CONVEX PROGRAMMING COLOUR CONSTANCY WITH A DIAGONAL-OFFSET MODEL

G. D. Finlayson, S. D. Hordley
School University of East Anglia
School of Computing Sciences
Norwich, UK

R. Xu*
Zilhka Neurogenetic Institute
University of Southern California
Los Angeles, USA

ABSTRACT
Gamut mapping colour constancy algorithms attempt to map image RGBs captured under an unknown light to corresponding RGBs under a reference light so as to render images colour constant. While the approach often works well, for a significant number of real images the algorithm delivers a null solution. We show that the null solution problem arises because of a failure of the diagonal model of illumination change on which the algorithm is based. We address the problem by proposing a new diagonal-offset model which has 6 rather than 3 parameters and which is able therefore to deal with a wider range of imaging conditions. Based on this model we formulate a convex programming solution to colour constancy and we show (by testing on real images) that the new algorithm is robust to the failures of the diagonal model and is capable of delivering very good colour constancy.

1. INTRODUCTION
Image RGBs depend on the reflective properties of the imaged objects and, in equal part, on the spectral properties of the light incident upon them. This implies that an image has an overall colour cast determined by the colour of the prevailing scene illumination. Identifying and removing this colour cast is called colour constancy and is important in computer vision because many vision tasks which make use of colour information do so under the assumption that colour is an intrinsic property of the imaged objects and thus independent of the prevailing illumination conditions. Despite many attempts to solve the colour constancy problem (see [?] for a review) no adequate solution has been obtained. The problem is difficult to solve, in part because it is strictly ill-posed. That is, for most images there are many combinations of surfaces and illuminant which give rise to the same image data so that there is an inherent uncertainty in the scene illuminant.

* R. Xu performed the work while at the University of East Anglia.

Forsyth’s [?] gamut mapping algorithm is one of the most successful colour constancy algorithms and its strength lies in the fact that it makes the uncertainty in the illuminant explicit. So, rather than solving for a single estimate of the scene light, the algorithm recovers the set of all illuminants (the feasible set) consistent with a given set of image data. To determine the feasible set Forsyth exploited the observation that the range, or gamut of colours we observe depends on the illumination. Thus, the gamut of observed image colours provides information about the colour of the scene light. To exploit this observation Forsyth began by defining a canonical gamut, denoted $C$: the set of all RGBs which can be observed under some reference illumination. Next, given an image recorded under an unknown illuminant, Forsyth posed the colour constancy problem as that of determining the mappings which take the image gamut (the set of all image RGBs) into the canonical gamut.

Key to solving the problem when it is posed in this way is to define the form of the mapping from image to canonical gamut. Forsyth modelled this illumination change using a set of three scale factors $[d_1, d_2, d_3]$ such that an observed RGB $[R^o, G^o, B^o]$ is mapped to its corresponding RGB under a reference light according to:

$$
\begin{bmatrix}
R^c \\
G^c \\
B^c
\end{bmatrix} =
\begin{bmatrix}
d_1 & 0 & 0 \\
0 & d_2 & 0 \\
0 & 0 & d_3
\end{bmatrix}
\begin{bmatrix}
R^o \\
G^o \\
B^o
\end{bmatrix}
$$

(1)

This model is referred to as the diagonal model of illumination change because $[d_1, d_2, d_3]$ parameterise a diagonal matrix. Now, $[d_1, d_2, d_3]$ is a valid mapping (a solution to the colour constancy problem) if and only if $\forall i D q_i \in C$ where $D$ is $[d_1, d_2, d_3]$ written as a diagonal matrix and $q_i$ represents the $i$th image RGB. For a given image many diagonal matrices will satisfy this constraint and Forsyth derived an algorithm to recover the whole set of feasible mappings. Once recovered, a single element of this set is chosen (according to some heuristic) as an estimate of the scene illuminant.

