Gamut Constrained Illuminant Estimation

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

I. Tastl
Hewlett Packard Incorporated, Palo Alto, USA

Abstract.
   This paper presents a novel solution to the *illuminant estimation* problem: the problem of how, given an image of a scene taken under an unknown illuminant, we can recover an estimate of that light. The work is founded on previous *gamut mapping* solutions to the problem which solve for a scene illuminant by determining the set of diagonal mappings which take image data captured under an unknown light to a gamut of reference colours taken under a known light. Unfortunately, a diagonal model is not always a valid model of illumination change and so previous approaches sometimes return a null solution. In addition, previous methods are difficult to implement. We address these problems by recasting the problem as one of illuminant classification: we define *a priori* a set of plausible lights thus ensuring that a scene illuminant estimate will always be found. A plausible light is represented by the gamut of colours observable under it and the illuminant in an image is classified by determining the plausible light whose gamut is most consistent with the image data. We show that this step (the main computational burden of the algorithm) can be performed simply and efficiently by means of a non-negative least-squares optimisation. We report results on a large set of real images which show that it provides excellent illuminant estimation, outperforming previous algorithms.

Keywords: Colour Constancy, Illuminant Estimation, Gamut Mapping

1. Introduction

It is well established that colour is a useful cue which helps to solve a number of classical computer vision problems such as object recognition [29], object tracking [27] and image segmentation [8]. However, implicit in this use of colour is the assumption that colour is an inherent property of an object. In reality an object's colour depends in equal measure on the object's physical properties and the characteristics of the light by which it is illuminated. That is, the image colours recorded by a camera change when the colour of the light illuminating the scene is changed. One way to deal with this problem is to derive so called *colour invariant* features from the image data which are invariant to the prevailing illumination (for some examples of such an approach see [16, 15, 17]). An alternative approach, which is the focus of this paper, is to correct images to account for the effect of the scene illuminant on recorded colours. This correction procedure is typically

achieved in a two-stage process. In the first stage the colour of the scene illumination is estimated from the image data. Then, the estimate of the scene illuminant is used to correct the recorded image such that the corrected image is independent of the scene illuminant. The first stage of the process – estimating the scene illuminant – is the more difficult of the two and it is this problem which is the subject of this paper.

Despite much research [22, 6, 24, 14, 9, 7, 10, 2, 3] there does not exist a satisfactory solution to the illuminant estimation problem. Early attempts at a solution sought to simplify the problem by making certain assumptions about the composition of a scene. For example it was assumed that a scene will contain a "white" (maximally reflective) surface [22] or that the average of all surfaces in the scene is neutral [6]. Other authors modelled lights and surfaces by low-dimensional linear models and derived algebraic solutions to the problem [24]. It is easy to understand why such approaches do not work in practice: the constraints they place on scenes are too strong. More recently a number of more sophisticated algorithms have been developed [14, 7, 10, 5, 28] and these approaches can often give reasonable illuminant estimation.

These methods work not by placing constraints on scenes but by exploiting prior knowledge about the nature of objects and surfaces. For example, the Neural Network (NN) approach of Funt et al [7] attempts to learn the relationship between image data and scene illuminant on the basis of a large number of example images. While it performs reasonably well this approach is unsatisfactory because it provides a black box solution to the problem which gives little insight into the problem itself. In addition neural networks are non-trivial to implement and they rely for their success on training data which properly reflects the statistics of the world. In practice neural networks often do not generalise well and this has been found to be the case for illuminant estimation [3].

Perhaps, a more promising approach is the Correlation Matrix (CM) algorithm of Finlayson *et al* [10] which has the advantage of being simple both in its implementation and in terms of the computations which must be carried out to estimate the illuminant. This method gains part of its simplicity from the observation that the set of possible scene lights is quite restricted and can be adequately represented by considering just a small discrete set of plausible lights. The method works by exploiting the fact that the statistical distribution of possible image colours varies as the scene illuminant changes. Using this fact, and given a set of image data, it is possible to obtain a measure of the probability that each of the set of plausible illuminants was the scene light. Clearly, the approach relies on having a reasonable statistical model of lights and surfaces: (i.e. it requires accurate training data to

work in practice) but unfortunately such data is difficult to obtain. A second disadvantage of the method is that it works not with 3-d RGB camera data but with 2-d chromaticity data: i.e. brightness information is discarded. There is evidence to suggest however [4, 30] that brightness information is important in estimating the scene illuminant and while brightness information can be incorporated into the correlation matrix algorithm [4], doing so destroys the simplicity of the approach.

A third approach which has also shown promise is the Gamut Mapping (GM) method first proposed by Forsyth [14]. Gamut Mapping also exploits knowledge about the properties of surfaces and lights to estimate the scene illuminant. However, it is not based on a statistical analysis of the frequency with which surfaces occur in the world but on the simple observation that the set of camera RGBs observable under a given illuminant is a bounded convex set. This observation follows directly from physical constraints on the nature of a surface (all surfaces reflect between 0 and 100% of light incident upon them) and the linear nature of image formation. It therefore exploits the minimal assumptions that can be made about the world.

Forsyth exploited this observation by characterising the gamut of possible image colours which can be observed under a reference light. He called this set the *canonical gamut*. The colours in an image whose scene illuminant it is wished to estimate can also be characterised as a set: the *image gamut*. Estimating the scene illuminant can then be cast as the problem of finding the mapping which takes the image gamut into the canonical gamut. To properly define the algorithm requires that the form of the mapping between the two sets be specified. That is, the relationship between sensor responses under two different illuminants must be defined. Forsyth modelled illumination change using a *diagonal model* [14] under which sensor responses for a surface viewed under two different lights are related by simple scale factors which depend on the sensors and the pair of lights but which are independent of the surface.

Using this diagonal model of illumination change Forsyth derived the CRULE algorithm which estimates the scene illuminant in two stages. First the set of all mappings (the *feasible set*) which take the image gamut to the canonical gamut are determined: usually many mappings are consistent with the image data. In the second stage a single mapping is selected from this feasible set according to some pre-defined selection criterion. The algorithm has been found to perform quite well [3] but it suffers from a number of limitations. It is these limitations which we aim to address in this paper. We do this by proposing a new algorithm which shares some of the fundamental ideas of the original work but which is implemented in such a way as to avoid its weaknesses.

The first weakness of CRULE is the fact that the algorithm is founded on the assumption that illumination change can be well modelled by a diagonal transform. This leads to an algorithm which is conceptually simple but which can nevertheless fail. One reason for this is that real images can contain an illuminant or (more realistically) one or more image colours which do not conform to the diagonal model of illumination change and unfortunately in such situations the theoretical foundation of the algorithm implies that there is no illuminant estimate which satisfies all the image data. That is, there does not exist a diagonal transform that maps the image gamut inside the canonical gamut. While this problem might be ameliorated by carrying out a change of sensor basis (and using a generalised diagonal transform [11]) it does not completely disappear: a linear model of illumination change is only approximate.

Gamut mapping is also a complex algorithm which is difficult to implement. Estimating the feasible set involves the intersection of many convex sets and numerical inaccuracies in this procedure can lead to a situation in which the result is a null intersection even if all the image colours satisfy the diagonal model. Finally, even when a feasible set of mappings is successfully determined, the mappings in that set may not correspond to realistic illuminants. This is because the range of illuminants under which images are captured is quite restricted whereas the set of mappings taking image colours into the canonical gamut can be relatively large.

The new algorithm we propose in this paper avoids the problem of unrealistic solutions by defining *a priori* a set of plausible illuminants. In this sense our algorithm is similar to the CM approach of Finlayson *et al* [10] who also restrict the set of plausible illuminants. Like Forsyth's original algorithm our algorithm works by exploiting the fact that under a given illuminant the set of observable image colours is a bounded set but we determine the gamut of possible image colours for each of the plausible illuminants rather than for just a single canonical light. Then, given an image whose illuminant we wish to estimate, we simply check whether the image data lies within the gamut of each plausible light. If the image data lies completely within the gamut for a given light, that light is a plausible estimate of the scene illuminant. More often the image data will fall only partially within the gamut for a light and so we propose a method to determine how consistent a given set of image data is with the gamut of each plausible illuminant. We show that a measure of consistency can be obtained by solving a simple non-negative least squares problem and we propose a number of different strategies for determining an estimate of the scene illuminant based on these consistency measures.

