectrometers, where measurements also may need to integrate over times at the scale of seconds for small light intensities.

5.4. Results

We measured reflective spectra which emit with 50-300 cd/m^2, corresponding to a well lit room. Figure 8 shows spectra of different patches from a color chart under halogen lighting. We can trivially measure the reflectivity by additionally measuring the light source and dividing the measured material spectrum by the light source spectrum.

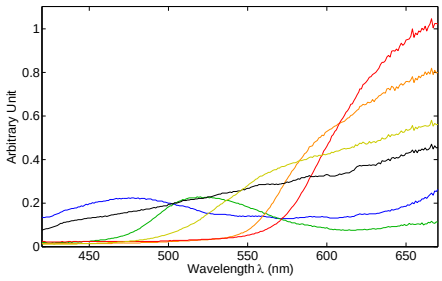

Figure 8: *Spectra of differently colored materials under halogen lighting. The color of the spectra corresponds to the color of the material.*

To test the accuracy of our calibrated CTIS device, we use the fluorescent spectrum used for the wavelength calibration because it consists of peaks and smooth as well as structured sections. We compare our result with a measurement of the fluorescent tube using an independently calibrated Ocean Optics USB 4000 non-imaging spectrometer together with an integrating sphere (see Figure 9). We can match the result of this professional spectrometer at a resolution of 0.8 nm in our configuration, with an average deviation of less than 1% from the USB 4000 measurement. Unfortunately, the Bayer filter of the Canon EOS 5D is very well designed and has a sharp cut-off at 410 nm and no infrared spill with a full cut-off at 700 nm. This limits the spectral range from 420 nm to 670 nm, as the signal degrades in the high blue and low red sections due to the very small spectral responses. However, this is just a limitation of the camera used in our prototype and not the method itself, and other cameras with different spectral responses may deliver a higher spectral range, or the Bayer filter could be removed from the camera, which however would require disassembly.

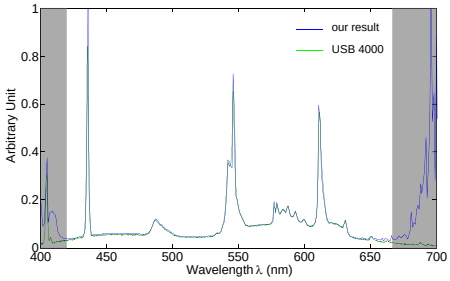

Figure 9: *Our result compared to an measurement taken with an Ocean Optics USB 4000. Sections with too inaccurate reconstructions are marked in grey.*

6. Hyperspectral Calibration and Imaging

Although the slit configuration provides high-resolution spectral sensing, we can only measure a column of a scene. By using a square aperture instead of a slit aperture, we create a computed tomography imaging configuration.

We use a 531.5 lines/mm cross grating that creates 4 usable first order diffractions in orthogonal directions. As we will later see, we are limited by the reconstruction rather than by the measurement resolution. Because of that, we use the re-imaging lens at 17mm focal length to maximize the light intensity per pixel, avoiding strong distortions as only a small area of projection in the center of the sensor is used. Still, we can reach a theoretical resolution of 4.59 nm and a final spectral image resolution of 124x124 Bayer pixels with our system, which is enough for most applications. A typical CTIS image acquisition can be seen in Figure 12.

6.1. CTIS Reconstruction Theory

The essential difference between the slit and square aperture configuration is that while we can directly extract single responses in the slit configuration, the responses of every Bayer pixel in the image of the square aperture (i.e. zero mode) are superimposed.

The parallel projection (see section 2.3) is described by a matrix H which operates on the vector $\vec{f}$ containing the serialized voxels of the data cube. The diffraction operation

produces the vector $\vec{g}$, representing the serialized version of the image pixels the CTIS objective acquires:

$$\vec{g} = H\vec{f}. \tag{3}$$

H is a $M \times N$ matrix, where M is the number of pixels in the image and N is the number of voxels in the data cube. Each column contains 5 entries at the position of the diffracted pixels at one wavelength and the zero mode. Therefore, the H-matrix is very large but also very sparse. To solve for the data cube, we need to formally invert equation 3. Since H is a non-square matrix, we need to apply an iterative process to solve for $\vec{f}$.

