Spectral sharpening: sensor transformations for improved color constancy

Graham D. Finlayson, Mark S. Drew, and Brian V. Funt

School of Computing Science, Simon Fraser University, Vancouver, B.C., Canada V5A 1S6

Received March 8, 1993; revised manuscript accepted October 28, 1993; accepted October 28, 1993

We develop sensor transformations, collectively called spectral sharpening, that convert a given set of sensor sensitivity functions into a new set that will improve the performance of any color-constancy algorithm that is based on an independent adjustment of the sensor response channels. Independent adjustment of multiplicative coefficients corresponds to the application of a diagonal-matrix transform (DMT) to the sensor response vector and is a common feature of many theories of color constancy, Land's retinex and von Kries adaptation in particular. We set forth three techniques for spectral sharpening. Sensor-based sharpening focuses on the production of new sensors as linear combinations of the given ones such that each new sensor has its spectral sensitivity concentrated as much as possible within a narrow band of wavelengths. Data-based sharpening, on the other hand, extracts new sensors by optimizing the ability of a DMT to account for a given illumination change by examining the sensor response vectors obtained from a set of surfaces under two different illuminants. Finally in perfect sharpening we demonstrate that, if illumination and surface reflectance are described by two- and three-parameter finite-dimensional models, there exists a unique optimal sharpening transform. All three sharpening methods yield similar results. When sharpened cone sensitivities are used as sensors, a DMT models illumination change extremely well. We present simulation results suggesting that in general nondiagonal transforms can do only marginally better. Our sharpening results correlate well with the psychophysical evidence of spectral sharpening in the human visual system.

Key words: spectral sharpening, color constancy, color balancing, lightness, von Kries adaptation.

1. INTRODUCTION

The performance of any color-constancy algorithm, whether implemented biologically or mechanically, will be strongly affected by the spectral sensitivities of the sensors providing its input. Although in humans the cone sensitivities obviously cannot be changed, we need not assume that they form the only possible input to the color-constancy process. New sensor sensitivities can be constructed as linear combinations of the original sensitivities, and in this paper we explore what the most advantageous such linear transformations might be.

We call the sensor response vector for a surface viewed under an arbitrary test illuminant an observation. The response vector for a surface viewed under a fixed canonical light is called a descriptor. We take as the goal of color constancy the mapping of observations to descriptors. Since a descriptor is independent of illumination, it encapsulates surface reflectance properties.[1]

In the discussion of a color-constancy algorithm there are two separate issues: the type of mechanism or vehicle supporting the transformation from observations to descriptors in general and the method used to calculate the specific transformation that is applicable under a particular illumination. In this paper we address only the former and therefore are not proposing a completely new theory of color constancy.

A diagonal-matrix transformation (DMT) has been the transformation vehicle for many color-constancy algorithms, in particular von Kries adaptation,[2] all the retinex/lightness algorithms,[3–5] and, more recently, Forsyth's gamut-mapping approach.[6] All these algorithms respond to changing illumination by adjusting the response of each sensor channel independently, although the strategies that they use to decide on the actual adjustments differ.

DMT support of color constancy is expressed mathematically in relation (1). Here $\mathbf{p}_i{}^e$ denotes an observation (a 3-vector of sensor responses), where i and e are index surface reflectances and illumination, respectively. The vector $\mathbf{p}_i{}^c$ represents a descriptor and depends on the single canonical illuminant. The diagonal transform $\mathcal{D}^e$ best maps observations onto descriptors. Throughout, the superscript e denotes dependence on a variable illuminant and the superscript c denotes dependence on the fixed canonical illuminant. Boldface indicates vector quantities.

$$\mathbf{p}_i{}^c \approx \mathcal{D}^e \mathbf{p}_i{}^e . \tag{1}$$

In general, there may be significant error in this approximation. Indeed West and Brill[2] and D'Zmura and Lennie[7] have shown that a visual system equipped with sensors having the same spectral sensitivity as the human cones can achieve only approximate color constancy through a DMT.

A DMT will work better with some sensor sensitivities than with others, as one can see by considering how the illumination, surface reflectance, and sensor sensitivities combine in the formation of an observation. An observation corresponds to

$$\mathbf{p} = \int_\omega E(\lambda) S(\lambda) \mathbf{R}(\lambda) \mathrm{d}\lambda , \tag{2}$$

where $E(\lambda)$, $S(\lambda)$, and $\mathbf{R}(\lambda)$ denote illumination, surface reflectance, and sensor sensitivities, respectively, and the

0740-3232/94/051553-11$06.00

integral is taken over the visible spectrum ω. For a DMT to suffice in modeling illumination change,[2] it must be the case, given a reference reflectance S_r, a secondary arbitrary reflectance S, and illuminants E^i and E^j, that,

$$\frac{\int_\omega E^i(\lambda)S(\lambda)\mathbf{R}(\lambda)\mathrm{d}\lambda}{\int_\omega E^i(\lambda)S_r(\lambda)\mathbf{R}(\lambda)\mathrm{d}\lambda} = \frac{\int_\omega E^j(\lambda)S(\lambda)\mathbf{R}(\lambda)\mathrm{d}\lambda}{\int_\omega E^j(\lambda)S_r(\lambda)\mathbf{R}(\lambda)\mathrm{d}\lambda}. \quad (3)$$

As others have observed, one way to ensure that this condition holds is to use extremely narrow-band sensors, which in the limit leads to sensors that are sensitive to a single wavelength (Dirac delta functions).[6] Our intuition when we began this work was that if we could find a linear combination of sensor sensitivities such that the new sensors would be sharper (more narrow band), then the performance of DMT color-constancy algorithms should improve and the error of relation (1) would be reduced. It should be noted, however, that Eq. (3) can be satisfied in other ways, such as by the placement of constraints on the space of illuminants or reflectances.[2] With the addition of sharpening, relation (1) becomes

$$\mathcal{T}\mathbf{p}_i{}^c \approx \mathcal{D}^e\mathcal{T}\mathbf{p}_i{}^e, \quad (4)$$

where $\mathcal{T}$ denotes the sharpening transform of the original sensor sensitivities. It is important to note that applying a linear transformation to response vectors has the same effect as application of the transformation to the sensor sensitivity functions.

The sharpening transform effectively generalizes diagonal-matrix theories of color constancy. Other authors[2,3,6] also have discussed this concept of an intermediate (or sharpening) transform. However, our work appears to be the first to consider the precise form of this transform.

The sharpening transform is a mechanism through which the inherent simplicity of many color-constancy algorithms can be maintained. For example, Land's retinex algorithm requires color ratios to be illumination independent (and hence implicitly assumes a diagonal-matrix model of color constancy), which, as can be seen from Eq. (3), they generally will not be. It seems difficult to improve the accuracy of retinex ratioing directly without making the overall algorithm much more complicated[8]; however, if a simple, fixed sharpening transformation of the sensors as a preprocessing stage is applied, the rest of the retinex process can remain untouched. Similar arguments apply to Forsyth's standards for coefficient rule (CRULE)[6] and Brill's[9] volumetric theory.

We initially present two methods for calculating $\mathcal{T}$: sensor-based and data-based sharpening. Sensor-based sharpening is a general technique for determining the linear combination of a given sensor set that is maximally sensitive to subintervals of the visible spectrum. This method is founded on the intuition that narrow-band sensors will improve the performance of DMT theories of color constancy. We apply sensor-based sharpening over three different ranges in order to generate three new sharpened sensors that are maximally sensitive in the long-wave, medium-wave, and shortwave bands. Figure 1 below contrasts the cone fundamentals derived by Vos and Walraven[10] (VW) before and after sharpening. Although the new sensitivity functions are sharper, they are far from meeting the intuitive goal of being very narrow band (i.e., with strictly zero response in all but a small spectral region); nonetheless, we perform simulations that show that they in fact work much better than the unsharpened cone fundamentals. In Section 2 we present the details of sensor-based sharpening.

