Robust Averaged Projections Onto Convex Sets

Ali Alsam and Casper Find Andersen
The National Gallery In London, London, United Kingdom
The Graphic Arts Institute of Denmark, Copenhagen, Denmark

Abstract

We [1] present a new robust method for recovering the spectral sensitivity of digital cameras and scanners. It is well known that the recovery of camera spectral sensitivities is an ill-posed problem. To stabilize the solution to the problem constraints are often imposed on the solution space. Among common constraints are: non-negativity, degree of smoothness, number of peaks, noise level bounding and that estimated curves result in the lowest possible error between predicted and measured data. These constraints are not always physically justified; and imposing them on the solution space can result in poor estimates that adhere only to our expectations of sensor curves.

Knowing that all previous methods result in perfect sensor prediction when the data is noise-free, we introduce a robust algorithm that enables the user to heavily dampen the impact of noise and outliers on the solution. By controlling the effect of noise we show that the only additional constraint needed is the physically feasible non-negativity. Despite being iterative the method is computationally fast and simple to implement.

To evaluate the new method, we used data from real trichromatic camera systems as well as simulated data. The results support our assertions that controlling the noise results in better sensor estimates.

Introduction

Camera sensor calibration is the problem of estimating the device's spectral sensitivities from its responses to a number of spectrally different surfaces. Generally, there are two approaches to solving the spectral calibration problem: one based on physical measurements and one based on a theoretical model. The physical approach, using a monochromator, gives an accurate estimate of the spectral sensitivities, but it is expensive and time consuming to use. The model-based approach is cheaper and provides insight into the characteristics of the camera system. It is based on solving a linear equation system of the form:

$$\mathbf{B}_\Upsilon = \mathbf{B} - \Upsilon = \mathbf{AX} \tag{1}$$

Here $\mathbf{A}$ is the $m \times n$ matrix of measured color signals, $\mathbf{X}$ is a $n \times l$ dimensional matrix whose elements are the spectral sensitivities, $\mathbf{B}$ is of size $m \times l$ and contains the measured camera responses and Υ is an acquisition noise matrix of size $m \times l$. The color signals are the component-wise products of the illumination spectrum and the reflectance spectra; l is the number of sensor sensitivity functions of the camera, m is the number of surfaces used and n the dimension of the spectral data. Typical values are; $n = 31$, corresponding to a $10nm$ sampling of the wavelength range $400nm$ to $700nm$ and $l = 3$ for an RGB-camera.

The goodness of the solution for the spectral sensitivities based on Equation (1) depends on two main factors: the noise level in the response data Υ and the statistical properties of the spectral data available from the calibration chart (the matrix $\mathbf{A}$).

[1] The authors' list is alphabetical.

Estimation of $\mathbf{X}$ from the RGB measurements $\mathbf{B}_\Upsilon$ and $\mathbf{A}$ is a typical inverse problem and standard methods from linear algebra are often used to solve it.

The findings in Reference [1] indicate that the uncertainty surrounding spectral recovery is related to the size of the recovered set and governed by factors such as: the noise level, the dimensionality of the spectral data, and the constraints imposed on the solution space. In Reference [2] the authors constrained the sensors to be positive, smooth, and to predict the responses within an acceptable noise bound. In Reference [3] the authors added a constraint on the number of peaks allowed in the recovered sensor, while the authors in Reference [4, 5, 6] constrained the sensor's magnitude to be small. All these methods in Reference [4, 5, 6, 3, 2] require that the recovered sensor minimizes the difference between the measured and estimated responses.

Although, these constraints improve upon the stability of the solution, their influence on the recovered sensor is determined by their physical justification. As an example: measured data clearly shows that a number of sensors exhibit more than one peak. Further, constraining the sensors to be smooth means that no sharp transitions are allowed in the solution: This is not always the case. Finally, imposing a constraint on the noise level requires a priori knowledge of the noise statistics and can result in false predictions in the case when a number of measured points are outliers. Thus we state that: Ideally no constraints that are not guaranteed to be true, should be imposed on the solution space. Thus in the case of sensor recovery, the only constraints that are always physically satisfied are non-negativity and that the sensor is orthogonal to the metamers plane, Reference [7].

