Multispectral imaging: How many sensors do we need?

David Connah, Ali Alsam, Jon Y. Hardeberg
Norwegian Color Research Laboratory, Gjøvik University college, Norway

Abstract

The surface reflectance functions of natural and man made surfaces are invariably smooth. It is desirable to exploit this smoothness in a multispectral imaging system by using as few sensors as possible to capture and reconstruct the data. In this paper we investigate the minimum number of sensors to use, whilst also minimising reconstruction error. We do this by deriving different numbers of optimised sensors, constructed by transforming the characteristic vectors of the data, and simulating reflectance recovery with these sensors in the presence of noise. We find an upper limit to the number of optimised sensors one should use, above which the noise prevents decreases in error. For a set of Munsell reflectances, captured under educated levels of noise, we find that this limit occurs at approximately 9 sensors.

Introduction

The information contained in a black and white image is insufficient to reproduce the scene's spectral information. For example, it is not possible to know if a shirt which appears grey in the image is red, green, blue or yellow. This means that surfaces with different reflectance properties are likely to integrate to the same grey shade. This phenomenon, whereby spectrally different surfaces integrate to the same camera response, is known as metamerism [1, 2].

It is possible to reduce metamerism by increasing the number of channels in the device. For example, most commercially available cameras employ three channels, which are commonly chosen to be red, green and blue. Three channel, or trichromatic, cameras significantly reduce the degree of metamerism encountered in black and white cameras. Unfortunately, trichromatic cameras are not able to fully eradicate metamerism [2]. Thus, like in the example of a grey shirt in a black and white image, many surfaces might integrate to the same trichromatic response, making surface separation an impossible task.

To further decrease the degree of metamerism it is necessary to make cameras with more than three colour channels, such cameras are known as multispectral cameras [3, 4, 5, 6]. Unfortunately, the drive to increase the number of sensors is restricted by the increased cost and memory requirements as well as manufacturing limitations. In light of these constraints and the need for an increased number of channels we are faced with the question that we address in this paper, namely, what is the minimum number of sensors needed in a multispectral imaging device such that the error in recovering the reflectances is minimal.

As a first approximation we might assume that the number of channels needed in a multispectral camera is limited, and relates to the underlying dimensionality of the captured data. Such an assumption is supported by a large body of research in spectral data dimensionality where it is agreed that a small number of basis functions is adequate to fully represent large data sets. From analyzing 150 out of 433 Munsell chips Jozef Cohen [7] concluded that their reflectance depends on three components. Among later studies of the Munsell colours, Eem *et al.* [8] proposed four, Maloney[9] proposes five to seven, Burns[3] proposes five to six, Lenz *et al.* [10] use six, Parkkinen *et al.* [11] and Wang *et al.* [12] argue that eight components are necessary and in a recent study Hardeberg [13] demonstrated that as many as 18 basis functions are needed. The reason behind the discrepancies between these studies is that different authors use different thresholds for the required similarity between the original and reconstructed data.

There are two main drawbacks with basing our estimate for the number of sensors needed in multispectral imaging devices on the aforementioned studies. Firstly, the basis functions derived in those studies do not correspond to physically feasible sensors [14]. Secondly, the data is assumed to be noise-free; an assumption which is not justified in an actual imaging system where many types of noise are known to corrupt the response data [15]. Hence, in this paper we present two methods to derive physically feasible sensors such that they are optimised to record and reproduce the spectral data. Using the spectral curves of these sensors we are able to synthesise their responses to a database of Munsell reflectance spectra [11]. Doing so allows us to add educated levels of quantisation and shot noise [15], which makes it possible to study the efficacy of increasing the number of sensors in the imaging device

without having to assume perfect noise-free conditions.

Finally, the sensor design methods presented in this paper are derived such that, in the absence of noise, an increasing number of sensors is guaranteed to improve the reflectance estimates. Choosing sensors with this property allows us to concentrate on the question of the minimum number needed rather than the spectral properties of the sensors. In other studies that include variable numbers of sensors, among other factors, the sensors are often chosen to have arbitrary characteristics [4, 16]. As a result the effect of sensor number is confounded by the particular sensor characteristics chosen.

