

Spectral sharpening with positivity

Mark S. Drew

School of Computing Science, Simon Fraser University, Vancouver, British Columbia, Canada V5A 1S6

Graham D. Finlayson

School of Information Systems, The University of East Anglia, Norwich, England NR4 7TJ

Received May 19, 1999; revised manuscript received January 19, 2000; accepted April 11, 2000

Spectral sharpening is a method for developing camera or other optical-device sensor functions that are more narrowband than those in hardware, by means of a linear transform of sensor functions. The utility of such a transform is that many computer vision and color-correction algorithms perform better in a sharpened space, and thus such a space can be used as an intermediate representation for carrying out calculations. In this paper we consider how one may sharpen sensor functions such that the transformed sensors are all positive. We show that constrained optimization can be used to produce positive sensors in two fundamentally different ways: by constraining the coefficients in the transform or by constraining the functions directly. In the former method, we prove that convexity can be used to constrain the solution exactly. In a sense, we are continuing the work of MacAdam and of Pearson and Yule, who formed positive combinations of the color-matching functions. However, the advantage of the spectral sharpening approach is that not only can we produce positive curves, but the process is “steerable” in that we can produce positive curves with as good or better properties for sharpening within a given set of sharpening intervals. At base, however, it is positive colors in the transformed space that are the prime objective. Therefore we also carry out sharpening of sensor curves governed not by positivity of the curves themselves but of colors resulting from them. Curves that result have negative lobes but generate positive colors. We find that this type of constrained sharpening generates the best results, which are almost as good as for unconstrained sharpening but without the penalty of negative colors. All methods discussed may be used with any number of sensors. © 2000 Optical Society of America [S0740-3232(00)00308-2]

OCIS codes: 330.1690, 330.1710, 330.1720, 330.1730, 040.1490.

1. INTRODUCTION

Spectral sharpening is a method of transforming pixel values of multispectral images produced by a color camera, scanner, or other optical device into new values that would have resulted from sensors with more-narrowband spectral sensitivities.¹ The utility of such a transform is that for many computer vision and color image processing algorithms, sharper sensors result in better performance. For example, consider the simplest form of color correction, the von Kries diagonal transform for correcting from red/green/blue (RGB) values under one illuminant to those under a second illuminant. Theoretical sensors that act as delta functions would exactly obey a diagonal transform, and it was shown in Ref. 1 that spectral sharpening could greatly benefit such a color-constancy strategy.

However, since spectral sharpening need not produce positive sensor curves, it is entirely possible that transformed RGB triples may include negative values. This presents no problem if a particular algorithm may use negative values in the transformed space, since one can then simply transform results back using an inverse transform.

Suppose $\rho^{E,S}$ denotes the camera triple $[RGB]^t$ for a surface $S(\lambda)$ viewed under $E(\lambda)$. We would like to find the RGB 3-vector ρ that would be produced under a second illuminant E' . The simplest approach to this problem involves a matrix transform:

$$\rho = \mathbf{T}^{-1} \mathbf{D}^{E,E'} \mathbf{T} \rho^{E,S}. \tag{1}$$

In Eq. (1), \mathbf{T} is a fixed 3×3 matrix and $\mathbf{D}^{E,E'}$ is an illuminant-dependent diagonal matrix. It is important to note that ρ need not equal $\rho^{E',S}$, though it will generally be quite similar. For our purposes, Eq. (1) can be usefully simplified by premultiplying both sides of it by \mathbf{T} , yielding

$$\mathbf{T} \rho = \mathbf{D}^{E,E'} \mathbf{T} \rho^{E,S}. \tag{2}$$

In Eq. (2) the linear transform can be thought of as defining new sensor functions $r'(\lambda)$, $g'(\lambda)$, and $b'(\lambda)$:

$$\begin{bmatrix} r'(\lambda) \\ g'(\lambda) \\ b'(\lambda) \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} r(\lambda) \\ g(\lambda) \\ b(\lambda) \end{bmatrix}. \tag{3}$$

If we denote the RGB response for $r'(\lambda)$, $g'(\lambda)$, and $b'(\lambda)$ as ρ' , it follows that

$$\rho' = \mathbf{D}^{E,E'} \rho'^{E,S}. \tag{4}$$

Thus the effect of the illumination can be modeled by simple scalar multipliers operating individually on each of the R , G , and B (the diagonal matrix has only three nonzero terms). In comparison, the relationship between corresponding ρ 's in the untransformed color space is much more complex.

However, many color-constancy algorithms are predicated on positive RGB values, e.g., those that utilize a maximum value in an image (see, e.g., Ref. 2), and these could therefore not be used directly in a transformed space. As well, computer vision algorithms that rely on a factor model of color (see, e.g., Ref. 3) often perform better using sharpened values yet rely on positive values. Perhaps the most compelling reason for sharpening with positivity is color-ratio stability: If an algorithm uses color ratios, as in Ref. 4, then if values fall close to zero or change sign, color ratios can vary substantially. As well, if we consider log-based homomorphic filtering for image enhancement, then the advantage of positive values is evident.

Therefore it is worthwhile to guarantee nonnegative output values of the camera or other device. In this paper we first approach the problem from the straightforward viewpoint that transformed sensors should be positive. This naturally forces RGB color values in the “sharpened” space to be positive. The most direct method for producing positive sensors is to insist that the scalar coefficients that multiply sensor functions in a linear transform be themselves positive. We investigate this premise both analytically and by use of numerical optimization techniques. This work is in a sense the natural completion of that begun by MacAdam and by Pearson and Yule.⁵ These authors formed linear combinations of the human color-matching functions, adding various proportions of the curves until negatives resulted. Here we use a straightforward optimization technique instead, but in addition we make the Pearson–Yule procedure steerable, as it were, by also insisting that the optimization concentrate each curve’s energy within a given sharpening interval. We show that it is possible to prove useful theorems about just where solutions must be found, and this narrows the range of numerical calculations that must be done.

Another approach that is possible is to constrain not the scalar coefficients that multiply sensors but the output sensors themselves. Here, only numerical methods can be brought to bear. We examine different objective functions and different norms for these tasks.

Finally, we relax the requirement that *sensors* be nonnegative and insist only that *output values* be nonnegative; i.e., we do not insist on positive curves but only on positive results. We show that it is possible to use only the convex hull of a representative set of channel values, over a sample database of surfaces and illuminants, to constrain the representative transformed values to the nonnegative range. It turns out (at least in the cases we have studied) that this less stringent optimization actually results in sharper, and hence more successful, sensor sets.

2. ANALYTICAL METHODS

We begin by developing sharpening transforms based on analytic methods and then go on in Section 3 to extend these methods to optimization techniques based on numerical methods. In the following, we investigate how sharpening proceeds in the cases where we minimize an optimization objective based on the sum of the sensor—an

L_1 objective—or based on the sum of squares of the sensor—an L_2 objective. In either case we may choose to normalize the new sensor by adding a Lagrange multiplier term that is itself based on an L_1 or an L_2 summation. An L_2 normalization is naturally associated with an L_2 objective, and an L_1 normalization with an L_1 objective; nevertheless we show below that it is useful to also consider an L_2 objective and an L_1 normalization.

