A Basis for Cones

Brian Funt and Weihua Xiong
School of Computing Science, Simon Fraser University,
8888 University Drive, Burnaby, B.C. Canada V5A 1S6

ABSTRACT

Why do the human cones have the spectral sensitivities they do? We hypothesize that they may have evolved to their present form because their sensitivities are optimal in terms of their ability to recover the spectrum of incident light. As evidence in favor of this hypothesis, we compare the accuracy with which the incoming spectrum can be approximated by a three-dimensional linear model based on the cone responses and compare this to the optimal approximations defined by models based on principal components analysis, independent component analysis, non-negative matrix factorization and non-negative independent component analysis. We introduce a new method of reconstructing spectra from the cone responses and show that the cones are almost as good as these optimal methods in estimating the spectrum.

Keywords: Cone spectral sensitivities, PCA, ICA, NNMF, NNICA, spectral reconstruction

1. INTRODUCTION

Techniques such as principal components analysis (PCA) or independent component analysis (ICA) have often been used to extract an optimal basis for spectra [1-5]. Although these methods have been very successful, there is no immediate connection between the use of these models and physically realizable sensors. In the case of the cones, their response to incoming light can be modeled mathematically as the projection of the incoming spectrum on to the cone sensitivity functions. In the case of a PCA, weighting coefficients are calculated by projection of the spectrum on to the pseudo-inverse of the PCA basis. This is analogous to projection on to sensors, but the problem is that the PCA 'sensors' contain many negative values and are therefore not physically realizable. This leads us to ask the question: Are the human cones optimal in terms of the information they collectively capture about the incoming spectrum? We will limit our investigation to trichromatic sensor systems operating within the visible spectrum.

Mollon [6] provides an evolutionary theory of the human cones in terms of a dichromatic ancient subsystem inherited from reptiles (roughly S and M cones) followed by the later addition of L cones in response to a particular type of fruit tree the yellow-orange fruit of which was distributed only by monkeys. This theory provides a reason for trichromacy and a rough specification of the spectral range of the sensor classes; however, it does not specify the cone sensitivities in detail. Osorio et. al. [7] consider how the demands for coding spectral versus spatial information affect the optimality of sensors. They also consider the discriminability of edible fruit against a background of leaves and the influence this has on the need for trichromatic vision[24] and how this might effect the spectral tuning in terms of where the peak sensitivity of each cone type might lie [25]. Wachtler et. al. [8] use independent component analysis to extract a non-orthogonal basis for chromatic and spatial components of hyperspectral images and find connections to color opponency.

The problem of reconstructing spectra from LMS or equivalently XYZ triples has been addressed previously in the context of mapping camera RGB to XYZ[9]. In that work, a spectrum for a given XYZ was constructed by finding the weighting coefficients of a 3-dimensional linear model of a metameric spectrum. Lenz et. al [10] address the issue of how to derive realizable sensors optimized to a set of reflectances; however, they do not relate these sensors to the cones.

Finite-dimensional models of spectra have been widely used in modeling spectra[1,2,4,5]; however, often they are used without reference to any particular sensor functions. In a finite-dimensional linear model (FDM), a spectrum is approximated in terms of a weighed sum of a set of basis spectra. The weights are determined by projecting the original spectrum onto the pseudo-inverse of the basis spectra. PCA finds an orthogonal basis, so in this case the pseudo-inverse is simply the transpose of the basis.

Our first approach to linking the FDM basis to the cones is to consider approaches that find non-negative basis vectors. In particular, non-negative ICA and non-negative matrix factorization both provide entirely non-negative basis vectors. However, the pseudo-inverse of a non-negative basis is not necessarily non-negative; nonetheless, it turns out to be mainly non-negative [11]. When the negative parts are set to zero, are these 'sensors' then similar to the cones? They turn out to be similar, but still quite different.

A second approach is to find the best linear transformation fitting the cone sensitivity functions to the pseudo-inverse of the basis functions. In this case, the fit turns out to be quite good. Therefore we conclude that the cones are capturing almost all the information required for an optimal 3-dimensional linear model of color signal spectra. Although a 3-dimensional model may not suffice when modeling arbitrary reflectances, Chiao et. al. [12] established that it is probably sufficient for natural forest and coral reef scenes. In any case, since the human visual system has only 3 cone types, any model linking it to surface reflectance is necessarily 3-dimensional. We are not arguing here whether or not a 3-dimensional model is sufficient, but rather whether the cones are optimized to extract the best 3-dimensional model.