In the first section of this paper we review the principles of reflectance recovery and examine the role of surface smoothness in choosing the number of sensors in an imaging system. In Section 3 we introduce methods for deriving physically feasible sensors that are optimised for spectral recovery. We use these sensors in computational experiments described in Section 4 to assess the effect of sensor number on reflectance recovery in the presence of noise. In Sections 5 and 6 we present results that suggest that sensor noise provides a natural limit to decide the best number of sensors.

Background

By assuming that all surfaces are Lambertian, and that there is no fluorescence, the response of a digital camera at a single pixel can be modelled by Equation 1:

$$q_i = \int_\lambda Q_i(\lambda) E(\lambda) R(\lambda) d\lambda \tag{1}$$

where q_i is the response of the i^{th} sensor (i = 1, ...,P), $Q_i(\lambda)$ is the i^{th} sensor response function, $E(\lambda)$ is the spectral power distribution of the illuminant and $R(\lambda)$ is the surface spectral reflectance function. Note that we are neglecting noise for the time being.

These continuous functions can be sampled at a number of discrete wavelength intervals n without a significant loss of accuracy, providing that the interval is sufficiently small [17]. In this work we sample functions on the range from 400 nm to 700 nm at 10 nm intervals, thus $n = 31$. With this in mind Equation 1 can be rewritten as:

$$q_i = \sum_\lambda Q_i(\lambda) E(\lambda) R(\lambda) \Delta\lambda \tag{2}$$

This discrete sum is more conveniently expressed in terms of matrix-vector notation, thus we write:

$$\mathbf{q} = \mathbf{Q}^T\mathbf{r}, \tag{3}$$

where $\mathbf{q}$ is a $p \times 1$ vector of sensor responses and $\mathbf{r}$ is an $n \times 1$ reflectance vector. For compactness we represent the product of each sensor response function $Q_i(\lambda)$ and the illuminant $E(\lambda)$ as a single vector which forms the i^{th} column of the $n \times p$ sensor matrix $\mathbf{Q}$.

The problem of recovering reflectance from camera responses can now be expressed as the problem of estimating the $n \times 1$ vector $\mathbf{r}$ given the $p \times 1$ vector of camera responses $\mathbf{q}$ and the matrix $\mathbf{Q}$. This is a system of p linear equations in n unknowns. For an exact solution it is sufficient to set the number of knowns equal to the number of unknowns, i.e to use $p = 31$ independent sensors in the imaging system. However, such a large number of sensors may not be necessary for reflectance recovery. Real reflectance spectra are known to be strongly constrained and may be represented accurately with fewer than 31 parameters [13]. A convenient way to express this is to write reflectance as the weighted linear sum of a small number of basis vectors [9], i.e.

$$\mathbf{r} = \sum_{i=1}^{m} \mathbf{b}_i \omega_i \tag{4}$$

where $\mathbf{b}_i$ are the basis vectors, ω_i are the respective weights and $m \ll n$. This relation can be expressed in matrix vector notation thus:

$$\mathbf{r} = \mathbf{B}\boldsymbol{\omega} \tag{5}$$

where the columns of $\mathbf{B}$ are the basis vectors and $\boldsymbol{\omega}$ is a vector of weights. Replacing Equation 5 into Equation 3 gives:

$$\mathbf{q} = \mathbf{Q}^T\mathbf{B}\boldsymbol{\omega}. \tag{6}$$

This is a system of p equations in m unknowns. To solve uniquely for ω, and therefore $\mathbf{r}$, it is sufficient to set the number of independent sensors $p = m$. Providing that $\mathbf{Q}^T\mathbf{B}$ is invertible we can solve for the ω as follows [18, 19]:

$$\boldsymbol{\omega} = \left(\mathbf{Q}^T\mathbf{B}\right)^{-1}\mathbf{q}. \tag{7}$$

The principal problem with this approach is that it is not straightforward to determine an objective value for m, and therefore p. In order to understand this it is necessary to consider how to derive m from a statistical analysis using the singular value decomposition.

