Exemplar-Based Colour Constancy and Multiple Illumination

Hamid Reza Vaezi Joze and Mark S. Drew

Abstract—Exemplar-based learning or, equally, nearest neighbour methods have recently gained interest from researchers in a variety of computer science domains because of the prevalence of large amounts of accessible data and storage capacity. In computer vision, these types of technique have been successful in several problems such as scene recognition, shape matching, image parsing, character recognition and object detection. Applying the concept of exemplar-based learning to the problem of colour constancy seems odd at first glance since, in the first place, similar nearest neighbour images are not usually affected by precisely similar illuminants and, in the second place, gathering a dataset consisting of all possible real-world images, including indoor and outdoor scenes and for all possible illuminant colours and intensities, is indeed impossible. In this paper we instead focus on *surfaces* in the image and address the colour constancy problem by unsupervised learning of an appropriate model for each training surface in training images. We find nearest neighbour models for each surface in a test image and estimate its illumination based on comparing the statistics of pixels belonging to nearest neighbour surfaces and the target surface. The final illumination estimation results from combining these estimated illuminants over surfaces to generate a unique estimate. We show that it performs very well, for standard datasets, compared to current colour constancy algorithms, including when learning based on one image dataset is applied to tests from a different dataset. The proposed method has the advantage of overcoming multi-illuminant situations, which is not possible for most current methods since they assume the colour of the illuminant is constant all over the image. We show a technique to overcome the multiple illuminant situation using the proposed method and test our technique on images with two distinct sources of illumination using a multiple-illuminant colour constancy dataset. The concept proposed here is a completely new approach to the colour constancy problem and provides a simple learning-based framework.

Index Terms—Colour Constancy, Exemplar Based Learning, Multiple Illuminants.

1 Introduction

Many computer vision applications as well as image processing problems for both still images and video can make use of colour constancy processing as a prerequisite to ensure that the perceived colour of the surfaces in the scene does not change under varying illumination conditions. The observed colour of the surfaces in the scene is a combination of the actual colour of the surface, i.e., the surface reflection function, as well as the illumination. Estimation of illumination is the main goal of the colour constancy task.

Recently, notwithstanding large amounts of accessible data, many problems can be simply solved by a search through data instead of applying sophisticated algorithms. Sometimes these methods make use of nearest neighbour methods. Such use of these techniques occurs in a variety of computer vision problems such as shape matching [1], character recognition [2], human pose estimation [3], image parsing [4], scene recognition [5] and object detection [6]. As an example, Torralba et al. [7] gathered a large data set of some 80 million tiny 32×32 images, each labelled with a word. They solve different computer vision problems such as scene recognition, object classification, person detection, object categorization, picture orientation determination and even colourization by nearest neighbour methods using this large dataset.

Learning based on a previously seen examples is not a new concept. This concept appears in different domains such as exemplar theory in psychology as a model of perception and categorization, case-based reasoning in artificial intelligence and instance-based methods [8] in machine learning. It has also been suggested that the familiarity of colours of some natural objects (memory colours) might help in estimating the colour of the illuminant in the human visual system [9], [10]. Thus even human colour constancy may learn by examples.

Many colour constancy algorithms have been proposed (see [11], [12] for an overview). The White-Patch, or Max-RGB, method estimates the light source colour from the maximum response of three different colour channels [13]. Another well-known colour constancy method is based on the Grey-World hypothesis [14], which assumes that the average reflectance in the scene is achromatic. Grey-Edge is a recent version which assumes that the average of the reflectance differences in a scene is achromatic [15]. Shades of Grey [16] is another grey-based method which uses the Minkowski p-norm instead of regular averaging. The Gamut Mapping algorithm [17], a more complex and more accurate algorithm, is based on the assumption that in real-world images, for a given illuminant one observes only a limited number of colours. As mentioned in [18], these methods deal with an image as a bag of pixels and the spatial relation

• *H. Vaezi Joze and M. Drew are with the school of Computing Science, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. E-mail: hrv1@sfu.ca , mark@cs.sfu.ca*

Manuscript received Month 00, 2013; revised Month 00, 2013.

between pixels is not considered.

Applying the concept of exemplar-based or instance-based learning to the colour constancy problem seems to be an odd idea at first glance since similar or nearest neighbour images are not usually affected by precisely similar illuminants and moreover gathering a dataset consisting of all possible real world images including indoor and outdoor scenes for all possible illuminant colours and intensities is indeed impossible. In contrast, what can we say about surfaces themselves? Every moderate sized dataset of real images includes thousands of surfaces under different viewing and lighting conditions. We can make these surfaces weakly invariant to illumination changes by simple colour constancy algorithms. Therefore, using the exemplar theory concept we can reduce our illumination estimation task down to the following steps: (1) finding surfaces in an image; (2) finding a similar surface or surfaces in the training dataset for each of our image surfaces; (3) estimating the illumination for each surface based on comparing the statistics of pixels belonging to similar surfaces with the target surface; (4) combining these estimated illuminants into a unique estimate.

In this paper we present a wholly new line of approach to the colour constancy problem, which we call Exemplar-Based Colour Constancy. We use both texture features and weakly colour-constant three-channel RGB colour values in order to find the nearest neighbour surfaces from training data for each surface. Then we estimate the possible illuminant for each surface based on histogram matching of each surface to its nearest neighbour surfaces from training data. The final step is integrating these estimates into a unique illuminant estimation for the whole image. Since we have no labelled or clustered data for our training process as would be the case for a semantic segmentation task or texture detection task, we lack information for providing confidence for our mapping (such as k-nearest neighbour). Nevertheless, although we find some amount of mismatching for surfaces, the illumination estimation process simply considers these cases as outliers compared to the other estimates. Operating on three standard colour constancy datasets, we show that exemplar-based colour constancy produces excellent results that are better than for previous colour constancy algorithms.

Most colour constancy algorithms assume that the spectral distribution of light source is uniform across the image and therefore that the colour of illuminant is constant all over the image, Hence, estimation of this uniform illuminant is the main goal of such colour constancy method as discussed below. Although this assumption works well in most cases and is widely used in commercial cameras, nevertheless there exist common cases in which this assumption is violated in real images, including: skylight from windows plus indoor light; in-shadow plus non-shadow lights; or two different light sources in an indoor room. This situation, multiple illuminants or multiple light sources with different colours, is a common source of failure for current colour constancy methods. Exemplar-based colour constancy has the advantage of succeeding even in the multiple illuminant situation, which is not possible for most current methods. Hence as another contribution we show a technique to overcome the multiple illuminant situation using our proposed method, and test our technique on standard images having two distinct sources of illumination.