Several options and variations are available. Like previous authors, we found that the expectation maximization algorithm (EM) [SV82] produces consistent results and converges relatively fast and is easy to implement. With a given H-matrix, we generate a first guess through $\vec{f}^0 = H^T\vec{g}$ and perform the iterative expectation-maximization step

$$\hat{f}_n^{k+1} = \frac{\hat{f}_n^k}{\sum_{m=1}^{M} H_{mn}} \sum_{m=1}^{M} H_{nm}^T \frac{g_m}{(H\hat{f}^k)_m} \tag{4}$$

until the system converges. We implicitly treat noise as part of the signal and do not pose any constraints on the system for simplicity, though more specialized methods can improve the convergence stability and speed. A simple implementation of the EM iteration in MATLAB is given in Appendix A in the additional materials.

6.2. Hyperspectral Spatial Calibration

Calibrating the CTIS configuration corresponds to determining the entries of the H matrix. Let us call (s_x, s_y, λ) the coordinates in the data cube, and (x, y) the coordinates on the image plane (see Figure 2). Let P_j be one of the 5 projections from the data cube to the image plane, i.e., $P_j(s_x, s_y, \lambda) = (x, y)$. If we know P_j, then the matrix H can be constructed easily. Let us assume the voxels are indexed as $\vec{f} = (v_1, \ldots, v_N)^T$, and the Bayer pixels are indexed $\vec{g} = (p_1, \ldots, p_M)^T$. Then for each data cube voxel $v_i = (s_{x_i}, s_{y_i}, \lambda_i)$, construct the 5 projections $P_j(s_x, s_y, \lambda) = (x, y)$, which gives 5 pixels $p_k = (x_k, y_k)$ with indices $k \in 1 \ldots M$. Then H can be constructed by setting the 5 corresponding entries $p_{k_1}, \ldots, p_{k_5}$ in the $i-th$ column of H to one.

In order to do this, we first need to determine the positions of the 5 projections of the data cube. Here we take a similar approach as in the slit calibration, although since we have two spatial axes, we need to localize our calibration light source in both spatial dimensions. We use a red, green and blue laser (see Figure 5) to excite different image positions in both spatial dimensions. As a first step, we determine the correct wavelengths of the red, green and blue laser by measuring them with the slit configuration. The goal of the calibration is to excite only 1–4 Bayer pixels of the respective color and find the excited pixels in the diffractions to define the H-matrix. To create a very localized light source, we shoot a laser pointer through a white translucent gel that is masked by a tiny pinhole (< 1 mm), see Figure 5. The translucent gel avoids two problems. First, we cannot point a laser directly into the camera due to its high intensity, and the gel mostly destroys the coherency of the laser light. We then photograph the pinhole over a distance of a few meters at different positions, sampling the complete image with all 3 laser pointers (see Figure 10). For each sampled laser

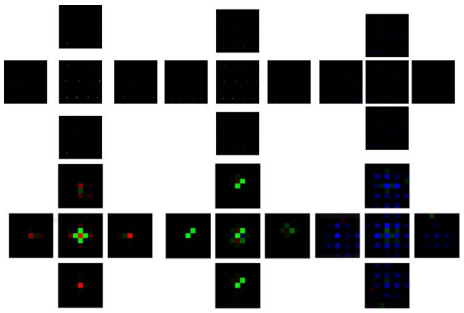

Figure 10: *Calibration points of the red, green and blue laser (top, cropped for clarity) and close-ups of a red, green and blue calibration point and diffractions (bottom).*

point, we obtain 5 projections $P_j(s_x, s_y, \lambda) = (x, y)$. We use the zero order diffraction, resulting from P_0, as the basis for the data cube coordinate system. Thus, for a laser point, we can directly recover the (s_x, s_y)-coordinate from the position of the laser point in the zero mode. For the other projections $P_{1..4}$, we take the position of the laser point in that projection and directly obtain the projection of the corresponding data cube voxel $P_{1..4}(s_x, s_y, \lambda) = (x, y)$.