We can represent a set of k reflectance spectra as the columns of an $n \times k$ matrix $\mathbf{R}$. The singular value decomposition of $\mathbf{R}$ is given by:

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \tag{8}$$

where the matrices $\mathbf{U}$ and $\mathbf{V}$ are both orthonormal, i.e $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$, and $\boldsymbol{\Sigma}$ is a matrix whose leading diagonal contains the singular values of $\mathbf{R}$ with zeros elsewhere. The columns of $\mathbf{U}$ are the eigenvectors of the matrix $\mathbf{R}\mathbf{R}^T$ and are referred to as characteristic vectors. The characteristic vectors are a set of basis vectors for $\mathbf{R}$ that are ordered such that the first vector accounts for the most

variability in the data, the second accounts for the most variability in the residual from the first vector, and so on. Thus by increasing the number of characteristic vectors in the linear model we are guaranteed to progressively improve the reflectance estimate. Furthermore, the first m characteristic vectors give the closest possible fit of a linear model for any given value of m.

Although, by increasing the number of bases m in the linear model, the approximation can always be improved, the inherent smoothness of reflectance spectra determines that there is a point when increasing m results in very small improvements in accuracy. Generally m is estimated as the point when the small improvement drops below some arbitrary threshold. However, if we intend to use m to determine the number of sensors in a multispectral imaging system, then we must consider the point when the improvement in the accuracy of the linear model is cancelled by the detrimental effect of imaging noise. In order to assess the role of noise we need to make a real set of sensors to capture the data in typical noise conditions. This requires that we choose the sensor functions carefully, according to objective criteria[20]. Different sensor characteristics capture different information, and hence result in different reflectance estimates for the same sensor number. By not choosing carefully, the effect of increasing the number of sensors, and hence parameters in the linear model, will be confounded with the sensor characteristics. We would therefore like to choose sensors that are optimised for spectral recovery and thus guarantee that, in the absence of noise, increasing the number of sensors results in decreasing error. In order to do this we should choose sensors $\mathbf{Q}$ whose columns span the same vector space as the first m characteristic vectors[21].

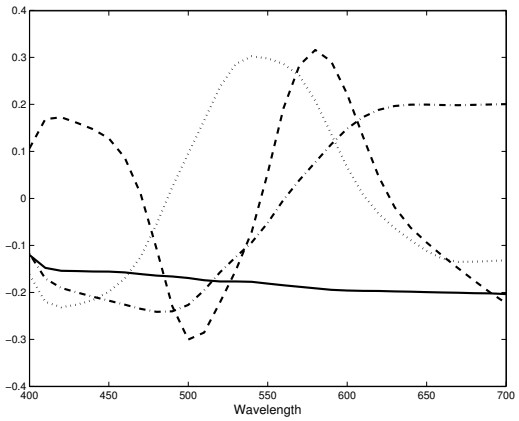

Figure 1: *The first four characteristic vectors of the Munsell reflectance data.*

Non-negative sensors

To guarantee that we choose the sensors to be within a linear transform of the basis vectors we must choose $\mathbf{Q}$ such that it satisfies the following relation:

$$\mathbf{Q} = \mathbf{BA} \tag{9}$$

where $\mathbf{A}$ is a linear transformation and $\mathbf{B}$ contains the first m characteristic vectors of $\mathbf{R}$ as columns. Initially, one might consider using the characteristic vectors themselves as sensors, i.e. let $\mathbf{A}$ be the identity matrix. However, as can be seen from Figure 1, the characteristic vectors contain many negative values, yet real sensors must be everywhere non-negative. Further, the modulation of the characteristic vectors is proportional to their order, i.e. additional vectors have an increasing number of peaks and troughs. It is therefore desirable to transform these vectors into a non-negative vector space, such that their individual sensitivities are concentrated in distinct regions of the visible spectrum. For example, in a trichromatic camera system the sensors are commonly chosen to be red, green and blue. Finally, the transformed vectors should ideally span the same space as the original.