Sensor-based sharpening does not take into account the characteristics of the possible illuminants and reflectances but considers only the sharpness of the resulting sensor. Our second sharpening technique, data-based sharpening, is a tool for validating the sensor-based sharpening method. Given observations of real surface reflectances viewed under a test illuminant and their corresponding descriptors, data-based sharpening finds the best, subject to a least-squares criterion, sharpening transform $\mathcal{T}$. Interestingly, data-based sharpening yields stable results for all the test illuminations that we tried, and in all cases the data-based-derived sensors are similar to the fixed sensor-based sharpened sensors. Data-based sharpening is presented in Section 3.

In Section 4 we present simulations evaluating diagonal-matrix color constancy for sharpened and unsharpened sensor sets. Over a wide range of illuminations, sensor-based sharpened sensors provide a significant increase in color-constancy performance.

Data-based sharpening is related to Brill's[9] volumetric theory of color constancy. This relationship is explored in Section 5. Through spectral sharpening, the volumetric theory is shown to be informationally equivalent to Land's[11] white-patch retinex.

The data-based sharpening technique finds the optimal sharpening transform for a single test illuminant. In Section 6 we investigate the problem of finding a good sharpening transform relative to multiple illuminants. We show that if surface reflectances are three dimensional and illuminants two dimensional there exists a sharpening transform with respect to which a diagonal matrix supports perfect color constancy. This analysis constitutes a third technique for deriving the sharpening transform.

In Section 7 we relate our work specifically to theories of human color vision. Sharpened spectral sensitivities have been measured in humans.[12–16] We advance the hypothesis that sharpened sensor sensitivities arise as a natural consequence of optimization of the visual system's color-constancy abilities through an initial linear transformation of the cone outputs.

2. SENSOR-BASED SHARPENING

Sensor-based sharpening is a method of determining the sharpest sensor, given an s-dimensional (s is usually 3) sensor basis $\mathbf{R}(\lambda)$ and wavelength interval $[\lambda_1, \lambda_2]$. The sensor $\mathbf{R}^t(\lambda)\mathbf{c}$, where $\mathbf{c}$ is a coefficient vector, is maximally sensitive in $[\lambda_1, \lambda_2]$ if the percentage of its norm lying in this interval is maximal relative to all other sensors. We can solve for $\mathbf{c}$ by minimizing:

$$I = \int_\phi [\mathbf{R}(\lambda)^t\mathbf{c}]^2\mathrm{d}\lambda + \mu\left\{\int_\omega [\mathbf{R}(\lambda)^t\mathbf{c}]^2\mathrm{d}\lambda - 1\right\}, \quad (5)$$

where ω is the visible spectrum, ϕ denotes wavelengths outside the sharpening interval, and μ is a Lagrange multiplier. The Lagrange multiplier guarantees a nontrivial solution for Eq. (5): the norm of the sharpened sensor is equal to 1. Moreover, this constraint ensures that the same sharpened sensor is recovered independent of the initial norms of the basis set $\mathbf{R}(\lambda)$.

By differentiating with respect to $\mathbf{c}$ and equating to the zero vector, we find the stationary values of Eq. (5):

$$\frac{1}{2}\frac{\partial I}{\partial \mathbf{c}} = \int_{\phi} \mathbf{R}(\lambda)\mathbf{R}(\lambda)^t \mathbf{c}\, d\lambda + \mu\left[\int_{\omega} \mathbf{R}(\lambda)\mathbf{R}(\lambda)^t \mathbf{c}\, d\lambda\right] = \mathbf{0}. \tag{6}$$

Differentiating with respect to μ simply yields the constraint equation $\int_{\omega}[\mathbf{R}(\lambda)^t\mathbf{c}]^2 d\lambda = 1$. The solution of Eq. (6) can thus be carried out, assuming that the constraint holds.

Define the $s \times s$ matrix $\Lambda(\alpha) = \int_{\alpha} \mathbf{R}(\lambda)\mathbf{R}(\lambda)^t d\lambda$ so that Eq. (6) becomes

$$\Lambda(\phi)\mathbf{c} = -\mu\Lambda(\omega)\mathbf{c}. \tag{7}$$

Assuming a nontrivial solution $\mathbf{c} \neq 0$, $\mu \neq 0$ and rearranging Eq. (7), we see that solving for $\mathbf{c}$ (and consequently for the sharpened sensor) is an eigenvector problem:

$$[\Lambda(\omega)]^{-1}\Lambda(\phi)\mathbf{c} = -\mu\mathbf{c}. \tag{8}$$

There are s solutions of Eq. (8), each solution corresponding to a stationary value, so we choose the eigenvector that minimizes $\int_{\phi}[\mathbf{R}(\lambda)^t\mathbf{c}]^2 d\lambda$. It is important that $\mathbf{c}$ be a real-valued vector, as it implies that our sharpened sensor is a real-valued function. That $\mathbf{c}$ is real valued follows from the facts that the matrices $[\Lambda(\omega)]^{-1}$ and $\Lambda(\phi)$ are positive definite and that eigenvalues of the product of two positive-definite matrices are real and nonnegative.[17] Solving for $\mathbf{c}$ for each of three wavelength intervals yields a matrix $\mathcal{T}$ for use in relation (4). The matrix $\mathcal{T}$ is not dependent on any illuminant and denotes the sensor-based sharpening transform.

We sharpened two sets of sensor sensitivities: the cone absorptance functions measured by Bowmaker and Dartnell[18] (BOW) and the cone fundamentals derived by Vos and Walraven[10] (VW), which take into account the spectral absorptions of the eye's lens and macular pigment. The BOW functions were sharpened with respect to the wavelength intervals (in nanometers) [400, 480], [510, 550], and [580, 650] and the VW fundamentals with respect to the intervals [400, 480], [520, 560], and [580, 650]. (All spectra used in this paper are in the range 400–650 nm measured at 10-nm intervals.) These intervals were chosen to ensure that the whole visible spectrum would be sampled and that the peak sensitivities of the resulting sensors would roughly correlate with those of the cones.

The results for the VW sensors are presented in Fig. 1 (those for the BOW sensors are similar), where it can be seen that the sharpened curves contain negative sensitivities. These need not cause concern in that they do not represent negative physical sensitivities but merely represent negative coefficients in a computational mechanism. Clearly, the sharpening intervals are somewhat arbitrary. They were chosen simply because the resulting sharpened sensors appeared, to the human eye, significantly sharper than the unsharpened sensors. Nevertheless, the intervals used are sensible, and their suitability is verified by the fact that they provide sharpened sensors that are in close agreement with those derived by data-based sharpening as described in Section 3. The actual values of the $\mathbf{c}$'s in Eq. (8) are given in Section 6 below.

Figure 1 contrasts the sharpened VW sensor set with the corresponding unsharpened set; the degree of sharpening is quite significant. The peak sensitivities of the new sensors are shifted with respect to the initial sensitivities; this shift is due to both the choice of sharpening intervals and the shape of the VW sensitivities. The sharpened long-wavelength mechanism is pushed further toward the long-wavelength end of the spectrum; in contrast, the medium-wavelength mechanism is shifted toward the shorter wavelengths; and the short-wavelength mechanism remains essentially the same. Intriguingly, field sensitivities of the human eye measured under white-light conditioning with long test flashes[13] are sharpened in an analogous manner.