From this sparse set of explicitly measured projections that define the entries in H, we reconstruct the full projections through linear interpolation and fill the other entries to complete the H matrix. Please note that we do not take the intensity of the diffractions into account, but use only the geometric positions, which introduces an irrelevant global scaling factor to all reconstructed responses. In contrast to the slit configuration, a linear interpolation is sufficient, because the smaller area covered on the sensor leads to only very light distortions.

6.3. Measuring Hyperspectral Images

With a fully defined H-matrix, we can now solve the data cube of any recorded CTIS image. To avoid unnecessary zeros and therefore unnecessary calculations in the EM iteration of equation 4, we separate out each channel from the Bayer array image, treating the two green sub-grids of the RGGB Bayer array as separate channels. We do the same with the H matrix to extract H matrices for each of the four channels, only solving for wavelengths where the corresponding response function is not zero (see Figure 7). In

our case, this results in 4 maps with a size of 402 × 402 pixels and a spectral resolution of 4.59 nm in each channel, defining H-matrices of the size of 161,604 × 215,818 for red, 161,604 × 223,260 for the green channels and 161,604 × 186,050 for blue. Depending on the image content, the 4 systems converge enough after 10-20 iterations. Our non-optimized but multithreaded MATLAB implementation can solve all 4 systems in 11 minutes on a 16 core (4 processors) AMD Opteron 2 GHz system.

6.4. Spectral Demosaicing and Reconstruction

Before we reconstruct the spectra from the response functions, we apply a demosaicing step similar to demosaicing RGB images. We use bilinear interpolation to get RGB responses for each Bayer pixel. Spectral demosaicing is principally different from standard demosaicing because there is a significant overlap of the responses which can be used to demosaic the spectral information. Therefore, spectral demosaicing constitutes an open problem, as many more sophisticated interpolation and vision-based algorithms can be applied spectrally. With a demosaiced data cube, we recover every pixel spectrum using equation 2. To calculate the spectral camera response at the lower resolution, we cannot simply re-sample the spectral camera response $S_{\mathrm{r,g,b}}(\lambda)$, but we need to take into account that each pixel integrates over a larger area. Therefore, we integrate the spectral response measured with the slit configuration (see Figure 7) and integrate over the respective range of each low resolution range r (4.59 nm in our case) to calculate the response $S^l_{\mathrm{r,g,b}}(\lambda_i)$ at lower resolutions:

$$S^l_{\mathrm{r,g,b}}(\lambda_i) = \int_{\lambda_i - \frac{r}{2}}^{\lambda_i + \frac{r}{2}} S_{\mathrm{r,g,b}}(\lambda)\,\mathrm{d}\lambda. \qquad (5)$$

6.5. Results

To show the accuracy of the CTIS configuration, we again use the fluorescent lamp spectrum from the slit calibration. In Figure 11, we reconstruct the fluorescent spectrum and compare it to a down-sampled version of the high-resolution spectrum following equation 5.

As can be seen, we can resolve all peaks and the structure of the spectrum is preserved. To compare the consistency of the CTIS configuration, the same spectra displayed in Figure 8, measured and reconstructed in the CTIS configuration are shown in Figure 11. The full hyperspectral reconstruction results in 54 spectral slices of the data cube at a resolution of 4.59 nm. Figure 12 shows the recorded image from which the data cube displayed slices have been reconstructed. Full data sets can be found in the additional electronic material.