In this paper, we present a new iterative method to solve a general linear system of equations with a view to solving the problem of sensor estimation from real camera data. The method has a number of advantageous characteristics including: It is mathematically rigorous and transparent, i.e. it is guaranteed to yield the true solution in the ideal noise-free case. It is also robust to noise by virtue of a formulation that is based on averaging the solutions obtained from the calibration data, i.e. if the noise is a random variable, then the solution will be very similar to that obtained in the noise-free case. The robustness to noise can be further enhanced by controlling the number of iterations. It is also possible to impose constraints on the solution space, however, in our implementation the only constraint used is non-negativity. Finally, the method is fast, simple and lends itself to hardware implementation.

The basic idea in the proposed method is that geometrically the solution vector, the sensor, is the intersection point of all the m planes that are orthogonal to the spectral functions (calibration surfaces). Thus at each iteration step, we find m sensors that satisfy the individual m equalities $ax = b$, i.e. for each surface a solution is obtained. The result at a given iteration step is the average of the m sensors.

The proposed method can be thought of as a special implementation of Projection onto Convex Sets **POCS**. In POCS,

however, a guess solution, at a given iteration level, is the consecutive projection onto each of the different sets. The solution is found when a convergence criterion is met. In our formulation, the guess solution, at a given iteration level, is projected onto each set independently and the result is the robust average value. It is this averaging that makes our formulation more robust to noise and ensures the uniqueness of the solution.

Method Description

We start our presentation of the proposed algorithm by a number of definitions. First we define a hyperplane P_i as:

$$P_i: \quad a_i x = b_i \tag{2}$$

where a_i is i*th* row vector of the matrix A, i.e. the measured spectrum of the i*th* calibration surface and b_i is the corresponding camera response at a single channel, red, green or blue. Mathematically, the solution to the linear system of the form $Ax = b$ is the intersection point of all the m hyperplanes P_i.

Secondly, we define a distance from an arbitrary guess value x to a hyperplane P_i as:

$$D(x, P_i) = \frac{b_i}{\| a_i \|} - n_i x \tag{3}$$

where $\| a_i \|$ is the second norm of a_i and n_i is defined as:

$$n_i = \frac{a_i}{\| a_i \|} \tag{4}$$

Thirdly, given a solution vector x^k at the k*th* iteration level we define, for a single spectrum a_i, a solution vector x_i^{k+1} as:

$$x_i^{k+1} = D\left(x^k, P_i\right) n_i + x^k \tag{5}$$

Knowing that x_i^{k+1} is the solution for each $a_i x_i^{k+1} = b_i$ we define the solution x at iteration level $k+1$ as:

$$x^{k+1} = \frac{1}{m} \sum_{i=1}^{m} x_i^{k+1} \tag{6}$$

The non-negativity constraint is implemented as a simple clipping of the negative values at each wavelength λ, i.e.:

$$x(\lambda)^{k+1} = \left\{ \begin{array}{ll} x(\lambda)^{k+1} & if \;\; x(\lambda)^{k+1} \geq 0 \\ 0 & if \;\; x(\lambda)^{k+1} < 0 \end{array} \right\} \tag{7}$$

To clarify the description of the proposed algorithm we outline the first three iterations $k = 0 \cdot \cdot 2$. To aid the discussion we make use of the 2-dimensional example depicted in Figure (1) where two vectors $n_1 = \frac{a_1}{\|a_1\|}$ and $n_2 = \frac{a_2}{\|a_2\|}$ are considered together with their corresponding orthogonal lines P_1 and P_2 (P_1 and P_2 are hyperplanes in the n-dimensional case). We start by setting $x^0 = \mathbf{0}$ where $\mathbf{0}$ is a vector whose elements are all zeros. Thus formally at iteration level $k = 0$ we have:

$$x_i^0 = \mathbf{0} \qquad ; \; x^0 = \frac{1}{m} \sum_{i=1}^{m} x_i^0 = \mathbf{0} \tag{8}$$

where m is 2 in the example of Figure (1).