A. Unconstrained L_2 – L_2 Optimization

Spectral sharpening means choosing three specific sharpening intervals within the visible spectrum, in which we would like energy in sensor curves to be concentrated. Spectral sharpening in its original form¹ makes use of an optimization based on an L_2 objective with an L_2 normalization term to set a scale, but it is not constrained by any inequality constraints.

Suppose that in general the number of sensors in use is p , with p not necessarily 3, and suppose that we measure s samples for each filter-plus-optical-system sensor function. For example, s could be 31 if the visible is from 400 to 700 nm and we measure every 10 nm. A three-filter camera or scanner would have $p = 3$. Let matrix \mathbf{Q} be the $s \times p$ matrix of sensor sensitivity values. We wish to determine the $p \times p$ matrix \mathbf{T} that maximizes the energy within a set of p chosen sharpening intervals ψ_k , $k = 1 \dots p$. Thus a sensor set \mathbf{Q} is transformed into a new set \mathbf{Q}' by

$$\mathbf{Q}' = \mathbf{Q}\mathbf{T}. \quad (5)$$

If the visible spectrum consists of wavelengths ω , then our objective is to decrease the amount of energy for wavelengths $\phi = \{\omega\} - \{\psi\}$ outside the sharpening interval ψ_k . We may choose a different sharpening interval for each of the p filters and hence carry out a separate minimization for each of the p color channels.

Thus spectral sharpening consists of finding a p -component vector \mathbf{t} , the k th column of matrix \mathbf{T} , that minimizes the least-squares summation

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]^2 + \mu \left\{ \sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}]^2 - 1 \right\}, \quad k = 1 \dots p, \quad (6)$$

where μ denotes a Lagrange multiplier and \mathbf{t} is a p -vector to be solved for.

Let us define an operator Δ_α that picks out wavelength indices in the sharpening interval α within any sum. For example, the operator Δ_{ψ_k} picks out wavelength indices in the sharpening interval ψ_k ; i.e., Δ_α is a projection operator.

When this operator is used, it is further useful to define a $p \times p$ matrix involving the summation

$$\Lambda(\alpha) = \sum_{\lambda \in \alpha} \mathbf{Q}^t(\lambda)\mathbf{Q}(\lambda) = \mathbf{Q}^t\Delta_\alpha\mathbf{Q}. \quad (7)$$

Then taking partial derivatives with respect to the unknown vector \mathbf{t} and equating to the zero vector produces the Euler equation, which can be written as

$$\Lambda(\phi_k)\mathbf{t} + \mu[\Lambda(\omega)\mathbf{t}] = \mathbf{0}. \quad (8)$$

Note that here $\Lambda(\omega)$ is just $\mathbf{Q}'\mathbf{Q}$.

Differentiating Eq. (6) with respect to μ simply sets the scale of the resulting sensor to unity in the L_2 norm:

$$\sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}]^2 = \mathbf{t}'\Lambda(\omega)\mathbf{t} = 1. \quad (9)$$

Rearranging Eq. (8), we see that solving for \mathbf{t} (and consequently for the sharpened sensor) is an eigenvector problem:

$$[\Lambda(\omega)]^{-1}\Lambda(\phi_k)\mathbf{t} = -\mu\mathbf{t}. \quad (10)$$

There are p solutions of the above equation, each solution corresponding to a stationary value, so we choose the eigenvector that minimizes $\sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]^2$. The vector \mathbf{t} derived in this way is always a real-valued vector since the matrix in Eq. (10) is the product of two symmetric positive definite matrices. We can take condition (9) into account by rescaling \mathbf{t} .

The preceding statement of spectral sharpening is denoted sensor-based sharpening, in that only the sensors themselves are used to determine the sharpened sensors; this is in contradistinction to data-based sharpening, which employs the idea of deriving a sensor transform from the behavior of a set of RGB data under an illumination change instead (see Section 4 below).

As an exemplar set of sensors, consider the sensors for the Kodak DCS-420 digital camera,⁶ shown in Fig. 1. Here we have reset the sensors' scaling such that each function sums to unity; i.e., the sensor is normalized in the L_1 norm. Since this camera has three sensors, for this device $p = 3$.

We can see that in fact this camera is sorely in need of sharpening: the blue channel in particular does not give a clear blue response. So we need to sharpen, but how do we know where to sharpen? Various studies have shown that the regions of the visible spectrum most important for the human visual system are the so-called prime intervals: around 450, 540, and 610 nm. These particular wavelengths were shown to be the set of spikes closest to the human color-matching functions for a collection of uniformly distributed spectra, and they have important efficiency characteristics for displays. The reasons these wavelengths are so important is not yet completely understood. The interested reader is referred to Refs. 7 and 8. It is reasonable to sharpen in these intervals since ultimately the output from a camera is used to drive a display that is viewed by an observer. We must therefore capture the information that is most relevant to the human visual system.

Results will depend to some degree on the specific sharpening intervals chosen. Here we choose [430–470] for blue, [520–560] for green, and [610–650] for red. Then sensor-based sharpening of the DCS-420 results in the sharpened sensors of Fig. 2. As can be seen, substantial negative lobes are introduced, although the sensor functions are indeed much more narrowband than the original set. As well, while the blue sensor retains a small bump in the green–red, it is much improved over

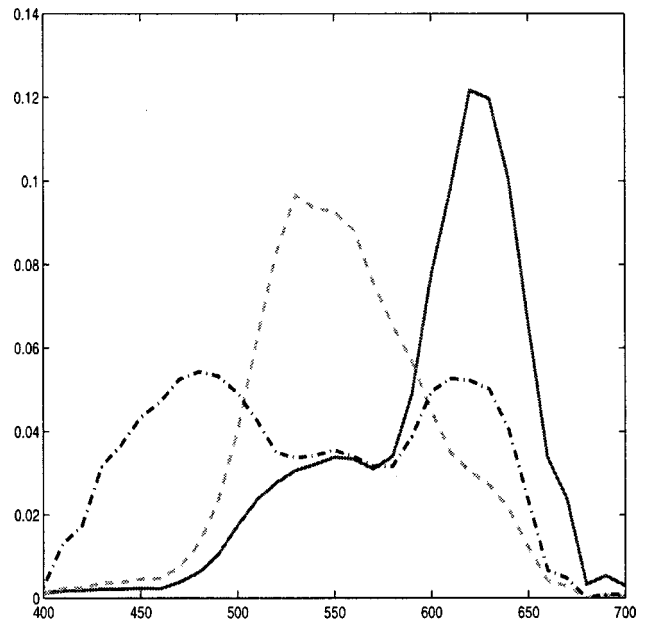


Fig. 1. Original sensors for Kodak DCS-420 digital camera.