The method has a number of advantages over the original gamut mapping approach. First, it is guaranteed to provide a realistic illuminant estimate since we constrain the set of plausible illuminants in advance. Second, we are guaranteed to obtain an illuminant estimate regardless of the image data. In particular, consider the situation in which the image data is only partially consistent with all of the plausible lights. In CRULE this could result in an empty set of feasible mappings and no illuminant estimate. However, in the new algorithm we determine the illuminant estimate based on the consistency measures and therefore do not require complete consistency to obtain a solution. In addition, we are no longer required to adopt the assumption of a diagonal model of illumination change: indeed no explicit model of illumination change is required.

The proposed algorithm also has a number of other advantages over other illuminant estimation algorithms. First, unlike the Neural Network or Correlation Matrix based approaches, the method is not so dependent on accurate knowledge of the statistics of surfaces and illuminants so that "training" the algorithm for practical use is easier than in those methods. A second advantage is that the new algorithm can be applied in essentially the same form whether image data is represented in 3-d sensor space or a 2-d chromaticity space. Thus unlike other methods which are forced into making a decision on colour representation based on implementation issues, the algorithm we develop in this paper makes it easy to choose a colour representation simply on the basis of what gives the best illuminant estimation. Finally, our non-negative least-squares solution is significantly simpler to implement than the original gamut mapping algorithm.

The rest of the paper begins (Section 2) with a formal definition of the illuminant estimation problem and a brief summary of Forsyth's Gamut Mapping solution. Then, in Section 3 we address the weaknesses of this approach by presenting a modified algorithm which we call *Gamut Constrained Illuminant Estimation* (GCIE). In Section 4 we evaluate the performance of the new algorithm on a large set of real images and we also discuss how best to set some of the free parameters in the algorithm's implementation. We conclude the paper in Section 5 with a brief summary.

2. Background

To better understand the illuminant estimation problem we first consider how an image is formed. We adopt a Lambertian model [20] of the image formation process and assume that our imaging device

samples the light incident on its image plane with three types of sensor. We represent the sensitivities of the device by $R_k(\lambda)$ $(k = 1, 2, 3)$ – functions of wavelength (λ) which characterise what proportion of energy at each wavelength a given sensor type absorbs. Light from an object is characterised by the colour signal $C(\lambda)$ which, under the assumptions of the Lambertian reflectance model, can be written as: $C(\lambda) = E(\lambda)S(\lambda)$ where $E(\lambda)$ is the spectral power distribution (SPD) of the ambient illumination, assumed constant across the scene, and $S(\lambda)$ is the surface reflectance function of the object. The colour signal is incident at a point on the image plane and the response of the kth sensor, p_k is given by:

$$p_k = \int_\omega E(\lambda)S(\lambda)R_k(\lambda)\mathrm{d}\lambda \tag{1}$$

where the integral is taken over ω the part of the electromagnetic spectrum over which the sensors have non-zero response. The response of a colour camera at a point on its image plane is thus: $\underline{p} = (p_1, p_2, p_3)^t$: a triplet of sensor responses. Throughout this paper we will refer to this triplet either as $\underline{p}$, or as RGB.

Equation (1) makes it clear that a camera's sensor responses strongly depend on $E(\lambda)$, the ambient illumination. We can pose the illuminant estimation problem as that of inverting Equation (1) to recover $E(\lambda)$. However, light is a continuous function of wavelength but is sampled by the imaging device at only three (broad) bands of the spectrum. So solving Equation (1) is impossible without additional constraints. Fortunately, in most applications we are interested not in the spectrum of the illumination *per se*, but in determining what the scene would look like when rendered under a reference light.

Let us represent an image taken under an unknown light o, by a set $\mathcal{I}^o$, of n sensor responses $\mathcal{I}^o = \left\{ \underline{p}_1^o, \underline{p}_2^o, \ldots, \underline{p}_n^o \right\}$. For this image we would like to determine the corresponding set of data $\mathcal{I}^c$ which would be observed under a reference, or *canonical light c*. Defining the problem in this way has the advantage that explicit knowledge of $E(\lambda)$ is not required. Rather, solving this problem amounts to determining a mapping $\mathcal{F}(\,\cdot\,)$ such that:

$$\underline{p}_i^c = \mathcal{F}(\underline{p}_i^o), \;\; i = 1 \ldots n \tag{2}$$

Forsyth [14] formulated the illuminant estimation problem in this way, and developed an algorithm to solve it. His solution is founded on the observation that due to physical constraints on the nature of lights (a light can emit no less than no energy at each wavelength) and surfaces (a surface can reflect no less than no light incident upon it, and

no more than all incident light), the set of image colours observable under a given light is restricted. That is, for a reference light c, we can define a *canonical gamut* representing all the image colours that can be observed under that light. Forsyth showed that this canonical gamut forms a bounded convex set; for a three sensor device this set is a convex polyhedron in RGB space. A change of illumination implies a different gamut of observable image colours and this new gamut is related to the first by a mapping $\mathcal{F}$. That is, if $\mathcal{C}$ denotes the gamut of possible image colours under the canonical light c, and $\mathcal{O}$, the gamut under the second light o then:

$$p^o \in \mathcal{O} \iff \mathcal{F}^{o,c}(\underline{p}^o) \in \mathcal{C} \tag{3}$$

where $\mathcal{F}^{o,c}$ is the mapping taking colours under light o, to their corresponding colours under light c.

To specify an algorithm to find this mapping we must first establish what form the mapping should take. The CRULE algorithm proposed and implemented by Forsyth is founded on the assumption that the mapping takes the form of a *diagonal matrix*. That is:

$$p^o \in \mathcal{O} \iff D^{o,c}\underline{p}^o \in \mathcal{C} \tag{4}$$

where D^o is a 3×3 diagonal matrix. The validity of this assumption was investigated by Finlayson *et al* [12] who found that the model is well justified for a large class of devices and illuminants. Furthermore, their work showed that for those devices which do not conform to the model it is usually possible to determine a fixed linear transform of the device's sensor responses which renders the diagonal model appropriate once more.

Under the diagonal model, estimating the scene illuminant becomes the problem of finding the three non-zero entries of $D^{o,c}$. Forsyth developed a two-stage algorithm to perform this task. The algorithm works by considering each image colour $\underline{p}_i^o \in \mathcal{I}$ in turn. $\underline{p}_i^o$ can be mapped to any point in the canonical gamut with a diagonal transform. That is, there is a whole set of transforms $\mathcal{D}_i^{o,c}$ which map $\underline{p}_i^o$ into the canonical gamut. A different set of diagonal transforms exists for each different image colour. The diagonal transforms which are consistent with all the image data are the transforms which are common to all the sets of all the individual image colours. That is, we can define a set of plausible transforms as:

$$\mathcal{D}^{o,c} = \bigcap_{i=1}^{n} \mathcal{D}_i^{o,c} \tag{5}$$

Thus the first step of Forsyth's algorithm is to determine the sets of mappings taking each image colour individually to the canonical

gamut. Then the feasible set of mappings taking all image colours to the canonical gamut is found by intersecting these individual sets. It is this intersection step of the algorithm which is most problematic when applied in practice. There are two reasons for this. First, each set of mappings is a bounded 3-d convex set and so Equation 5 represents many computationally expensive 3-d intersections. Computational cost is reduced by choosing only those points on the convex hull of the set of image colours which Forsyth showed leads to the same overall intersection set being calculated. The second problem is that the intersection in Equation 5 might be empty.

There are two reasons why the intersection might be null. The first case is that the diagonal model of illumination change is invalid for one or more colours in the image. In this case it is not possible to find a diagonal mapping which is consistent with all image colours. The second case is that there exists one or more diagonal mappings consistent with all the image data but these mappings are not found. This happens in practice because it is difficult to intersect the individual mapping sets with complete accuracy. When the set of feasible mappings is small numerical inaccuracies are sometimes enough to result in an erroneous null intersection. One of the primary aims of the new algorithm we propose in the next section is to overcome this null intersection problem.