At iteration level $k = 1$, we then have:

$$x_i^1 = \frac{b_i}{\| a_i \|} \qquad ; \; x^1 = \frac{1}{m} \sum_{i=1}^{m} x_i^1 \tag{9}$$

and finally, at iteration level $k = 2$ we have:

$$x_i^2 = D\left(x^1, P_i\right) n_i + x^1 \qquad ; \; x^2 = \frac{1}{m} \sum_{i=1}^{m} x_i^2 \tag{10}$$

For $k > 2$ Equations (5) and (6) are consecutively employed to obtain the subsequent iteration level. Notice that the non-negativity constraint stated in Equation (7) is employed at each step.

As the algorithm converges towards a solution a stop criterion must end the number of iterations. A stop criterion can be constructed in a number of ways. Firstly, the maximum number of iterations k_{max} can be decided upon by the user directly. Secondly, it is possible to define a convergence criterion such as:

$$k_{max}: \quad \| x^{k_{max}} - x^{k_{max-1}} \| \leq \varepsilon \tag{11}$$

where ε is small positive value. Alternatively, it is possible to check the goodness of the fit between the estimated responses $b^{est} = Ax^k$ and those recorded by the camera b.

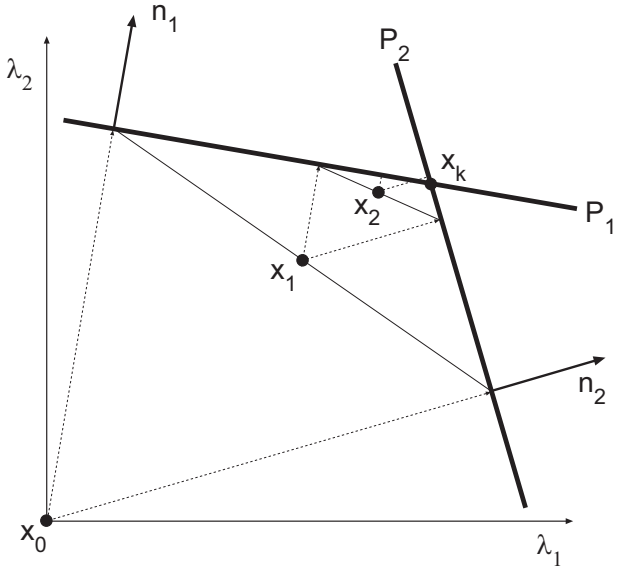

Figure 1. *In a two dimensional space (λ_1, λ_2), two intersecting planes P_1 and P_2 are depicted. The planes have unit length normal vectors n_1 and n_2 respectively. x_k is the guessed value of the intersection point after k iterations. The iteration starts at x_0 in origo and converges towards x_k through x_1, x_2,...,x_k. x_0 is projected to the planes P_1 and P_2 perpendicularly. The resulting projection points on the planes are averaged and equals the guess value of the intersection at the first iteration level. x_1 is projected to P_1 and P_2 and the resulting projection point are averaged and so forth. The iterations stop when the stop criteria is met.*