Table 1 contrasts the percentage-squared norm lying in the sharpening intervals for the original versus the sharpened curves. For both the VW and the BOW sensors the degree of sharpening is significant. Furthermore, from

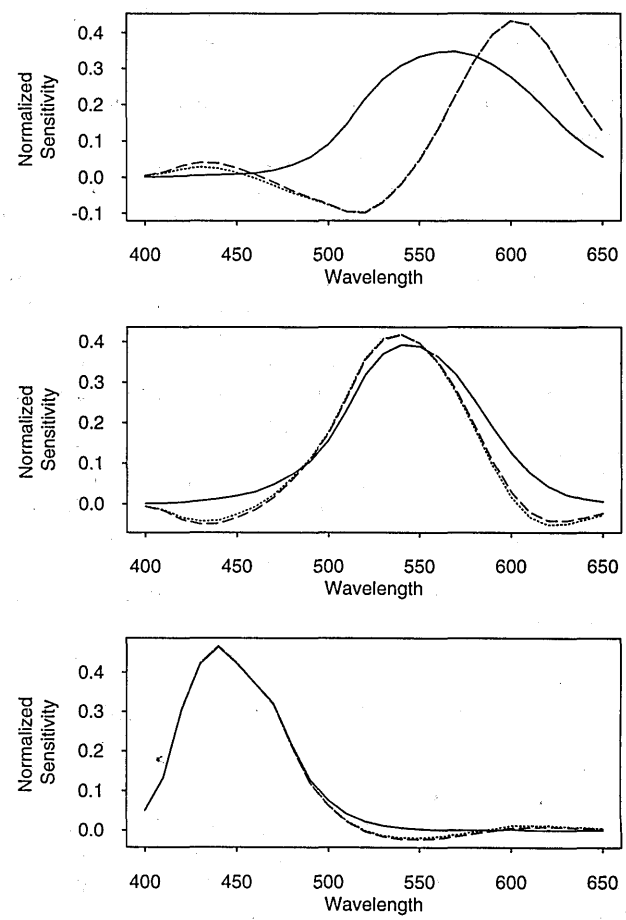

Fig. 1. VW fundamentals (solid curves) are contrasted with the sharpened sensitivities derived through sensor-based (dotted curves) and data-based (dashed curves) sharpening. Top, long-wavelength mechanism; middle, medium-wavelength mechanism; bottom, short-wavelength mechanism.

Table 1. Percentage of Total Squared Norm in the Sharpening Intervals

Sensors	% Squared Norm [400, 480]	[510, 550]	[580, 650]
BOW, original	98.9	51.4	30.1
BOW, sharpened	99.4	66.3	76.2
		[520, 560]	
VW, original	97.5	67.6	40.3
VW, sharpened	97.8	74.7	89.1

Fig. 1 it is clear that the sharpening effect is not limited to the sharpened interval.

3. DATA-BASED SHARPENING

It could be the case that the best sensors for DMT algorithms might vary substantially with the type of illumination change being modeled. If so, sensor-based sharpening, which does not take into account any of the statistical properties of collections of surfaces and illuminants, might perform well in some cases and poorly in others. To test whether radically different sharpening transformations are required in different circumstances, we explore a data-based approach to deriving sharpened sensitivities, in which we optimize the sensors for DMT algorithms by examining the relationship between observations, obtained under different test illuminants, and their corresponding descriptors.

Let W^c be a $3 \times n$ matrix of descriptors generated from a set n of surfaces observed under a canonical illuminant E^c. Similarly, let W^e be the matrix of observations of n surfaces imaged under another test illuminant E^e. To the extent that DMT-based algorithms suffice for color constancy, W^c and W^e should be approximately equivalent under a DMT:

$$W^c \approx \mathcal{D}^e W^e . \tag{9}$$

The diagonal transform is assumed to be an approximate mapping and will have a certain degree of error. The idea of spectral sharpening is that this approximation error can be reduced if W^c and W^e are first transformed by a matrix $\mathcal{T}^e$;

$$\mathcal{T}^e W^c \approx \mathcal{D}^e \mathcal{T}^e W^e . \tag{10}$$

$\mathcal{D}^e$ will in fact be optimal in the least-squares sense if it is defined by the Moore–Penrose inverse:

$$\mathcal{D}^e = \mathcal{T}^e W^c [\mathcal{T}^e W^e]^+ , \tag{11}$$

where + denotes the Moore–Penrose inverse. (The Moore–Penrose inverse of the matrix A is defined as $A^+ = A^t[AA^t]^{-1}$.) Now $[\mathcal{T}^e]$ must be chosen to ensure that $[\mathcal{D}^e]$ is diagonal. To see how to do this, carry out some matrix manipulation to obtain

$$[\mathcal{T}^e]^{-1} \mathcal{D}^e \mathcal{T}^e = W^c [W^e]^+ . \tag{12}$$

Since the eigenvector decomposition of the matrix on the right-hand side of Eq. (12), $W^c[W^e]^+ = \mathcal{U}^e \mathcal{D}^e [\mathcal{U}^e]^{-1}$, is unique, its similarity to the left-hand side implies that $\mathcal{T}^e$ is unique also, always exists, and equals $[\mathcal{U}^e]^{-1}$, for diagonal $[\mathcal{D}^e]$.

It is interesting to compare Eq. (12) with the relation for the problem of finding the best general transform $\mathcal{G}^e$ that maps observations obtained under a test illuminant to their corresponding descriptors:

$$W^c \approx \mathcal{G}^e W^e . \tag{13}$$

Such optimal fitting effectively bounds the possible performance within a linear model of color constancy. When the approximation of relation (13) is to be optimized in the least-squares sense, $\mathcal{G}^e$ is simply

$$\mathcal{G}^e = W^c [W^e]^+ . \tag{14}$$

Equation (12) can be interpreted, therefore, as simply the eigenvector decomposition of the optimal general transform $\mathcal{G}^e$. Of course, it is obvious that the optimal transform could be diagonalized; what it is important is that if one knows a sharpening transform $[\mathcal{T}^e]$, then the best least-squares solution relating $\mathcal{T}^e W^c$ and $\mathcal{T}^e W^e$ is precisely the diagonal transformation $\mathcal{D}^e$; that is, $\mathcal{T}^e W^c [\mathcal{T}^e W^e]^+ = \mathcal{D}^e$. In other words, when one is using the sharpened sensors, the optimal transform is guaranteed to be diagonal, so finding the best diagonal transform after sharpening is equivalent to finding the optimal general transform. Therefore sharpening allows us to replace the problem of determining the nine parameters of $\mathcal{G}^e$ with that of determining only the three parameters of $\mathcal{D}^e$.

Data-based sharpening raises two main questions: (1) Will the resulting sensors be similar to those obtained by sensor-based sharpening, and (2) will the derived sensors vary substantially with the illuminant? To answer these questions requires the application of data-based sharpening to response vectors obtained under a single canonical illuminant and several test illuminants. For illuminants we used five Judd daylight spectra[19] and CIE standard illuminant A,[10] and as reflectances we used the set of 462 Munsell spectra.[20] We arbitrarily chose Judd's D55 (55 stands for 5500 K) as the canonical illuminant; descriptors are the response vectors for surfaces viewed under D55. For each of the other illuminants E^e, we derived the data-based sharpening transform $\mathcal{T}^e$ in accordance with Eq. (12).

Figure 2 shows for the VW sensors the range of the five sets of data-based sharpened sensors obtained for mapping between each of the five illuminants and D55. For these five illuminants the results are remarkably stable and hence relatively independent of the particular illuminant, so the mean of these sharpened sensors characterizes the set of them quite well. Referring once again to Fig. 1, we can see that these mean sensors are very similar to those derived through sensor-based sharpening.