Though an elegant theoretical solution to the colour constancy problem, gamut mapping has a serious shortcoming:
on a significant number of real images it returns a null feasible set, i.e. no estimate of the scene illuminant. The underlying cause of this failure of the algorithm is a failure of the diagonal model to accurately model illumination change for all surfaces in the given image. In this paper we address this failure by proposing a modified model of illumination change: a six parameter diagonal-offset model. Adopting this model we show that a simple, convex programming solution to colour constancy can be derived and we demonstrate that this novel algorithm is robust (it always returns a colour constancy solution) and has performance comparable to the original gamut mapping approach. We motivate and introduce the new model of illumination change in the next section. In §3 we formulate a convex programming solution to colour constancy in the context of this model. We present an empirical evaluation of the algorithm and summarise our findings in §4.

2. THE DIAGONAL-OFFSET MODEL

A diagonal model of illumination change is strictly valid for narrow-band camera sensors (or narrow-band illuminants). Of course, real camera sensors are not narrow-band, but even for many broad-band sensors the model is applicable for most surfaces and illuminants. However, the model does break down, particularly for saturated colours (i.e. surfaces near the gamut boundary) and this can lead to a failure of the gamut mapping algorithm. Failure can also occur because an image contains surfaces which are not represented in the canonical gamut. In either case, gamut mapping algorithms can return a null feasible set.

Such problems can be avoided by modifying the canonical gamut. For example Finlayson [?] proposed incrementally extending the size of the canonical gamut until a non-null feasible set is found. A similar approach was proposed by Barnard et al [?] but in this case error is added to the observed image data incrementally until the gamut mapping algorithm returns a non-empty feasible set. Finally, in Colour by Correlation [?], a discrete implementation of the gamut mapping idea is implemented. That approach effectively counts the number of image colours successfully mapped into the canonical gamut for a particular mapping (i.e. a particular illuminant). If there is no mapping which successfully maps all image colours into the canonical set the mapping which successfully maps most colours is chosen. All of these approaches however, detract from the inherent elegance of gamut mapping. Modifying the canonical or observed gamut is at best heuristic and at worst weakens the gamut constraint so that poor colour constancy performance is obtained while discretising the problem means that we must discretise the number of mappings. In this paper we seek to re-formulate gamut mapping so that it remains a continuous domain estimation problem while continuing to work with a conservatively sized canonical gamut.

![Diagram](image)

Fig. 1. (a) Example of Gamut Mapping Failure. (b) Failure corrected by applying a translation to the image gamut.

If the gamut mapping algorithm returns a null feasible set this implies that there is no diagonal mapping which can simultaneously map all colours in the image gamut into the canonical gamut. However, we might expect that there exists a mapping which maps the image gamut into the canonical set. Figure 1a illustrates this idea for the case of 2-d sensor responses. In this example, if in addition to the diagonal mapping, we also allow a translation of all sensor responses, then we can map the image gamut into the canonical set (see Figure 1b). That is, we can find a 3-parameter diagonal transform plus a 3-parameter offset term which maps any image colour \([R^o \ G^o \ B^o]\) to a corresponding colour \([R^c \ G^c \ B^c]\) in the canonical set:

\[
\begin{bmatrix}
R^c \\
G^c \\
B^c
\end{bmatrix} =
\begin{bmatrix}
d_1 & 0 & 0 \\
0 & d_2 & 0 \\
0 & 0 & d_3
\end{bmatrix}
\begin{bmatrix}
R^o \\
G^o \\
B^o
\end{bmatrix} +
\begin{bmatrix}
o_1 \\
o_2 \\
o_3
\end{bmatrix}
\]

(2)

In practice we expect departures from the diagonal model to be quite small and as a result, the offset term \([o_1 \ o_2 \ o_3]\) will be small relative to the diagonal part \([d_1 \ d_2 \ d_3]\). Thus in effect, we continue to model illumination change using a diagonal model (which we know often works well) but in addition we introduce some “slack” into the model in the form of the offset term to avoid the null solution problem. In the ideal case in which the diagonal model is valid, the offset term will be zero. However, if no such valid solution can be found, a non-zero offset allows for a solution and thus provides an illuminant estimate even when the diagonal model fails.

