Paper: An improved optimization method for finding a color filter to make a camera more colorimetric

Graham D. Finlayson *, ***Yuteng Zhu*** *
Colour and Imaging Laboratory, University of East Anglia, Norwich, UK
** ***Joint First Authors***

Abstract

Recently, an iterative optimization method was proposed that determines the spectral transmittance of a color filter which, when placed in front of a camera, makes the camera more colorimetric [1]. However, the performance of this method depends strongly on the filter (guess) that initializes the optimization. In this paper, we develop a simple extension to the optimization where we systematically sample the set of possible initial filters and for each initialization solve for the best refinement.

Experiments demonstrate that improving the initialization step can result in the effective 'camera+filter' imaging system being much more colorimetric. Moreover, the filters we design are smoother than previously reported (which makes them easier to manufacture).

Introduction

An imaging device is colorimetric if it meets the so-called Luther condition [2], i.e. its spectral sensitivities curves are a linear transform of the spectral sensitivity curves of the human visual system (or equivalently linearly related to the CIE XYZ color matching functions [3]). Under the Luther condition, the imaging device records exactly the same triplets of the scene—after a linear mapping—sensed by a standard human observer [4]. Cameras have sensors with spectral sensitivities that do not meet the Luther conditions both because such sensors are difficult to manufacture and the sensors deployed need to consider other issues such as image noise [5].

Finlayson *et al.* [6] proposed making a camera more colorimetric by designing a spectrally precise color filter that when placed in front of the camera the new effective spectral sensitivities (of 'camera+filter') meet or approximately meet the Luther condition. The Luther condition is a strong requirement in that a camera which meets this condition will always measure color correctly for all possible spectra. But, of course not all spectra in the world are equally likely. In [1], Finlayson and Zhu developed a **Data-driven** filter design where the color filter is designed to minimize the least-squares errors between the camera responses and the XYZ tristimulus values of a collection of measured lights and surfaces.

The **Data-driven** optimization seeks to find a filter and a per illuminant based linear transform such that the filtered camera responses corrected with the corresponding linear transform are as close to the target XYZs as possible, as illustrated in Fig. 1. As we review later, the optimization is solved using an alternating least-squares algorithm (ALS). In ALS, we solve for the per illuminant correction matrices assuming a known filter and then with these matrices in hand we solve for the filter. We iteratively solve for

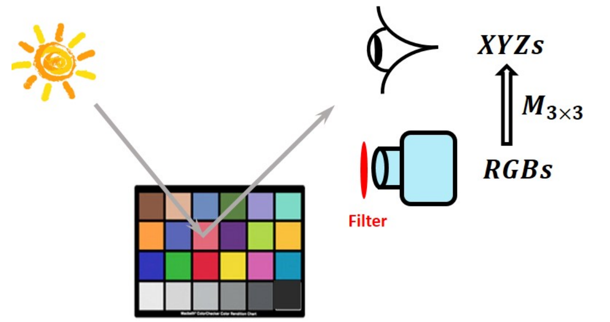


Figure 1: A camera and the eye measure triplets (RGBs and cone responses, or equivalently XYZs). In this paper we find a filter so that the camera with the filter measures RGBs that are approximately linearly related to XYZs.

the filter then the correction matrices until the process converges (convergence is guaranteed in ALS).

An important issue not considered in the original method is the choice of filter initialization. Indeed, while the ALS method converges, there is no guarantee that it converges to the global optimum and moreover it will converge to a different solution depending on the initialization condition. We have found that, empirically, the choice of initial filter has a strong impact on the extent which the solved-for filter makes a camera colorimetric.

Our solution to this problem is simple. We simply run the optimization many times for many filter initializations. Of course we can only sample so many filters and we would like to be confident that we are sampling the set of plausible filters reasonably finely. Our solution to this problem is to revisit the filter design problem assuming the filters belong to a set of smooth filters. With respect to the smooth set we find a relatively small number of -systematically sampled - samples suffices to sample the set. Additionally, we add the constraint that filters must transmit more than a minimum percentage of the incident light.

Background

Optimization Formulation

We wish to find a filter which, when placed in front of the camera - for a given collection of illuminants and reflectance spectra - makes the camera colorimetric. That is, the filtered RGBs multiplied by a per illuminant based color correction matrix are close to the corresponding CIE XYZs [1].