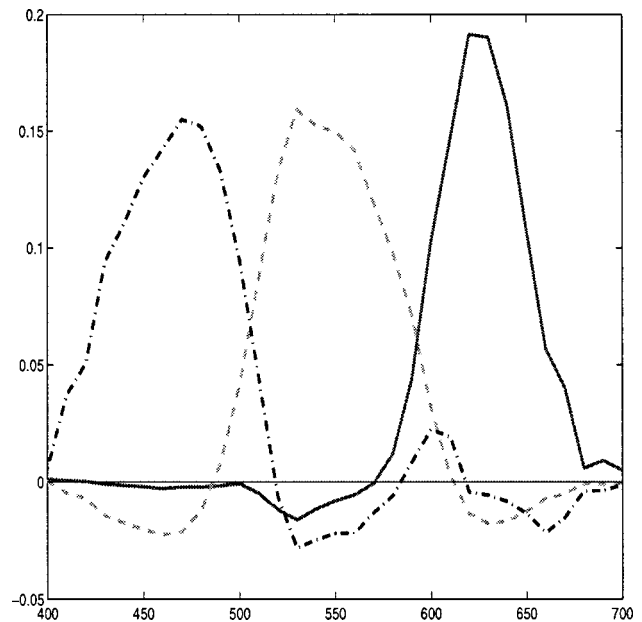


Fig. 2. Sensors sharpened by sensor-based L_2 – L_2 unconstrained sharpening.

the original. Here again, the new sensors are scaled such that each sums to unity.

We can define a degree-of-sharpness goodness measure ϵ showing how much energy is concentrated in the sharpening interval ψ . Each sensor $k = 1 \dots 3$ will employ a different ψ_k . If ϵ measures the amount of energy contained in ψ relative to in the entire visible spectrum ω , we may define

$$\epsilon = 100 \frac{\sum_{\lambda \in \psi_k} |q_k(\lambda)|^2}{\sum_{\lambda \in \omega} |q_k(\lambda)|^2} \quad (11)$$

for each of the $k = 1 \dots p$ sensors, where $q_k(\lambda)$ is the λ th component of vector \mathbf{q}_k . Table 1 shows how unconstrained sensor-based spectral sharpening with an L_2 objective and L_2 norm behaves with respect to the goodness measure. We note that spectral sharpening greatly improves the energy concentration.

Another useful feature of sharpened sensors is that any cross talk between sensors is usually diminished. Let us define cross talk κ between channels i and j by the angle

$$\kappa = \cos^{-1} \left\{ \frac{|\mathbf{q}_i^t \mathbf{q}_j|}{\|\mathbf{q}_i\| \|\mathbf{q}_j\|} \right\}, \quad (12)$$

where \mathbf{q}_i is the i th column of \mathbf{Q} and $\|\cdot\|$ is the L_2 norm.

The idea of cross talk is important since it captures one essential aspect of sharpness. In particular, sensors that are perfectly sharp (delta functions) have a cross talk of 90° . That is, the sensors measure orthogonal (hence independent) parts of the visible spectrum.

Thus the ideal value for κ is 90° ; however, a small penalty on κ would be acceptable for a good increase in ϵ . Table 2 shows that the value for the cross talk between channels is actually improved substantially with sharpening.

B. Unconstrained L_2 - L_1 Optimization

Since the first term in Eq. (6) is equivalent to a sum over all wavelengths minus those inside the sharpening interval, $\phi_k = \{\omega\} - \{\psi_k\}$, it is equivalent to extremizing (maximizing) the energy within the sharpening interval. The Lagrange multiplier term serves to set a scale for the new sensor, given via a column of Eq. (5).

Thus it is reasonable to explore how the L_2 minimization behaves if instead of using Eq. (5) we substitute an L_1 normalization condition $\sum_{\lambda \in \omega} \mathbf{Q}(\lambda) \mathbf{t} = 1$. A problem with this notion, however, is that since we are aiming at maximizing the ratio ϵ defined in Eq. (11), every candidate transformed sensor will have a different L_2 norm, and therefore although arriving at the optimum that extremizes our objective, we may not derive the best possible value of ϵ .

Nonetheless we consider this L_2 - L_1 case because it is tractable and in fact produces less-negative or entirely nonnegative solutions.

Such an L_2 - L_1 optimization consists of minimizing

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda) \mathbf{t}]^2 - \mu \left[\sum_{\lambda \in \omega} \mathbf{Q}(\lambda) \mathbf{t} - 1 \right], \quad k = 1 \dots p. \quad (13)$$

Notice that now we are using a sum, rather than a sum of squares, for the normalization condition.

Taking partial derivatives with respect to vector \mathbf{t} , we arrive at the Euler equation

$$\Lambda(\phi_k) \mathbf{t} = \mu \mathbf{f}(\omega), \quad (14)$$

Table 1. Energy Concentration in Sharpening Intervals for Original DCS-420 Sensors and for Sensors Sharpened with Unconstrained L_2 - L_2 Optimization^a

Algorithm	Energy Concentration		
	R	G	B
ϵ_0 : No Sharpening	74.34%	61.21%	21.96%
ϵ : L_2 - L_2 Sensor-Based	87.78%	71.93%	57.67%
ϵ/ϵ_0	1.181	1.156	2.626

^aRatios ϵ/ϵ_0 are sharpened over unsharpened.

Table 2. Cross talk κ for Original DCS-420 Sensors and for Sensors Sharpened with Unconstrained L_2 - L_2 Optimization

Algorithm	Cross Talk ($^\circ$)		
	(R,G)	(R,B)	(G,B)
No Sharpening	53.20	42.91	43.53
L_2 - L_2 Sensor-Based	85.81	89.13	81.57

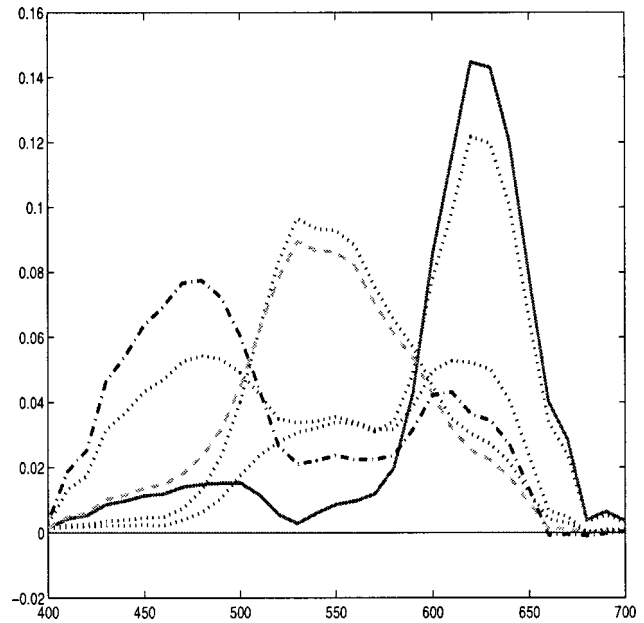


Fig. 3. Sensors sharpened by sensor-based L_2 - L_1 unconstrained sharpening. Original sensors are also shown, by the dotted curves.