Robust Averaging

As shown in Equation (6), the resultant sensor estimate at each iteration level is the average value of all the estimates obtained based on the individual surfaces. This aspect of the algorithm represents the fundamental difference between our approach and that in projection onto convex sets where the estimate is projected from one place P_i onto another. The importance of averaging is crucial in a number of situations and these can be classed as follows: When the planes do not intersect in which case POCS will result in a point on one plane, while our method will result in a sensor, that is the average between the two guesses. When the intersection of the planes forms a region, a solution space, in which case POCS will result in a point on one of the planes or the intersection point between two planes, while our method will result in a point inside the solution space. Finally, and most importantly, when there are outliers in the data, i.e. some of the planes have been shifted by a large distance

from their original position; then the worst case scenario for POCS would be to terminate when the estimate is on an outliers plane. Unfortunately, though averaging would reduce the effect, it would not resolve this problem. Thus we propose to make use of a robust averaging technique to identify and dampen the influence of the outliers. The approach followed in this paper, to estimate a robust mean, is based on clustering the data x_i at iteration level k into r random clusters with each including $s\%$ of the calibration points. The final average is taken as the median of all the average values. Formally, we define a set Q such that:

$$Q[x_i^{k+1}] = \{x_i \in R^n \wedge Ax_i^{k+1} = b_i\} \tag{12}$$

Further, we define the clusters as subsets S of Q where:

$$S[x_h^{k+1}] \subseteq Q \tag{13}$$

where h is a random element between $i = 1$ and $i = m$. As and example we can think that the sum of h is one fifth of m.

The robust averaging is then calculated as the median of the values of the mean of all the subsets S; and the more sets we have the more robust the average is.

Results

To test the performance of the proposed algorithm robust averaged projections onto convex sets **RAPOCS** and compare it with standard methods we spectrally calibrated a MegaVision camera and the Nikon D70 digital SLR camera. For the *Nikon D70* the actual sensitivities were not available while for the MegaVision the sensor curves were measured using a monochromator.

In the first experiment, MegaVision, the spectral data was that of the *Macbeth Color Checker Chart* measured under a daylight simulator. Further, because the actual sensitivities were available, it was possible to compare the sensor estimate with the measured sensitivities in the spectral space. The similarity between the estimated sensors and the measured curves was based on the Vora value Reference [8], which is the ratio between the estimated sensor and its projection element in the space of the measured curves.

In the second experiment, Nikon D70, we used the Esser chart with 282 colored patches. The spectral data was measured using a Minolta CS-1000 spectroradiometer under the daylight simulator of the Macbeth Verda viewing booth. The camera responses were captured in the Nikon raw image format; and the response data was checked for linearity and the dark noise was subtracted.

For numerical data comparison we used the absolute error between the estimated and measured responses. This is defined as:

$$AE = |a_i\widetilde{x} - b_i| \tag{14}$$

where $\widetilde{x}$ is the estimated sensor. To allow meaningful comparison in terms of the absolute error metric; the data in b and $\mathbf{A}\widetilde{x}$, for all channels, were normalized such that the maximum value was set to 100. Thus a difference of one is equivalent to 1% error.

MegaVision

In this experiment, we used three estimation methods: The truncated singular value decomposition **TSVD**, projection onto convex sets POCS, and the proposed robust averaged projections onto convex sets RAPOCS. To ensure robustness, we divided the solution set Q into $r = 100$ subsets S where each subset contained a random selection of surfaces corresponding to one third $s = 0.33$ of the available 24 calibration patches. The estimated sensor set is shown using dashed lines in Figure (2) where we have superimposed the measured sensors in solid lines. We note that although not perfect, the estimate is visually close to the measured. Further, for all the methods, we calculated the channel

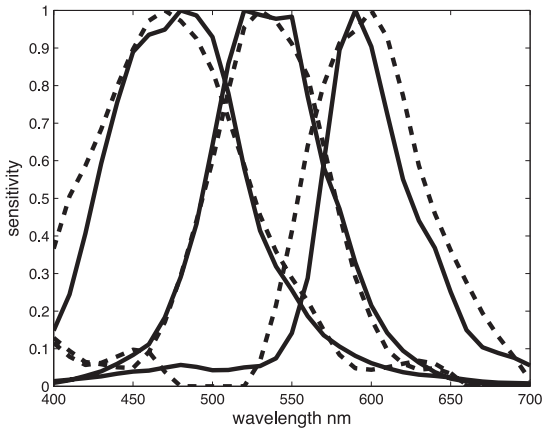

Figure 2. *The sensors set estimated with the proposed method are shown in dashed lines. The measured sensors set is shown as the solid lines. The data was sampled at 10nm intervals between 400 and 700nm and the sensors were normalized such that the maximum value of each is unity.*