The approach of Lewis and Zhaoping [13] shares the same goal as the present paper. They consider whether shifting the wavelength of the peak sensitivity of each of the long-, medium- and short-wavelength cones would lead them to recover more information about reflectance. They measure the change in terms of the change in the amount of mutual information between reflectances and cone outputs.

We will now proceed to describe the technical details of our method.

2. OPTIMAL BASES

A color signal, $C(\lambda) = S(\lambda)E(\lambda)$, is the light entering the eye from a surface with reflectance $S(\lambda)$ illuminated by light with spectral power distribution $E(\lambda)$. Let the cone sensitivity functions be $\Phi_j(\lambda)$. For uniform, finite sampling of wavelength at *n* intervals, the functions can be represented as column vectors, in which case the cone quantum catches are

$$\phi_j = \bar{c}^T \Phi_j$$

The $\phi_j, j = 1,2,3$ will also be referred to as LMS.

For a collection of *m* color signals $\bar{c}_i$ arranged as the rows of a matrix *C*, a 3-dimensioanl linear model representation of the color signals expressed in terms of a matrix *B* of basis vectors and a matrix *W* of weights is a matrix product:

$$C = WB$$

C is *m-by-n, W* is *m-by-3 and B* is *3-by-n*. *W* is sometimes referred as the mixing matrix, and each row W_i is the corresponding weights vector of $\bar{c}_i$. In terms of finding an optimal basis for color signals, the central issue is to find *B* minimizing:

$$\sum_{i=1}^{m} \left\| \bar{c}_i - W_i B \right\|$$

A standard solution to this last equation is PCA, which finds a basis which is uncorrelated and orthogonal. However, PCA relies on the assumption that the data have a Gaussian distribution. ICA is another technique for solving it. The basis vectors derived via ICA are not only uncorrelated but also independent; however, they are not orthogonal.

The last equation can also be solved subject to the constraint that all elements of B be non-negative. Non-negative Matrix Factorization (NNMF) employs an iterative computation to minimize it using

$$\begin{cases} B_{ij}^{n+1} = B_{ij}^{n} \dfrac{((W^n)^T C)_{ij}}{((W^n)^T W^n B^n)_{ij}} \\ W_{ij}^{n+1} = W_{ij}^{n} \dfrac{(C(B^n)^T)_{ij}}{(W^n B^n (B^n)^T)_{ij}} \end{cases}$$

For initialization, B^0 and W^0 are given uniformly distributed, random elements.

ICA has also been extended to non-negative independent component analysis (NNICA) with the addition of a non-negativity constraint on the elements of B [14,16,17].

3. EXPERIMENTAL PRELIMINARIES

3.1. Data Preparation

Color signals were computed using the 1781 surface reflectances and 102 illuminant spectra available for download from [20]. The wavelength range is from 380nm to 780nm sampled at a 4nm interval. This results in a set of 181,662 color signals. We then randomly divide this into two subsets, one for training and the other for testing.

The Vos et. al. cone fundamentals [21] formulated as sensitivity vectors are normalized so each vector has unit norm. We also experimented with the Stockman-Sharpe fundamentals [22] but they made no significant difference to our results.

3.2. Implementations Used

For PCA we used 'princomp' from Matlab Version 7.0.0 [23]. For ICA, we used the JADE [18] method based on the implementation available from [19]. For NNMF and NNICA, we implemented them in Matlab based on the algorithms provided in [15] and [17] respectively.

3.3. Error Measures

First of all, we need a measure of the difference between two spectra. For any test color signal, $\bar{c}$, and its corresponding reconstructed spectrum is $\tilde{c}$ the root mean square distance will be used to measure their difference:

$$RMS_{\bar{c}\_\tilde{c}} = \sqrt{\frac{1}{n}\sum_{k=1}^{n} | \bar{c}(\lambda_k) - \tilde{c}(\lambda_k) |^2}$$