The implications for a useful discounting or regularizing for light in images are substantial in various tasks in computer vision such as image retrieval, colour reproduction and object detection [19]

The outline of the paper is as follows: we discuss related work in §2 and then in §3 we introduce the proposed method by explaining our surface model, the process of learning surface models for training images, and the proposed illumination estimation procedure. In §4 we apply our proposed method to three standard colour constancy datasets, comparing performance to current colour constancy methods. In §2.1, we discuss our proposed method in the multiple illuminant situation, and carry out experiments. Finally, we conclude the paper and discuss future work in §6.

2 Related Work

Illumination estimation methods can categorized into two groups: (1) static methods which try to estimate the illuminant for each image based on its statistical or physical properties and (2) learning-based methods which try to estimate the illuminant using a model that is learned on training images. Grey-based methods, which form a main part of static methods, have been formalized into a single framework [16], [20]:

$$\left(\int \left\| \frac{\partial^n I_k(x)}{\partial x^n} \right\|^p dx \right)^{\frac{1}{p}} = e_k \tag{1}$$

where e is estimated illuminant colour, k denotes R, G or B, p denote the Minkowski norm and n is grey-edge order. If $n = 0$, for $p = 1$ the equation is equal to the grey-world assumption, for $p = \infty$ it is equal to colour constancy by White-Patch and it is Shades of Grey and for $1 < p < \infty$. For higher n it is Grey-Edge.

Static colour constancy methods also include some physics-based methods such as methods that use specularity to estimate illuminant chromaticity [21], [22]. Drew et al. [22] present an effective physics-based colour constancy method, called the Zeta-Image, which makes use of a log-relative-chromaticity planar constraint involving specular reflection. This method is fast and requires no training or tunable parameters.

One of the first colour constancy methods which estimates the illuminant by a model that is learned on training images is the Gamut Mapping algorithm [17]. It is based on the assumption that in real-world images, for a given illuminant one observes only a limited number of colours; therefore, colours forming a

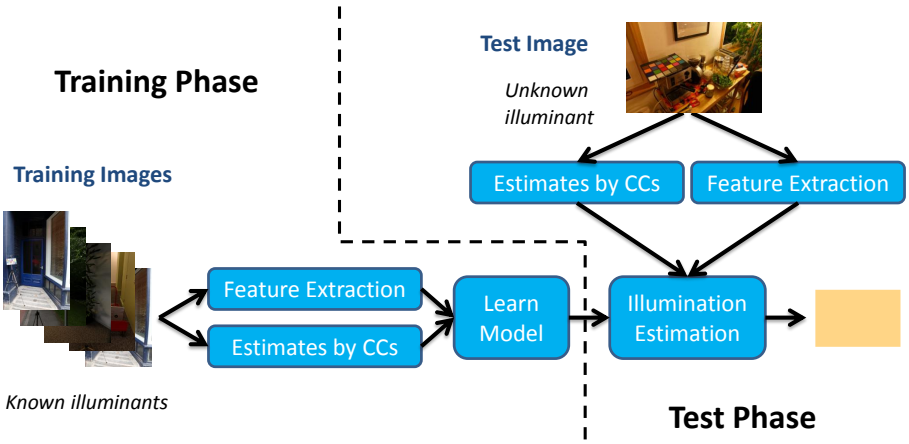

Fig. 1. The common procedure of learning-based colour constancy methods that try to find the best algorithm or a combination of algorithms for each image using extracted features in both training and test phases.

"canonical" gamut which contains possible colours can be observed under a canonical illumination in a training phase, and an estimate of a test-image illuminant can be derived by mapping current pixel colours to that canonical gamut. Several extensions have been proposed for gamut mapping algorithms. Several extensions to gamut mapping algorithms have been proposed [23], [24], [25]. Colour-By-Correlation [23] is a discrete implementation of gamut mapping, where the canonical gamut is replaced by a correlation matrix.

Another learning-based approach to the illumination estimation problem is the Bayesian approach [26], [27], in which the variability of reflectance and illuminant is modeled as independent random variables. These methods estimate illuminant colour from the posterior distribution condition learned from training images. Here the illuminant prior could be uniform over a subset of illuminants [26] or could be an empirical distribution of illuminants in training images [27]. Other machine learning techniques includes using neural networks [28], in which binarized chromaticity histograms are used to estimate 2D illuminant chromaticity via a neural network system, or support vector regression [29]. In a different learning approach, a maximum likelihood approach presented by [18], [30] uses spatial information along with standard statistical knowledge by developing a statistical model for the spatial distribution of colours in colour constant images.

Besides static colour constancy methods such as Max-RGB, Grey-World, Grey-Edge and Shades-of-Grey, which as mentioned above are based on simple assumptions, recently efforts at fusing these algorithms have generated better performance than for the individual algorithms. One of the first attempts in this direction was carried out by Cardei and Funt [31], who applied a weighted committee mechanism over several of these methods. More complex methods try to learn to select the best algorithm or combination of algorithms for each image using pixel information as well as spatial information, and hence they do not deal with the image as simply a bag of pixels.

As mentioned in [12], these learning-based colour constancy methods that try to find the best or a combination of algorithms for each image using extracted features go through a similar procedure. They extract texture, shape or colour features from sets of training images, and estimate the colour of the illuminant for each image using several statistical illumination estimation algorithms. They then learn a model based on extracted features as well as the error of these estimates compared to known ground truth. This type of model could e.g. learn the set of weights associate with estimates of these illumination estimation algorithms [32], [33] or directly learn the colour of the illuminant [34], [35]. Figure 1 shows this procedure in both the training and test phases. It can be stated that the main differences amongst this kind of algorithm are in the feature extraction blocks, where the feature could be simple, such as a colour histogram [36], [34], or edge direction histogram [34], or more complex features such as Weibull features [32], [35], [33], Wiccest features [33], or Wavelet statistics [34].

As an example, Gijsenij and Gevers [32] clustered the images by a k-means algorithm using natural image statistics to characterize the images on the basis of Weibull distribution parameters. They then correspond each cluster with the best single algorithm for training images for that cluster. To estimate the illuminant of a test image, they select the algorithm according to its cluster or combination of the individual algorithms according to the distances to neighbouring clusters.

In a different approach to selecting best algorithms, Wu et al. [33] introduce a multi-resolution texture descriptor based on an integrated Weibull distribution to extract texture information. They used an image similarity measure derived from the Wiccest feature and spatial pyramid matching to find the K most similar images for a test image from amongst training images, and with these neighbouring images they provide a

combination for uniting the data-driven strategy and prior knowledge.