3. SOLVING FOR COLOUR CONSTANCY

We formulate the colour constancy problem as in the diagonal model case: we seek the mapping (or, more generally, the set of mappings) which takes the image gamut \(T\) into the canonical gamut \(C\). The canonical gamut is defined as \(C = C(\mathcal{P})\) where \(\mathcal{P} = \{p_1^c, p_2^c, \ldots, p_n^c\}\) are RGBs recorded for a representative set of surfaces viewed under the reference light. The function \(C(\cdot)\) returns the set of all convex
combinations of \( \mathcal{P} \) so that any convex combination of points in \( \mathcal{P} \) belongs to the canonical gamut. Without loss of generality \( \mathcal{P} \) is chosen such that \( \forall \ i \ \mathbf{p}_i \notin C(\mathcal{P} - \mathbf{p}_i) \) where \( \mathcal{P} - \mathbf{p}_i \) denotes the set \( \mathcal{P} \) with the \( i \)th RGB removed. This ensures that \( \mathcal{P} \) is the set of points on the convex hull of the reference gamut. Equation (4) defines a convex polyhedron which Finlayson et al.\footnote{A plane in 3-dimensional space divides the space into two halves. The set of points which lie to one side of a plane is called a half-space.} observed can equally well be characterised as the intersection of a set of half-spaces\footnote{A plane in 3-dimensional space divides that space into two halves. The set of points which lie to one side of a plane is called a half-space.} rather than by the set of hull points \( \mathcal{P} \). In 3-dimensions half-spaces are defined by plane equations of the form \( a_1 R + a_2 G + a_3 B \leq b \).

If the 3-dimensional canonical gamut in sensor RGB space has \( n \) vertices then it can be shown that it also has \( N \) faces where \( N = O(n) \). Each face lies in a plane and only points which lie to one side of the plane (in one half-space) are in \( C \). The intersection of all such half-spaces defines the canonical gamut. That is, if a point \( [R \ G \ B] \) is in \( C \) it must satisfy the \( N \) half-space inequalities:

\[
\begin{align*}
 a_{11} R + a_{12} G + a_{13} B &\leq b_1 \\
 a_{21} R + a_{22} G + a_{23} B &\leq b_2 \\
 &\vdots \\
 a_{N1} R + a_{N2} G + a_{N3} B &\leq b_N
\end{align*}
\]

We can summarise this as: \( \mathbf{p} \in C \Leftrightarrow \mathbf{A} \mathbf{p} \leq \mathbf{b} \) where \( \mathbf{A} \) is an \( N \times 3 \) matrix whose \( i \)th row is \( [a_{1i} \ a_{2i} \ a_{3i}] \) and \( \mathbf{b} \) is an \( N \times 1 \) vector whose \( i \)th component is equal to \( b_i \).

To define the image gamut let \( \mathbf{Q} = \{q_1, q_2, \ldots, q_m\} \) denote the \( m \) points on the convex hull of the set of image RGBs. Then we define the image gamut to be \( \mathcal{I} = C(\mathbf{Q}) \). Now, suppose that the \( j \)th element of \( \mathbf{Q} \), \( q_j \), is to be mapped by a diagonal-offset transform to the canonical gamut where \( \mathbf{D} = \text{diag}(d) \) and \( \mathbf{g} \) denote the diagonal and offset parts of the transform respectively. It follows that the mapped image RGB \( \mathbf{D} q_j + \mathbf{g} \) must satisfy the inequalities in (3) which define the canonical gamut. That is:

\[
\mathbf{A} \left( \mathbf{D} q_j + \mathbf{g} \right) \leq \mathbf{b}
\]