Assuming that the null intersection problem does not arise, the first stage of Forsyth's algorithm results in a set of feasible mappings each of which maps the image data to the canonical gamut. To complete the solution a single mapping is chosen from this set as an estimate of the actual mapping $D^{o,c}$. Of course if many lights are feasible, choosing one of them may result in a wrong illuminant estimate. Thus, when selecting an illuminant care must be taken to choose an answer which is broadly representative of the whole set. This problem is made more difficult by the fact that often the feasible set will contain mappings which, while consistent with the image data, do not correspond to realistic scene illuminants. The algorithm we propose also addresses this problem.

Since Forsyth's original work a number of authors [9, 13, 1] have proposed modifications to the gamut mapping algorithm which aim to address some of the weaknesses we have pointed out so far. However, some or all of the weaknesses of Forsyth's original work remain for these modified algorithms. In particular all gamut mapping algorithms proposed so far rely on being able to accurately intersect a number of convex sets. Since this is difficult to achieve in practice all these algorithms are prone to returning a null solution.

3. Gamut Constrained Illuminant Estimation

The new algorithm which we call *Gamut Constrained Illuminant Estimation* (GCIE) is simple to describe in outline. We begin by defining *a priori* a discrete set of M plausible scene illuminants and we restrict our estimate of the scene illuminant to be a member (or combination of members) of this set. Restricting illumination in this way allows us to exploit the observation that the range of illuminants which occur in the world is quite restricted [9]. The range and number of illuminants we choose may vary depending on the application: in certain situations we may have more prior information about possible scene illuminants and we can exploit this information by changing the plausible set. Similarly we can adjust the number and range of illuminants in this set according to how precisely we wish to estimate the scene illuminant.

For each plausible illuminant we determine the gamut of colours which are observable under it. We can define the gamut for each illuminant in the same way as the canonical gamut is defined in the original gamut mapping algorithm. So, assuming an RGB representation of image colours, the gamut for a given illuminant is the set of RGBs which can be observed under that light. As Forsyth showed, such a gamut will have the form of a closed convex set in RGB space. We explore later in the paper how best to define the gamut for each illuminant both in terms of how we choose the set of surfaces on which the gamut is based and also in terms of whether we use an RGB or some other representation of colour.

We then use these gamuts as a basis for estimating the scene illuminant in an arbitrary image $\mathcal{I}$. We do this by hypothesising each illuminant in turn as the scene illuminant. If we hypothesise the ith plausible illuminant as the scene illuminant we would expect, if this hypothesis is true, that all the image colours fall within the gamut for that light. On the other hand, if the hypothesis is false, then some of the image colours will fall outside the gamut. In either case we can measure the error in our hypothesis according to how far outside the gamut the image colours are. In this way we can determine an error measure for each plausible illuminant which tells us something about how consistent with the image data is each of the plausible illuminants. We can then estimate the scene illuminant by choosing the plausible light which is most consistent with the image data. In some cases two or more lights might be equally consistent with the image data, that is we obtain a set of feasible illuminants. In this case we require a strategy for choosing a single estimate from this feasible set. We note that this case is similar to the situation which arises in the original gamut mapping algorithm. Let us now look at each stage of the algorithm in more detail.

3.1. Characterising the Gamut of an illuminant

There are two important factors that must be considered when characterising the gamut for each plausible illuminant: on which set of surfaces should the gamut be based and in which colour space should we represent those surfaces? When defining the canonical gamut in his CRULE algorithm Forsyth started from the minimal assumptions which can be made about surfaces. Thus, in theory an illuminant's gamut should be based on all physically realisable reflectance functions. In practice, the canonical gamut is based on a set of real surface reflectances which form a gamut representative of surfaces which are encountered in the world. When the CRULE algorithm is applied in practice large gamuts are often used since this reduces the chances of obtaining a null intersection when computing feasible mappings. This is not a consideration in our case and in fact, choosing a large gamut is likely to reduce the effectiveness of the algorithm as it will increase the number of plausible illuminants consistent with the image data and thus reduce the algorithm's ability to discriminate between lights. This suggests that smaller gamuts are more appropriate and in practice the optimum gamut size will be determined empirically. In Section 4.2.1 of this paper we propose a method for varying gamut size in a well founded way and we investigate the effect that gamut size has on the performance of the algorithm.

The second factor we must address is the space in which we represent the gamuts. Given a set of surface reflectances we can calculate RGB values according to our model of image formation (Equation 1). Suppose we base an illuminant gamut on N surface reflectance functions and let $G = \{\underline{p}_1^i, \underline{p}_2^i, \ldots, \underline{p}_N^i\}$ be the set of sensor responses of a device to the N surfaces under the ith illuminant. We can represent these responses by $\Gamma(G)$, the convex hull [26] of G since each element of G can be represented as a convex combination of $\Gamma(G)$. In fact it can be shown [14] that if the N surfaces in G are observable under the ith illuminant then so too are all convex combinations of the elements of $\Gamma(G)$:

$$\underline{p}_k^i = \sum_{\underline{p}_j^i \in \Gamma(G)} \alpha_j \underline{p}_j^i, \quad \forall \ \alpha_j \geq 0, \ \sum \alpha_j = 1 \qquad (6)$$

The gamut of the ith illuminant in RGB space is therefore the set of all convex combinations of $\Gamma(G)$ which we denote $\mathcal{G}$. Figure 1a illustrates a typical gamut in RGB space.

When constructing gamuts in RGB space we make an assumption about the overall intensity of the illuminant. In practice however, an illuminant $E(\lambda)$ can have an arbitrary intensity. This implies that if

the sensor response $\underline{p}^i_j$ is observable under illuminant i then so too is any scalar multiple of this response $s\underline{p}^i_j$. We can allow for an illuminant of arbitrary intensity in two ways. The first way is to use an intensity independent representation of colour. For example, given a sensor response vector $\underline{p}^i_j$, we can factor out intensity with a chromaticity [21] transform:

$$c^i_{j,1} = \frac{p^i_{j,1}}{\sum_{k=1}^3 p^i_{j,k}}, \;\; c^i_{j,2} = \frac{p^i_{j,2}}{\sum_{k=1}^3 p^i_{j,k}} \qquad (7)$$

The chromaticity co-ordinates defined in Equation (7) provide a 2-d representation of image colours. Let $G' = \{\underline{c}^i_1, \underline{c}^i_2, \ldots, \underline{c}^i_N\}$ represent the set of N chromaticity co-ordinates corresponding to the surfaces from which gamuts are constructed. As in the 3-d RGB case, it can be shown that if the N chromaticities in G' are observable under illuminant i then so too are any convex combinations of these chromaticities. Thus, in 2-d chromaticity space the gamut of the ith illuminant is represented by $\mathcal{G}'$ the set of all convex combinations of the elements of $\Gamma(G')$, the convex hull of G'. Figure 1b illustrates a typical gamut in a 2-d chromaticity space.

The second way to allow for an illuminant of arbitrary intensity is to apply a transform to factor out intensity but to continue to represent gamuts in 3-dimensions. In this case we represent colours using the co-ordinates:

$$q^i_{j,1} = \frac{p^i_{j,1}}{\sum_{k=1}^3 p^i_{j,k}}, \;\; q^i_{j,2} = \frac{p^i_{j,2}}{\sum_{k=1}^3 p^i_{j,k}}, \;\; q^i_{j,3} = \frac{p^i_{j,3}}{\sum_{k=1}^3 p^i_{j,k}} \qquad (8)$$

The set of all $\underline{q}^i_j$ lie on a plane in three-dimensional sensor space. We can thus define the response of a device to the set of surface reflectances imaged under illuminant i as: $G'' = \{\underline{q}^i_1, \underline{q}^i_2, \ldots, \underline{q}^i_N\}$. Now, if all of the elements of G'' are observable under illuminant i, then so is any linear combination (with positive coefficients) of those elements:

$$\underline{q}^i_k = \sum_{\underline{q}^i_j \in \Gamma(G'')} \alpha_j \underline{q}^i_j, \quad \forall \;\; \alpha_j \geq 0 \qquad (9)$$

The gamut of the ith illuminant, which we denote $\mathcal{G}''$, is thus an infinite cone in sensor space whose vertex is the origin and whose extreme rays are defined by the convex hull of G''. Figure 1c is an illustration of this unbounded conical gamut. This cone defines the set of observable chromaticities, but places no restriction on how bright a colour can be.