Table 3. Ratio of Energy Concentration in Sharpening Intervals for Sensors Sharpened with Unconstrained L_2 - L_1 Optimization Compared with Original DCS-420 Sensors

Algorithm	Energy Concentration Ratios		
	R	G	B
L_2 - L_2 Unconstrained	1.181	1.156	2.626
L_2 - L_1 Unconstrained	1.136	0.957	1.826

where the i th component of $\mathbf{f}(\omega)$ is defined as the L_1 norm (the sum) of the i th column of sensors $\mathbf{Q}(\lambda)$ over all wave-

lengths ω . (For sensors that are already L_1 normalized, every component of $\mathbf{f}(\omega)$ is 1.)

Solving, we have

$$\mathbf{t} = \mu[\Lambda(\phi_k)]^{-1}\mathbf{f}(\omega). \quad (15)$$

Taking the derivative with respect to the Lagrange multiplier μ , we get the auxiliary condition

$$\mathbf{f}(\omega)^t \mathbf{t} = 1. \quad (16)$$

Thus a solution is obtained by using Eq. (15) with μ set equal to 1 and then rescaling that solution for \mathbf{t} by simply dividing every component by $\mathbf{f}(\omega)^t \mathbf{t}$ so that Eq. (16) is satisfied. An extremely useful benefit of having analytic solutions available is as a check on numerical optimization methods, which we examine below.

Figure 3 shows the results of using this method. We note that the sensors arrived at are similar to the original ones, shown dotted, except that the blue sensor is somewhat improved. We also note the encouraging feature that these sharpened sensors are nearly nonnegative. The second line of Table 3 shows the results for the energy concentration statistic (the first line repeats the last line in Table 1). We see that, in comparison with the results for L_2 - L_2 unconstrained sharpening, we have not done as well as using sensors with negative lobes. However, the energy concentration has at least improved over the original sensors in the red and blue, while it has decreased in the green.

3. OPTIMIZATION METHODS

With the guiding background of analytic methods, we can now turn to consideration of numerical optimization techniques for solving for a spectral sharpening transform with positivity. The methods introduced in this paper can be solved using either linear or quadratic programming: Quadratic optimization with quadratic constraints can proceed by iterated quadratic programming.⁹

We can ensure positivity of spectrally sharpened sensors in two different ways, and each of these ways gives rise to a different approach to sharpening. First, since we start with positive sensor curves, the simplest approach to developing a transform with positivity is to constrain the optimization to a solution with positive, or nonnegative, weights.

A second approach is to relax the above condition by allowing positive or negative weights but directly constraining the optimization so that the resulting sensors themselves (i.e., the spectral curves) are nonnegative.

We refer to the first approach as an optimization method with constrained coefficients and the second approach as an optimization method with constrained sensors.

Here, using numerical optimization schemes, we investigate the effect of using an L_1 objective with an L_1 constraint, and an L_2 objective with both L_1 and L_2 constraints.

A. L_1 - L_1 Optimization: Constrained Coefficients

In this case our objective is to carry out a numerical optimization with the objective function

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]$$

with constraints

$$\begin{cases} \sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}] = 1 & L_1 \text{ normalization} \\ \mathbf{t} \geq \mathbf{0} & \text{nonnegative coefficients} \end{cases}. \quad (17)$$

1. Proof of Convexity

First, it is useful to show that in fact the above minimization need not be carried out throughout the \mathbf{t} -space.

Theorem 1. Convexity implies that the solution for L_1 - L_1 sharpening, with coefficients constrained to be nonnegative, lies on the boundary of the set of possible vectors \mathbf{t} .

Proof: Without loss of generality we may assume that each sensor sums to unity. Then we seek a linear combination of sensors that also sums to unity, and is all nonnegative, with nonnegative weights. This is a convex set. If there are p sensors, then the convex set has $p - 1$ degrees of freedom.

The valid range of components of p -vector \mathbf{t} is $[0 \dots 1]$, since the composed sum of spectra whose sum is unity times these (nonnegative) weights must also sum to unity; as well, the components of each p -vector \mathbf{t} must sum to 1.

We wish to prove that the solution yielding the largest sum in the sharpening interval must lie on the boundary of possible p -vectors \mathbf{t} . Suppose that \mathbf{q}_1 and \mathbf{q}_2 are boundary functions (all nonnegative s -vectors) such that $\Sigma \mathbf{q}_1 = 1$ and $\Sigma \mathbf{q}_2 = 1$. That is, if the coefficient vectors for \mathbf{q}_1 and \mathbf{q}_2 are \mathbf{t}_1 and \mathbf{t}_2 , and if the $s \times p$ collection of sensors is matrix \mathbf{Q} , then $\mathbf{q}_1 = \mathbf{Q}\mathbf{t}_1$, $\mathbf{q}_2 = \mathbf{Q}\mathbf{t}_2$. Furthermore, let the sum in the sharpening interval be S_1 and S_2 . Any convex combination of \mathbf{q}_1 and \mathbf{q}_2 also sums to 1 (since the weighting coefficients themselves must sum to 1). It follows that the sum in the sharpening interval is a convex sum of S_1 and S_2 and so is no larger than S_1 and S_2 (we are basically calculating a weighted average). ■

What this is saying is that a convex contribution $\alpha \mathbf{q}_1 + (1 - \alpha) \mathbf{q}_2$ takes on a maximum sum S in the sharpening interval at the end $\alpha = 0$ or $\alpha = 1$. Figure 4 illustrates the situation, in the special case $p = 3$; the figure shows how S changes, over a convex combination of \mathbf{q}_1 and \mathbf{q}_2 (the situation might well be reversed, with S maximized at $\alpha = 0$).

Thus one can see that S must necessarily be maximized by continuing along the $(\mathbf{t}_1 - \mathbf{t}_2)$ direction until \mathbf{t} lies on the boundary of possible p -vectors.

Since the same situation applies along each of the boundary lines around the polygon $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, one sees that L_1 - L_1 optimization must necessarily drive solutions \mathbf{q} to the original sensors themselves. (A special case would obtain if two or more of the original sensors had equal contributions in a sharpening interval.)

Therefore one arrives at the following lemma:

Lemma 1. The solution for L_1 - L_1 sharpening, with coefficients constrained to be nonnegative, must be one of the original sensors themselves.

2. Implementation

Using the same choices for sharpening intervals as in Subsection 2.A, we can validate Lemma 1 by implementing optimization (17), using linear programming with an equality constraint (the L_1 normalization). The nonnegativity constraint on coefficients \mathbf{t} is inserted by means of a lower bound of $(0,0,0)^t$.

The resulting transformed sensors are indeed on the boundary on the convex set of linear combinations, as postulated in Theorem 1; moreover, the set of solution vectors are simply $(1,0,0)^t$, $(0,1,0)^t$, and $(0,0,1)^t$ for $k = 1,2,3$, so we find that the optimization is in fact best satisfied by the original sensors themselves.