We can re-write (4) as: \( \mathbf{A} \{Q_j | I\} d_o \leq \mathbf{b} \) where \( \mathbf{I} \) is the \( 3 \times 3 \) identity matrix and \( Q_j \) is \( q_j \) written as a diagonal matrix. Or equivalently, with some algebraic manipulation:

\[
\mathbf{A} \{Q_j | I\} d_o \leq \mathbf{b}
\]

where \( d_o = [d_1 \ d_2 \ d_3 \ o_1 \ o_2 \ o_3] \). The set of inequalities in (5) delimits a set of 6-parameter Diagonal-Offset mappings \( \mathcal{M}_j \) which map \( q_j \) into the reference gamut:

\[
\mathcal{M}_j = \{d_o \mid \mathbf{A} \{Q_j | I\} d_o \leq \mathbf{b}\}
\]

If there are \( N \) image hull points then we will have \( N \) similar sets of inequalities, one for each hull point. The set of mappings which take all image points into the reference gamut is thus those mappings which are in all \( \mathcal{M}_j \):\footnote{A plane in 3-dimensional space divides that space into two halves. The set of points which lie to one side of a plane is called a half-space.}

\[
\mathcal{M}' = \{d_o \mid \mathbf{A}' d_o \leq \mathbf{B}\}
\]

where \( \mathbf{A}' \) has the matrices \( \mathbf{A} \{Q_j | I\} \) (\( j = 1 \ldots m \)) stacked one on top of the other and \( \mathbf{B} \) is an \( NM \times 1 \) vector: \( M \) copies of the vector \( \mathbf{b} \) stacked one on top of another. Thus \( \mathcal{M}' \) defines the set of feasible mappings. Each member of this set is a 6-dimensional vector defining a diagonal-offset mapping taking the image gamut into the canonical gamut.

To complete the gamut mapping solution we must select a single mapping from \( \mathcal{M}' \) as the scene illuminant estimate. In his original formulation Forsyth chose the mapping which maximises the volume of the image gamut when it is mapped to the reference gamut. This is achieved by maximising the determinant \( d_1 d_2 d_3 \) of the diagonal mapping. In our case, in addition to the diagonal mapping we also have to fix the offset term. However, we expect that in many cases the diagonal part of the model will account for the illumination change and the offset term will be small (or ideally zero). Thus a possible selection strategy is to maximise the determinant of the diagonal part of the mapping (as Forsyth did) whilst minimising the offset:

\[
\min (-d_1 d_2 d_3 + o_1^2 + o_2^2 + o_3^2)
\]

where any solution must also satisfy the constraints in (7). Since the objective function is non-linear, finding its minimum involves quite significant computational cost. Nevertheless robust algorithms for solving such minimisation problems do exist\footnote{Another alternative is to leave the offset terms as free parameters in the optimisation and construct objective functions based only on the diagonal part of the mapping. This gives rise to two further optimisation schemes. The first} and are well understood.

In\footnote{Another alternative is to leave the offset terms as free parameters in the optimisation and construct objective functions based only on the diagonal part of the mapping. This gives rise to two further optimisation schemes. The first} Finlayson et al. proposed a selection method for the original gamut mapping problem in which they maximise \( d_1 + d_2 + d_3 \) (the \( L_1 \) norm of the diagonal model). This delivers similar performance to that obtained by maximising the determinant and has the advantage that the algorithm can be formulated as a linear programming problem. In our case we would like to maximise the \( L_1 \) norm of the diagonal part of the mapping whilst minimising the offset term. We can do this provided that the offset terms are guaranteed always to be positive. In practice this is not the case but we can force positivity by adding extra linear constraints. The solution can then also be posed as a linear program:

\[
\min (-d_1 - d_2 - d_3 + o_1 + o_2 + o_3)
\]

subject to: \( \mathbf{A}' d_o \leq \mathbf{B}, \ d_o \geq 0 \).
method maximises the $L_1$ norm and results in a linear program:

$$\min (-d_1 - d_2 - d_3) \text{ subject to: } A'd_o \leq B$$

(10)

Alternatively we can maximise the determinant of the diagonal part of the mapping:

$$\min (-d_1 d_2 d_3) \text{ subject to: } A'd_o \leq B$$

(11)

Equation (11) is similar to Forsyth’s original solution except that the constraints on the possible mappings are different.