wise absolute error. These results are tabulated in Table (1), in which we observe that the RAPOCS performs better than TSVD and POCS on the red and green sensor and slightly worse than TSVD on the blue sensor. We notice, however, that RAPOCS performance is clearly superior to POCS. Further, by definition the RAPOCS is derived to avoid outliers and dampen the effect of noise in the data. This characteristics means that we should expect the rgb fit of the calibration data to be less than that of the TSVD which is derived to minimize the square difference between the estimate and the measured data. To judge the goodness of the sensor estimates we made use of the Vora value. For all the methods the values are reported in Table (2). Here we clearly, notice that the proposed algorithm performs favorably to the other methods.

Nikon D70

As previously stated, there were no measured sensors available in this experiment. Thus we are reliant on the level of rgb data fit in our judgement of the recovery goodness. This data is reported in Table (3). The recovered sensors are plotted in Figure (3). From Table (3), we notice that the performance of RAPOCS is equivalent to that of the TSVD, however, when we compare the recovered sensor from Figure (3) we notice that the sensors estimated by TSVD include negative values as well as a degree of ringing where the sensitivity of the sensor is not high, i.e. the red part of the blue sensor. Thus we can conclude that the new method results in a data fit comparable to TSVD while stratifying positivity, avoiding the ringing effects and not constraining the estimated sensors to lie in a reduced dimensional space.

Noise robustness

Finally, we conducted an experiment based on synthetic data where the rgb responses were calculated based on the MegaVision sensors and the spectral data of the Esser calibration chart. The responses were then normalized such that the maximum value at each channel was set to 100 and random noise with bi-modal distribution was added. The noise level was increased from 1-10% from the maximum value. For each noise level we

Table 1: Mega Vision Camera error estimation.

method	TSVD		
Abs Error	red	green	blue
mean	0.580	0.574	0.464
median	0.439	0.499	0.424
max	1.995	1.936	1.347
method	POCS		
Abs Error	red	green	blue
mean	1.576	2.431	2.244
median	1.296	1.930	1.737
max	4.992	5.834	5.572
method	RAPOCS		
Abs Error	red	green	blue
mean	0.530	0.479	0.671
median	0.510	0.324	0.505
max	1.271	2.018	3.867

The absolute error between the measured and estimated responses for the red, green and blue channels of the *Mega Vision Camera*. The results are based on the truncated singular value decomposition TSVD, projection onto convex sets POCS and the proposed algorithm RAPOCS. The calibration data was that of the Macbeth Color Checker.

Table 2: Vora value for TSVD, POCS and RAPOCS.

method	TSVD	POCS	RAPOCS
Vora Value	0.947	0.920	0.970

The Vora value which is a measure of similarity between the actual sensors and the estimates. The results are based on the methods TSVD, POCS and RAPOCS.

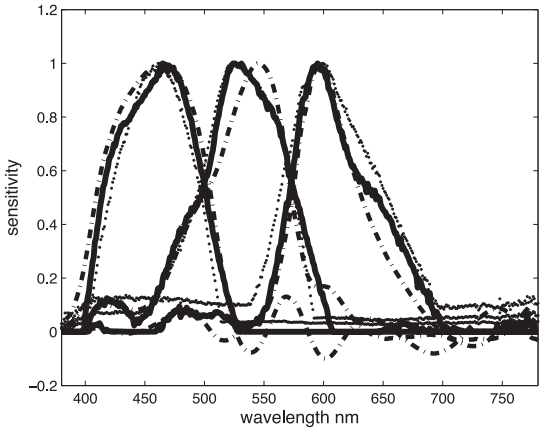

Figure 3. *The sensors set estimated with the proposed RAPOCS method are shown using lines. The TSVD sensors' set is shown using dash dotted lines and that estimated with POCS as a dotted line. The data was sampled at 2nm intervals between 380 and 780nm and the sensors were normalized such that the maximum value of each is unity.*