Image formation for Lambertian surfaces [7], which we consider here, is written as:

$$\underline{\rho} = \int_{\omega} E(\lambda)S(\lambda)\underline{Q}(\lambda)\,\mathrm{d}\lambda \tag{1}$$

where $E(\lambda)$, $S(\lambda)$ and $\underline{Q}(\lambda)$ respectively denote the spectral power distribution of the light, the surface spectral reflectance and the R, G and B spectral sensitivities of a camera ($\underline{Q}$ is a vector function). The corresponding RGB triplets, $\underline{\rho}$, result from integrating the product of these functions over the visible spectrum ω. It will be convenient to use color signals, $C(\lambda) = E(\lambda)S(\lambda)$, in the following derivations.

We can well approximate this integral by a discrete matrix-vector product:

$$\underline{\rho} = Q^T \underline{C} \tag{2}$$

where Q is an $n \times 3$ matrix with n denoting the number of sampled points over the visible spectrum ω. Each column in Q denotes a discretely sampled camera spectral sensitivity curve. The vector $\underline{C}$ is $n \times 1$ and denotes a sampled color signal spectrum. In this paper, we sample the visible spectrum from 400 to 700 nanometrs at an interval of 10 nanometers, which makes $n = 31$.

Let us now consider how to formulate the color filter design problem. First let X denote the sampled CIE 1931 standard observer functions [3]. Let the vector $\underline{f}$ represent the spectral transmittance of a color filter sampled over the visible spectrum. When placing a filter in front of a camera, the new effective spectral sensitivities of the camera system equal, at each wavelength, the filter transmittance multiplied by the camera sensitivities. This is written mathematically as $diag(\underline{f})Q$ where $diag()$ operator turns the filter vector into a diagonal matrix. Per 10 nanometer sampling, the size of matrices Q and X are 31×3 and $\underline{f}$ as a 31-dimensional vector. Let us denote a $31 \times N$ color signal matrix C (N color signals in N columns).

Returning to Figure 1, we would like to solve for a filter and a color correction matrix that minimize:

$$C^T diag(\underline{f})QM \approx C^T X. \tag{3}$$

Actually, we wish to optimize the filter for a more general condition. Let us assume we have *cnt* illuminants. We are going to solve for a single filter but a different color correction matrix for each illuminant. The color signals for a set of measured reflectance spectra under the jth illuminant is denoted as C_j and the corresponding color correction matrix is denoted as M_j. The data-driven filter design problem is formulated as the following minimization

$$\sum_{j=1}^{cnt} \min_{\underline{f}, M_j} \| C_j^T diag(\underline{f})QM_j - C_j^T X \|_F^2 \;\; \text{s.t. } \underline{f} > 0 \tag{4}$$

where the superscript ^{T} denotes the matrix transpose and $\| \, \|_F$ denotes the Frobenius norm. Note that the filter vector $\underline{f}$ is solved subject to positive values since physically a color filter must have non-negative transmittance.

In the minimization formula, the first term $C_j^T diag(\underline{f})QM_j$ denotes the camera responses with a color filter in place while the second term $C_j^T X$ denotes the corresponding XYZ tristimuli. Mathematically, the objective of the optimization is to look for the best filter and correction matrices solutions that minimize the cost function defined by the squared errors between the effective camera responses and the reference XYZ tristimuli for a set of color signals.

Algorithm 1 ALS algorithm for solving the optimization problem

filter initialization:
$k = 0, \underline{f}^0 = \underline{f}^{seed}$
calculate the best correction matrices for the initial filter:
$\min_{M_j^0} \| C_j^T diag(\underline{f}^0)QM_j^0 - C_j^T X \|_F^2, \; j = 1,2,...,cnt$
update the effective sensitivity matrix after linear transform:
$Q_j^0 = diag(\underline{f}^0)QM_j^0, \; j = 1,2,...,cnt$
for $k = 1$ to K **do**
  # refine the filter solution:
  $\min_{\underline{f}^k} \sum_{j=1}^{cnt} \| C_j^T diag(\underline{f}^k)Q_j^{k-1} - C_j^T X \|_F^2$ s.t. $\underline{f} = \prod_{s=0}^{k} \underline{f}^s > 0$
  # refine the correction matrices:
  $\min_{M_j^k} \| C_j^T diag(\underline{f}^k)Q_j^{k-1}M_j^k - C_j^T X \|_F^2, j = 1,2,...,cnt$
  # update the effective sensitivity matrix:
  $Q_j^k = diag(\underline{f}^k)Q_j^{k-1}M_j^k, \; j = 1,2,...,cnt$
end for
$\underline{f} = \prod_{k=0}^{K} \underline{f}^k$ and $M_j = \prod_{k=0}^{K} M_j^k, \; j = 1,2,...,cnt$