We have proposed three possible representations of the gamut of possible surfaces under a given illuminant. In the next section we detail

how we use each of these representations to estimate the scene illuminant for an arbitrary image. In Section 4.2.2 we investigate empirically which of these three representations gives the best performance.

3.2. CHECKING GAMUT CONSISTENCY

We consider first the case in which we represent the gamut for the ith illuminant by the unbounded cone $\mathcal{G}_i''$ in 3-d sensor space. Let us represent the image whose illuminant we wish to estimate by the set of all its 3-d sensor responses:

$$\mathcal{I}^o = \{\underline{p}_1^o, \underline{p}_2^o, \ldots, \underline{p}_n^o\} \tag{10}$$

Now let us hypothesise that the ith illuminant is the scene illuminant. If this hypothesis is true then any image colour, $\underline{p}_k^o$ will fall within the cone of observable image colours for the ith illuminant. That is, we can find a set of α_j s.t. $\underline{p}_k^o$ is represented as a linear combination (with positive coefficients) of the elements of $\Gamma(G'')$:

$$\underline{p}_k^o = \sum_{\underline{q}_j^i \in \Gamma(G'')} \alpha_j \underline{q}_j^i \quad \alpha_j \geq 0, \ \forall \ j \tag{11}$$

On the other hand, if the hypothesis is not true then an image colour might fall outside the gamut of the ith illuminant. In this case we cannot find a set of α_j to satisfy Equation (11). However, we can find the linear combination of the elements of $\Gamma(G'')$ which is as close as possible to $\underline{p}_k^o$ by finding the α_j which minimise the error term below:

$$e_k^i = \left\| \underline{p}_k^o - \sum_{\underline{q}_j^i \in \Gamma(G'')} \alpha_j \underline{q}_j^i \right\|^2 \tag{12}$$

Solving for the α_j which minimise the error e_k^i is a least-squares problem. However, we have the added constraint that all the α_j should be positive, so we have in fact a *non-negative* least squares problem - a problem for which a known, fast and simple solution exists [23]. We point out that readers who are familiar with optimisation will understand that minimising (12) might be carried out using a variety of numerical algorithms. The non-negative least squares algorithm is chosen because of its very fast operation.

In practice, an image colour might fall outside the gamut of the ith illuminant for reasons other than the fact that it was recorded under a light other than i. First, in real images we might encounter surfaces which we did not consider when constructing the gamuts for

each light. Indeed, this case will arise partly by design since, for reasons discussed in Section 4, we construct gamuts using a subset of all possible reflectance functions. The result of this is that an image colour captured under illuminant i might still fall outside that illuminant's gamut. Second, when constructing image gamuts we assume a Lambertian model of image formation (Equation 1) however, in practice the physical processes governing colour formation can be quite different to this model and it is possible that an image colour formed by a non-Lambertian process may fall outside the gamut of the illuminant under which it was captured. An example of such a situation would be image colours corresponding to specular reflections. We make no attempt to identify such cases, but simply consider that any image colour which falls outside the gamut of the ith illuminant is inconsistent with that light. The performance of the algorithm (see Section 4) suggests that neither of these situation occurs frequently enough to significantly affect algorithm performance. It is also important to understand that in the original gamut mapping algorithm either situation could lead to algorithm failure since the presence of erroneous image colours of these types can make it impossible to find a diagonal transform consistent with all the image data.

Returning to the details of the algorithm, we can express the total error in hypothesising the ith illuminant as the scene light as the sum of all the errors (regardless of their source) for each image colour:

$$e_{total}^i = \sum_{k=1}^{n} e_k^i \qquad (13)$$

If the image data falls completely within the gamut of a light then the corresponding error for this light will be zero and we can say that the hypothesis accounts perfectly for the data. Such a situation will be rare in practice firstly, because we construct gamuts using only a subset of all possible surfaces and secondly, because real images can contain colours resulting from a non-Lambertian reflectance process. In general, one or more of the colours in the image will fall outside the gamut for any given light and the corresponding error for that light will be non-zero. The greater the error, the less well the hypothesis accounts for the data.

We can repeat this procedure for each plausible illuminant and obtain an error measure for each of these lights, which we represent in a vector $\underline{e}$. In general, we expect that the most of the image colours will fall within the gamut of the scene light so that this light will have smaller error. Thus, we can use $\underline{e}$ to estimate the scene illuminant: for example, by choosing the illuminant corresponding to the element of $\underline{e}$ with minimum error. We discuss this step of the algorithm in more

detail in Section 3.3. Next, we consider two different variations of the algorithm which are derived by adopting one or other of our alternative representations of gamuts.

3.2.1. *3-d Gamuts*

Let us consider the case in which we represent gamuts in 3-d sensor space so that the gamut for the ith illuminant is a bounded convex set $\mathcal{G}_i$. Hypothesising that the ith plausible illuminant is the scene illuminant implies that we can represent sensor responses in the image as a *convex* combination of the elements of $\Gamma(G_i)$:

$$\underline{p}_k^o = \sum_{\underline{p}_j^i \in \Gamma(G_i)} \alpha_j \underline{p}_j^i, \ \ \alpha_j \geq 0, \sum \alpha_j = 1 \tag{14}$$

As in the first case, the hypothesis might be false in which case there will be an error associated with representing an image colour as a convex combination of the elements of G_i. This error is given by:

$$e_k^i = \left\| \underline{p}_k^o - \sum_{j=1}^N \alpha_j \underline{p}_j^i \right\|^2 \tag{15}$$

Once again we want to find the set of α_j which minimise the error e_k^i in Equation (15). In this case though we have the constraint that the α_j should be positive but also the additional constraint that the α_j sum to one, because we are looking for a convex combination of the elements of $\Gamma(G'')$.

Thus, the problem can no longer be solved using the method of non-negative least-squares. However we can convert this problem to a non-negative least-squares optimisation in the following way. Let us convert three-dimensional sensor responses $\underline{p}$ into four-dimensional vectors $\underline{r}$:

$$\underline{r} = (p_1, p_2, p_3, W)^t \tag{16}$$

where W is some constant. If we do this for all sensor responses in both the illumination gamut and the image then we can define a new error term $\hat{e}_k^i$:

$$\hat{e}_k^i = \left\| \underline{r}_k^o - \sum_{j=1}^N \alpha_j \underline{r}_j^i \right\|^2 \tag{17}$$

and with some manipulation of the terms in Equation (17) we can write:

$$\hat{e}_k^i = e_k^i + W \left(1 - \sum_{j=1}^N \alpha_j \right) \tag{18}$$

Now, if we choose W to be sufficiently large then the term $W\left(1 - \sum \alpha_j\right)$ will be the dominant one in Equation (18). This error term is minimised when the α_j sum to one. Thus if we minimise the error in Equation (18) by the method of non-negative least squares we will obtain a set of positive α_j whose sum is approximately one. In this way we can use the method to determine the convex combination of points in $\Gamma(G_i)$ which best represent the sensor response $\underline{p}_k^o$. In all other ways the algorithm is the same as that presented above.