recovered the sensor set using RAPOCS and the TSVD. The sensors results of the recovery are shown in Figure (4) and Figure (5) for the propose and TSVD respectively. Though it is not possible, from the figures, to clearly indicate which set corresponds to which noise level; the figures serve to show that RAPOCS results in much more stable estimates even when the noise levels are unreasonably high. To give a numerical comparison we calculate the Vora value for each set at the different noise levels. These values are reported in Table (4) where we notice that the new methods outperforms TSVD especially when the noise level is high.

Table 3: Nikon D70 error estimation.

method	TSVD		
Abs Error	red	green	blue
mean	1.287	1.006	0.950
median	0.863	0.643	0.704
max	8.630	7.962	6.653
method	POCS		
Abs Error	red	green	blue
mean	3.240	4.940	3.784
median	2.933	4.425	2.683
max	12.098	20.243	12.637
method	RAPOCS		
Abs Error	red	green	blue
mean	1.288	0.996	1.039
median	0.786	0.618	0.765
max	8.252	7.856	6.770

The absolute error between the measured and estimated responses for the red, green and blue channels of the Nikon D70 Camera. The results are based on the truncated singular value decomposition TSVD, projection onto convex sets and the proposed algorithm. The calibration data was that of the Macbeth Color Checker.

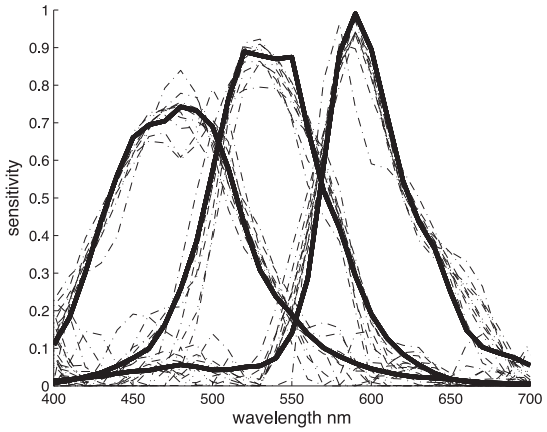

Figure 4. *The sensors set estimated with the proposed RAPOCS method. Sensors are recovered from Esser chart data with 10 percentage levels of simulated noise in steps of 1 percent. The solid curves are the measured MegaVision sensors and the 10 recovered sensors in dotted curves.*

Discussion and Conclusions

In this paper we have presented a new robust method of spectral sensor recovery based on averaged projection onto convex sets. The only constraint that was imposed on the solution space is non-negativity, making the algorithm independent of a priori knowledge of the sensor curves. When the data is noise-free the new method, like previous methods, results in perfect sensor solutions. Real data with unknown noise distribution is the real challenge in sensor recovery. Thus, we have introduced a robust algorithm that enables the user to heavily dampen the

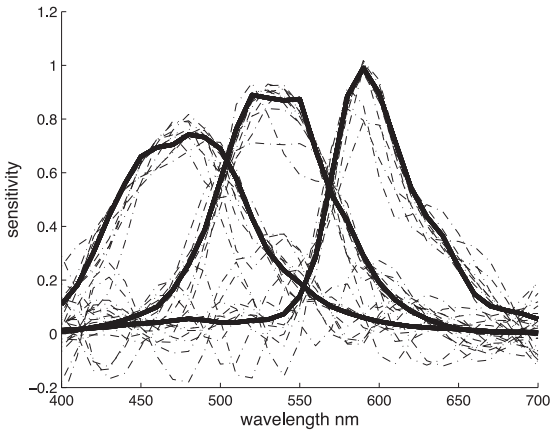

Figure 5. *The sensors set estimated with the TSVD method. Sensors are recovered from Esser chart data with 10 percentage levels of simulated noise in steps of 1 percent. The solid curves are the measured MegaVision sensors and the 10 recovered sensors in dotted curves.*

Table 4: Vora values from simulated noise on Esser data.