Van de Weijer et al. [37] extend the grey world hypotheses to say: the average reflectance of semantic classes in an image is equal to a constant colour, rather than being just grey. Therefore, for each of the semantic classes present in an image they compute that illuminant that transforms the pixels assigned to that class into the average reflectance colour of that semantic class in the training images. They call this a top-down approach as opposed to bottom-up approaches in many other colour constancy methods. They also make use of high-level visual information to select the best illuminant out of a set of possible illuminants generated by other methods. In a similar approach [38], the special visual object categories (called here memory-colour categories) which have a relatively constant colour such as sky or grass and foliage (which were used in their experiment) are detected using the Bag-of-Features machine learning method. Then the initial estimate provided by a statistical colour constancy method can be adjusted to map the observed colour of the category to its actual colour which is determined in the training phase. The main difference between this work and [37] is that the visual object categories are known and hand labelling and tagging with the category label is required for training images.

2.1 Multiple Illuminants

Most colour constancy algorithms assume that the spectral distribution of light is uniform in the image and therefore the colour of the illuminant is constant across the image, Estimation of this single illuminant is the main goal of most colour constancy methods, and very few methods explicitly focus on local illuminant estimation. Although this assumption works well in most cases and is widely used in commercial cameras, there exist common cases in which this assumption is violated in real images. These include: skylight from windows plus indoor light; in-shadow pixels plus out-of-shadow pixels; ambient light and flash-light in photography; and two different light sources in an indoor room. This situation, which we call multiple illuminants or multiple light sources with different colours is a common failure case for current colour constancy methods.

Figure 2 shows a scene with two distinct illuminants: outdoor sky-light and indoor luminaire light. Assuming uniform illumination in the scene, the image on the left is colour-corrected using the outdoor sky-light and the image on the left is colour-corrected using indoor light. It is obvious that the colour constancy task, with a uniform-illumination assumption, fails for this scene even when we use the ground truth as our colour-correction illuminant.

In early research, it is shown in [39], [40] that a difference in illumination, once it has been identified correctly, provides additional constraints which can be applied to obtain better colour constancy; however these

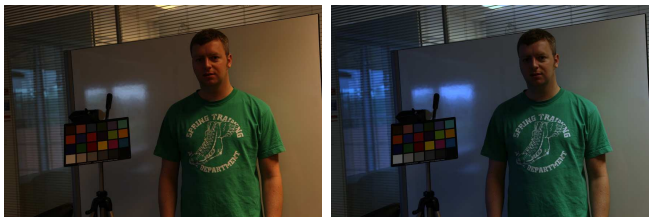

Fig. 2. A scene with two distinct illuminants. The image on the left is colour-corrected using outdoor sky-light and the image on the left is colour-corrected using indoor light.

works do not provide algorithms to detect multiple illuminants automatically. Ebner [41] assumes a grey-world assumption works locally, and not just globally, and attempts a diffusion based methodology for pixel intensities. However a local grey-world can be noisy and inaccurate, especially for near uniform-colour scenes. In another approach, Bleier et al. [42] propose a method to overcome the multiple illumination problem in which the image is segmented into a set of superpixels based on colour; then a collection of colour constancy algorithms (Bayesian, plus different versions of grey-world, grey-edge and gamut mapping) is applied to each superpixel independently. The illumination estimate for each superpixel is computed using different fusion techniques such as the average, or a machine learning method such as Gradient tree boosting or Random forest regression. A local colour constancy algorithm is also presented by [43], which adjusts colours pixel-by-pixel based on its local area to solve the multi-illuminants for High-Dynamic-Range images. This estimates the illumination for each pixel from the colour information of its neighbouring pixels, weighted by the spatial distance, luminance intensity difference, and chromaticity.

A specific case of the multiple-illuminant scenario is encompassed in research on in-shadow and out-of-shadow regions. That is, the shadow removal problem [44], [45], [46] can be considered as a colour constancy problem involving two light sources.

Recently, Gijsenij et al. [47] proposed a colour constancy method for multiple light sources. They obtain image patches by grid-based, keypoint-based, or segmentation-based sampling, and then estimate the illuminant for each image patch by some of the grey-based methods in eq. (1), assuming a uniform illuminant. Focusing on scenes with two distinct light sources and their combination, their final estimation of two distinct illuminant colours is either by clustering the illuminant estimates for patches or taking into account the spatial relations between these estimates by applying segmentation.

2.2 Evaluation

In order to evaluate performance of a colour constancy method we need to calculate the error of our estimation. Considering the actual chromaticity of illuminant $\mathbf{e}$ and the estimated chromaticity of illuminant $\mathbf{e_{est}}$ by any of

of the aforementioned illumination estimation methods, there are different measures used to calculate the error. Two measures commonly used to calculate this error are angular error, which is the angle between 3-vector $\mathbf{e}$ and 3-vector $\mathbf{e_{est}}$, and Euclidean error, i.e., Euclidean distance in r, g chromaticity space; note that these two measures are highly correlated to each other [48]. Here we could also mention perceptual Euclidean distance [49], which uses perceptually optimized weights for Euclidean distance, and also the RMS error for all pixels in r, g, b or R, G between the corrected image and the ground truth image. In this paper, we use angular error as our measurement for computing estimation error because of its frequent use in the literature [48], [50]. It is also shown by [49] that the angular error is a reasonably good indicator of the human perceptual performance of colour constancy methods.

$$err_{angle}(e, e_{est}) = \text{acos}(\frac{e \cdot e_{est}}{\|e\|\|e_{est}\|}) \quad (2)$$

Having a set of images with known colour of illuminant, we can calculate the error of estimation (using either of the generally used error metrics) for each image. Then the overall metric of performance of an algorithm for that set of images can be the mean of errors. However the mean by itself is not a good index for evaluating performance of methods [48]. The median or trimean of errors is usually preferred in the literature [48], [49] because the mean is sensitive to outliers. The median indicates the performance of the methods for half of the images or equally the 50^{th} percentile error. The trimean of a distribution is defined as a weighted average of its median and its two quartiles. It is also important to minimize the worst-case errors or, equally, consider the maximum of errors as the measure [51]. In this paper we calculate both mean and median as well as the 75th percentile error as our measurement to compare different illumination estimation algorithms.