The stability of the results for the five cases and the similarity of these results to the sensor-based results is reassuring; nonetheless, it would be nice to find the sharpening transformation that optimizes over all the illuminants simultaneously. This issue is addressed in Section 6, where we show that, given illuminant and reflectance spectra that are two and three dimensional,

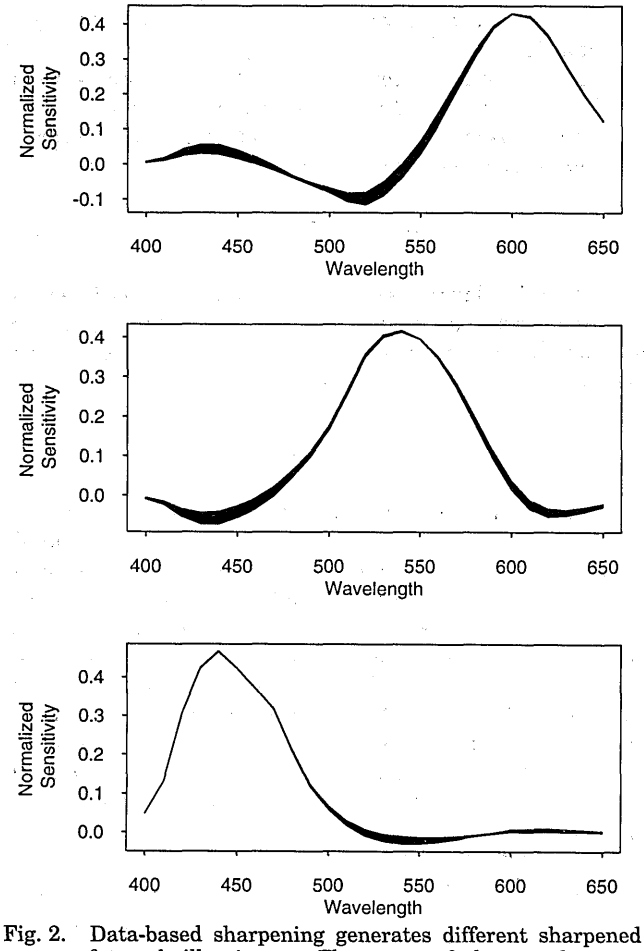

Fig. 2. Data-based sharpening generates different sharpened sensors for each illuminant. The range of sharpened curves over all the test illuminants (CIE A, D48, D65, D75, and D100) mapped to D55 is shown for the VW cone mechanisms. Top, long-wavelength mechanism; middle, medium-wavelength mechanism; bottom, short-wavelength mechanism.

respectively, there exists a unique optimal sharpening transform.

4. EVALUATING SHARPENED SENSORS

Since the sensor-based and data-based sharpened sensors are similar, we restrict our further attention to the evaluation of sensor-based sharpened sensors alone. The results for VW and BOW sensors are similar as well, so we include figures only for the VW case.

For each illuminant we generated sensor responses for our 462 test surface reflectances, using both the sharpened and the unsharpened VW sensors so that we could determine how much sharpening improves DMT performance.

To measure the performance of a DMT in mapping test observations to canonical descriptors, we compare fitted observations (observations mapped to the canonical illuminant by means of a diagonal matrix) with corresponding (canonical) descriptors. The Euclidean distance between a fitted observation $\mathbf{q}^e$ and its descriptor $\mathbf{p}^c$, normalized with respect to the descriptor's length, provides a good error metric, given the definition of color constancy that we are using. This percent normalized fitted distance (NFD) metric is defined as

$$\mathrm{NFD} = 100 * \frac{\|\mathbf{p}^c - \mathbf{q}^e\|}{\|\mathbf{p}^c\|}. \tag{15}$$

Let W^c be a 3×462 matrix of descriptors corresponding to the 462 surfaces viewed under the canonical illuminant. Similarly, let W^e denote the 3×462 matrix of observations for the 462 surfaces viewed under a test illuminant. Relation (9) can then be solved to yield the best diagonal transformation in the least-squares sense; this procedure will be called simple diagonal fitting. Since $\mathcal{D}^e$ is a diagonal matrix, each row of W^e is fitted independently. The components of $\mathcal{D}^e$ are derived as follows:

$$\mathcal{D}_{ii}{}^e = W_i{}^c[W_i{}^e]^+ = W_i{}^c[W_i{}^e]^t / W_i{}^e[W_i{}^e]^t, \tag{16}$$

where the single subscript i denotes the ith matrix row and the double subscript ii denotes the matrix element at row i column i.

Given a fixed set of sensor functions, Eq. (16) yields the best diagonal transformation that takes observations under the test illuminant onto their corresponding descriptors. Simple diagonal fitting, therefore, does not include sharpening but rather for a fixed set of sensors finds the best diagonal matrix $\mathcal{D}^e$ for that set of sensors.

For performance comparisons we are also interested in the NFD that results under transformed diagonal fitting. Transformed diagonal fitting proceeds in two stages:

1. $\mathcal{T}W^c \approx \mathcal{D}^e\mathcal{T}W^e$, where $\mathcal{T}$ is the fixed sharpening transform and $\mathcal{D}^e$ is calculated by means of Eq. (16);
2. $\mathcal{T}^{-1}\mathcal{T}W^c \approx \mathcal{T}^{-1}\mathcal{D}^e\mathcal{T}W^e$.

Application of $\mathcal{T}^{-1}$ transforms the fitted observations back to the original (unsharpened) sensor set so that an appropriate comparison can be made between the fitting errors: the performance of sharpened diagonal-matrix constancy is measured relative to the original unsharpened sensors.

Figure 3 shows NFD cumulative histograms for diagonal fitting of VW observations (solid curves) and diagonal fitting of sharpened VW observations (dotted curves). For each illuminant the sharpened sensors show better performance than the unsharpened ones, as indicated by the fact that in the cumulative NFD histograms the sharpened sensor values are always above those for the unsharpened ones. In general the performance difference increases as the illuminant color becomes more extreme: from D55 to D100 the illuminants become progressively bluer and toward CIE A, redder.

Figure 4 shows the cumulative NFD histograms for diagonal fitting of VW observations (solid curves) and transformed diagonal fitting of sharpened VW observations (dashed curves). Once again it is clear that the sharpened sensors perform better; however, the performance difference is greater, which shows that the question of sensor performance is linked to the axes in which color space is described.

Finally, Fig. 5 contrasts the cumulative NFD histograms for optimal fitting of VW observations [the unrestricted least-squares fit of relation (13) (solid curves)] and transformed diagonal fitting of sensor-based sharp-

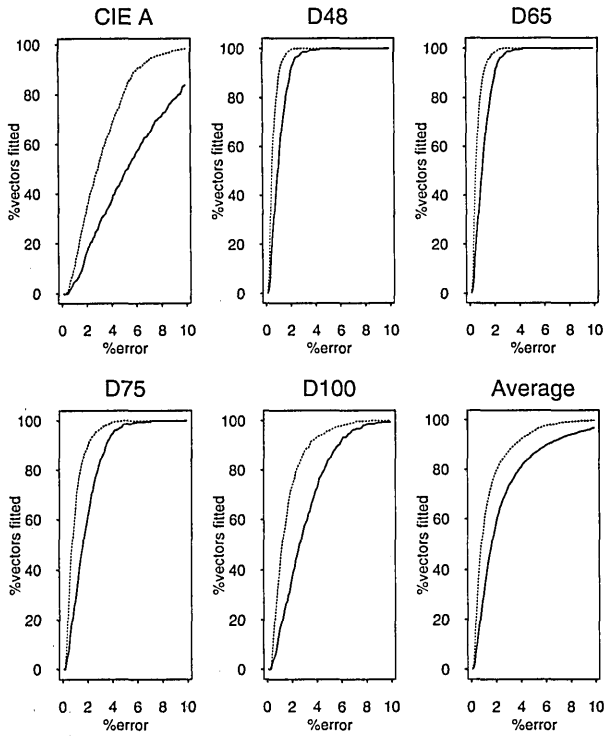

Fig. 3. Cumulative NFD histogram obtained with each test illuminant (CIE A, D48, D65, D75, and D100) for diagonal fitting of VW observations (solid curves) and diagonal fitting of sensor-based sharpened VW observations (dotted curves). The sixth cumulative NFD histogram shows the average fitting performance.

ened VW observations (dashed curves). For these cases, with the exception of CIE A, a DMT achieves almost the same level of performance as the best nondiagonal transform.