% Noise	RAPOCS	TSVD
1	0.9941	0.9962
2	0.9931	0.9930
3	0.9903	0.9766
4	0.9844	0.9816
5	0.9764	0.9725
6	0.9755	0.9614
7	0.9651	0.9451
8	0.9591	0.9031
9	0.9626	0.8561
10	0.9423	0.8882

The Vora values for recovered red, green and blue sensors from simulated Esser chart data responses with 1, 2 ... 10 percent noise added. The proposed method and the TSVD is compared. The Vora value is increasing with decreasing noise for both methods, while the proposed method systematically results in a higher value, indicating a better set of recovered sensors. Furthermore the proposed method appears to yield relatively better estimates with increasing noise.

impact of noise and outliers on the solution. Controlling and dampening the effect of noise is obtained through robust averaging and employment of non-negativity. The experiments with noise simulated data show that the method yields stable results with high similarity, in terms of Vora value, between the measured sensor and the estimated, even when the noise level is unusually high. Further, the algorithm avoids the ringing effects and does not constrain the estimated sensors to lie in a reduced dimensional space. The method is computationally fast and simple to implement.

References

[1] A. Alsam and G. D. Finlayson, Recovering spectral sensitivities with uncertainty, In The First European Conference on Color in Graphics, Imaging and Vision, pg. 22-26, 2002.

[2] G. Sharma and H. Trussell, Characterization of scanner sensitivity, IS&T/SID Color Imaging Conference, pg. 103-107, 1993.

[3] G. Finlayson, S. Hordley, and P. Hubel, Recovering device sensitivities with quadratic programming, IS&T/SID Sixth Color Imaging Conference: Color Science, Systems and Applications, 1998.

[4] Bob Dyas, Robust sensor response characterization, IS& T/SID Eighth Color Imagining Conference Arizona, USA, pg. 144-148, 2000.

[5] Kobus Barnard and Brian Funt, Camera characterization for color research, Color Research and Application, Vol. 27, No. 3, pg. 153-164, 2002.

[6] Kobus Barnard and Brian Funt, Camera calibration for colour vision research, SPIE Conference on Electronic Imaging, Human Vision and Electronic Imaging IV, SPIE Vol. 3644, pg. 576-585, 1999.

[7] A. Alsam and R. Lenz, Calibrating color cameras using metameric blacks, Journal of the Optical Society of America A, 24 Issue 1, pg. 11-17, January 2007.

[8] P. L. Vora and H. J. Trussell, Measure of goodness of a set of color scanning filters, Journal of the Optical Society of America A, 10(7), pg. 1499-1508, 1993.

Author Biography

Ali Alsam *is a HP research fellow at the National Gallery in London. Until 2006 he was an Associate Professor at the Norwegian Colour Laboratory in Gjøvik University College, Norway. He received a PhD degree in computational colour science from the University of East Anglia. His research interest include: colour science, computational colour, image processing, metamerism, vision, inverse problems, optimisation and convex analysis.*

Casper Find Andersen *M.Sc. 1993 from the Danish Technical University DTU. Specialized in Computational Fluid Dynamics and Descriptive Geometry. Employed by DTU as research assistant specializing in color theory and color management until 1998. From 1998 to 2001 senior researcher at R&D department Phase One dealing with color management and image manipulation. From 2001 working as teacher, consultant and researcher at the Graphic Arts Institute of Denmark. Currently, working on a Phd-project on "characterization of digital color cameras" with Prof. Graham D. Finlayson at the University of East Anglia.*