Given these criteria, we would like to find the best transform $\mathbf{A}$ to solve for the sensors $\mathbf{Q}$. In this paper we propose to solve this problem using the varimax rotation algorithm described in[22, 23]. Starting from the $n \times m$ bases matrix $\mathbf{B}$, with elements b_{jk}, the varimax criterion is given by:

$$V(\mathbf{B}) = \sum_k \left(\frac{1}{n} \sum_j b_{jk}^4 - \left(\frac{1}{n} \sum_j b_{jk}^2 \right)^2 \right) \tag{10}$$

Verbally, Equation 10 is the columnwise variances of the squared elements of $\mathbf{B}$. Given the varimax criterion in Equation 10, the optimal transform $\mathbf{A}$ in Equation 9 is any orthogonal rotation of $\mathbf{B}$ that maximizes the varimax criterion among all other orthogonal rotations. Constraining $\mathbf{A}$ to be an orthogonal transform means that the resultant sensors $\mathbf{Q}$ are themselves orthogonal. Maintaining the orthogonality of the sensors is important, as it makes the recovery of reflectance from their output maximally robust to sensor noise[24]. The sensors generated by this procedure are shown in Figure 2.

The sensitivity of each sensor is clearly focussed in a different region of the visible spectrum. However, the rotated vectors still contain some negative lobes, which means that they cannot be used as sensors. We therefore choose everywhere non-negative sensors that are as close as possible to the rotated sensors, denoted $\widehat{\mathbf{Q}}$, but still within a linear transform of $\mathbf{B}$. Denoting $\widehat{\mathbf{Q}}_i$ as the i^{th} column of $\widehat{\mathbf{Q}}$, we can do this sequentially for the i^{th} sensor by solving

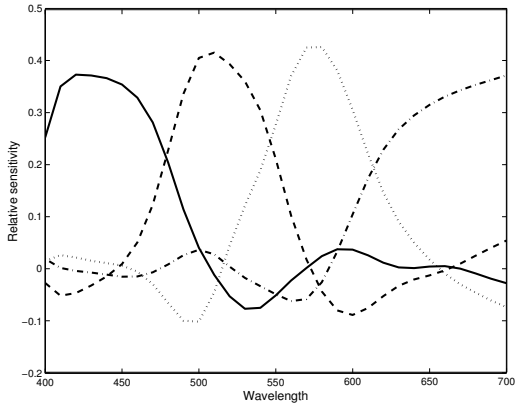

Figure 2: *The first four characteristic vectors rotated by the varimax algorithm to be maximally positive*

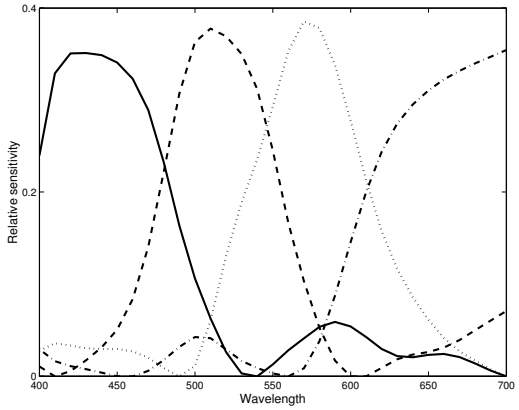

Figure 3: *Non-negative sensors formed by varimax rotation with added positivity constraint*

the following optimisation problem:

$$\min_{\mathbf{g}_i} \|\widehat{\mathbf{Q}}_i - \widehat{\mathbf{Q}}_i\mathbf{g}_i\|^{\mathbf{2}} \text{ subject to } \widehat{\mathbf{Q}}_i\mathbf{g}_i \geq \mathbf{0} \qquad (11)$$

where $\mathbf{g}_i$ is a $p \times 1$ vector and $\mathbf{0}$ is a vector of zeros. This results in the sensors shown in Figure 3, termed here the varimax sensors.

In earlier work Piché[25] generates non-negative combinations of the characteristic vectors that explicitly maximise the mutual orthogonality of the sensors. He points out that the orthogonality of the sensors can be measured directly by the *condition number* of $\mathbf{Q}$, where the condition number is given by:

$$cond\,(\mathbf{Q}) = \|\mathbf{Q}\|_2\|\mathbf{Q}^\dagger\|_2. \qquad (12)$$

Here $\|\cdot\|_2$ denotes the spectral norm of a matrix, which is given by its largest singular value, and † denotes the pseudoinverse operation. Given that the characteristic vectors are guaranteed to be orthogonal, the condition number of $\mathbf{Q}$ is determined solely by the condition number of $\mathbf{A}$. Piché therefore generates transformations of the characteristic vectors that explicitly attempt to minimise the condition number of $\mathbf{A}$. That is, he minimises the following objective function:

$$\min_{\mathbf{A}} \|\mathbf{A}\|_2\|\mathbf{A}^\dagger\|_2 \text{ subject to } \mathbf{Q} = \mathbf{BA} \geq \mathbf{0} \qquad (13)$$

where $\mathbf{0}$ is a matrix of zeros. This optimisation problem that can be tackled directly by using iterative non-linear optimisation methods. Sensors generated using this procedure, termed here Piché sensors, are shown in Figure 4. Note the similarity between the sensors in Figure 4 and those generated by the varimax procedure in Figure 3. We also find that the varimax and Piché procedures generate sensors with equally low condition numbers, even though Piché minimises condition number explicitly.

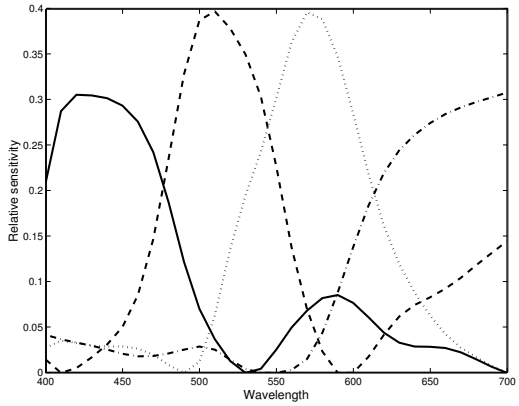

Figure 4: *Non-negative sensors formed by Piché's procedure*

Method

To generate each set of p sensors we choose the first p characteristic vectors of a set of 1269 Munsell reflectance spectra [11] and transform them into non-negative sensors using both the Piché and varimax procedures. We use the synthesised responses of these sensors to assess the effect of increasing sensor number, from 3 to 15 sensors, on reflectance recovery performance in the presence of noise. Synthetic camera responses are generated according to the following camera model:

$$\mathbf{q} = \mathbf{Q}^T\mathbf{r} + \mathbf{n}_{shot} + \mathbf{n}_{quant} \qquad (14)$$

where the vectors $\mathbf{n}_x$ denote sources of noise. Shot noise $\mathbf{n}_{shot}$ arises from the inherent uncertainty in the generation, reflection and capture of light. This is a Poisson process[15], thus the variance of the shot noise component

increases with increasing input intensity. We have modelled this using multiplicative Gaussian noise thus:

$$\mathbf{n}_{shot} = [\zeta_1 \mathbf{q}_1, \zeta_2 \mathbf{q}_2, \ldots, \zeta_p \mathbf{q}_p]^T \quad (15)$$

where each of the ζ_i is a pseudorandom variable taken from a Gaussian distribution with zero mean and variable standard deviation and $\mathbf{q}_i$ represents the i^{th} sensor response. Quantisation noise $\mathbf{n}_{quant}$ is incorporated by directly quantising the simulated responses after the application of shot noise. Other sources of noise, such as dark noise, are assumed to be negligible or corrected for.

In all calculations the equal energy illuminant E is used and the columns of $\mathbf{Q}$ all sum to 1, thus ensuring a camera response of 1 to a perfect reflecting diffuser. Reflectance is estimated from camera responses using Equations 7 and 5. The difference between original and estimated spectra is measured in terms of absolute route-mean-squared error, given by:

$$rms = \sqrt{\frac{(\mathbf{r} - \hat{\mathbf{r}})^T (\mathbf{r} - \hat{\mathbf{r}})}{n}} \quad (16)$$

where $\hat{\mathbf{r}}$ is the reflectance estimate and $\mathbf{r}$ is the original.

Simulation results

Results for different sensor numbers with 1% shot noise and 12 bit quantisation are shown in Figure 5 along with noise-free estimation results. When there is noise in the sensor responses the recovery error does not decrease monotonically with increasing sensor number as it does when there is no noise. Minimum error is reached at 11 sensors and 9 sensors for the varimax and Piché methods respectively, although the varimax sensors show little improvement beyond 9 sensors. Recent measurements of noise levels in a trichromatic camera suggest shot noise to be around 1-2% of the overall signal[26] thus results with 2% shot noise are shown in Figure 6. The overall RMS error is high, relative to the 1% case. However, the point at which minimum error occurs is still around 9 sensors. Clearly now, increasing the number of sensors beyond 9 is detrimental to performance. In results not shown here we also found that this limit holds using a least squares minimiser to recover reflectance, instead of Equations 7 and 5.