In [52], we introduced the exemplar-based colour constancy approach, and in this paper we go on to investigate the method much more substantially. Firstly, the method's details are set out considerably more comprehensively, and with illustrations. The method is first delineated in the context of previous learning-based approaches. Then entirely new tests are carried out, as well as challenging the method by applying colour constancy via exemplars obtained from one dataset to test images for a different image dataset – an inter-dataset test which is very demanding. Moreover here we go on to investigate how to structure the new algorithm in the face of the very difficult multiple-illuminant situation, and an entirely fresh set of tests on multiple-illuminant images with ground truth is carried out.

3 Proposed Method

The proposed method falls into the learning based colour constancy category, in which a model needs to be learned from training images. The main distinctions between this work and other learning based colour constancy methods that use spatial information by local feature descriptors, such as [32], [33], [37], is that they use this information to determined the best or combination of best possible illumination estimation algorithms (the procedure is shown in Fig. 1), while we use selected instances for illumination estimation. Compared to the top-down approach [37], [38] in which they assign a semantic classes (or memory-colour categories) to each patch of an image based on models learned in the training phase, in our proposed model we assign to segmented regions we call "surfaces", from training images to each surface of the test image. As well, [37] used the extended version of the grey world assumption to estimate the illuminant whilst we use the ground truth of corresponding surfaces for illumination estimation.

On the other hand, scenes with a single or a just few number of surfaces (such as images captured of grass or sky) are a common failure for grey-based methods which form the core of most recent learning based methods, since the grey assumption is not satisfied for these images. The Gamut Mapping method also fails for these images since only a limited number of colours are seen in the image and that is not enough to map the input gamut to the canonical gamut. We will see that our exemplar-based method can overcome this problem since the exemplar-based method estimates the illuminant based on similar surfaces and there is no assumption that more than one surface is needed (although more surfaces do make the estimate more robust).

3.1 Surface Model

We find surfaces for both training and test images by mean-shift segmentation, implemented via [53]. Since the pixels in the margin of segmented areas affect texture information, we remove margin pixels of segments by dilating segment edges as well as small segments. In order to define a model for each surface we use both texture features and colour features. For the purpose of texture features, the MR8 filter bank [54] on three channels is selected for use because of its good performance in texture classification applications [55] and also its fast implementation.

The MR8 filter bank consists of 38 filters (6 orientations at 3 scales for 2 oriented filters, plus 2 isotropic) but only 8 filter responses. The filter bank contains filters at multiple orientations but it records only the maximum filter response across all orientations. We use the normalized histogram of frequency of appearance in that particular surface for each colour channel as our colour features. Here we store the colour constant diagonal transformations generated by the Max-RGB method $\mathcal{M}$ for each surface model, which will be used in illumination estimation process. Since we deal with illumination variation, we apply this transformation (divide each channel by its maximum value) before computing each

histogram. This makes our surface model weakly colour constant [56]. This means that we are not interested in specific colours for our surface matching, but instead on its relative distribution.

In the learning stage, training images are convolved with a filter bank to generate dense filter responses. Exemplar filter responses are chosen as textons via K-Means clustering (with $K = 1000$) and are collected into a dictionary. The histogram of frequency of textons belonging to this dictionary is a common description for texture detection [57] although other local descriptors such as scale-invariant feature transform (SIFT) [58] may used instead of the MR8 filter bank.

Given a surface in a training image, its corresponding model is generated by first convolving it with the filter bank and then labelling each filter response with the Euclidean nearest neighbour texton in the texton dictionary. The histogram of textons, i.e. the frequency with which each texton occurs in that surface, forms the first histogram in the corresponding model for that training surface. We then add a coarse three-channel histogram to that surface's model (we use only 10 bins for each channel to be robust again noise). In order to make our model weakly invariant to variation in illuminant colour, we stretch the histogram for each channel to have maximum equal to 1 or, equally, divide the values of each colour channel by its maximum value among pixels within that surface (thus making it colour constant for Max-RGB via using a diagonal transformation). Therefore, each surface model includes four normalized histograms that are then stored in a single vector. We also need to store some meta-data for each model, consisting of ground truth illumination colour for that image as well as the maximum response for each channel used for stretching histograms. Fig. 3 shows a surface and its normalized histogram of textons and three weakly colour constant normalized histograms of colour channels. In summary, the training phase for exemplar-based colour constancy is expressed in Algorithm 1. Here the $MaxRGB$ function makes the input pixels colour constant by dividing by the illuminant as estimated by Max-RGB method.

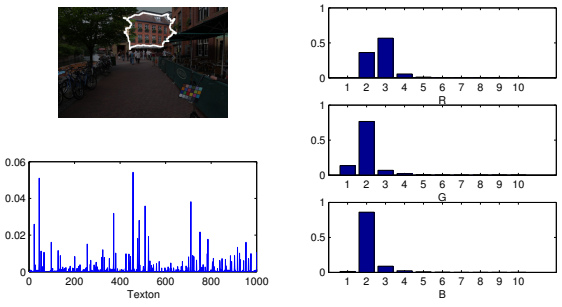

Fig. 3. A surface and its normalized histogram of textons and three weakly colour constant normalized histograms of colour channels.

In the test stage, the same procedure is followed to build the model (one histogram of textons and three histograms of colour channels) corresponding to each surface in the test image. This model is then compared with the models corresponding to training surfaces by nearest neighbour classifier with the chi-squared statistic employed to measure distances. We select M nearest neighbours from training surfaces (with $M = 10$). Fig. 4 shows some test image surface examples and their eight nearest surface models from training data. We carry out our experiment on the re-processed version of the dataset [27], [59] (denoted "ColorChecker") (refer to § 4 for details). Since we have no labelled or clustered data for our training process as would be the case for a semantic segmentation task or texture detection task, we lack information for providing confidence for our mapping. Nevertheless, although we find some amount of mismatching for surfaces as shown in Fig. 4, the illumination estimation process simply considers these cases as outliers compared to the other estimates.