Therefore this constrained optimization method, with constrained coefficients, does not help us. However, constraining every component of the sensors does produce new curves.

B. L_1 - L_1 Optimization: Constrained Sensors

We may continue to use only linear programming methods and still constrain the entire sensor function result, in the L_1 - L_1 case. That is, we may allow coefficients \mathbf{t} to take negative values but constrain the resulting sensor function itself to nonnegative values.

In this case Eq. (17) is modified. We no longer use a lower-bound constraint on \mathbf{t} but instead constrain $\mathbf{Q}\mathbf{t}$:

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]$$

with constraints

$$\begin{cases} \sum_{\omega} [\mathbf{Q}(\lambda)\mathbf{t}] = 1 & L_1 \text{ normalization} \\ \mathbf{Q}(\lambda)\mathbf{t} \geq 0 & \text{nonnegative sensor result} \end{cases} \quad (18)$$

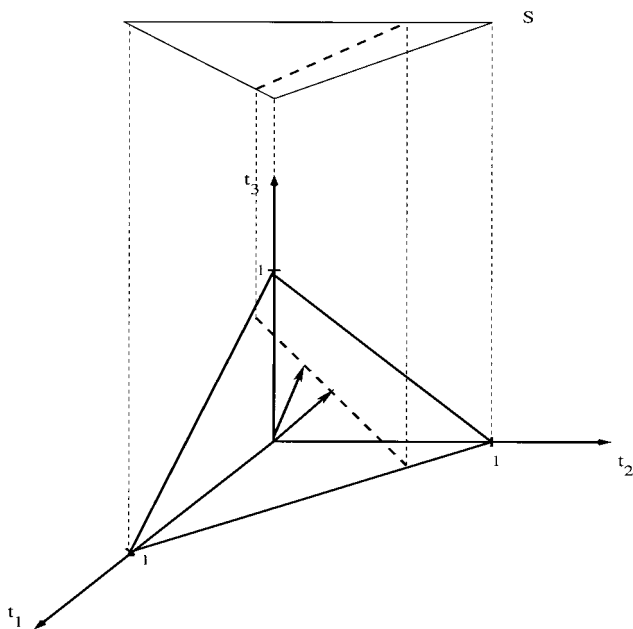


Fig. 4. Convexity drives an L_1 - L_1 constrained-coefficient solution to the boundary of the set of possible coefficients and thence to the original sensor itself. Shown here is the simplest, $p = 3$, case. The surface S represents the contribution of a combination of sensors to the sharpening interval.

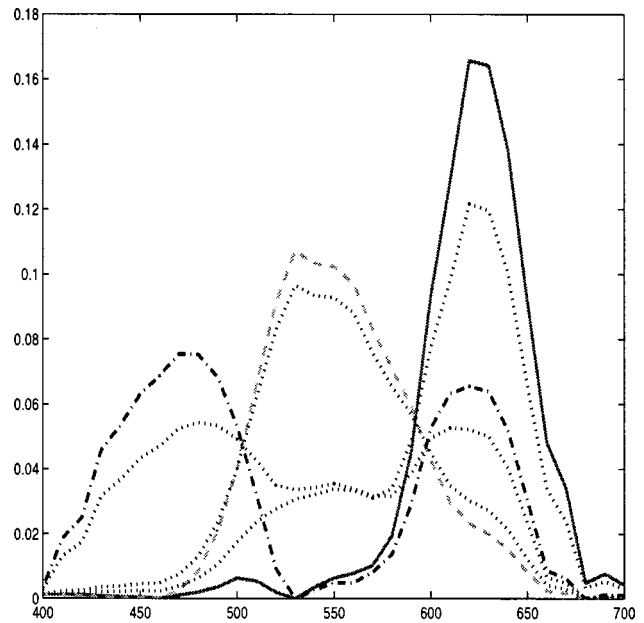


Fig. 5. Sensors sharpened by constrained L_1 - L_1 sharpening with the sensor result constrained to nonnegativity. Original sensors are also shown, by the dotted curves.

Table 4. Ratio of Energy Concentration in Sharpening Intervals for Sharpened Sensors Compared with Original DCS-420 Sensors

Algorithm	Energy Concentration Ratios		
	R	G	B
L_1 - L_1 Constrained Sensors	1.167	1.062	1.612
L_2 - L_2 Constrained Sensors	1.167	1.062	1.612
L_2 - L_1 Constrained Coefficients	0.976	0.927	1.000
L_2 - L_1 Constrained Sensors	1.136	0.941	1.242
L_2 - L_2 Data-Driven	1.181	1.152	2.596

This is a linear programming problem with one equality constraint and s inequality constraints (e.g., s may be 31).

The results of this type of sharpening are shown in Fig. 5. The energy concentration is shown in the first line of Table 4. We can see that while we have increased the energy concentration, Table 1 shows that unconstrained optimization that allows negative lobes does better.

C. L_2 - L_2 Optimization: Constrained Coefficients

The L_2 - L_2 case is the same as original spectral sharpening but makes use of constrained optimization. In this case, as in Theorem 1 for the L_1 - L_1 case, we have a convexity result that allows us to examine only the boundary of possible values of vector \mathbf{t} .

1. Proof of Convexity

Theorem 2. Convexity implies that the solution for L_2 - L_2 sharpening lies on the boundary of the set of possible vectors \mathbf{t} , if those coefficients are constrained to the nonnegative range.

Proof: Let $\mathbf{q}_1 = \mathbf{Q}\mathbf{t}_1$ and $\mathbf{q}_2 = \mathbf{Q}\mathbf{t}_2$ be two s -vectors such that each has norm 1 in the L_2 norm. That is, each of \mathbf{t}_1 and \mathbf{t}_2 is on the p -dimensional ellipsoid with quadratic form

$$\Lambda(\omega) = \mathbf{Q}^t \mathbf{Q}. \quad (19)$$

It is helpful to also recall that the index function Δ_{ψ_k} picks out wavelength indices in the sharpening interval ψ_k .