Discussion and conclusions

In these experiments we have used carefully chosen sensors that are both optimised for recovering the Munsell reflectances and are maximally robust to noise. We have used these to find the minimum number of sensors such that they provide minimal reconstruction error. Using fewer sensors leaves potential room for improvement, whereas using more sensors does not decrease RMS error due to the effect of noise. In the typical noise environment investigated here we find that the limit occurs at approximately 9 sensors, although we expect this value to change for different noise levels. This value corresponds to previous estimates of the dimensionality of this dataset made using different decision criteria by Parkkinen *et al.*[11] and Wang *et al.* [12]. In further work we will extend this analysis to a greater number reflectance datasets.

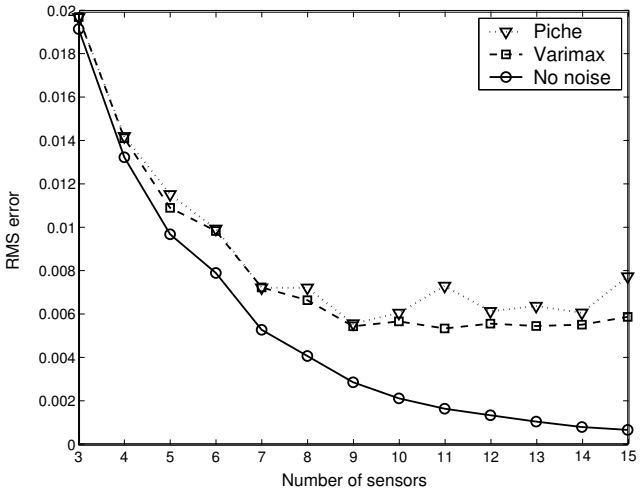

Figure 5: *Effect of increasing sensor number with 12 bit quantisation and 1% shot noise*

References

1. G. Wyszecki and W. S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae., John Wiley & Sons, 2 nd edn (1982).
2. M. S. Drew and B. V. Funt, Natural metamers, Computer Vision, Graphics, and Image Processing. Image Understanding, 56(2), 139 (1992).
3. P. D. Burns, Analysis of image noise in multispectral color acquisition, Ph.D. thesis, Center for Imaging Science, Rochester Institute of Technology (1997).
4. F. König and W. Praefcke, A multispectral scanner, in L. W. MacDonald and M. R. Luo, editors, Colour Imaging: Vision and Technology, John Wiley and Sons Ltd, pp. 129–144 (1999).
5. H. Sugiura, T. Kuno, N. Watanabe, N. Matoba, J. Hayashi and Y. Miyake, Development of high accurate multispectral cameras, in Proceedings of the International Symposium on Multispectral imaging and Color Reproduction for Digital Archives, Chiba University, Japan, pp. 73–80 (1999).
6. J. Y. Hardeberg, F. Schmitt and H. Brettel, Multispectral color image capture using a liquid crystal tunable filter, Optical Engineering, 41(10), 2532 (2002).
7. J. Cohen, Dependency of the spectral reflectance curves of the munsell color chips, Psychonomic Science, 1, 369 (1964).
8. J. K. Eem, H. D. Shin and S. O. Park, Reconstruction of surface spectral reflectances using characteristic vectors of Munsell colors, in Proceedings of IS&T and SID's 2nd Color Imaging Conference: Color Science, Systems and Applications, Scottsdale, Arizona, pp. 127–131 (1994).