We also need a way to compare the difference between two sets of sensors since we will be comparing the cone sensitivities to sensors approximating the cone sensitivities. We will compare them in terms of the difference in their response to color signals. Projecting color signal $\bar{c}$ onto the actual cone functions versus the simulated sensitivities gives two sets of response triples (L, M, S) and $(\tilde{L}, \tilde{M}, \tilde{S})$. The distance and angular difference between them are given by:

$$e_{dis} = \sqrt{(L - \tilde{L})^2 + (M - \tilde{M})^2 + (S - \tilde{S})^2}$$

$$e_{ang} = COS^{-1}\left[\frac{(L, M, S) \bullet (\tilde{L}, \tilde{M}, \tilde{S})}{\sqrt{L^2 + M^2 + S^2}\,\sqrt{\tilde{L}^2 + \tilde{M}^2 + \tilde{S}^2}}\right] \times \frac{2\pi}{360}$$

For a test set of N color signals $\bar{c}_i$ the mean RMS spectral difference is defined as:

$$M\,RMS = \frac{1}{N}\sum_{i=1}^{N} RMS_{\bar{c}_i\_\tilde{c}_i}$$

The RMS of LMS distance and LMS angular errors are:

$$RMS_{dis\tan ce} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} e_{dis\_i}}$$

$$RMS_{angular} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} e_{ang\_i}}$$

4. EXPERIMENTS AND RESULTS

4.1. Experiment I

Given a set of basis vectors, a color signal spectrum $\overline{C_i}$ can be expressed in terms of the weighting coefficients of a linear model by projection on to the inverse of the basis, $w = \overline{c}_i^{\,T} B^{-1}$. The PCA basis vectors are orthogonal so $B^{-1} = B^T$. However, for the other methods the basis vectors are not orthogonal so B^+, the pseudo-inverse of matrix B, is used to obtain the coefficients instead. Given the weighting coefficients the color signal can be estimated as $\overline{C_i} = (wB)^T = B^T w^T$. We first consider whether the pseudo-inverse, B^+, of the basis vectors, B, immediately yields non-negative sensors that are similar to the human cone sensitivities Φ. To the extent that they are similar, the cone quantum catches would then be the weighting coefficients needed to reconstruct an incident color signal in terms of the B. As such, the cone quantum catches would represent the optimal information about the color signal.

We proceed as follows:

(1) For each analysis method (PCA, ICA, NNMF and NNICA) determine the matrix B of basis vectors.

(2) Calculate the pseudo-inverse B^+, (We used Matlab pinv(B)). In the case of PCA the basis vectors are mutually orthogonal so $B^+=B^T$.

(3) Set all negative values in B^+ to zero. Call the zeroed result B_Z^+. We might expect there to be less need for truncation for NNMF and NNICA, but it is possible that PCA and ICA might have mainly non-negative pseudo-inverses.

(4) Evaluate how closely B_Z^+ matches the cone sensitivities Φ graphically (Figure 1) and numerically (Table 1)

As can be seen from Figure 1, the curves are quite similar but far from identical. In particular the long wavelength basis vector does not overlap very well with the L cone sensitivity. For a numerical comparison, we use the distance and angular measurements from (8). The results are tabulated in Table 1. From these results, we conclude that the fit is perhaps not accurate enough to sustain the argument that the cones are effectively equivalent to B_Z^+.