Let us denote the inner product of the sensor vectors by

$$\chi \equiv \mathbf{q}_2^t \mathbf{q}_1 = \mathbf{t}_2^t \Lambda(\omega) \mathbf{t}_1. \quad (20)$$

In order for an intermediate vector

$$\mathbf{q} = a\mathbf{q}_1 + b\mathbf{q}_2 \quad (21)$$

to be length 1, we must choose scalar b such that

$$a^2 + b^2 + 2ab\chi \equiv 1. \quad (22)$$

The values a and b are in $[0 \dots 1]$. Also $\chi \leq 1$ since it is the dot product of two unit-length vectors, so we have that

$$a, b, \in [0 \dots 1], \quad a + b \leq 1. \quad (23)$$

Let us denote by $\mathbf{u}_1, \mathbf{u}_2$ the parts of the two sensors (which lie on the convex set boundary) that fall in the sharpening interval:

$$\mathbf{u}_1 = \Delta_{\psi_k} \mathbf{q}_1, \quad \mathbf{u}_2 = \Delta_{\psi_k} \mathbf{q}_2. \quad (24)$$

Then we need to show that

$$\|a\mathbf{u}_1 + b\mathbf{u}_2\|^2 \leq (a\|\mathbf{u}_1\| + b\|\mathbf{u}_2\|)^2, \quad (25)$$

but this follows from the Cauchy–Schwartz inequality and the proof is complete. ■

Once again, the theorem implies that a linear combination of the original camera sensitivities will not result in a projection in the sharpening interval that is larger than the projection of the original curves, and we have the following lemma:

Lemma 2. The solution for L_2 – L_2 sharpening, with coefficients constrained to be nonnegative, must be one of the original sensors themselves.

2. Implementation

Constraining the coefficient vector \mathbf{t} to nonnegativity implies the following optimization:

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]^2$$

with constraints

$$\begin{cases} \sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}]^2 = 1, & L_2 \text{ normalization} \\ \mathbf{t} \geq \mathbf{0}, & \text{nonnegative coefficients} \end{cases} \quad (26)$$

Once again, as in Subsection 3.A.2, we find that for the data used here the original sensors themselves are the optimum solution, and Lemma 2 is validated.

D. L_2 – L_2 Optimization: Constrained Sensors

If we constrain the sensor curves themselves, rather than the transform coefficients, we wish to optimize

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]^2$$

with constraints

$$\begin{cases} \sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}]^2 = 1, & L_2 \text{ normalization} \\ \mathbf{Q}(\lambda)\mathbf{t} \geq \mathbf{0}, & \text{nonnegative sensor result} \end{cases} \quad (27)$$

The results are shown in Fig. 6 and in the second line of Table 4. The energy concentration is the same in this L_2 – L_2 constrained-sensor case as in the L_1 – L_1 situation in Subsection 3.B—the sensors found are virtually the same. However, we see that notwithstanding the sharpening, the blue sensor is made more bimodal, not less, than the original.

E. L_2 – L_1 Optimization: Constrained Coefficients

As a useful check on methods, and for consistency (removing the constraints allows us to check results versus those of Subsection 2.B), we again examine the L_2 – L_1 case:

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]^2$$

with constraints

$$\begin{cases} \sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}] = 1, & L_1 \text{ normalization} \\ \mathbf{t} \geq \mathbf{0}, & \text{nonnegative coefficients} \end{cases} \quad (28)$$

Results are shown in Fig. 7 and the third line of Table 4 and are seen to be poor.

F. L_2 – L_1 Optimization: Constrained Sensors

Constraining the sensor curves themselves, the minimization

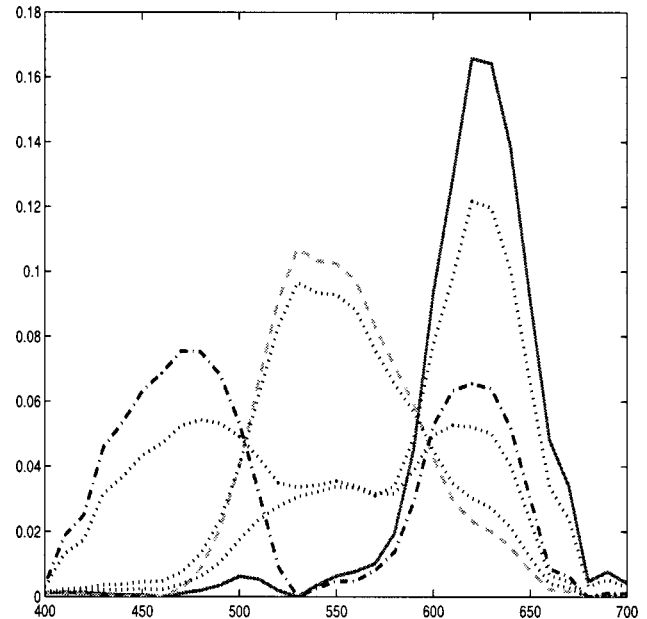


Fig. 6. Sensors sharpened by L_2 – L_2 sharpening with the sensor result constrained to be nonnegative. Original sensors are also shown, by the dotted curves.

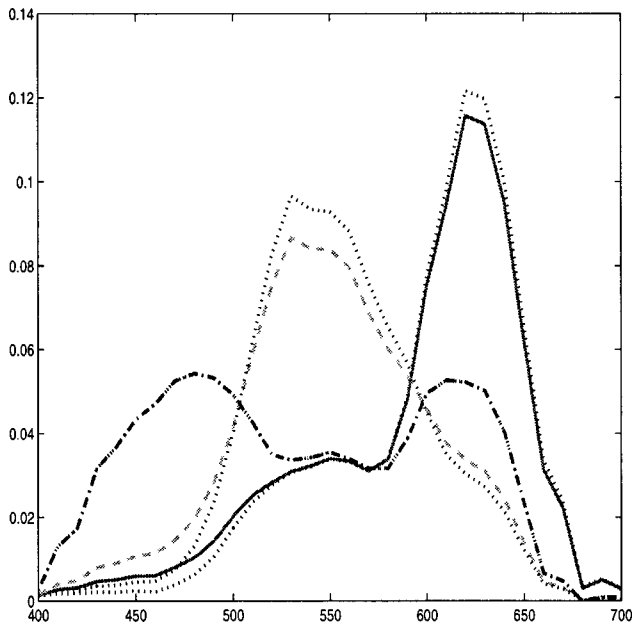


Fig. 7. Sensors sharpened by constrained L_2-L_1 sharpening with coefficients constrained to be nonnegative. Original sensors are also shown, by the dotted curves.

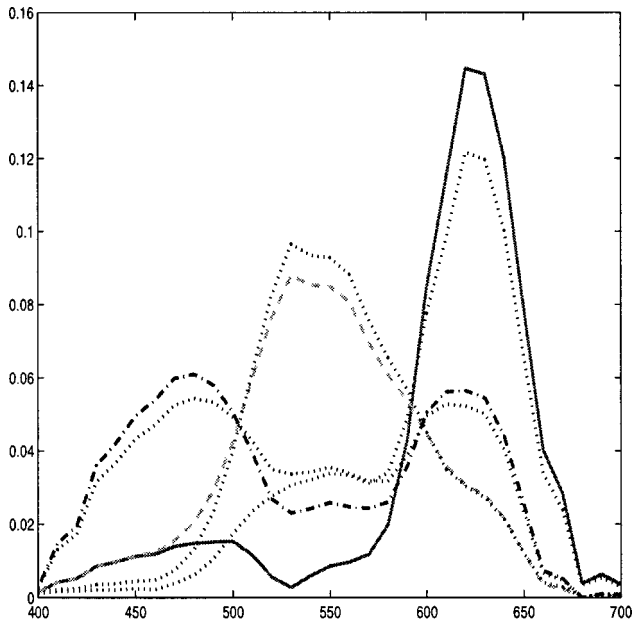


Fig. 8. Sensors sharpened by L_2-L_1 sharpening with the output sensor constrained to nonnegativity. Original sensors are also shown, by the dotted curves.

$$\min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}]^2$$

with constraints

$$\begin{cases} \sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}] = 1, & L_1 \text{ normalization} \\ \mathbf{Q}(\lambda)\mathbf{t} \geq 0, & \text{nonnegative sensor result} \end{cases} \quad (29)$$

again results in poor sensors, as shown in Fig. 8 and line four in Table 4.