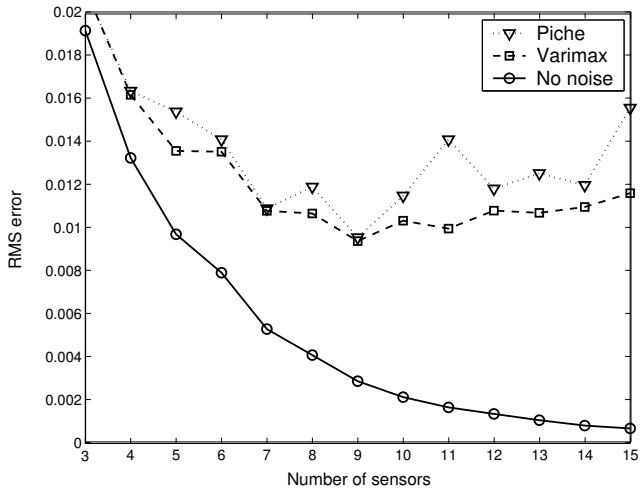

Figure 6: *Effect of increasing sensor number with 12 bit quantisation and 2% shot noise*

9. L. T. Maloney, Evaluation of linear models of surface spectral reflectance with a small number of parameters, Journal of the Optical Society of America - A, 3(10), 1673 (1986).
10. R. Lenz, M. Österberg, J. Hiltunen, T. Jaaskelainen and J. Parkkinen, Unsupervised filtering of color spectra, Journal of the Optical Society of America - A, 13(7), 1315 (1996).
11. J. Parkkinen, J. Hallikainen and T. Jaaskelainen, Characteristic spectra of Munsell colors, Journal of the Optical Society of America - A, 6, 318, see *http://cs.joensuu.fi/ spectral/* (1989).
12. W. Wang, M. Hauta-Kasari and S. Toyooka, Optimal filters design for measuring colors using unsupervised neural network, in Proceedings of the 8th Congress of the International Colour Association, AIC Color 97, Kyoto, Japan, vol. I, pp. 419–422 (1997).
13. J. Y. Hardeberg, On the spectral dimensionality of object colors, in Proceedings of CGIV'2002, First European Conference on Colour in Graphics, Imaging, and Vision, Poitiers, France, pp. 480–485 (2002).
14. G. Sharma and H. J. Trussell, Characterization of scanner sensitivity, in IS&T/SID Color Imaging Conference, pp. 103–107 (1993).
15. G. Healey and R. Kondepudy, Radiometric ccd camera calibration and noise estimation, IEEE transactions on Pattern Analysis and Machine Intelligence, 16(3), 267 (1994).
16. T. Keusen, Multispectral color system with an encoding format compatible with the conventional tristimulus model, Journal of Imaging Science and Technology, 40(6), 510 (1996).
17. H. J. Trussell and M. S. Kulkarni, Sampling and processing of color signals, IEEE Transactions on Image Processing, 5(4), 677 (1996).
18. L. T. Maloney and B. A. Wandell, Color constancy: A method for recovering surface spectral reflectance, Journal of the Optical Society of America - A, 3(1), 29 (1986).
19. G. Buchsbaum, A spatial processor model for object colour perception, Journal of the Franklin Institute, 310, 1 (1980).
20. J. Hernández-Andrés, J. L. Nieves, E. M. Valera and J. Romero, Spectral-daylight recovery by use of only a few sensors, Journal of the Optical Society of America - A, 21(1), 13 (2004).
21. Y-L Chang, P. Liang and S. Hackwood, Unified study of color sampling, Applied Optics, 28(4), 809 (1989).
22. H. F. Kaiser, The varimax criterion for analytical rotation in factor analysis, Psychometrika, 23, 187 (1958).
23. R. A. Reyment and K. G. Joreskog, Applied Factor Analysis in the Physical Sciences, Cambridge University Press (1993).
24. M. J. Vrhel and H. J. Trussell, Optimal color filters in the presence of noise, IEEE Transactions on Image Processing, 4(6), 814 (1995).
25. R. Piché, Nonnegative color spectrum analysis filters from principal component analysis characteristic spectra, Journal of the Optical Society of America - A, 19(10), 1946 (2002).
26. K. Barnard and B. Funt, Camera characterization for color research, Color Research and Application, 27(3), 153 (2002).

Biography

David Connah obtained a BSc in Biology in 1997 and an MSc in Machine Perception and Neurocomputing in 1998, both from the University of Keele, England. In 2004 he completed his PhD in Colour Science at the Colour and Imaging Institute, University of Derby, and since then has worked as a researcher in the Norwegian Color Research Lab at Gjvik University College. His primary research topic is theoretical and practical aspects of multispectral imaging.