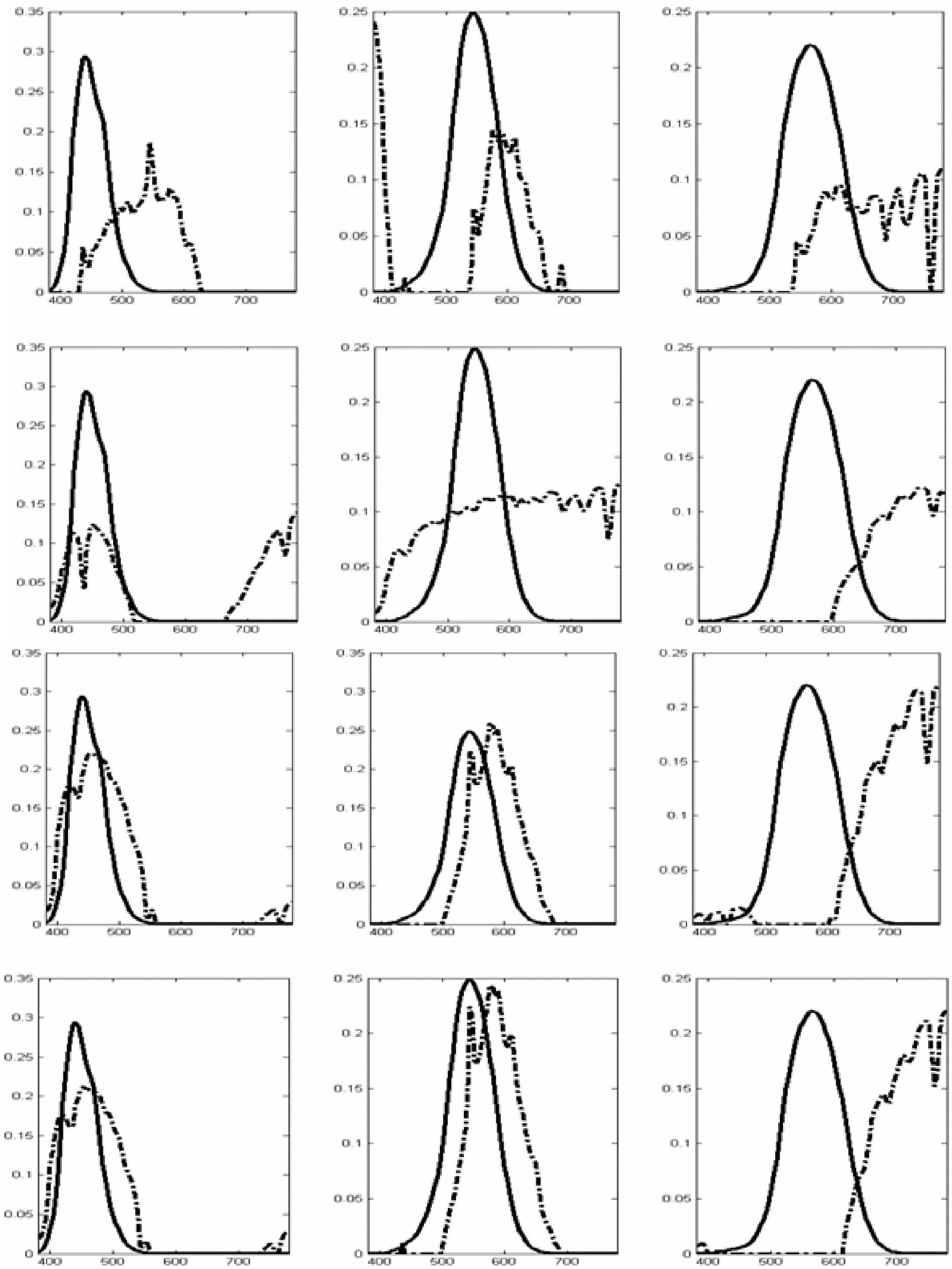

Figure 1: Cone sensitivities (solid line) compared to the truncated pseudo-inverse B_Z^+ (broken line) for ICA, PCA, NNMF and NNICA in top-to-bottom order. The non-negative methods result in the best fit; nonetheless, the long-wavelength sensor (right hand column) does not match very well.

	Distance		Angular	
	Max	RMS	Max	RMS
ICA	4.284	0.863	58.82	23.2193
PCA	4.435	1.097	67.98	30.4996
NNMF	4.173	0.809	56.77	22.8229
NNICA	4.259	0.802	57.49	22.6264

Table 1: Distance and angular errors between projecting spectra onto B_Z^+ versus the cones.

4.2. Experiment II

In our second approach, we introduce a linear transformation Q mapping cone sensitivities Φ to a best fit of B^+. In Experiment I, we simply took B^+ as the candidate cone sensitivities. In this second approach, we in essence are asking how similar the information that B^+ captures about the color signal is to the information the cones capture. If there is a linear mapping of the Φ to B^+, then it will also map LMS cone responses directly to the weighting coefficients of a 3-dimensional linear model. Similar transformations occur in the human visual, for example in color opponency.

We proceed as follows:

(1) and (2) are as in Experiment I

(3) Find a 3x3 linear transformation Q such that
$Q\Phi = B^+$, that is $Q = B^+ \Phi^+$.

(4) The linear transformed human cones are $\Phi' = Q\Phi$. Φ' should approximate B$^+$.