Again the energy concentration is reduced, and we may conclude that an L_2-L_1 optimization is not successful in either a coefficient-constrained or sensor-constrained form.

G. Constrained Sensors: Equivalence of Linear Programming and Quadratic Programming Results

With positive transformed sensors, best results were obtained with either L_2-L_2 or L_1-L_1 optimization. It is indeed remarkable that the two methods arrive at nearly the same results. Generally, we would prefer the L_1-L_1 , linear programming approach over the L_2-L_2 , quadratic programming one since the former has less complexity. Nevertheless, we obtain better results with the data-driven method, described next, using quadratic programming.

4. DATA-DRIVEN OPTIMIZATION

Armed with the results of Subsection 3.G, we might inquire whether there is any relationship between sharpening based on the sensor curves alone and that based on the requirement of nonnegativity of RGB data. We ask this because in Ref. 1 it was shown that the least-squares transform from a set of RGB's under one illuminant to the set under another illuminant yields approximately the same sharpening transform as recapitulated in Subsection 2.A if that least-squares matrix is diagonalized. In Ref. 1 this phenomenon was referred to as "data-based sharpening." We can understand why this correspondence arises by recalling from Eq. (2) that sharpening makes a diagonal model of illuminant change more accurate. Then the diagonalization transform for the best least-squares 3×3 matrix is that which produces sharpened sensors.

Here we wish to investigate whether the requirement that the sensors be sharpened can be combined with the idea that under the new sensors we wish to have only nonnegative RGB values. We shall see that this scheme, which may yield sensors with negative lobes, leads to sharper sensors than those formed under the assumption of strict nonnegativity of the curves themselves.

Suppose that we consider an RGB triple ρ formed from a color signal $C(\lambda)$ arriving at the camera sensors: If $E(\lambda)$ is the illuminant and $S(\lambda)$ is the surface spectral reflectance function corresponding to a particular pixel, then

$$C(\lambda) = E(\lambda)S(\lambda),$$

$$\rho = \sum_{\lambda \in \omega} C(\lambda)\mathbf{Q}(\lambda). \quad (30)$$

Suppose that we collect all such RGB triples ρ into an $n \times 3$ array \mathcal{R} and also collect all the color signals into an $n \times s$ array \mathcal{C} . Then we have

$$\mathcal{R} = \mathcal{C}\mathbf{Q}. \quad (31)$$

If the sensors themselves are changed to \mathbf{Q}' via a matrix transform, then we obtain

$$\mathbf{Q}' = \mathbf{Q}\mathbf{T}, \quad (32)$$

where the column \mathbf{t}_k pertains to sharpening the k th interval. The collection of RGB values changes to those seen under the new sensors:

$$\mathcal{R}' = \mathcal{C}\mathbf{Q}\mathbf{T} = \mathcal{R}\mathbf{T}. \quad (33)$$

Now, we could set up a minimization to achieve sharpening in a given interval, subject to nonnegativity of *all* such sharpened RGB values \mathcal{R}' . However, this presents an unworkable set of constraints, and in fact we can make use of convexity to work with convex hull points only.

Suppose that we form the color signal collection \mathcal{C} by using the 462 reflectance spectra of the set of Munsell paint chips measured by Newhall *et al.*,¹⁰ along with the 170 reflectances of natural objects measured by Vrhel *et al.*¹¹ To form color signals from these reflectances, let us use the standard illuminant spectra A, C, D48, D55, D65, D75, and D100 and the fluorescent illuminant F2.¹² These illuminants represent incandescent lighting (illuminant A), a variety of standard daylights at correlated color temperatures from 4800° to 10,000°, and a standard fluorescent. Thus, with the reflectances, the set \mathcal{C} has 5056 members.

Under each of the illuminants the set of measured RGB's forms a convex set.¹³ Since we are not concerned here with brightness, we can form a two-dimensional convex set for each illuminant by forming the chromaticities $r = R/(R + G + B)$, $g = G/(R + G + B)$ and form the convex hull of the convex hulls of (r, g) values (see Ref. 14). These chromaticities can then be turned back into RGB triples.

Let us impose the reasonable constraint that the nonnegativity of RGB points corresponding to the overall convex hull of the set \mathcal{R} be maintained under a transform [Eq. (32)].

Suppose that the boundary set of RGB values is $\tilde{\mathcal{R}}$, with $\tilde{\mathcal{R}}$ an $\tilde{n} \times s$ matrix, where \tilde{n} is the number of samples in the boundary set. Our data-driven minimization is thus

$$\begin{aligned} \min \sum_{\lambda \in \phi_k} [\mathbf{Q}(\lambda)\mathbf{t}_k]^\nu \\ \text{with constraints} \\ \left\{ \begin{array}{l} \sum_{\lambda \in \omega} [\mathbf{Q}(\lambda)\mathbf{t}_k]^\nu = 1, \quad L_\nu \text{ normalization} \\ \tilde{\mathcal{R}}\mathbf{t}_k \geq \mathbf{0}, \quad \text{nonnegative RGB values} \end{array} \right. \end{array} \quad (34)$$

and we shall be interested in an L_1 norm, with $\nu = 1$, or an L_2 norm, with $\nu = 2$.

We found that an L_2 norm gives slightly better results for the DCS-420 data, and the resulting sharpened sensors are displayed in Fig. 9.

The energy concentration is shown as line five of Table 4 and is seen to be the best found for this camera. Interestingly, the sensors themselves are not all positive. The data-driven sharpening results in sensors that are similar to the unconstrained ones of Fig. 2 but produce positive RGB's.