Alternating Least-Squares

The minimization in Eq. (4) can be solved using the Alternating Least-Squares (ALS) technique, see Algorithm 1. As we have two unknown parts in the minimization equation, the filter and correction matrices are tackled in the one-after-the-other way. The algorithm starts by initializing the filter $\underline{f}^0$, the corresponding color correction transforms, $M_j^0 (j = 1,2,...,cnt)$, for each illumination condition can be calculated. After the initialization step, the filter and correction matrices will be refined at each iteration. Solving for the best correction matrices and filter are both, individually, simple least-squares problems and are readily solved (in closed-form) [1]. In previous work, the uniform vector $\underline{f}^{seed} = \underline{1}$ (i.e. a filter that is 100% transmissive at each wavelength) was used to seed the minimization.

The reader will note that Algorithm 1 runs for a fixed number of iterations. Alternately, a 'while' loop can be used where we keep iterating until the solution at step k differs from the one in step $k-1$ by less than a criterion amount. We found the process converges quickly and the simple fixed number of iterations works well.

The final step in Algorithm 1 returns the filter solution by multiplying (element-wise multiplication) all the refined filters obtained from each iteration, $\underline{f} = \prod_{k=0}^{K} \underline{f}^k$. Similarly, the final color correction matrix for the jth illuminant is also calculated by multiplying all the refined matrices, $M_j = \prod_{k=0}^{K} M_j^k$.

While the ALS method is guaranteed to converge, the overall global minimum may not be found [8]. In Algorithm 1, empirically we found that different initializations (i.e. not initializing with the uniform vector $\underline{f}^{seed} = \underline{1}$) could result in different filters being found. Consequently, the *goodness* (the extent they supported low colorimetric error) of the discovered filters varied significantly.

Improving the Optimization

The most important contribution of this paper is to present a deterministic way to evaluate different initializations to the filter design problem. However, a key substep to achieving this is to constrain the shape of the filter. Indeed, as we are representing filters as discrete vectors (in a 31-dimensional space), we cannot plausibly sample all vectors. Rather we will constrain our filters to be sufficiently smooth that we can sensibly sample the filter space. A great advantage of this approach is that smooth filters should be easier to manufacture. Another concern is that the filters should transmit enough light. So, the optimization will also be modified to incorporate a lower-bound on the filter transmittance.

Filter Constraints

Let us constrain $\underline{f}$ as a linear combination of an m-dimensional *basis* set of filters - denoted by a $31 \times m$ matrix B - and to have the minimum and maximum transmittance thresholds:

$$\underline{f} = B\underline{c} \qquad \text{s.t. } f_{min} \le \underline{f} \le f_{max} \tag{5}$$

where f_{max} is set to 1 as fully transmissive and f_{min} is a positive value between 0 and 1. In this paper, B denotes the first m basis of the discrete Cosine series (specifically we adopt the 2nd variant of the discrete Cosine basis vectors [9]).

Now we rewrite the filter design optimization as:

$$\sum_{j=1}^{cnt} \min_{\underline{c}, M_j} \| C_j^T diag(B\underline{c}) Q M_j - C_j^T X \|_F^2 \ \text{ s.t. } f_{min} \le B\underline{c} \le f_{max} \tag{6}$$

In the current minimization, we are looking for the basis coefficient vector $\underline{c}$ to form a bounded smooth filter that returns the least error.

This new minimization formulation can also be solved by the same paradigm shown in Algorithm 1 except that in each iteration, we refine the coefficient vector to satisfy the basis and threshold constraints (and to do this Quadratic programming is used for our minimizations [10]).

Initialization Set by Sampling

We have found that in Algorithm 1, whether finding an unconstrained or a constrained filter, the solved filter depends strongly on the initialization condition. Given the filters are represented as a linear combination of the first m terms in the Cosine series, we would like to sample this space and then find the best filter using each sample as the initialization (e.g. for $\underline{f}^{seed}$ in Algorithm 1). Algorithm 2 sets forth an algorithm for finding $\#filters$ (number of initial filters) by uniformly and randomly sampling the filter space subject to smoothness and minimum transmittance constraints.