(5) From the cone quantum catches, weighting coefficients $\widetilde{W}_{Q\_i}$ representing color signal $\overline{c}_i$ in terms of basis B are estimated as $\widetilde{W}_{Q\_i} = Q\overline{\phi}_i$, where $\overline{\phi}_i$ is the 3-tuple of LMS response to color signal $\overline{c}_i$

(6) Given the weights $\widetilde{W}_{Q\_i}$, the color signal is reconstructed as $\widetilde{c}_{Q\_i} = \widetilde{W}_{Q\_i} B$

In other words, the cone sensitivities are mapped to fit the pseudo-inverse of the basis functions. Since this mapping and the projection of color signal onto sensors are both linear operations, projection on to the transformed sensitivity functions is equivalent to applying the mapping to the sensor responses directly. Hence, the mapping takes LMS directly to the weighting coefficients of the linear model defined by basis *B*. If the fit of the cones to the basis pseudo-inverse in step (3) were perfect, then the cones would be capturing all the information required to make an optimal estimate of the incident spectrum.

To evaluate how good the fit is of the cones to B^+, we first compare the error in reconstructing spectra via the cones to the error using B^+ directly. The standard linear model reconstruction of spectrum $\overline{c}_i$ using B^+ is $\widetilde{c}_{R\_i} = \overline{c}_i^T B^+ B$.

We also compare the reconstruction error to the metamer method [15], which also estimates spectra in terms of basis B from cone quantum catches. In this method, weights $\widetilde{W}_N$ are calculated that produce a metamer to the incident color signal. The reconstructed color signal is then $\widetilde{c}_N = \widetilde{W}_N B$.

Figure 2 shows how well the basis functions map to the cone sensitivity functions. Table 2 compares the errors of the three methods in terms of mean and maximum RMS error in spectrum estimation. The best case is for the combination of the Experiment II method and the ICA basis. The results, of course, depend on the training set of color signal spectra used. The error estimating the color signal from the cones is surprisingly close to the error with any of the 4 possible 3-dimensional linear models given that the cones are restricted to being physically realizable. For the other models, the color signal is projected onto 'sensors' (i.e., B^+) that are not restricted to being physically realizable. In terms of physically realizable sensors, the cones are very effective in characterizing the incident color signal spectrum.

		Distance Mean RMS	Distance Max RMS
ICA	C vs. $\widetilde{C}_R$	0.029347	0.1815
	C vs. $\widetilde{C}_Q$	0.049312	0.2497
	C vs. $\widetilde{C}_N$	0.070862	0.5800
PCA	C vs. $\widetilde{C}_R$	0.030171	0.1839
	C vs. $\widetilde{C}_Q$	0.075304	0.3919
	C vs. $\widetilde{C}_N$	0.087495	0.8036
NNMF	C vs. $\widetilde{C}_R$	0.030324	0.1841
	C vs. $\widetilde{C}_Q$	0.080972	0.4600
	C vs. $\widetilde{C}_N$	0.088564	0.8173
NNICA	C vs. $\widetilde{C}_R$	0.032166	0.1993
	C vs. $\widetilde{C}_Q$	0.081558	0.4673
	C vs. $\widetilde{C}_N$	0.130542	1.2803

Table 2: A comparison of the mean and maximum RMS error of the various combinations of reconstruction method ($\widetilde{c}_Q$ cones fit to basis pseudo-inverse from Experiment II; $\widetilde{c}_N$ metamer method) and basis type when estimating a color signal C from LMS quantum catches. The errors are also compared to the best possible reconstruction $\widetilde{c}_R$ obtained using a 3-dimensional linear model when the 'sensors' are not restricted to being physically realizable.