5. DATA-BASED SHARPENING AND VOLUMETRIC THEORY

Data-based sharpening is a useful tool for validating our choice of sensor-based sharpened sensors. However, more than this, data-based sharpening can also be viewed as a generalization of Brill's[9] volumetric theory of color constancy. In that theory Brill develops a method for generating illuminant-invariant descriptors that is based on two key assumptions:

1. Surface reflectances are well modeled by a three-dimensional basis set and are thus defined by a surface weight vector $\boldsymbol{\sigma}$. For example, $S(\lambda) = \sum_{i=1}^{3} S_i(\lambda)\sigma_i$.
2. Each image contains three known reference reflectances. In the discussion that follows, $\mathcal{Q}^e$ denotes the 3×3 matrix of observations for the reference patches seen under $E^e(\lambda)$.

Given the first assumption, observations for surfaces viewed under $E^e(\lambda)$ are generated by application of a lighting matrix (a term first used by Maloney[21]) to surface weight vectors:

$$\mathbf{p}^e = \Lambda^e \boldsymbol{\sigma}, \tag{17}$$

where the ijth entry of Λ^e is equal to $\int R_i(\lambda)E^e(\lambda)S_j(\lambda)\mathrm{d}\lambda$.

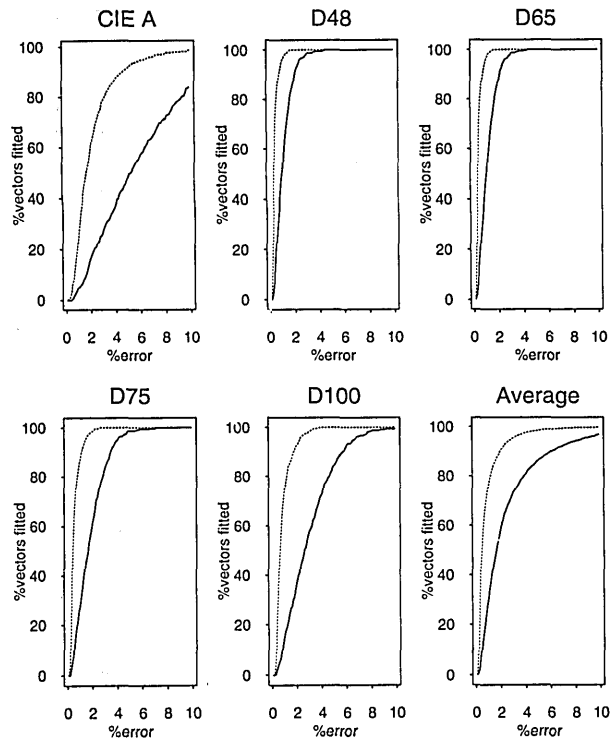

Fig. 4. Cumulative NFD histogram obtained with each test illuminant (CIE A, D48, D65, D75, and D100) for diagonal fitting of VW observations (solid curves) and transformed diagonal fitting of sensor-based sharpened VW observations (dotted curves). The sixth cumulative NFD histogram shows the average fitting performance.

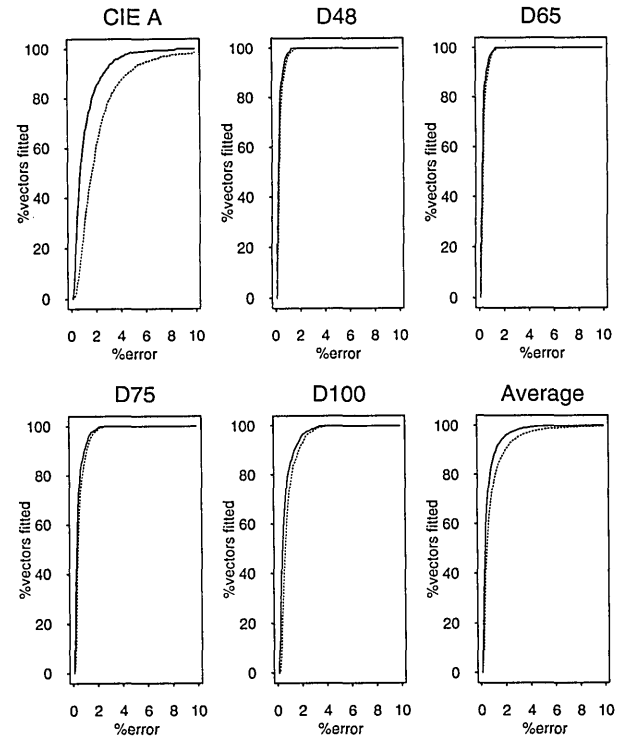

Fig. 5. Cumulative NFD histogram obtained with each test illuminant (CIE A, D48, D65, D75, and D100) for optimal fitting of VW observations (solid curves) and transformed diagonal fitting of sensor-based sharpened VW observations (dotted curves). The sixth cumulative NFD histogram shows the average fitting performance.

It follows immediately that $\mathcal{Q}^e$ is a fixed linear transform of the lighting matrix:

$$\mathcal{Q}^e = \Lambda^e \mathcal{A}, \tag{18}$$

where the columns of $\mathcal{A}$ correspond to the surface weight vectors of the reference reflectances and are independent of the illuminant. Now, given any arbitrary response vector $\mathbf{p}^e$, one can easily generate an illuminant-invariant descriptor by premultiplying with $[\mathcal{Q}^e]^{-1}$:

$$[\mathcal{Q}^e]^{-1}\mathbf{p}^e = \mathcal{A}^{-1}[\Lambda^e]^{-1}\Lambda^e\boldsymbol{\sigma} = \mathcal{A}^{-1}\boldsymbol{\sigma}. \tag{19}$$

The color-constancy performance of Brill's volumetric theory is linked directly to the dimensionality of surface reflectances. Real reflectance spectra are generally not three dimensional (Maloney[21] suggests that a basis set of between three and six functions is required), leading to inaccuracies in the calculated descriptors.

Data-based sharpening, like volumetric theory, aims to generate illuminant-invariant descriptors by applying a linear transform. If we impose the extremely strong constraint that all surface reflectances appear in each image, then data-based sharpening can be viewed as an algorithm for color constancy. As a color-constancy algorithm, data-based sharpening has a distinct advantage over volumetric theory in that it is optimal with respect to the least-squares criterion and consequently is guaranteed to outperform the volumetric method. Unfortunately, this performance increase is gained at the expense of the extremely strong requirement that all surface reflectances appear in each image.

In practice we can weaken this constraint and assume that there are k known reference patches per image, where k is small. Novak and Shafer[22] develop a similar theory called supervised color constancy, which is based on the assumption that there are 24 known reference patches in each image; however, unlike data-based sharpening, their constancy transform is derived by examination of the relationship between measured responses and finite-dimensional models of reflectance and illumination. Certainly for the reference patches themselves the data-based sharpening method will outperform Novak's supervised color constancy since, for these patches, data-based sharpening finds the optimal least-squares transform. However, further study is required for comparing overall color-constancy performance. In the limiting case, where $k = 3$, data-based sharpening reduces to the volumetric theory.