Algorithm 1 Training Exemplar Based Model

Generate Texton Dictionary

1: *features* ← *convolve all training images with MR8 filter*
2: *textons* ← *K-Means clustering of features* ($k = 1000$)

Finding Surfaces

1: *surfaces* ← *mean-shift seg. of all training images*

Generate Surface Models

1: ***for all*** S *in surfaces* ***do***
2: *features* ← *convolve* S *with MR8 filter*
3: *label* ← *NN(features,textons)*
4: *texture hist* ← *normalized histogram of labels*
5: $S_{cc} \leftarrow MaxRGB(S)$
6: *colour hist* ← *normalized histogram of each colour channel in* S_{cc} *(10 bins)*
7: $trainmodel_S$ ← *(texture hist, colour hist)*
8: ***end for***

3.2 Illumination Estimation

Given a test surface model and its nearest neighbour surface model from training models, we can transfer the test surface's colours to its corresponding training surface's colours linearly by a 3×3 matrix. We can approximate this matrix by a diagonal matrix as discussed in [60] and solve the transformation for each channel separately based on their channel histograms. Therefore, we can write this matrix which transforms test surface to training surface as follows:

$$D = \mathcal{M}_{test}^{-1} D_H \mathcal{M}_{train} \tag{3}$$

where $\mathcal{M}_{test}$ and $\mathcal{M}_{train}$ are the weakly colour constant diagonal transformations of the test and train surface colours via the Max-RGB division method and D_H is the transformation of the test surface's histograms to training surfaces' histograms. Since we use this histogram

Fig. 4. Surfaces from test images (on the left) and their 8 nearest surface models from training images.

to find similar models these histograms are usually approximately identical or not far from each other; this means that D_H is approximately the identity matrix (we use $D_H = I$ in our experiments). Finally, since the illumination colour of a training surface e_{train} is known, the estimation for test surface illumination colour is:

$$e_{test} = De_{train} = \mathcal{M}_{test}^{-1} D_H \mathcal{M}_{train} e_{train} \quad (4)$$

Given a test image, we will have n large enough surfaces and M nearest neighbour surfaces from training data, or equally M illumination estimates by eq. (4) corresponding to each. The final estimate can be the median or the mean after removing outliers of all of these estimates in rg chromaticity space ($\{r, g\} = \{R, G\}/(R+G+B)$). We can also use weighted averaging by defining weights for each estimated illuminant according to the confidence of estimation for each surface, which we compute based on the standard deviation of estimates for that single surface and also similarity which we compute based on chi squared distance between their normalized histograms. Experiments show that none of these techniques outperforms the others, and therefore for simplicity we estimate the final illuminant by finding the median over all estimates for the three channels separately.

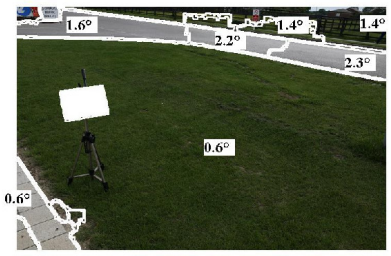

Fig. 6. A test image and angular errors of estimated illuminant for its surfaces (computation is not carried out for small segments, and the ColorChecker is masked off).

A scene with a single surface constitutes a failure case for most non-learning methods since their assumptions are not satisfied. In case of textured surfaces, the proposed method overcomes the problem by considering texture histograms. Note that a completely uniform surface is obviously an ill-posed problem (uniform yellow wall under white light or uniform white wall under yellow light) and no method can solve this problem. Fortunately since the proposed method uses many estimates we find experimentally that it performs well with a limited number of uniform surfaces in the scene, with their estimates considered as outliers.

Figure 5 shows the procedure for illuminant estimation for a test image using exemplar-based color constancy. Figure 6 shows a test image and the angular error of estimated illuminant for its surfaces; as mentioned, we do not form any estimate for small segments. We see that more textured surfaces such as grass or textured road have more precise estimates compared to smooth road. In summary, the proposed illumination estimation for a test image is expressed in Algorithm 2.

Algorithm 2 Illumination Estimation for image I

```
surfaces ← mean-shift segment of I
for all S in surfaces do
    features ← convolve S with MR8 filter
    label ← NN(features, textons)
    texture hist ← normalized histogram of labels
    S_cc ← MaxRGB(S)
    colour hist ← normalized histogram of each colour
    channel in S_cc (10 bins)
    model_S ← ( texture hist , colour hist )
    for all i in KNN(model_S,trainmodels) do
        estimates_Si ← eq. (4)
    end for
end for
return median(estimates)
```

3.3 Colour Correction

Once the colour of the light has been estimated, the input image, which was captured under an illuminant

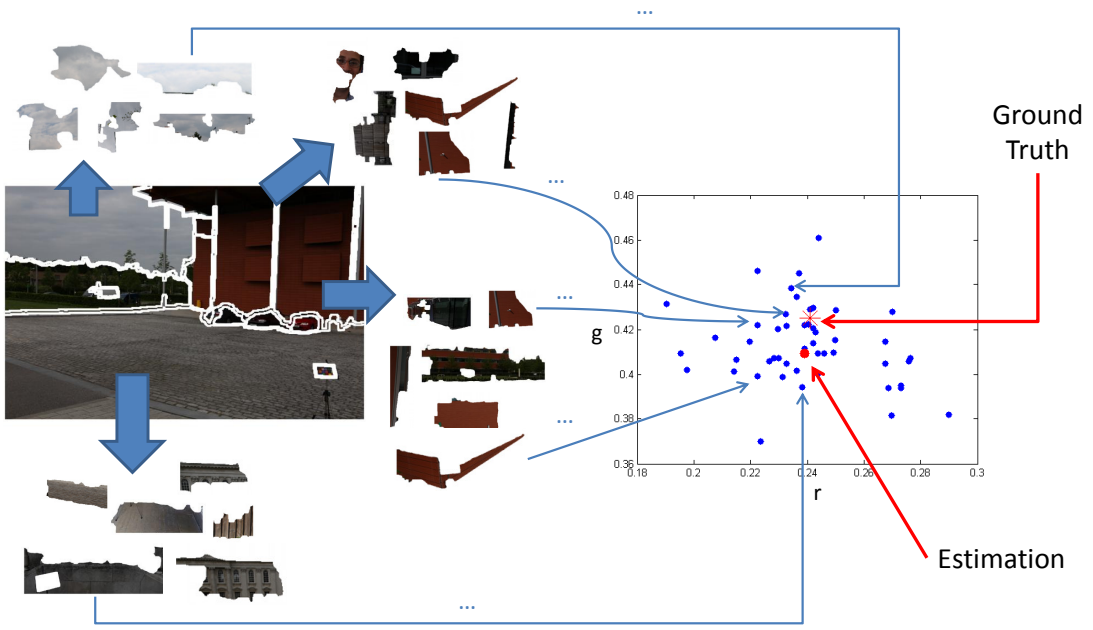

Fig. 5. The procedure of estimating illuminant for a test image using exemplar-based colour constancy. A test image and its nearest neighbour surface models from training images on the left and estimated illuminants according to each model in rg chromaticity space on the right.

as estimated, should be transferred back to an image as it would appear under the canonical illuminant. This procedure is usually done using the diagonal model. Although exemplar based colour constancy estimates a unique illumination for an image, and this can be used for colour correction by a diagonal model, the method actually estimates a distinct illuminant for each surface, and this can be used for colour correction for each surface separately. The diagonal model for pixels that belong to surface edges or to small surfaces can be constructed as a weighted linear combination of those for its neighbour pixels.