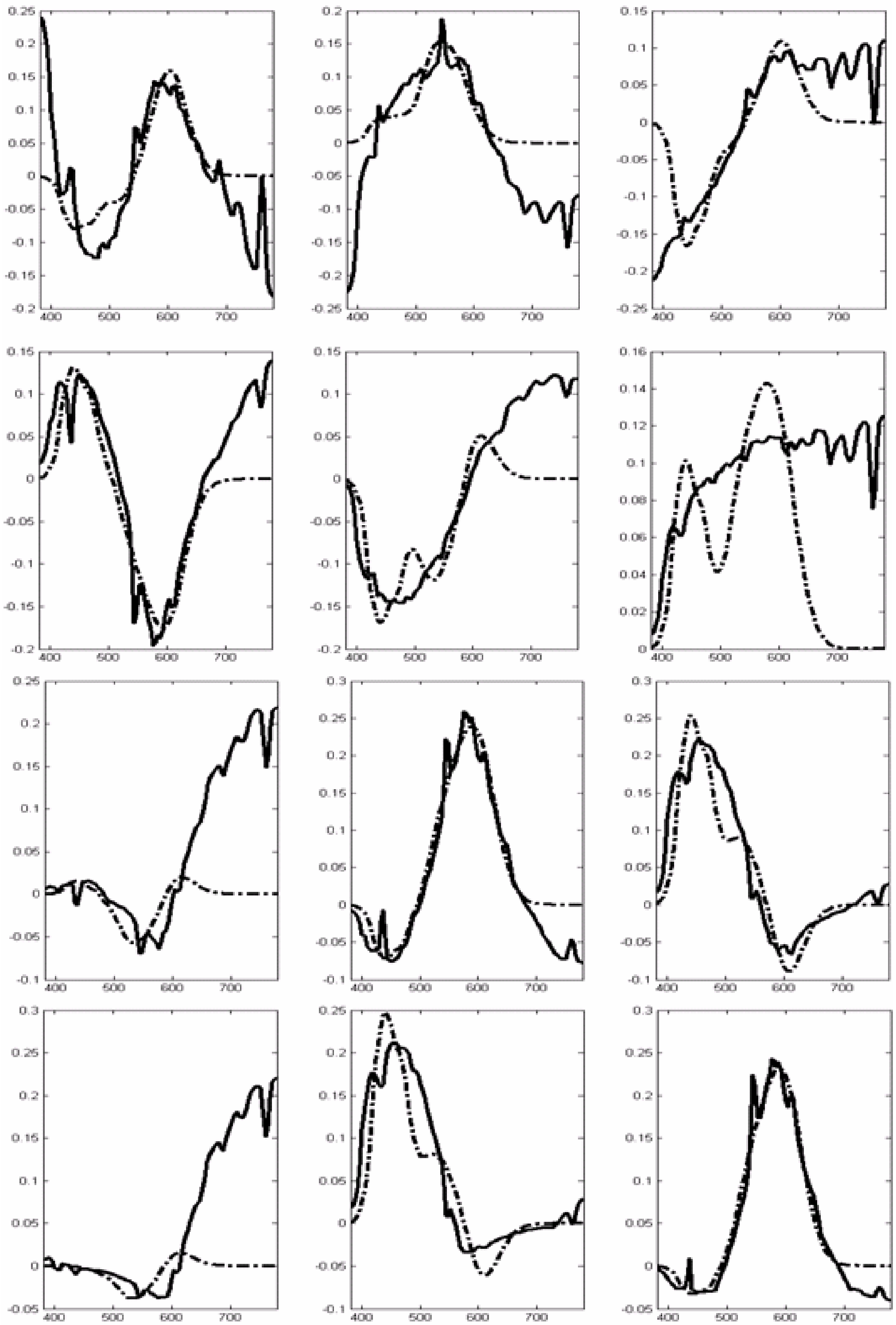

Figure 2: A comparison of the best fit mapping of the cone sensitivities Φ (broken line) to B^{+} (solid line) the pseudo-inverse of the basis from each of ICA, PCA, NNMF and NNICA in top-to-bottom order.

5. CONCLUSION

Our hypothesis was that the cone sensitivity functions are perhaps optimized in terms of recovering information about the spectrum of the light they receive. As evidence for this hypothesis we consider 2 new methods of reconstructing spectra from the cone responses. The best method finds the linear transformation mapping the cone sensitivities to the pseudo-inverse of the basis functions derived from independent component analysis. Our results show that the cones are almost as effective at modeling a spectrum as the best possible (but not physically realizable) trichromatic system. This does not prove that the cones are the best physically realizable trichromatic sensor system, but it does demonstrate that they must be at least close to the best.

ACKNOWLEDGMENT

Financial support was provided by the Natural Sciences and Engineering Council of Canada.