Volumetric theory requires three reference patches in order to recover the nine parameters of $[\mathcal{Q}^e]^{-1}$ and thereby achieve color constancy. As shown by the performance tests of the preceding section, however, after a fixed sharpening transformation of DMT models illumination change almost as well as does a nondiagonal matrix. Since only three parameters instead of nine need to be determined for specification of the diagonal matrix when sharpened sensors are being used, only one reference patch is required instead of three for achieving color constancy. This follows because a single response vector seen under a test illuminant $E^e(\lambda)$ can be mapped to its canonical appearance by a single diagonal matrix:

$$\mathbf{p}^c = \mathcal{D}^e\mathbf{p}^e, \tag{20}$$

$$\mathcal{D}_{ii}{}^e = \frac{p_i^c}{p_i^e}. \tag{21}$$

If we choose our reference patch to be a white reflectance, then, through sharpening, volumetric theory reduces to Land's white-patch retinex.[11] Similarly, West and Brill[2] consider white-patch normalization to be consistent with von Kries adaptation.

We performed a simulation, called transformed white-patch normalization, to evaluate the quality of color constancy obtainable with use of a single reference patch. For each illumination (CIE A, D48, D55, D65, D75, and D100) we

1. generated a matrix W^e of observations of Munsell patches for VW sensors;
2. transformed observations to the (sensor-based) sharpened sensors, $\mathcal{T}W^e$;
3. calculated $\mathcal{D}^e\mathcal{T}W^e$, where $\mathcal{D}_{ii}{}^e = 1/p_i{}^w$ (the reciprocal of the ith sharpened sensor's response to the white patch; the Munsell reflectance closest to the uniform white, in the least-squares sense, was chosen as the white reference patch);
4. transformed white-patch normalized observations back to VW sensors, $\mathcal{T}^{-1}\mathcal{D}^e\mathcal{T}W^e$.

Again D55 was the canonical light. Thus we evaluated constancy by calculating the NFD between the white-patch-corrected observations seen under D55 (the descriptors) and the white-patch-corrected observations under each other illuminant. In Fig. 6 we contrast the cumulative NFD histograms for white-patch normalization (dotted curves) with those for the optimal fitting performance (solid curves). White-patch normalization yields very good constancy results that are generally comparable with the optimal fitting performance.

6. SHARPENING RELATIVE TO MULTIPLE ILLUMINANTS

Data-based sharpening was introduced primarily to validate the idea of sensor-based sharpening and to ensure that our particular choice of sharpening parameters led to reasonable results. Figure 2 shows that the optimal sensors, as determined by means of data-based sharpening for each illuminant, closely resemble one another and furthermore that they resemble the unique set of sensor-based sharpened sensors as well. Although all the sensors are similar, the question remains regarding whether there might be an optimal sharpening transform for the entire illuminant set.

In Ref. 23 we derive conditions for perfect DMT color constancy, using sharpened sensors. Because the sharpening transform applies to a whole space of illuminants, it is in essence a type of global data-based sharpening.

The theoretical result is based on finite-dimensional approximations of surface reflectance and illumination, and what is shown is that, if surface reflectances are well modeled by three basis functions and illuminants by two basis functions, then there exists a set of new

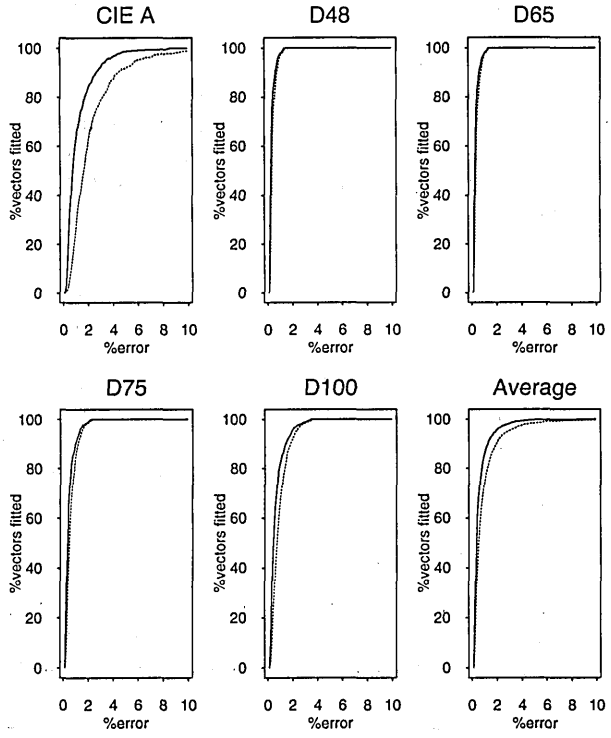

Fig. 6. Cumulative NFD histogram obtained with each test illuminant (CIE A, D48, D65, D75, and D100) for optimal fitting of VW observations (solid curves) and transformed white-patch normalization of VW observations (dotted curves). The sixth cumulative NFD histogram shows average color constancy performance.

sensors for which a DMT can yield perfect color constancy. These restrictions are quite strong; nonetheless, statistical studies have shown that a three-dimensional basis set provides a fair approximation to real surface reflectance[21] and that a two-dimensional basis set describes daylight illumination[19] reasonably well. Moreover, Marimont and Wandell[24] have developed a method for deriving basis functions that is dependent on how reflectance, illuminant, and sensors interact to form sensor responses [i.e., Eq. (2) above is at the heart of their method]. A three-dimensional model of reflectance and a two-dimensional model of illumination is shown to provide a good model of actual response vectors.

Given these dimensionality restrictions on reflectance and illumination, cone response vectors of surfaces viewed under a canonical illuminant, that is descriptors, can be written as

$$\mathbf{p}^c = \Lambda^c \boldsymbol{\sigma}, \tag{22}$$

where the superscript c denotes the canonical illuminant. Since illumination is two dimensional, there is necessarily a second illuminant $E^2(\lambda)$ linearly independent with $E^c(\lambda)$ (together they form the span). Associated with this second illuminant is a second linearly independent lighting matrix Λ^2. It follows immediately that Λ^2 is some linear transform $\mathcal{M}$ away from Λ^c:

$$\Lambda^2 = \mathcal{M}\Lambda^c, \tag{23}$$

$$\mathcal{M} = \Lambda^2[\Lambda^c]^{-1}. \tag{24}$$

Since every illuminant is a linear combination of $E^c(\lambda)$ and $E^2(\lambda)$, lighting matrices in turn are linear combinations of Λ^c and $\mathcal{M}\Lambda^c$, as a result of Eq. (24). Consequently an observation vector obtained for any surface under an illuminant $E^e(\lambda) = \alpha E^c(\lambda) + \beta E^2(\lambda)$ can be expressed as a linear transform of its descriptor vector:

$$\mathbf{p}^e = [\alpha I + \beta\mathcal{M}]\Lambda^c\boldsymbol{\sigma} = [\alpha I + \beta\mathcal{M}]\mathbf{p}^c, \tag{25}$$

where I is the identity matrix. Calculating the eigenvector decomposition of $\mathcal{M}$,

$$\mathcal{M} = \mathcal{T}^{-1}\mathcal{D}\mathcal{T}, \tag{26}$$

and expressing the identity matrix I in terms of $\mathcal{T}$,

$$I = \mathcal{T}^{-1}I\mathcal{T}, \tag{27}$$

enables us to rewrite the relationship between an observation and a descriptor, Eq. (25), as a diagonal transform:

$$\mathcal{T}\mathbf{p}^e = [\alpha I + \beta\mathcal{D}]\mathcal{T}\mathbf{p}^c. \tag{28}$$

Writing $\mathbf{p}^c$ in terms of $\mathbf{p}^e$ leads directly to

$$\mathcal{T}\mathbf{p}^c = [\alpha I + \beta\mathcal{D}]^{-1}\mathcal{T}\mathbf{p}^e. \tag{29}$$

The import of Eq. (29) is that when the appropriate initial sharpening transformation $\mathcal{T}$ is applied, a diagonal transform supports perfect color constancy, subject of course to the restrictions imposed on illumination and reflectance.