The algorithm has a preprocessing step where we find two bounding vectors $\underline{c}^{min}$ and $\underline{c}^{max}$ for the coefficient vectors. Over all filters that can be written as $B\underline{c}$, for the ith component in vector $\underline{c}$, the minimum and maximum values the coefficient can take are denoted respectively c_i^{min} and c_i^{max},

$$c_i^{min} \le c_i \le c_i^{max}, i = 1, 2, ..., m. \tag{7}$$

Given we have an m-dimensional coefficient vector, $\underline{c}^{min}$ and $\underline{c}^{max}$ effectively delimit a hypercube in m-dimensional space. Not all coefficient vectors in this hypercube satisfy the upper and lower thresholds on transmittance but all filters that are parameterized by coefficients outside of the hypercube do not satisfy the transmittance bounds.

Algorithm 2 Algorithm for generating an initial filter subset

$\mathscr{F} = \{\}$
for $i = 1$ to m **do**
  $c_i^{min} = \arg\min c_i$ s.t. $f_{min} \le B\underline{c} \le f_{max}$
  $c_i^{max} = \arg\max c_i$ s.t. $f_{min} \le B\underline{c} \le f_{max}$
end for
while $cardinality(\mathscr{F}) < \#filters$ **do**
  $c_i \sim U\left(c_i^{min}, c_i^{max}\right), i = 1, 2, ..., m$
  $\underline{f} = B\underline{c}$
  if $f_{min} \le \underline{f} \le f_{max}$ & $\{\forall \underline{q} \in \mathscr{F} : angle(\underline{f}, \underline{q}) > \theta\}$ **then**
    $\mathscr{F} \leftarrow \mathscr{F} \cup \{\underline{f}\}$
  end if
end while

Within the hypercube we uniformly and randomly choose coefficient vectors. For each selected vector we check if the constructed filter satisfies the transmittance threshold constraints. If it does, and if it is sufficiently distinguishable from those already in the set, we add it to the initial filters set, $\mathscr{F}$, that we wish to use as initializations for the optimization. We keep adding to the set $\mathscr{F}$ until the number of members reaches $\#filters$ (formally $cardinality(\mathscr{F}) = \#filters$). In the pseudocode of Algorithm 2 we denote the action of choosing every component of the coefficient vector $\underline{c}$ randomly and uniformly from the coefficient hypercube as $c_i \sim U\left(c_i^{min}, c_i^{max}\right), i = 1, 2, ..., m$. We say that a filter is sufficiently far from those already in the set if the angle between these two filter vectors is larger than θ degrees.

Results

For our camera, we use a Canon 5D Mark II DLSR digital camera with known spectral sensitivity functions [11]. The Data-driven optimization is carried out on a collection of measured spectral data of 102 illuminants and 1995 surface reflectances [12].

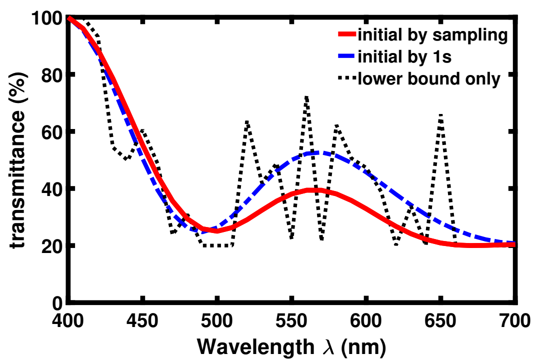
The color reproduction results of our experiments are evaluated using CIELAB color difference metric ΔE_{ab}^* [3]. It is a perceptual color difference metric computing a single number of Euclidean distance between two colors in the CIELAB color space. One ΔE_{ab}^* unit corresponds approximately to the 'Just Noticeable Difference' to a standard human observer.

We are going to solve for the best filter for the camera using 6 and 8 Cosine basis functions. In both cases we seek a filter that has a lower bound transmittance of 20% and we generate a set of $\#filters = 20,000$ initial filters that drive our optimization. Every filter in the initialization set is at least 1 degree ($\theta = 1^\circ$) from its closest neighboring filters in the set. For each basis condition (6 or 8), we run our optimization 20,000 times. Then, we simply choose the filter that delivers the least mean DeltaE error overall.