REFERENCES

1. H. Laamanen, T. Jaaskelainen, and J.P.S. Parkkinen, "Comparison of PCA and ICA in color recognition," *Proceedings of Intelligent Robots and Computer Vision*, SPIE vol. 4197, Boston, USA, pp367-377, November 7-8, 2000.
2. H.S. Fairman and M.H. Brill, "The Principal Component of Reflectances," *Color Research and Application*, Vol. 29, Issue 2, pp.104-110, April 2004
3. R .Ramanath, R.G. Kuehni, W.E. Snyder, and D.Hinks, "Spectral Spaces and Color Spaces," *Color Research and Application*, Vol. 29, Issue 1, Feb. 2004
4. D.H. Marimont and B.A. Wandell, "Linear Models of Surface and Illuminant Spectral," *Journal of the Optical Society of America A*, Vol. 9, No. 11, Nov. 1992
5. L.T. Malony and B.A. Wandell, "evaluation of linear models of surface spectral reflectance with small numbers of parameters," *Journal of the Optical Society of America A*, Vol. 3, pp1673-1683, 1986
6. J. D. Mollon "Cherries among the leaves: The evolutionary origins of colour vision." In *Colour Perception: Philosophical, Psychological, Artistic, and Computational Perspectives, S. Davis (Ed)*, Oxford University Press, pp. 10-30, 2000.
7. D.Osorio, and M. Vorobyev, "Photoreceptor spectral sensitivities in terrestrial animals: adaptations of luminance and colour vision," *Proc. R. Soc. B.* Vol. 272, pp. 1745-1752, 2005
8. T. Wachtler, TW Lee, TJ Sejnowski , "Chromatic Structure of Natural Scenes*," J. Opt. Soc. Am. A,* Vol. 18, No.1, pp. 65-77, 2001
9. M. Drew, and B. Funt, "Natural Metamers," CVGIP: Image Understanding, Vol.56, No.2, pp.139-151, 1992
10. R. Lenz., M. Österberg, J. Hiltunen, and T.J.J. Parkkinen, "Unsupervised Filtering of Color Spectra," *J. Opt. Soc. Am. A, Vol.* 13, pp. 1315-1324, 1996)
11. W.H. Xiong, and B. Funt, "Independent Component Analysis and Non-negative Linear Model Analysis of Illuminant and Reflectance Spectra," *Proceedings AIC Colour of 10th Conference of the International Colour Association*, Spain, May 2005.
12. CC. Chiao, TW. Cronin, D. Osorio. "Color Signals in natural Scenes: Characteristics of Reflectance Spectra and Effects of Natural Illuminants," *J. Opt. Soc. Am. A, Vol.* 17, No. 2, pp. 218-224, 2000
13. A. Lewis, L. Zhaoping, "Efficient Coding, Natural Colour Statistics and Photoreceptor Responses, *Proc. AIC Colour*, pp. 65-68, 2005
14. M. D. Plumbley, "Optimization using Fourier Expansion over a Geodesic for Non-Negative ICA," *5th International Conference on Independent Component Analysis and Blind Signal Separation (ICA 2004)*, Granada, Spain, pp 22-24, Sep., 2004.
15. D.D. Lee and H.S.Seung, "Algorithms for Non-negative Matrix Factorization," *Advanced Neural Information Processing System*, Vol. 13, pp43-56, 2001
16. M.D. Plumbley, "A 'Nonnegative PCA' Algorithm for Independent Component Analysis," *IEEE Transactions On Neural Networks*, Vol. 15, No 1, Jan., 2004

17. M.D. Plumbley, “Algorithm for Non-Negative Independent Component Analysis,” *IEEE Transactions On Neural Networks*, Vol. 14, pp. 534-543, May, 2003
18. J.-F. Cardoso and A. Souloumiac, “Blind Beamforming for non Gaussian signals,” *IEEE Transactions on Signal Processing*, Vol. 46, Issue 7, pp.1878-1885, Jul. 1998.
19. http://sig.enst.fr/~cardoso/icasentral/Algos/cardoso
20. http://www.cs.sfu.ca/~colour/data/
21. J.J. Vos, O. Estévez, & P.L. Walraven, “Improved color fundamentals offer a new view on photometric additivity,” *Vision Research*, 30, pp. 936-943, 1990.
22. W, A. Stockman and L. T. Sharpe, “The Spectral Sensitivities of the Middle- and Long-wavelength-sensitive Cones Derived from Measurements in Observers of Known Genotype,” *Vision Research,* Vol. 40, pp. 1711–1737, 2000.
23. www.mathworks.com
24. D. Osorio and M. Vorobyev, "Colour vision as an adaptation to frugivory in primates," *Proc. R. Soc. Lond. B*, Vol. pp. 263: 593-599, 1996
25. D. Osorio, “Eye design for coding natural spectra,” *Proc. AIC Colour 05 10th Congress of the International Colour Association*, pp. 61-65, Granada 2005