These restrictions compare favorably with those employed by D'Zmura,[25] who showed that, given three-dimensional reflectances and three-dimensional illuminants, one can obtain perfect color constancy, given two images of three color patches under two different illuminants and using a nondiagonal transform.

From the Munsell reflectance spectra and our six test illuminants, we used principal-component analysis to derive the basis vectors for reflectance and illumination. Using these vectors, we constructed lighting matrices and then calculated the sharpening transform by means of Eq. (26). The formulas for the perfect sharpened sensors are given in Eqs. (30)–(32), where they are contrasted with the corresponding formulas obtained with respect to sensor-based and data-based sharpening. The symbols R, G, and B denote the VW red, green, and blue cone mechanisms, respectively, scaled to have unit norms, and the superscripts p, s, and d denote perfect, sensor-based, and data-based sharpening, respectively:

$$\begin{aligned} R^p &= 2.44R - 1.93G + 0.110B, \\ R^s &= 2.46R - 1.97G + 0.075B, \\ R^d &= 2.46R - 1.98G + 0.100B, \end{aligned} \tag{30}$$

$$\begin{aligned} G^p &= 1.55G - 0.63R - 0.16B, \\ G^s &= 1.58G - 0.66R - 0.12B, \\ G^d &= 1.52G - 0.58R - 0.14B, \end{aligned} \tag{31}$$

$$\begin{aligned} B^p &= 1.0B - 0.13G + 0.08R, \\ B^s &= 1.0B - 0.14G + 0.09R, \\ B^d &= 1.0B - 0.13G + 0.07R. \end{aligned} \tag{32}$$

It is reassuring that the perfect sharpened sensors are almost identical to those derived through sensor-based and data-based sharpening. Therefore, even though neither the sensor-based sharpened sensors nor the data-based sharpened ones are optimized relative to a whole set of illuminants, sharpening in all cases generates sensors that are similar to those that work perfectly for a large, although restricted, class of illuminants. This theoretical result provides strong support for the hypothesis, already confirmed in part by the consistency of the data-based sharpening results, that a single sharpening transformation will work well for a reasonable range of illuminants.

7. SPECTRAL SHARPENING AND THE HUMAN VISUAL SYSTEM

If the human visual system employs a DMT for color constancy, our results show that we should expect it to use sharpened sensors, since doing so would optimize its performance. In this section we briefly examine some of the psychophysical evidence for sharpened sensitivities.

Sharpened sensitivities have been detected in both field and test spectral sensitivity experiments (for a review of these terms see Foster[26]). Sperling and Harwerth[13] measured the test spectral sensitivities of human subjects conditioned to a large white background and found, consistent with our findings for sharpened sensors, sharpened peaks at 530 and 610 nm, with no sharpening of the blue mechanism.

These authors propose that a linear combination of the cone responses accounts for the sharpening. They found that the sharpened red sensor can be modeled as the red cone minus a fraction of the green and that the sharpened green sensor can be modeled as the green cone minus a fraction of the red. This finding corresponds well with our theoretical results in that our sharpening transformations, Eqs. (30) and (31), basically involve red minus green and green minus red, with only a slight contribution from the blue.

More recently, Foster[12] observed that field and test spectral sensitivities show sharpened peaks when they are derived in the presence of a small monochromatic auxiliary field coincident with the test field. Foster and Snelgar[27] extended this work by performing a hybrid experiment with a white, spatially coincident auxiliary field, and sharpened sensitivities again were found. In both cases these experimentally determined, sharpened sensitivities agree with our theoretical results. Foster and Snelgar[28] verified, as did Sperling and Harwerth,[13] that the sharpened sensitivities were a linear combination of the cone sensitivities.

Krastel and Braun[29] measured spectral field sensitivities under changing illumination when, as in the experiments by Sperling and Harwerth,[13] a white conditioning field is employed. Illumination color was changed by placement of colored filters in front of the eye. The same test spectral-sensitivity curve results under both a reddish and a bluish illuminant, suggesting that the eye's sharpened mechanisms are unaffected by illumination. More recently Kalloniatis and Harwerth[14] measured cone spectral sensitivities under white adapting fields of different intensity and found the sharpened sensitivities to be independent of the intensity of the adapting field.

Poirson and Wandell[16] developed techniques for measuring the spectral sensitivity of the eye with respect to the task of color discrimination. For color discrimination among briefly presented targets, the spectral sensitivity curve has relatively sharp peaks at 530 and 610 nm.

Although the general correspondence between our sharpened sensors and the above psychophysical results does not imply that sharpening in humans exists for the purpose of color constancy, at least the evidence that a linear combination of the cone responses is employed somewhere in the visual system lends plausibility to the idea that sharpening might be used in human color-constancy processing. On the other hand, our finding that sharpening could improve the performance of some color-constancy methods suggests an explanation for the existence of spectral sharpening in humans.

8. CONCLUSION

Spectral sharpening generates sensors that improve the performance of methods based on color constancy theories (e.g., von Kries adaptation, Land's retinex) that use

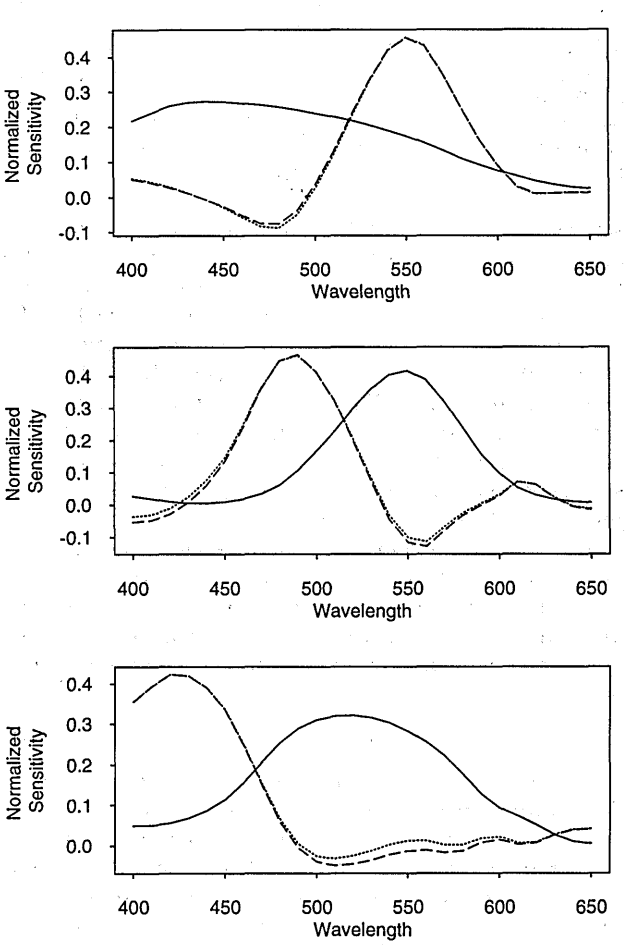

Fig. 7. Solid curves show results for the Kodak Wratten filters #66, #52, and #38; dotted curves show the results of sensor-based sharpening; dashed curves show the mean of the data-based sharpened sensors obtained for the five test illuminants (CIE A, D48, D65, D75, and D100). Top, long-wavelength mechanism; middle, medium-wavelength mechanism; bottom, short-wavelength mechanism.

diagonal-matrix transformations. Data-based sharpening finds sensors that are optimal with respect to a given set of surface reflectances and illuminants. Sensor-based sharpening finds the most-narrow-band sensors that can be created as a linear combination of a given set of sensors. Finally, for restricted classes of illuminants and reflectance (they are constrained to be two and three dimensional, respectively) we have shown that there exists a sharpening transform with respect to which a diagonal matrix will support perfect color constancy. The sharpening transform derived by means of this analysis is in close agreement with the sensor-based sharpening and data-based sharpening transforms. In all three cases sharpened sensors substantially improve the accuracy with which a DMT can model changes in illumination. The sensor-based and the data-based sharpening techniques are quite general and can be applied to visual systems that have more than three sensors.