4. EMPIRICAL EVALUATION

We evaluated the performance of our new algorithm on the Simon Fraser calibrated test set [?] which comprises 321 real images. We found that performance for all algorithms was improved by using a simple pre-processing strategy in which we first we smooth images using a Gaussian kernel and then sub-sample images by a factor of 5. Finally we discard image pixels which have either an R, G, or B response of 2 or less (the image data is in the range 0 to 255).

We formed the canonical gamut used by the algorithm following Barnard et al [?]. We estimate the performance of an algorithm in terms of how well its (3- or 6-parameter) mapping predicts $q_w$, the RGB of a white surface under the scene light. For example, if $p_w$ is the RGB of a white surface under the reference light and the 3-d gamut mapping algorithm returns a map $\hat{D}$ then $p_w \approx \hat{D} q_w$. Or equivalently $q_w \approx \hat{D}^{-1} p_w$. Defining $q_w = \hat{D}^{-1} p_w$ we measure the error between this estimate and the RGB of the actual scene illuminant white $q_w$. In common with other work [?] we use an intensity independent error measure: the angle between the actual and estimated white.

Table 1 summarises the results. We compare performance to Max RGB (row one) since in the limiting case of a cube gamut the gamut mapping algorithm is equivalent to Max RGB [?]. We also test two variants of the diagonal model gamut mapping algorithm implemented in the convex programming framework [?] (rows 2-3). Rows 4-7 show the results for the four variants of the new algorithm introduced in this paper. This table shows the Root Mean Square (RMS), mean, median, and maximum angular error taken over all images for which the algorithms returned a solution, together with the number of images for which no solution was found. We note that given the nature of the underlying error distributions, the median statistic is the most reliable single summary statistic for judging relative algorithm performance [?].

The results of the experiments reveal a number of important points. First, all gamut algorithms deliver good colour constancy performance supporting the findings in [?] where gamut mapping algorithms were found to be amongst the best performing algorithms. However, the experiment also highlights the fact that the null intersection problem of the original gamut mapping algorithm is a real problem: for a significant number of images (44) the 3-parameter algorithms return no solution. Importantly the results for the 6-parameter algorithms show that adopting a diagonal-offset model addresses the problem of diagonal model failure. All variants of this algorithm return a solution for all 321 images tested and performance is as good as or better than that of the 3-parameter variants.

Table 1. Summary of performance (details in text).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>RMS</th>
<th>Mean</th>
<th>Med</th>
<th>Max</th>
<th>Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max-RGB</td>
<td>8.77</td>
<td>6.30</td>
<td>4.02</td>
<td>26.19</td>
<td>0</td>
</tr>
<tr>
<td>D. M. Max $L_1$</td>
<td>6.68</td>
<td>4.70</td>
<td>2.75</td>
<td>23.87</td>
<td>44</td>
</tr>
<tr>
<td>D. M. Max Det</td>
<td>6.36</td>
<td>4.49</td>
<td>2.75</td>
<td>22.02</td>
<td>44</td>
</tr>
<tr>
<td>Eq (8)</td>
<td>6.75</td>
<td>4.90</td>
<td>2.99</td>
<td>22.04</td>
<td>0</td>
</tr>
<tr>
<td>Eq (9)</td>
<td>7.29</td>
<td>5.19</td>
<td>3.11</td>
<td>24.17</td>
<td>0</td>
</tr>
<tr>
<td>Eq (10)</td>
<td>6.78</td>
<td>4.63</td>
<td>2.71</td>
<td>23.16</td>
<td>0</td>
</tr>
<tr>
<td>Eq (11)</td>
<td>6.78</td>
<td>4.60</td>
<td>2.67</td>
<td>23.16</td>
<td>0</td>
</tr>
</tbody>
</table>

5. REFERENCES