Convergence Errors

It is not guaranteed that the expectation maximization converges to an accurate solution, as sub-pixel structures of the

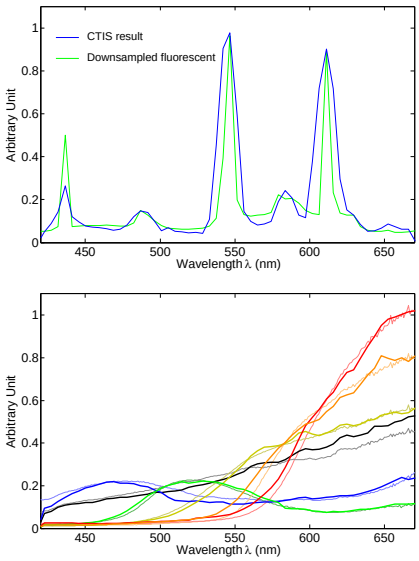

Figure 11: *The fluorescent spectrum measured with the CTIS configuration in comparison to a downsampled version from Figure 9 (top) and the previously measured spectra shown in Figure 8 measured with the CTIS configuration (bottom).*

photographed scene can cause artifacts in the reconstruction. This is a problem inherent to all discrete sampling and reconstruction methods, including all CTIS methods and devices. Because of the superposition of all spectra, the error is transported through the whole reconstruction. Though the optical low-pass filter of the camera still reduces aliasing, it may not be sufficient to fully suppress convergence artifacts. In such cases, we found that photographing the scene slightly out of focus can minimize the convergence artifacts at the expense of effective spatial resolution.

7. Conclusion and Future Work

We have shown that consumer cameras together with a low-cost custom camera objective can deliver both high-resolution single spectra via a slit configuration as well as full hyperspectral imaging. Both the imaging spectrometer objective construction and the measurements presented are at the lower end of what is possible. To improve on the construction, we intend to build a camera lens extension that directly mounts to a camera, using optical tube systems and custom lenses similar to macro extension tubes. This allows for a far shorter and lighter build and a significantly increased system aperture, avoiding long exposure times while the cost is still kept well under $ 1,000 compared to the cost of ten thousands dollars for dedicated devices. We would also like to explore applications for mobile devices as well as video cameras through size reduction of the camera lens,

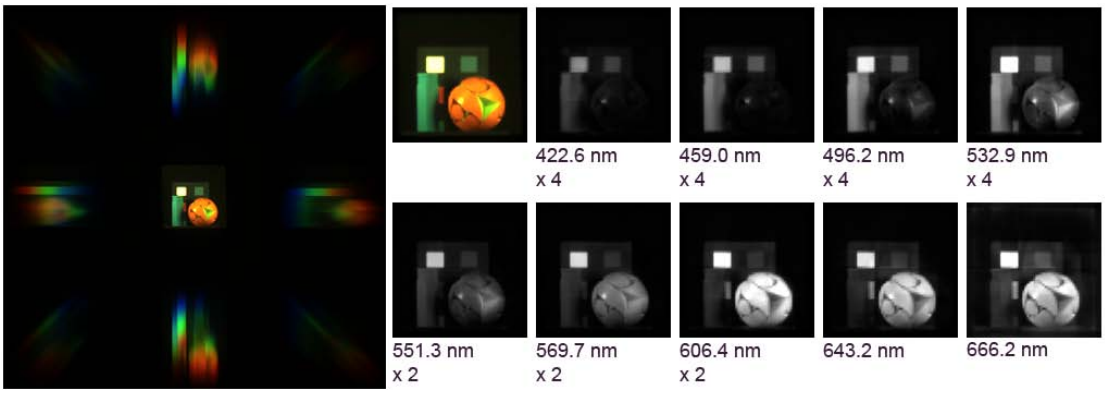

Figure 12: *The cropped result of a CTIS measurement with exposure adjusted zero mode (left) and a selection out of the 54 data cube slices (intensity scaled for clarity).*

since even low-quality signals can give significant information in a spectrum.

But the biggest challenge is to apply and develop mathematical methods that can solve the data cube faster and with increased stability to make large format CTIS spectrometry possible, applying methods and constraints specialized to the shown approach. Only recently, novel directions to this problem have been proposed [HDS07, Hor10]. With the help of GPUs and specialized constraints and preconditioning, we believe that a real-time solution even for large formats is possible, making a full real-time acquisition and reconstruction using consumer video and photographic cameras feasible.