As with the cone sensitivities, sharpening of a color camera's sensitivities can also have a significant effect. Use of overlapping, broadband filters such as Wratten #66, #52, and #38 (Ref. 30) could be advantageous, since from an exposure standpoint they filter out less light, but their use could be disadvantageous from a color-constancy standpoint. Sharpening such filters, as shown in Fig. 7, can provide a good compromise among the competing requirements.

Spectral sharpening is not, in itself, a theory of color constancy, in that it makes no statement about how to choose the coefficients of the diagonal matrix. Instead, we propose sharpening as a mechanism for improving the theoretical performance of DMT algorithms of color constancy regardless of how any particular algorithm might calculate the diagonal-matrix coefficients to be used in adjustment for an illumination change.

Since the performance of DMT algorithms improves significantly when sharpened sensitivities are employed, and, furthermore, since it then compares favorably with the performance of the best possible nondiagonal transform, our results suggest that if a linear transform is a central mechanism of human color constancy, then after an appropriate, fixed sharpening transformation of the sensors there is little to be gained through the use of anything more general than a DMT.

In general, our results lend support to DMT-based theories of color constancy. In addition, since spectral sharpening aids color constancy, we have advanced the hypothesis that sharpening might provide an explanation for the psychophysical finding of sharpening in the human visual system.

ACKNOWLEDGMENTS

This research has been supported by the Simon Fraser University Center for Systems Science and the Natural Sciences and Engineering Research Council of Canada under grant 4322.

REFERENCES

1. J. Beck, *Surface Color Perception* (Cornell U. Press, Ithaca, N.Y., 1972).
2. G. West and M. H. Brill, "Necessary and sufficient conditions for von Kries chromatic adaption to give colour constancy," J. Math. Biol. **15,** 249–258 (1982).
3. E. H. Land and J. J. McCann, "Lightness and retinex theory," J. Opt. Soc. Am. **61,** 1–11 (1971).
4. B. K. P. Horn, "Determining lightness from an image," Comput. Vision Graphics Image Process. **3,** 277–299 (1974).
5. A. Blake, "Boundary conditions for lightness computation in Mondrian world," Comput. Vision Graphics Image Process. **32,** 314–327 (1985).
6. D. Forsyth, "A novel algorithm for color constancy," Int. J. Comput. Vision **5,** 5–36 (1990).
7. M. D'Zmura and P. Lennie, "Mechanisms of color constancy," J. Opt. Soc. Am. A **3,** 1662–1672 (1986).
8. B. V. Funt and M. S. Drew, "Color constancy computation in near-Mondrian scenes using a finite dimensional linear model," in *Computer Vision and Pattern Recognition Proceedings* (Institute of Electrical and Electronics Engineers, New York, 1988), pp. 544–549.
9. M. H. Brill, "Computer-simulated object-color recognizer," Tech. Rep. 122 (MIT Research Laboratory of Electronics, Cambridge, Mass., 1980).
10. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, 2nd ed. (Wiley, New York, 1982).
11. E. H. Land, "The retinex theory of color vision," Sci. Am. **237,** 108–129 (1977).
12. D. H. Foster, "Changes in field spectral sensitivities of red-, green- and blue-sensitive colour mechanisms obtained on small background fields," Vision Res. **21,** 1433–1455 (1981).
13. H. G. Sperling and R. S. Harwerth, "Red–green cone interactions in the increment-threshold spectral sensitivity of primates," Science **172,** 180–184 (1971).
14. M. Kalloniatis and R. S. Harwerth, "Spectral sensitivity and adaptation characteristics of cone mechanisms under white-light adaptation," J. Opt. Soc. Am. A **7,** 1912–1928 (1990).
15. D. C. Hood and M. A. Finkelstein, "A case for the revision of textbook models of colour vision," in *Colour Vision: Physiology and Psychophysics*, J. D. Mollon and L. T. Sharpe, eds. (Academic, New York, 1983), pp. 385–398.
16. A. B. Poirson and B. A. Wandell, "Task-dependent color discrimination," J. Opt. Soc. Am. A **7,** 776–782 (1990).
17. J. H. Wilkinson, *Algebraic Eigenvalue Problem*, Monographs on Numerical Analysis (Oxford U. Press, Oxford, 1965).
18. J. K. Bowmaker and H. J. A. Dartnell, "Visual pigments of rods and cones in the human retina," J. Physiol. **298,** 501–511 (1980).
19. D. B. Judd, D. L. MacAdam, and G. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," J. Opt. Soc. Am. **54,** 1031–1040 (1964).
20. J. Cohen, "Dependency of the spectral reflectance curves of the Munsell color chips," Psychon. Sci. **1,** 369–370 (1964).
21. L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," J. Opt. Soc. Am. A **3,** 1673–1683 (1986).
22. C. L. Novak and S. A. Shafer, "Supervised color constancy using a color chart," Tech. Rep. CMU-CS-90-140 (Carnegie–Mellon University School of Computer Science, Pittsburgh, Pa., 1990).
23. G. D. Finlayson, M. S. Drew, and B. V. Funt, "Enhancing von Kries adaptation via sensor transformations," in *Human Vision, Visual Processing, and Digital Display IV*, J. P. Allebach and Bernice E. Roqowitz, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **1913,** 473–484 (1993).
24. D. H. Marimont and B. A. Wandell, "Linear models of surface and illuminant spectra," J. Opt. Soc. Am. A **9,** 1905–1913 (1992).
25. M. D'Zmura, "Color constancy: surface color from changing illumination," J. Opt. Soc. Am. A **9,** 490–493 (1992).
26. D. H. Foster, "Colour vision," Contemp. Phys. **25,** 477–497 (1984).
27. D. H. Foster and R. S. Snelgar, "Test and field spectral sensitivities of colour mechanisms obtained on small white backgrounds: action of unitary opponent-colour processes?" Vision Res. **23,** 787–797 (1983).

28. D. H. Foster and R. S. Snelgar, "Initial analysis of opponent-colour interactions revealed in sharpened field sensitivities," in *Colour Vision: Physiology and Psychophysics*, J. D. Mollon and L. T. Sharpe, eds. (Academic, New York, 1983), pp. 303–312.
29. W. Jaeger, H. Krastel, and S. Braun, "An increment-threshold evaluation of mechanisms underlying colour constancy," in *Colour Vision: Physiology and Psychophysics*, J. D. Mollon and L. T. Sharpe, eds. (Academic, New York, 1983), pp. 545–552.
30. *Kodak Filters: for Scientific and Technical Uses*, 2nd ed. (Eastman Kodak, Rochester, N.Y., 1981).