We call the matrix in which every pixel is assigned to its own illuminant estimate "back-projection", where this usage derives from [47]. Obviously, in the case when the uniform illumination assumption is made, in our experiment in §4, the back-projection is a constant. The colour correction process using estimated back-projections, $\mathcal{D}$, which have equal dimension to the input image I, is as follows:

$$I^*_{ijk} = I_{ijk} * e^*_k / \mathcal{D}_{ijk} \quad , \; k = \{R, G, B\} \tag{5}$$

where I^* is colour-transferred images under canonical illuminant e^* and ij is pixel address. Figure 7 shows an input image, its estimated back-projection using uniform and surface illumination, and uniform illumination and colour transferred output images using these two back-projections. The estimated back-projection for each surface is computed based on similar known surfaces (similarity for both colour and texture); therefore even if they are not accurate they seem reasonable, as seen in the colour of the grass in Fig. 7(d). Even though this colour is not necessarily accurate compared to the known illumination, nonetheless it seems reasonable to us since it is similar to some other grass surfaces in the training images.

4 Experiments

We applied our proposed method to four standard colour constancy datasets of real images of indoor and outdoor scenes. The first dataset is the Gehler colour constancy dataset [27], denoted the ColorChecker dataset, This dataset consists of 568 images, both indoor and outdoor. The illuminant ground truth for these images is known because each image has a Macbeth ColorChecker placed in the scene — which must masked off in tests. The images are captured by auto white balance setting of the camera. For this dataset, we used three-fold cross-validation to learn our models using this original dataset, as used by other learning based methods we compared to. The second dataset is the re-processed version of the above ColorChecker dataset, provided by Shi and Funt [59]. This dataset which includes the same number of images, but contains raw image data of ColorChecker dataset in an attempt to recover linear sensor values, which is in principle critical for our 3×3 matrix transformation.

Table 1 indicates the accuracy of the proposed methods for the ColorChecker dataset and its re-processed version, in terms of the mean, median and 75th percentile of angular errors, for several colour constancy algorithms applied to this dataset. For those methods which need tunable parameters (refer to [61]), we utilize

(a) (b) (c) (d) (e)

Fig. 7. (a) An input image, (b) constant back-projection (c) back-projection by surface illuminant estimates (d) colour corrected image via b and (e) colour corrected image via c.

TABLE 1
Angular errors for the ColorChecker dataset [27] in terms of mean, median and 75th percentile errors for several colour constancy algorithms.

Method	Original ColorChecker			Reprocessed ColorChecker		
	Median	Mean	75th	Median	Mean	75th
Do nothing	6.8°	9.5°	13.6°	13.5°	13.4°	14.8°
White-Patch	6.0°	8.1°	10.8°	5.7°	7.4°	11.7°
Grey-World	7.3°	9.8°	14.6°	6.3°	6.4°	8.4°
Grey-Edge	5.2°	7.0°	9.5°	4.5°	5.3°	7.0°
Zeta-Image [22]	5.0°	6.9°	9.0°	2.8°	4.1°	5.6°
Bayesian [27]	4.7°	6.7°	8.8°	3.5°	4.8°	6.7°
Gamut Mapping	4.9°	6.9°	8.9°	2.5°	4.1°	6.0°
Gamut Mapping 1jet [24]	4.9°	6.9°	9.0°	2.5°	4.1°	6.0°
Spatio-spectral Statistics ML [30]	-	-	-	3.0°	3.7°	4.9°
Bottom-up+Top-down [37]	4.5°	6.4°	8.8°	2.5°	3.5°	4.1°
Natural Image Statistics	4.5°	6.1°	8.2°	3.1°	4.2°	5.8°
Exemplar-Based	**3.7°**	**5.2°**	**7.0°**	**2.3°**	**3.1°**	**3.9°**

optimal parameters for this dataset and for learning-based methods we utilize standard testing setup used in [12] which include similar training-testing sets. Here dash in table cells means that the result for that dataset was not reported by their authors. We see that our exemplar-based method makes a substantial improvement over previous algorithms.

Another dataset, which contains lower quality real images (image resolution 360×240), is the GreyBall dataset of Ciurea and Funt [62]; this contains 11346 images extracted from video recorded under a wide variety of imaging conditions. The ground truth was acquired by attaching a grey sphere to the camera, displayed in the bottom-right corner of the image (again masked for experiments). In order to learn our models for this dataset, we use 15 folds each of which represents a recorded video as provided by the dataset itself, as for the other learning based methods we compared to. The ground truth of the GreyBall dataset is obtained using the original images (colour model is NTSC-RGB). Therefore, the ground truth was recomputed by [12] by converting the images from NTSC-RGB to linear RGB assuming gamma is equal to 2.2. This modified dataset is named linear GrayBall.

Table 2 shows the performance of our proposed method for the GreyBall dataset and its linear version in terms of the mean, median and 75th percentile of angular errors, for several colour constancy algorithms applied to this dataset. Again we utilize optimal parameters for this dataset for those methods which need tunable parameters (refer to [61]) and for learning-based methods we utilize a standard testing setup as in [12]. Here again, we see a substantive improvement over previous approaches, even those using complex methods.

To our knowledge, for these four standard datasets, widely used for evaluating colour constancy methods, Exemplar-Based Colour Constancy does best in terms of all of mean, median and 75th percentile angular error compared to any reported colour constancy methods, even those using a combination of algorithms such as Natural Image Statistics [32] or Bottom-up+Top-down [37] . Table 1 and 2 indicate that not only does Exemplar-Based Colour Constancy generally perform better (smaller mean and median angular errors) than existing methods but also it produces significantly less failure, indicated by a smaller 75th percentile angular angular error, for all datasets.

In Figure 8 we show colour-corrected images from GreyBall dataset based on Exemplar-based method compare to Grey-World, Grey-Edge, gamut mapping and Zeta-Image methods, along with their angular error compared to ground truth as obtained from the grey sphere mounted onto the video camera. The proposed exemplar-based method works well overall, and moreover also works significantly better compared to other methods for images of a single surface or a few surfaces, such as a grass scene as in Fig. 8: this constitutes a failure case for grey-based methods since the grey assumption is not satisfied, as well as for gamut mapping since only a limited number of colours are seen in the image.