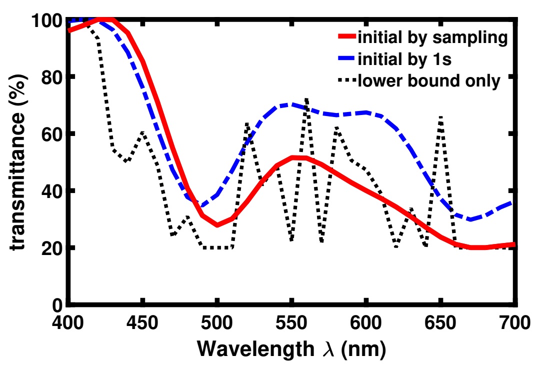
Table 1 shows the color reproduction results. '**NAT**' gives the baseline color correction performance when no color filter is used. Here the least-squares optimal 3×3 matrix is used, per illuminant, to map the Canon recorded RGBs (for the 1995 reflectances) to the corresponding XYZs. When we initialize the filter as a 100% transmitting filter (which we did in [1]), we label the results as '**DATA_1s**'. Note for this set we present both results from filters constrained and unconstrained by the basis functions. Finally, '**DATA_sampling**' denotes the results found by the cur-

Table 1: Comparison of Color Reproduction Results

	Mean	median	90%	95%	99%	max
NAT	1.72	1.03	3.68	5.12	12.94	28.39
a minimum transmittance of 20%						
DATA_1s	0.69	0.42	1.47	2.11	4.69	19.48
6 Cosine basis with a minimum transmittance of 20%						
DATA_1s	0.81	0.49	1.80	2.54	5.21	18.85
DATA_sampling	**0.59**	**0.35**	**1.30**	**1.83**	**3.77**	**14.19**
8 Cosine basis with a minimum transmittance of 20%						
DATA_1s	0.71	0.38	1.60	2.38	5.42	19.25
DATA_sampling	**0.45**	**0.25**	**1.02**	**1.41**	**3.10**	**10.63**



(a) 6-basis series



(b) 8-basis series

Figure 2: (a) Solid line (in red) shows the filter solved for by our *sampled initialization* method constraining the filter to be well described by a linear combination of the first 6 terms in a Cosine series. Dashed curve (in blue) shows the filter that is found without sampling (previous method) and the dotted line (in black) is for no sampling and no constraint on filter smoothness. (b) repeats the experiment for an 8-dimensional Cosine basis.

rent method, using the best filter where we search over a large set of possible initializations.

In Table 1, for these 6 different methods, we show the color reproduction in terms of mean, median, 90%-quantile, 95%-quantile, 95%-quantile, and maximum ΔE^*_{ab} errors. The spectral transmittance distribution of the best filters of 6 basis and 8 basis are shown in Figs. 2a and 2b respectively (see solid red lines denoted 'initial by sampling'). We also plot the spectral transmittance of the filters found where the 100% transmitting filter is used as the initialization condition (see dashed blue lines denoted by 'initial by 1s'). For reference, the filter using the uniform vector as initialization but constrained only by the minimum transmittance is also given (see black dotted lines in Figs. 2a and 2b denoted 'lower bound only'). It is evident that the discovered filters are quite different under different initialization conditions.

From the table, we can see that our new method (**DATA_sampling**) provides significantly better color reproduction accuracy: it reduces by about two-thirds of the mean, median, 90-quantile, 95-quantile, 99-qunatile error metrics and over half of the maximum color error, comparing to those by **NAT**ive color correction. Surprisingly, the new method also outperforms the results of non-smoothed filter (only bounded by the minimum transmittance) using a single fixed uniform initialization. When we compare **DATA_1s** to **DATA_sampling** under the smoothness constraints (constructed by basis functions), the latter method delivers errors about a 1/3 smaller. That is, the best initialization leads to (across all error measures) over 30% reduction in error. Finally, we see there is a modest improvement in the error statistics when 8 as oppose to 6 basis functions are used.

Of course the need to sample adds complexity to the filter design problem. So, let us investigate the number of sampled smooth filters we need for finding a good approximation to the optimal solution, e.g. within acceptable deviation (for the data at hand). We test on the 6-Cosine basis condition and evaluate the average and the standard deviation of the mean color errors by varying $\#filters$ in the initialization set from 100, 200, 500, 1000, 2000, 5000, 10,000 to 20,000 (where these filters are selected using Algorithm 2).