Thus we are left with the outcome that, for the sensors examined here, a data-driven sharpening that does not insist on nonnegative sensors but only on nonnegative

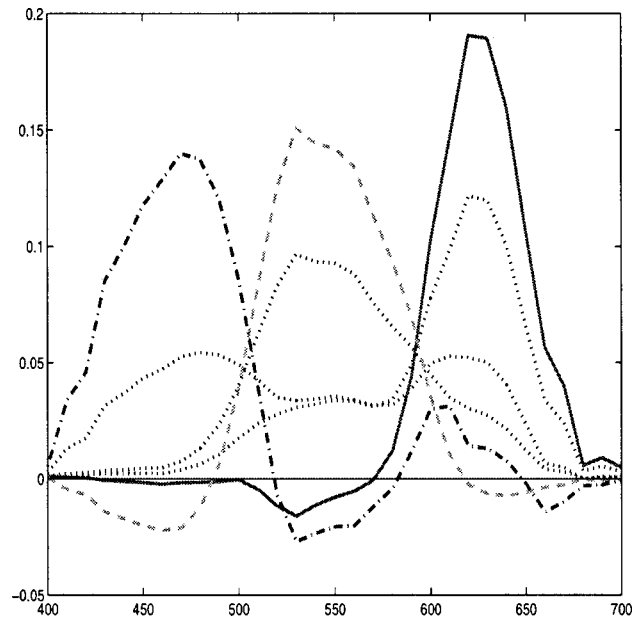


Fig. 9. Sensors sharpened by a data-driven L_2 - L_2 sharpening with transformed convex hull points constrained to nonnegativity. Original sensors are also shown, by the dotted curves.

sensor response values, gives the best sharpening. Comparing with the last line of Table 1, we see that the energy concentration is nearly as good as for the best possible, unconstrained, sharpened sensor set.

Of course, a drawback of the data-driven approach is that if a signal is out of the database's gamut, we may still arrive at a negative color. However, by choosing a wide enough range of surfaces and illuminants, we can likely avoid this problem.

5. CONCLUSIONS

In this paper we applied techniques involving both L_2 and L_1 objectives and norms to a three-band set of digital camera sensors. Unconstrained optimization gives the best energy concentration but results in curves with negative lobes. An optimization based on constraining only transformed RGB's rather than the curves themselves does best for a constrained sharpening for the camera studied, delivering almost as good results as for unconstrained sharpening but without the penalty of negative colors. Remarkably, the sharpened curves found are indeed quite close to the best curves that result from unconstrained sharpening.

We proved two theorems showing that for constrained coefficients, in both the L_1 - L_1 and the L_2 - L_2 cases, the optimum is guaranteed to be found on the boundary of possible values for coefficients for sensor sets of any dimensionality. As a consequence, if coefficient multipliers of sensor sets are constrained to be nonnegative, then the solution with the most concentration of energy in a sharpening interval is necessarily one of the original sensors themselves.

We determined that optimization with constrained sensors did better than optimization with constrained coefficients for any of the L_1 - or L_2 -based schemes.

A somewhat more complex approach to constrained-sensor optimization than that used here can in fact give sensors that are guaranteed to be positive, by means of searching the boundary of a coefficient set without insisting on positive coefficients. In such an approach we can enforce the condition that the sensors are all positive by constructing the set of all positive linear combinations of the original sensors. This set is convex with a finite number of points on the boundary. We can then find the coefficient solution relative to this set, which of course must be one of the boundary points. Then the computation can be carried out non-iteratively: We simply consider the points on the boundary of the all-positive sensor set, and this can be done with the usual tools of computational geometry. However, the method used here is likely to be less complex than the required construction of the set of all-positive sensors.

In a sense, the constrained-coefficient and constrained-sensor techniques presented here are a natural completion to the work of MacAdam, and of Pearson and Yule.⁵ The main advantage of using an optimization, with positivity, that maximizes energy concentration in desired sharpening intervals is that the process of making positive linear combinations of sensor curves is guided not by simply decreasing cross talk or making the narrowest curves but by the practical necessity of sharpening within specific areas of the visible spectrum.

Note that throughout this work we have in principle considered sensor systems of any dimensionality, although experiments were performed only for a three-band system. Note that further testing on higher-dimensional sensor systems may not necessarily corroborate the conclusion that the data-driven approach seems to be best compared with either the constrained-coefficient or constrained-sensor techniques.

Nevertheless, the results developed here do show that sharpening with positivity not only is possible but indeed leads to more well-behaved sensor functions than do the raw curves used in hardware. Moreover, we found that it generally made little difference whether an L_1 norm or an L_2 norm was used. This fact may point to the existence of a type of global optimum for sharpening with positivity. Further work will be aimed at exploring systems with more sensors and also at applying the methods set out here to color correction and other human-perception-motivated applications.

ACKNOWLEDGEMENTS

The authors are indebted to the anonymous referees for suggestions that markedly improved the paper. In par-

ticular, a reviewer pointed out that the theorems imply the much more restrictive derived lemmas given in Subsections 3.A.1 and 3.C.1.

The authors may be contacted by e-mail as follows: mark@cs.sfu.ca or graham@sys.uea.ac.uk.

REFERENCES AND NOTES

1. G. D. Finlayson, M. S. Drew, and B. V. Funt, "Spectral sharpening: sensor transformations for improved color constancy," *J. Opt. Soc. Am. A* **11**, 1553–1563 (1994).
2. K. Barnard and B. V. Funt, "Experiments in sensor sharpening for color constancy," in *6th Color Imaging Conference: Color, Science, Systems and Applications* (Society for Imaging Science & Technology/Society for Information Display, Springfield, Va., 1998), pp. 43–46.
3. M. S. Drew and M. H. Brill, "Color from shape from color: a simple formalism with known light sources," *J. Opt. Soc. Am. A* **17**, 1371–1381 (2000).
4. B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Trans. Pattern Anal. Mach. Intell.* **17**, 522–529 (1995).
5. M. L. Pearson and J. A. Yule, "Transformations of color mixture functions without negative portions," *J. Color Appearance* **2**, 30–35 (1973). Inspiration for their work is attributed to MacAdam.
6. P. L. Vora, J. E. Farrell, J. D. Tietz, and D. H. Brainard, "Digital color cameras. 2. spectral response," Tech. Rep. HPL-97-54 (Hewlett-Packard Laboratories, Palo Alto, Calif., 1997). <http://www.hpl.hp.com/techreports/97/HPL-97-54.html>.
7. G. D. Finlayson and B. V. Funt, "Coefficient channels: derivation and relationship to other theoretical studies," *Color Res. Appl.* **21**, 87–96 (1996).
8. M. H. Brill, G. D. Finlayson, P. M. Hubel, and W. Thornton, "Prime wavelengths and color imaging," in *6th Color Imaging Conference: Color, Science, Systems and Applications* (Society for Imaging Science & Technology/Society for Information Display, Springfield, Va., 1998), pp. 33–42.
9. P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, New York, 1981).
10. S. M. Newhall, D. Nickerson, and D. B. Judd, "Final report of the OSA subcommittee on the spacing of the Munsell colors," *J. Opt. Soc. Am.* **33**, 385–418 (1943).
11. M. J. Vrhel, R. Gershon, and L. S. Iwan, "Measurement and analysis of object reflectance spectra," *Color Res. Appl.* **19**, 4–9 (1994).
12. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, 2nd ed. (Wiley, New York, 1982).
13. D. A. Forsyth, "A novel approach to color constancy," in *Proceedings of the International Conference on Computer Vision '88* (IEEE Computer Society Press, Los Alamitos, Calif., 1998), pp. 9–18.
14. G. D. Finlayson, "Color in perspective," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**, 1034–1038 (1996).