TABLE 2
Angular errors for GreyBall dataset [62] in terms of mean, median and 75th percentile errors for several colour constancy algorithms.

	Original GrayBall			Linear GrayBall		
Method	Median	Mean	75th	Median	Mean	75th
Do nothing	6.7°	8.3°	14.0°	14.0°	15.6°	26.5°
White-Patch	5.3°	6.8°	10.4°	10.5°	12.7°	19.5°
Grey-World	7.0°	7.9°	10.8°	11.0°	13.0°	20.2°
Grey-Edge	4.7°	5.9°	8.6°	8.8°	10.6°	15.0°
Zeta-Image [22]	4.6°	5.9°	8.6°	9.0°	10.8°	15.0°
Gamut Mapping	5.8°	7.1°	10.2°	8.9°	11.8°	18.0°
Gamut Mapping 1jet [24]	5.8°	6.9°	9.6°	8.9°	11.8°	17.5°
Spatio-spectral Statistics ML [30]	-	-	-	8.9°	10.3°	13.9°
Bottom-up+Top-down [37]	-	-	-	7.7°	9.7°	13.3°
Natural Image Statistics [32]	3.9°	5.2°	7.4°	7.7°	9.9°	13.8°
Exemplar-Based	**3.3°**	**4.4°**	**6.1°**	**6.5°**	**8.0°**	**10.8°**

4.1 Inter Dataset Cross Validation

Although we have shown excellent performance for exemplar-based colour constancy, outperforming existing methods for the standard colour constancy data sets studied, we would also like to investigate whether our proposed methods also work well for any arbitrary images using a *fixed* learned model, inter-datasets. A three-fold cross validation for the ColorChecker datasets and 15-fold cross validation for the GreyBall dataset were already designed for this purpose, intra-dataset, and we applied them in the tests for all learning based colour constancy methods. However images from the same dataset may be to some degree correlated to each other because of the limitation of gathering image data, and therefore doing cross validation between different datasets is a more challenging task and has not been considered in most papers on learning-based illumination estimation methods. For this purpose, we run our proposed method, exemplar-based colour constancy, for the GreyBall dataset, but using surface models learned by images from ColorChecker dataset. The median and mean of angular errors of our illumination estimation for these 11346 images are respectively $5.3°$ and $6.6°$, which is acceptable.

As well, we also ran our proposed method on the ColorChecker dataset, but using surface models learned by images from the Greyball dataset. The median and mean of angular errors of our illumination estimation for these 568 images are respectively $5.1°$ and $6.5°$, again quite acceptable.

Although the results for inter dataset cross validation are not as good as intra dataset cross validation, shown above, they are good enough to convince us that a general surface model is sufficient for estimating the colour of light for any arbitrary image using an arbitrary camera, considering the fact that in comparison either static methods such as Grey-Edge or learning-based methods such as gamut mapping or Bayesian have parameters which must be tuned for each dataset separately. Especially for the two datasets examined, the images differ significantly: the images in the GreyBall dataset were captured from a video recorder and contain approximately 0.1 megapixels (360×240), while the images in the ColorChecker dataset were captured by two high quality DSLR cameras (Canon 5D and 1D) and contain approximately 5 megapixels (813×541 or 874×583)

5 Proposed Method for Multiple Illumination

Exemplar-based colour constancy has the advantage of working in the multiple illuminant situation, which is not possible for most current methods. White patch finds the brighter illuminant, grey-based method as well as grey-edge methods may find combination of illuminant colours, and gamut mapping approximately finds the dominant illuminant.

As discussed above, given a test image with n surfaces we will have M nearest neighbour surfaces from the training set for each of them, and then finally we have nM different estimates for the colour of illuminant that will end up as our final estimation assuming uniform illumination for the test image. If we have more than one illuminant in the scene, however, we then need to carry out an estimation procedure for such illuminants using our nM estimates. As we already seen, since we are dealing with estimates for each of the test image surfaces, we can have separate estimates for each of these surfaces via the median of the estimates from M nearest neighbour surfaces from the training data. This is in the general case in which we have no knowledge about the number of distinct illuminants in the scene. However, knowing the number of illuminants can make our final estimate more robust and incorporate more resistance to incorrect surface matching, which in itself is inevitable as we showed above for the assumption of a single illuminant in the scene.

As in [47], this task can be done either by clustering the illuminant estimates or by taking into account the spatial relations between these estimates. Knowing the number of illuminants, K, we cluster our estimates into K clusters and make the procedure more accurate by