3.2.2. *2-d Gamuts*

In the third variant of the algorithm we represent the gamut for each illuminant in a 2-d chromaticity space. In this case we also represent the image data in chromaticity space so that each image colour is represented by a 2-d vector $\underline{c}_j^o$:

$$c_{j,1}^o = \frac{p_{j,1}^o}{p_{j,1}^o + p_{j,2}^o + p_{j,3}^o}, \quad c_{j,2}^o = \frac{p_{j,2}^o}{p_{j,1}^o + p_{j,2}^o + p_{j,3}^o} \qquad (19)$$

In this case the algorithm has the same form as the version based on 3-d gamuts so that the first step in estimating the illuminant is to convert all 2-d chromaticities to 3-d vectors. So a 2-d chromaticity $\underline{c}$ becomes a 3-d vector $\underline{s}$:

$$\underline{s} = (c_1, c_2, W)^t \qquad (20)$$

where, as in the 3-d case W is a constant value. We now hypothesise each plausible illuminant in turn as the scene illuminant and calculate an error measure for each plausible light. Calculating the error in this case follows exactly the same procedure as the 3-d case. Hypothesising the ith illuminant as the scene illuminant is equivalent to stating that each image colour can be represented by a convex combination of elements in G'. The error in this hypothesis for the kth image colour is given by:

$$\hat{e}_k^i = \left\| \underline{s}_k^o - \sum_{j=1}^{N} \alpha_j \underline{s}_j^i \right\|^2 \qquad (21)$$

As before we solve for the minimum value of $\hat{e}_k^i$ using the method of non-negative least squares, choosing a sufficiently large value of W to ensure that the α_j sum to one. In all other respects the algorithm is identical to that presented above.

3.3. Choosing an illuminant estimate

The algorithm as we have described it so far provides a vector $\underline{e}$ whose elements contain a measure of the error associated with representing

the image data by each of the plausible lights. We can use this error vector to obtain an estimate of the scene illuminant, but first we need to specify the form that this illuminant estimate should take. Since we define the set of plausible lights in advance, in theory we know the actual illuminant spectrum for each light. However, our characterisation of each light is a 3-d or 2-d gamut in RGB or chromaticity space. This implies that in practice we can at best recover an estimate of the RGB value of the scene illuminant. If we discard intensity information then we can only recover 2 degrees of freedom in the form of an estimate of the scene illuminant chromaticity. In this paper we will restrict ourselves to the latter case and define our estimate of the scene illuminant to be a chromaticity value $\hat{\underline{c}}^o$: the chromaticity value that the camera records when a uniformly reflecting surface is viewed under the scene light. Below we propose two different methods of estimating $\hat{\underline{c}}^o$ based on $\underline{e}$.

3.3.1. *Minimum Error Estimate*

The simplest approach is to choose the plausible illuminant whose error measure is smallest. If we define a set $\mathcal{E}$ whose elements are the chromaticities of the M plausible illuminants then:

$$\hat{\underline{c}}^o = \operatorname*{mean}_j \; \left(\underline{c}^j\right) \;\; \text{s.t.} \;\; \underline{c}^j \in \mathcal{E} \;\; e_j \equiv \min(\underline{e}) \tag{22}$$

where $\min(\cdot)$ is a function which takes a vector argument and returns the value of the minimum element of that vector and $\mathrm{mean}(\cdot)$ is a function which takes the mean of a number of vector arguments. If the vector $\underline{e}$ has a unique minimum then this estimation procedure corresponds to the chromaticity of the plausible illuminant with minimum error. If more than one plausible light have an error equal to the minimum error then this method averages the chromaticities of all these plausible lights.

3.3.2. f^{th} *Quantile Error Estimate*

An alternative is to choose a feasible set of illuminants (a subset of the plausible lights) based on some quantile value of the error values for all lights. Estimation based on more than one illuminant makes sense if a number of illuminants have a similarly small error. In this case we can choose an estimate which is representative of all these illuminants: for example the average of the illuminants' chromaticity values:

$$\hat{\underline{c}}^o = \mathrm{mean}_j \left(\underline{c}^j\right) \;\; \text{s.t.} \;\; \underline{c}^j \in \mathcal{E}, \;\; e_j \leq \mathrm{quant}\left(\underline{e}, f\right) \tag{23}$$

where $\mathrm{quant}(\underline{v}, f)$ is a function which returns the value q such that a fraction f of the elements of $\underline{v}$ are less than or equal to q. In the

next section we assess the relative merits of these different estimators as part of a thorough experimental evaluation of the new algorithm's performance.

4. Experimental Evaluation

In evaluating the new algorithm we want to answer a number of questions. We have described an algorithm which can be implemented in a variety of different ways: for example, we have proposed three different representations of the gamuts for each plausible light and have also suggested three different ways of estimating the scene illuminant based on the error measures for the set of plausible lights. There are also a number of free parameters in the algorithm's implementation whose settings will affect algorithm performance. For example, how many and which plausible lights we choose as well as which set of surface reflectances we use to define the gamuts. We investigate the effect of all these variations and parameters in our experiments. In addition we compare the performance of the new algorithm to existing illuminant estimation algorithms. Of particular interest is its performance relative to CRULE since it is as an improvement to this approach that the algorithm was developed.

4.1. Experimental Method

All empirical investigations are based on an experiment described by Barnard *et al* [3] who recently published a comparison of many different illuminant estimation algorithms, assessing their performance on a set of real images. This set consists of images of 32 scenes captured under 11 different lights[1] and were captured specifically for testing the performance of illuminant estimation algorithms[2]. Scene content is varied (ranging from a standard photographic test chart to just a few green apples) and thus represents different levels of difficulty for illuminant estimation. The 11 scene illuminants encompass a wide range of illuminant colours that are typically encountered in the world.

This set of images allows us to obtain a robust indicator of the relative performance of the different variations of GCIE and also enables us to easily compare its performance with that of the algorithms

---

[1] In total 321 images were used in the experiment because Barnard *et al* considered that a small number of the captured images were unsuitable for evaluation purposes.

[2] We thank Drs. Kobus Barnard *et al* for making this data publicly available from www.cs.sfu.ca/∼colour/data.

tested by Barnard *et al* in [3]. The experimental paradigm is simple: the RGB image data for each image in turn is used as input into an estimation algorithm which returns an estimate of the scene illuminant. GCIE returns a 2-d chromaticity vector: the algorithm's estimate of the camera's response to a perfectly reflecting surface when viewed under the scene light. We convert this chromaticity value into an RGB triplet: $\hat{\underline{p}}^o = [\hat{c}_1^o \ \hat{c}_2^o \ (1 - \hat{c}_1^o - \hat{c}_2^o)]^t$. Some of the other algorithms to which we compare return an estimate of the RGB value of a perfectly reflecting surface. Algorithm estimates are compared to the RGB of a white tile viewed under the actual scene illuminant[3]. We use an intensity independent measure of algorithm accuracy by calculating the angular distance between an algorithm's estimate of the RGB ($\hat{\underline{p}}_w$) and the RGB of the actual illuminant ($\underline{p}_w$):

$$e_{\text{ang}} = \cos^{-1}\left(\frac{\hat{\underline{p}}_w \cdot \underline{p}_w}{\|\hat{\underline{p}}_w\| \|\underline{p}_w\|} \right) \tag{24}$$

We compare the performance of different algorithms by looking at the distribution of e_{ang} over the set of 321 images. One way to compare these distributions is to use single summary statistics such as the average angular error. In their original experiment Barnard *et al* used the Root Mean Square error statistic to summarise results. This statistic is appropriate when the error measure (angular error in this case) is normally distributed. However, as has been shown in [19] this is not the case for the algorithms tested here: the distribution of angular error is significantly skewed. In this case a more appropriate summary statistic is median error. So, we quote both RMSE and median error statistics in the results below. In addition we investigate the statistical significance of any differences we find between algorithms using the Wilcoxon Sign Test [18]. Given two samples of random variables A and B the Wilcoxon sign test is used to test the null hypothesis: $H_o: \ p = P(A > B) = 0.5$ i.e. the probability that A has a value larger than B 50% of the time. In our case the random variables A and B are the angular error results for two different illuminant estimation algorithms and we use the test to determine whether the performance of the algorithms is the same (the null hypothesis is true) or whether one algorithm performs significantly better than another (the null hypothesis is rejected). In this test the decision to accept or reject the null hypothesis is made on the number of times the random variable A are greater than the corresponding values of B and not on the magnitude of any differences. This test is

---

[3] This RGB was obtained from a white tile placed in each scene: images were obtained with and without the tile in the scene and testing was done on images without the tile.

appropriate when the underlying distributions are unknown, or cannot be well modelled using, for example, a Normal distribution (as is the case here).