In Figure 3 the x-axis represents the number of filters in our initialization set and the y-axis represents the average mean color error in terms of ΔE^*_{ab} (for the corresponding optimized filter which is refined for that set of initializations). The dashed line denotes the best results obtained from the overall 20,000 trials. One standard deviation error bar is also shown.

From the figure, we can see that a set of 1000 filter initializations suffices for our method: we obtain almost the same mean color error performance (compared to larger initializations) and the error bar is small. While we do not show the plot here, 1000 filters are also sufficed for filters that are described with an 8-dimensional Cosine basis.

Conclusion

Previous work has shown that a specially designed - via numerical optimization - transmittance filter can - when the filter is placed in front of a camera - make a camera significantly more colorimetric [1]. In this paper we extended that method. We showed that the performance of the optimized filter could be improved if initialization conditions of the optimization were considered. Specifically, we set forth a method to enumerate - to a criterion accuracy - the set of possible initialization filters when a 6- and 8-dimensional Cosine basis are used (to construct the filters). Our new method - across a variety of error metrics - reduces the recorded errors by at least a further 30%. Compared to using a simple 3×3 color correction matrix (per illuminant), our new method designs a filter which to be placed in front of the camera and using a 3×3 matrix can result in just 1/3 of the original error.

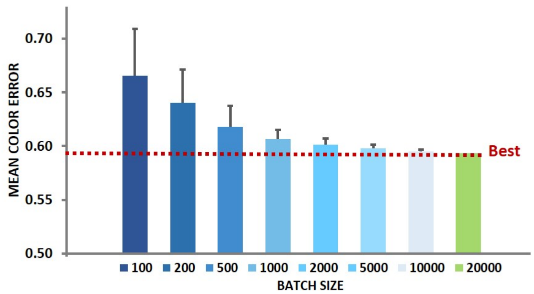


Figure 3: The effect of $\#filters$ in the initialization set on the color reproduction results in terms of mean color error assuming filters fall within the 6-dimensional Cosine basis. The magnitude of the bars represents the average color error from groups of varying $\#filters$. The upper 1 standard deviation error bar is also shown. The red dash line denotes the best result obtained from the overall 20,000 initializations.

References

[1] G. D. Finlayson and Y. Zhu, "Finding a Colour Filter to Make a Camera Colorimetric by Optimisation," In *International Workshop on Computational Color Imaging, Springer*, pp. 53-62, 2019.

[2] R. Luther, "Aus dem Gebiet der Farbreizmetrik," *Zeitschrift fur Technische Physik*, vol. 8, pp. 540-558, 1927.

[3] G. Wyszecki and W. S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Ed., Wiley: New York, 2000.

[4] B. K. Horn, "Exact reproduction of colored images," *Computer Vision, Graphics, and Image Processing*, vol. 26, no. 2, pp. 135-167, 1984.

[5] J. Nakamura, Image sensors and signal processing for digital still cameras, CRC press, 2016.

[6] G. D. Finlayson, Y. Zhu, and H. Gong, "Using a Simple Colour Pre-filter to Make Cameras More Colorimetric," In *26th Color Imaging Conference*, no. 1, pp. 182-186, 2018.

[7] N. Ohta and A. Robertson, Colorimetry: fundamentals and applications, John Wiley & Sons, 2006.

[8] T. Zhang and G. H. Golub, "Rank-one approximation to high order tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 23, no. 2, pp. 534-550, 2001.

[9] G. Strang, "The discrete cosine transform," *SIAM review*, vol. 41, no. 1, pp. 135-147, 1999.

[10] D. G. Luenberger, and Y. Ye, Linear and nonlinear programming, 3rd Ed., Springer-Verlag, 2008.

[11] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk, "What is the space of spectral sensitivity functions for digital color cameras?" *Applications of Computer Vision, IEEE Workshop*, pp. 168-179, 2013.

[12] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Research & Application*, vol. 27, no. 3, pp. 147-151, 2002.

Author Biography

Yuteng Zhu is currently a PhD candidate in the Colour and Imaging Laboratory, University of East Anglia. She works on filter design for color measurement, particularly color filters for digital cameras to improve its color reproduction accuracy.

Graham Finlayson is a computer scientist whose work in the fields of physics-based computer vision and color imaging has enabled computers to mimic important aspects of our visual system. His research has contributed to the theoretical underpinnings and the practical engineering that drive the processing pipelines found in embedded camera systems, including smart phones. Many of his research contributions including in automatic white balance, dynamic range compression, color correction and tone mapping have been deployed commercially.