References

[CYBE10] CHI C., YOO H., BEN-EZRA M.: Multi-spectral imaging by optimized wide band illumination. *Int. J. Comput. Vision 86* (January 2010), 140–151. 2

[DDD95] DESCOUR M. R., DERENIA E. L., DUBEY A. C.: Mine detection using instantaneous spectral imaging. In *Detection Technologies for Mines and Minelike Targets* (1995), pp. 286–304. 2

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co., pp. 369–378. 5

[DTCL09] DU H., TONG X., CAO X., LIN S.: A prism-based system for multispectral video acquisition. In *ICCV* (2009), pp. 175–182. 2

[Edm11] Edmund optics, stock no. nt54-509,nt40-267, 2011. `http://www.edmundoptics.com`. 4

[GKHT10] GAO L., KESTER R. T., HAGEN N., TKACZYK T. S.: Snapshot image mapping spectrometer (ims) with high sampling density for hyperspectral microscopy. *Opt. Express 18*, 14 (Jul 2010), 14330–14344. 2

[HBG*00] HARVEY A. R., BEALE J. E., GREENAWAY A. H., HANLON T. J., WILLIAMS J. W.: Technology options for hyperspectral imaging. In *Proc. SPIE Imaging Spectrometry VI* (2000), vol. 4132, pp. 13–24. 2

[HDS07] HAGEN N., DERENIAK E. L., SASS D. T.: Fourier methods of improving reconstruction speed for ctis. In *Imaging Spectrometry* (2007), vol. 6661, pp. 666103.1–666103.11. 10

[Hec01] HECHT E.: *Optics (4th Edition)*, 4 ed. Addison Wesley, Aug. 2001. 2

[Hor10] HORTON M. D.: *A Novel Technique for CTIS Image Reconstruction*. PhD thesis, Tenessee, Knoxville, 2010. 1, 10

[HOWJD03] HEGE K., O'CONNELL D., W. JOHNSON S. B., DERENIAK E. L.: Hyperspectral imaging for astronomy and space surviellance. In *Imaging Spectrometry IX, Proc. SPIE* (2003). 2

[JWF*07] JOHNSON W. R., WILSON D. W., FINK W., HUMAYU M., BEARMA G.: Snapshot hyperspectral imaging in ophthalmology. *Journal of Biomedical Optics*, 12 (2007), 014036. 2

[KF09] KRISHNAN D., FERGUS R.: Dark flash photography. In *SIGGRAPH '09* (2009), SIGGRAPH '09, pp. 96:1–96:11. 2

[Kil81] KILPEL E.: Compensation of systematic errors of image and model coordinates. *Photogrammetria 37* (1981), 15 – 44. 5

[Mal86] MALONEY L. T.: Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *J. Opt. Soc. Am. A 3*, 10 (Oct 1986), 1673–1683. 2

[MRT08] MOHAN A., RASKAR R., TUMBLIN J.: Agile spectrum imaging: Programmable wavelength modulation for cameras and projectors. *Computer Graphics Forum 27*, 2 (2008), 709–717. 2

[OTY93] OKAMOTO T., TAKAHASHI A., YAMAGUCHI I.: Simultaneous acquisition of spectral and spatial intensity distribution. *Appl. Spectrosc. 47*, 8 (Aug 1993), 1198–1202. 1

[PLGN07] PARK J., LEE M., GROSSBERG M. D., NAYAR S. K.: Multispectral Imaging Using Multiplexed Illumination. In *IEEE International Conference on Computer Vision (ICCV)* (Oct 2007). 2

[SN02] SCHECHNER Y., NAYAR S.: Generalized Mosaicing: Wide Field of View Multispectral Imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 10 (Oct 2002), 1334–1348. 2

[SV82] SHEPP L. A., VARDI Y.: Maximum Likelihood Reconstruction for Emission Tomography. *Medical Imaging, IEEE Transactions on 1*, 2 (1982), 113–122. 8