4.2. Experimental Results

Before considering the performance of the algorithm with respect to the various factors discussed above we need to address the issue of what is actually the input to GCIE. We could simply use each individual image pixel as input to the algorithm but this is undesirable for two reasons. First, regions in the image of large spatial extent might bias the error measures for each plausible light and thus degrade algorithm performance. Second, the main computational step in the algorithm is to check the consistency of a set of image data with the gamuts of each of the candidate illuminants. Checking the consistency, while a simple step, would be burdensome if repeated for all pixels in an image. Instead we segment the image and use the average RGB value of each segment as input into the algorithm. This implies that each surface in the scene has equal weight in the estimation process, regardless of its spatial extent. We used the Mean Shift segmentation method of Meer *et al* [8] which we found to give reasonably accurate segmentations of all the images, producing on average approximately 15 regions per image. There was no particular reason for choosing this segmentation algorithm beyond the fact that the authors had made available an implementation of their work. In particular, it was not chosen to optimise the performance of the illuminant estimation algorithm. We note that, as other authors have pointed out [3], pre-processing of image data can have a significant effect on algorithm performance. This implies that the results we give below are not necessarily representative of the optimal performance that can be obtained with a given algorithm.

An important parameter in GCIE is the choice of plausible illuminants. This choice will be decided by what prior knowledge we have about the range of scene illuminants we expect to encounter and will therefore vary with the application of the algorithm. In these experiments we report results for three different sets of plausible lights. The chromaticities of these three sets are shown in Figure 2. The first set (red asterisks in Figure 2) is a set of 87 lights representing a large gamut of illuminants obtained from measurements of real world illuminants. These illuminants were measured by Barnard *et al* [2] and represent the case in which we have only limited prior knowledge about the scene illuminant. The second set of lights (blue circles in Figure 2) consists of the 11 lights under which the test images were captured. This set represents the situation in which we have maximal prior knowledge about

the scene illuminant. Finally, the third set (green squares in Figure 2) is intended to represent the situation in which we have some, but not perfect, prior knowledge about the scene illuminant. This set consists of 28 lights and includes a variety of the most commonly occurring lights. For example, there are a range of daylight illuminants, various fluorescent sources as well as a range of Planckian blackbody radiators. The first and third sets contain at least one light which is similar (in chromaticity) to each of the 11 test lights.

4.2.1. *Investigation of Gamut Size*

The first factor in the performance of GCIE which we investigate is how the gamuts for each plausible light are best constructed. We report results for the variation of the algorithm in which we construct gamuts in 3-d sensor space. Similar trends in performance are observed when gamuts are unbounded cones in 3-d space and when they are represented in 2-d chromaticity space.

In developing the algorithm we insisted only that gamuts be constructed from physically realisable reflectances. This is a very weak constraint and by itself will lead to very large gamuts. Such gamuts are likely to be ineffectual since while the gamuts for two different lights will differ at their extremes, these extremes are likely to correspond to reflectances which occur only rarely in the world. For example, while a reflectance with response at only a single wavelength of light can occur in theory, such reflectances never occur in nature. This implies that the discriminating power of GCIE will be limited since most images will be similarly consistent with each plausible light. To avoid this we base our gamuts on reflectance functions which are in a statistical sense "likely" to occur in scenes. That is, we define gamuts which tell us something about which RGBs often occur under a particular light, rather than defining all the RGBs that might ever occur.

To achieve this we construct gamuts by first calculating the chromaticity co-ordinates of a set of measured surface reflectance functions. We then model the distribution of these chromaticities using a bi-variate Normal distribution and construct gamuts using only those reflectances whose chromaticities fall within some number of standard deviations of the mean chromaticity. Suppose we begin with a set of N measured surface reflectances. Let $\mathcal{P}^i$ represent the set of all sensor responses for these N reflectances under the ith plausible light and denote by $\mathcal{C}^i$ the corresponding set of chromaticity values. Suppose further that the elements of $\mathcal{C}^i$ can be represented by a bi-variate normal distribution with mean $\underline{\mu}_i$ and covariance matrix Σ_i. Then define G^i as

follows:

$$G_\sigma^i = \left\{ \underline{p}_j^i \ \mid \ \underline{p}_j^i \in \mathcal{P}, \ \ (\underline{c}_j^i - \underline{\mu}_i)^t \Sigma_i^{-1} (\underline{c}_j^i - \underline{\mu}_i) \leq \sigma^2 \right\} \qquad (25)$$

Finally, let $\mathcal{G}_\sigma^i = \Gamma(G_\sigma^i)$, the convex hull of the elements of G_σ^i. Then the gamut for the ith illuminant is the set of all convex combinations of $\mathcal{G}_\sigma^i$.

Representing gamuts in this way provides a well principled way of constructing a gamut based on statistically likely colours (chromaticities) and allows us to vary the size of the gamut in a controlled way simply by changing the value of σ. For example, by choosing $\sigma = 2.5$ we construct the gamut based on image colours within 2.5 standard deviations of the mean. It is important to note that the validity of this approach depends on whether or not the elements of $\mathcal{C}^i$ are properly modelled by a Normal distribution. This depends to a degree on the choice of chromaticity space used to calculate the elements of $\mathcal{C}^i$ and also (but to a lesser degree) on the initial choice of surface reflectance functions. We found that a Normal distribution is a good model in a log-chromaticity space:

$$c_1^i = \log\left(\frac{p_1^i}{p_1^i + p_2^i + p_3^i}\right), \ \ c_2^i = \log\left(\frac{p_2^i}{p_1^i + p_2^i + p_3^i}\right) \qquad (26)$$

In this investigation (and all subsequent experiments in this paper) we take as our initial set of surface reflectance functions, a combination of three different sets of published reflectance data. These are a set of 462 Munsell chips [32], a set of 219 natural reflectances measured by [25] and a 170 different object reflectances measured by [31].

Figure 3 summarises the effect that varying gamut size in this way has on the performance of GCIE. In these figures the median (left), mean (middle) and RMS (right) angular error of GCIE (using the minimum error estimator explained in Section 3.3.3) over the 321 images is plotted as a function of the standard deviation parameter σ. Results are shown for GCIE based on 11 (solid line), 28 (dashed line), and 87 (dotted line) plausible lights. All three graphs show the same trend: performance varies as a function of gamut size with best results obtained at a standard deviation of 2. This result is observed for each of the three different plausible illuminant sets. As might be expected using the 11 lights under which the images were taken as the plausible set gives the best results. However, results obtained using the set of 28 lights are similar, with those for the set of 87 lights somewhat worse, but still good.

Table 1 summarises the performance of the optimal gamut size. The first three columns shown median, mean and RMS angular error

over the 321 images. The last three columns summarise the statistical significance of these results. A "+" in the ij^{th} column implies that constructing gamuts using the number of plausible lights and the value of σ in the ith row gives significantly better (at a significance level of 0.01) performance than using the values in the jth column. A "−" implies that the values in row i gives significantly worse performance than those in column j while an empty box implies that the two pairs of values give the same performance. These results are based on the Wilcoxon Sign Test (explained above) with a significance level of 0.01. The results show that using 11 or 28 lights gives significantly better performance than using 87 lights but there is statistically no difference between the 11 and 28 light results. If we compare results for different gamut sizes using the Wilcoxon Sign Test we find that gamut sizes between 1.5 and 2.5 standard deviations are statistically equivalent with a gamut size of 2 standard deviations being significantly better than gamut sizes outside this range.

4.2.2. *Investigation of Gamut Dimension*

In the definition of the GCIE algorithm given in Section 3 above we formulated three different versions of the algorithm. In the first version, gamuts are represented by 3-d cones in sensor RGB space and image data is input to the algorithm in its original 3-d RGB form. In the second version both gamuts and image data are represented in a 2-d chromaticity space. Versions 1 and 2 of the algorithm then, both discard intensity information but in slightly different ways. The third version of the algorithm retains intensity information both when constructing gamuts and in the image data. Table 2 summarises the performance of the three different algorithm versions. In the case of Version 3 of the algorithm the results are obtained using gamuts constructed exactly as described above with $\sigma = 2$. The 2-d chromaticity gamut and the 3-d cone gamut are constructed from exactly the same set of chromaticities as in the 3-d case by the method explained in Section 3.1.1 and 3.1.2 respectively.

Table 2 summarises the results for the three different versions of the algorithm in terms of RMS, mean, median, and maximum angular error. These results are obtained using 11 plausible lights and the minimum error illuminant estimator described in Section 3.3.1. Similar relative performance is obtained in the case that 28 or 87 plausible lights are used. The results in Table 2 show that a clear advantage is obtained by using 3-d RGB gamuts and keeping brightness information when processing an image. For example, the median error for Version 3 of the algorithm is 1.31 as compared to 5.02 and 5.82 for Versions 1 and 2 respectively. Table 2 also shows (last 3 columns) the results of a Wil-

coxon Sign Test (at a significance level of 0.01) from which we conclude
that Version 3 of the algorithm is significantly better performing than
either Version 2 or 1. This test also suggests that retaining brightness
information in the image but discarding it when constructing gamuts
(Version 1) brings no advantage over discarding brightness both in the
image and the gamuts (version 2). We note however, that the error
statistics suggest a slight advantage for Version 1 over Version 2 and
we found that if the Wilcoxon test is computed at a significance level
of 0.05 (we accept a 5% chance that we draw the wrong conclusion)
then the difference between version 1 and 2 is significant.

4.2.3. *Investigation of Estimators*

In Section 3.3 we proposed two different ways of estimating the scene
illuminant based on the measures of gamut consistency which the GCIE
algorithm returns. In this section we investigate the relative perform-
ance of these different estimators. The fth-Quantile error estimate has
a single free parameter. Figure 4 shows the performance of the fth-
Quantile Estimate as its free parameter is varied. The three plots in
this figure show median angular error as a function of f for (from left
to right) 11, 28, and 87 lights. The dashed line in these plots show the
result for the minimum error estimate. For the smaller sets of plausible
lights (11 and 28 lights) there is no advantage in this estimator over
the minimum error estimate. However, when using 87 plausible lights
a slight advantage is obtained by using this estimate: optimal perform-
ance is achieved with $f = 0.05$. Table 3 summarises the results for the
best performance of this estimator.

4.2.4. *Comparison with previous algorithms*

The experimental evaluation thus far reveals that the optimal perform-
ance of the GCIE algorithm is obtained by using a full 3-d implement-
ation of the algorithm where gamuts are constructed using statistically
likely colours. Good performance can be obtained using a range of dif-
ferent plausible illuminant sets. In general the minimum error estimator
provides the best performance except for the case when we have a large
set of plausible lights in which case a slight advantage is gained by using
the fth-quantile estimator.

In the final section of this experimental evaluation we compare its
performance to a number of other illuminant estimation algorithms.
The most interesting comparison is with Forsyth's CRULE algorithm
to which the new algorithm is most closely related and which was also
found to perform very well in the original experiments by Barnard *et
al* [3]. We also compare to two benchmark simple algorithms: *Max-*

RGB and *Grey-World* which are widely used and estimate the scene illuminant using the image maximum and image average respectively.

Table 4 summarises the comparative performance of the different algorithms. Here we show the best performing version of GCIE for three different sets of plausible lights. The choice of estimator varies depending on the number of lights in the plausible set. It is clear that the GCIE performs very well in comparison with other algorithms. All three versions of the algorithm outperform Max-RGB and Grey-world as is indicated by both the summary statistics and the results of the significance testing which is shown in the last six columns of Table 4. A plus sign ('+') in a column implies that the algorithm in the corresponding row performs significantly better (at a significance level of 0.01) than the algorithm in the corresponding column. Most interesting is the performance of GCIE compared to the original gamut mapping algorithm. The conclusions we draw vary depending on which error measure we choose: looking at RMS or mean error suggests that GCIE performs worse than gamut mapping. However, looking at the median error statistic leads to the opposite conclusion.

5. Conclusions

In this paper we have presented a new Gamut Constrained Illuminant Estimation algorithm (GCIE). The algorithm we have proposed has clear similarities to Forsyth's original gamut mapping approach however it is formulated in such a way that it avoids the limitations of that original method. First, by hypothesising a set of plausible illuminants we avoid the need to adopt an explicit model of illumination change which as we have discussed, sometimes led to a failure of CRULE to provide an estimate of the scene illuminant. Defining *a priori* a set of feasible illuminants also ensures that the illuminant estimate we recover will correspond to a real, physically plausible light: a property also absent from Forsyth's formulation. Furthermore and in contrast to CRULE, GCIE is guaranteed to return an illuminant estimate regardless of the image data which is input to it. Finally, GCIE is much simpler to implement than CRULE being in essence a repeated application of the method of non-negative least squares for which there exist fast, robust and computationally simple algorithms.

A thorough experimental evaluation of the algorithm on a set of real images has shown how best to set some of the free parameters in the implementation of GCIE and has demonstrated that the algorithm is capable of very good illuminant estimation. The evaluation also showed that gamut mapping algorithms in general, and GCIE in particular give

significantly better (in a strict statistical sense) performance than other benchmark algorithms. This paper has also highlighted the important role that brightness information plays in estimating the scene illuminant. Finally on the basis of the work presented here it is reasonable to conclude that gamut mapping approaches to colour constancy are perhaps the most promising methods currently proposed.

Acknowledgements

References

1. K. Barnard. *Practical Colour Constancy*. PhD thesis, Simon Fraser Univ., School of Computing Science, 2000.
2. Kobus Barnard, Vlad Cardei, and Brian Funt. A comparison of computational color constancy algorithms; part one: Methodology and experiments with synthetic images. *IEEE Transactions on Image Processing*, 11(9):972–984, 2002.
3. Kobus Barnard, Lindsay Martin, Adam Coath, and Brian Funt. A comparison of computational color constancy algorithms; part two: Experiments with image data. *IEEE Transactions on Image Processing*, 11(9):985–996, 2002.
4. Kobus Barnard, Lindsay Martin, and Brian Funt. Colour by correlation in a three dimensional colour space. In *6th European Conference on Computer Vision*, pages 275–289. Springer, June 2000.
5. David H. Brainard and William T. Freeman. Bayesian Method for Recovering Surface and Illuminant Properties from Photosensor Responses. In *Proceedings of the IS&T/SPIE Symposium on Electronic Imaging Science & Technology*, volume 2179, pages 364–376, 1994.
6. G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310:1–26, 1980.
7. Vlad C. Cardei, Brian Funt, and Kobus Barnard. Estimating the scene illuminant chromaticity by using a neural network. *Journal of the Optical Society of America, A*, 19(12):2374–2386, 2002.
8. Dorin Comaniciu and Peter Meer. Mean shift analysis and applications. In *Proceedings of the 7th International Conference on Computer Vision*, pages 1197–1203. IEEE, 1999.
9. G. D. Finlayson. Color in Perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):1034–1038, 1996.
10. G. D. Finlayson, S. D. Hordley, and P. M. Hubel. Color by correlation: A simple, unifying framework for color constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, 2001.
11. Graham D. Finlayson, Mark S. Drew, and Brian V. Funt. Color constancy: generalized diagonal transforms suffice. *Journal of the Optical Society of America, A*, 11(11):3011–3019, 1994.

12. Graham D. Finlayson, Mark S. Drew, and Brian V. Funt. Spectral Sharpening: sensor transformations for improved color constancy. *Journal of the Optical Society of America, A*, 11(5):1553–1563, 1994.

13. Graham D. Finlayson Steven Hordley. Improving gamut mapping color constancy. *IEEE Transactions on Image Processing*, 9(10):1774–1783, October 2000.

14. D. A. Forsyth. A Novel Algorithm for Colour Constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990.

15. Brian V. Funt and Graham D. Finlayson. Color Constant Color Indexing. *EEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):522–529, 1995.

16. T. Gevers and A. W. M. Smeulders. Color Based Object Recognition. *Pattern Recognition*, 32: 453–464, 1999.

17. Glenn Healey and David Slater. Global color constancy: recognition of objects by use of illumination-invariant properties of color distributions. *Journal of the Optical Society of America, A*, 11(11):3003–3010, 1994.

18. Robert V. Hogg and Elliot A. Tanis. *Probability and Statistical Inference.* Prentice Hall, 2001.

19. S. D. Hordley and G. D. Finlayson. Re-evauating colour constancy. In *Proceedings of the 17th International Conference on Pattern Recognition*. IEEE, pages 76-79, August 2004.

20. Berthold K. P. Horn. *Robot Vision.* MIT Press, 1986.

21. R.W.G. Hunt. *The Reproduction of Colour.* Fountain Press, 5th edition, 1995.

22. Edwin H. Land. The Retinex Theory of Color Vision. *Scientific American*, pages 108–129, 1977.

23. Charles L. Lawson and Richard J. Hanson. *Solving Least Squares Problems.* Prentice Hall, 1974.

24. Laurence T. Maloney and Brian A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *Journal of the Optical Society of America, A*, 3(1):29–33, 1986.

25. J. Parkkinen, T. Jaaskelainen, and M. Kuittinen. Spectral representation of color images,. In *IEEE 9th International Conference on Pattern Recognition*, volume 2, pages 933–935, November 1998.

26. F.P. Preparata and M.I. Shamos. *Computational Geometry: An Introduction.* New York:Springer Verlag, 1985.

27. Yogesh Raja, Stephen J. McKenna, and Shaogang Gong. Colour model selection and adaptation in dynamic scenes. In *5th European Conference on Computer Vision*, pages 460–474. Springer, June 1998.

28. Guillermo Sapiro. Bilinear voting. In *ICCV98*, pages 178–183, November 1998.

29. Michael J. Swain and Dana H. Ballard. Color Indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991.

30. S Tominaga, S. Ebisui, and B.A. Wandell. Scene illuminant classification - brighter is better. *Journal of the Optical Society of America, A*, 18(1):55–64, 2001.

31. M.J. Vrhel, R. Gershon, and L.S. Iwan. Measurement and analysis of object reflectance spectra. *Color Research and Application*, 19(1):4–9, 1994.

32. G. Wyszecki and W.S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulas.* New York:Wiley, 2nd edition, 1982.
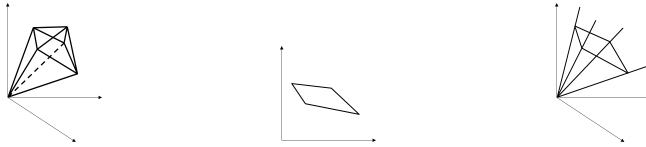
Figure 1. An illustration of a closed 3-d gamut in RGB space (left), a gamut in 2-d chromaticity space (middle) and an unbounded conical gamut in RGB space.
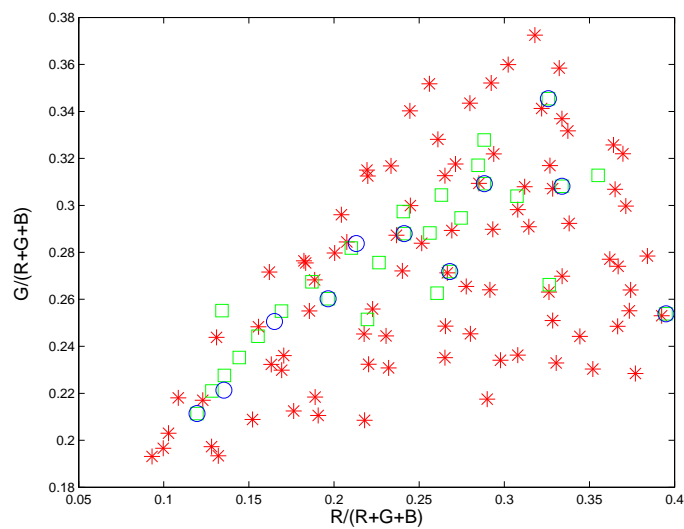

Figure 2. Chromaticities of the three different plausible sets of illuminants used in the experiments. Asterisks: 87 lights, squares: 28 lights and circles: 11 lights.
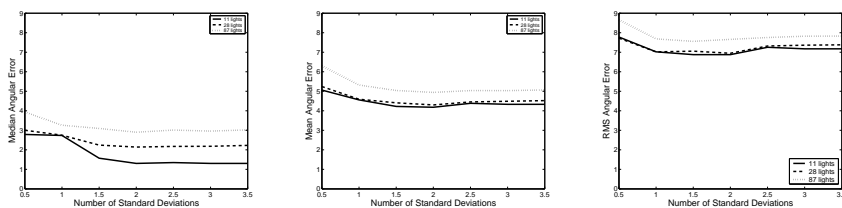

Figure 3. Summary of performance of the Minimum Error Estimate as a function of Gamut Size. From left to right: median, mean and RMS error as a function of number of standard deviations. Each plot shows results for 87 lights (dotted line), 28 lights (dashed lines) and 11 lights (solid lines).
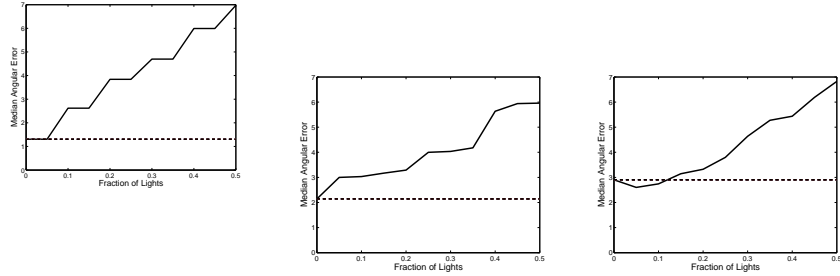
Figure 4. Summary of performance of the fth-quantile error estimator. From left to right: median angular error as a function of the fraction of illuminants for 11, 28 and 87 lights. Dotted lines in each plot shows performance of the minimum error estimate.

Table I. A summary of algorithm performance for the optimal gamut size (2 standard deviations). See main text for an interpretation.

	Median AE	Mean AE	RMS AE	11 lights	28 lights	87 lights
11 lights	1.31	4.18	6.88			+
28 lights	2.14	4.30	6.95			+
87 lights	2.90	4.94	7.66	-	-	

Table II. A summary of angular error results over 321 real images for the three different versions of the GCIE algorithm described in the text.

	Median	Mean	RMS	Max	Version 1	Version 2	Version 3
Version 1, 11 lights	5.02	6.46	9.05	29.16			-
Version 2, 11 lights	5.82	7.00	9.60	29.16			-
Version 3, 11 lights	1.31	4.18	6.88	28.40	+	+	

Table III. A summary of the results for the optimal parameter of
the fth-quantile estimator.

	Median	Mean	RMS	Max
11 lights, Quantile $f = 0.01$	2.62	4.81	7.00	21.21
28 lights, Quantile $f = 0.01$	2.96	4.63	6.54	19.70
87 lights, Quantile $f = 0.05$	2.60	4.75	7.11	19.43

Table IV. A summary of the performance of the new algorithm with respect to a number of
other algorithms.

	Median	Mean	RMS	Max	(1)	(2)	(3)	(4)	(5)	(6)
1. GCIE 11 lights	1.31	4.18	6.88	27.64			+	+	+	+
2. GCIE 28 lights	2.14	4.30	6.95	27.64			+	+	+	+
3. GCIE 87 lights	2.60	4.75	7.11	19.43	-	-		+	+	
4. Max-RGB	4.02	6.30	8.77	26.19	-	-	-		+	-
5. Grey-World	8.85	11.27	14.32	40.16	-	-	-	-		-
6. Gamut Mapping	2.92	4.17	5.6	23.19	-	-		+	+	