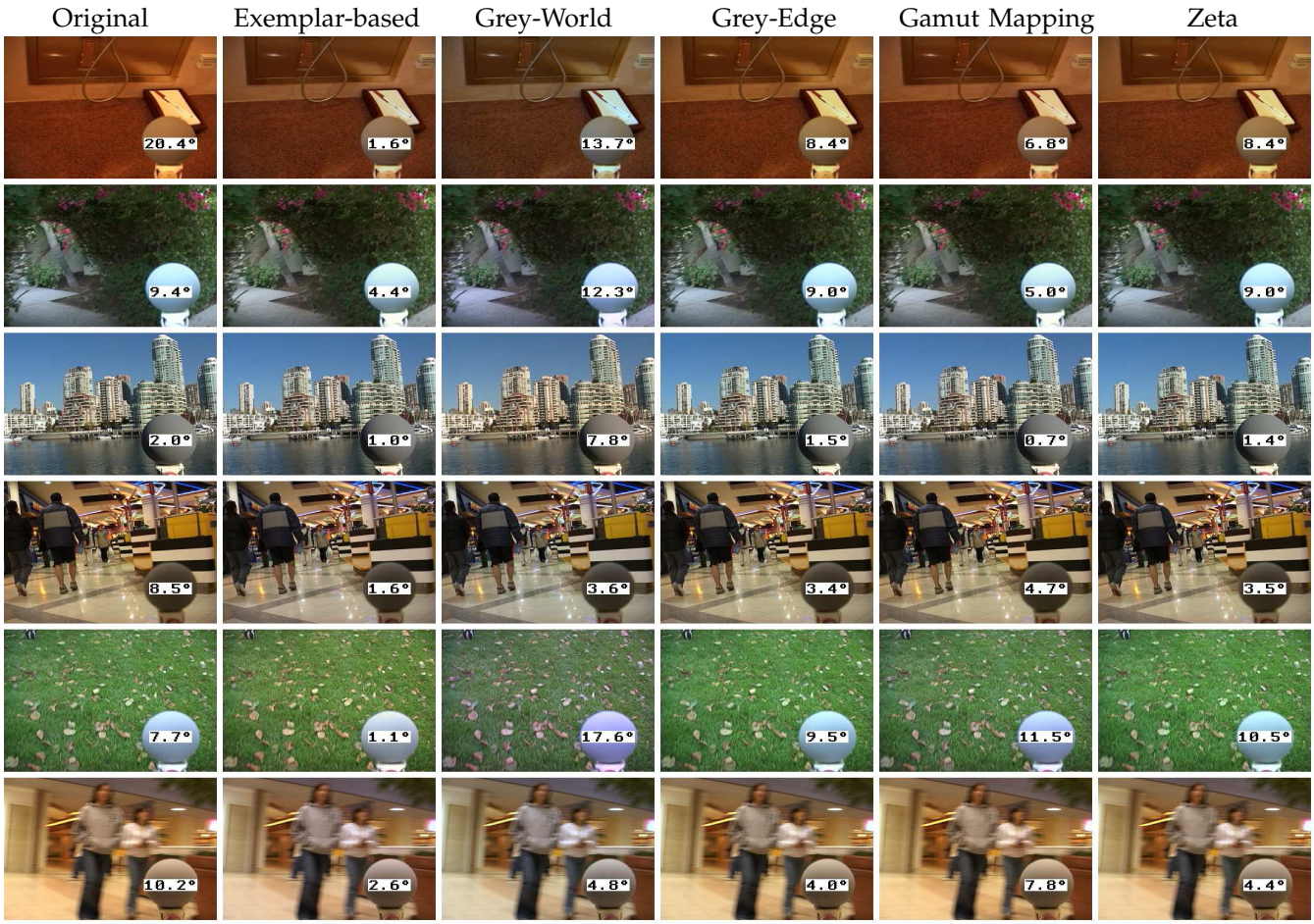

Fig. 8. Examples of colour-corrected images from GreyBall dataset based on the Exemplar-based method, compared to Grey-World, Grey-Edge, Gamut Mapping and Zeta methods, along with their angular error compared to the ground truth as obtained from the grey sphere mounted onto the video camera. Here the first column, labelled "Original", is the error for the video frame as compared to that under the canonical illuminant, and is the same as the "Do Nothing" entries in the Tables.

finding the median of each cluster as our final illumination estimation. If two of the illuminants are close to each other it is likely that they become clustered into the same cluster, with extra clusters containing some other estimates. In this case we simply remove the extra clusters by removing small sized clusters; so therefore we may end up having fewer than K final illuminant estimates. In summary, the proposed illumination estimation for a test image with K illuminants is expressed in Algorithm 3. The final step is colour correction, in which the RGB value for back-projection of each surface will be the nearest one to the separate estimates for that surface. Figure 9 displays an image with two distinct illuminants, all illuminant estimates by exemplar-based method in rg chromaticity space, and two final estimates using clustering.

Here we assume that each surface is illuminated with a distinct illumination and there is no surface with varying illumination; therefore we make a hard decision on which of the N illuminants each surface is assigned to. Fortunately, in the case of a sharp transmission the mean-shift-segmentation will assign separate segments for each light. However the case of a smooth transmission from one illuminant to another could be a problem for the illumination estimation process for that surface. However, in a scene with multiple illuminants there is usually only a small number of segments under varying illumination, and we can simply consider them as outliers, thereby still generating a valid estimation of the colour of multiple illuminants. In order to make the colour corrected output image perfectly consistent for these segments, we would need to estimate an illuminant for each pixel for these surfaces as a mixture of our

Algorithm 3 Estimation of K Illuminants

1: $estimates \leftarrow$ *Algorithm 2*
2: *clusters* $\leftarrow$ *k-means(* $estimates$*,* K *)*
3: *remove the clusters with size less than a threshold:*
clusters $\leftarrow$ *k-means(* $estimates$*, updated* K *)*
4: *illums* $\leftarrow$ *median of each clusters*
5: ***return*** $illums$

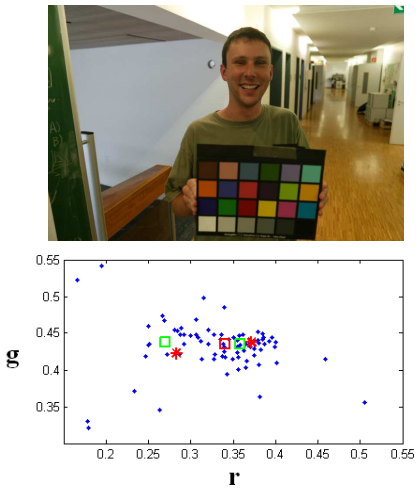

Fig. 9. An input image with multiple illuminants. All illuminant estimates, using our exemplar-based method, are shown as blue dots in rg chromaticity space. Two red stars are the ground truth for two distinct illuminants in the scene. The red square shows the illuminant as estimated by the proposed method for the single illuminant assumption. And two green squares indicate the two illuminants estimated by the proposed method assuming two illuminants.

estimated illuminants. Thus here we could apply light mixture techniques [63], [64], given knowledge of our estimated illuminant colours which is not the main concern of this paper.

5.1 Experimental Results

We evaluate our proposed method with the multi-illuminant dataset provided in [47], which includes 9 outdoor low-quality images with two distinct illuminants for each scene. The ground truth for the light sources is provided by several grey balls placed in the scene and is manually annotated for each image. The images differ in size and are of quite low resolution, containing approximately 40 kilopixels. Since we are dealing with scenes with varying illuminations, we need to find the error across the scene. Therefore the angular error compared to ground truth for each pixel is computed and the average angular error throughout the image is considered as our measure.

First we run our proposed method as enunciated in section 3 to estimate a single illuminant for each image. Then we again run exemplar-based colour constancy but estimating the illuminant for each surface estimate separately via the median of estimates from their nearest neighbour surface models; thus each surface can have a distinct illuminant and there is no limitation on the number of illuminants in the scene. Finally, we assume there are exactly two illuminants in the scene and run exemplar-based colour constancy for multiple illumination as described in the last subsection, clustering the illuminants into two clusters with the estimates for each surface assigned to one of these two estimates. For all of these three experiments we used the texton dictionary and surface models learned from the GreyBall dataset, for an inter-dataset test.

TABLE 3
Median angular errors for a 9 outdoor image dataset [47] in terms of mean and median angular error, for colour constancy algorithms using a one- or two-illuminant assumption.

No. of Illuminants	Method	Median Err.
One	White-Patch	7.8°
	Grey-World	8.9°
	Grey-Edge (n=1)	6.4°
	Grey-Edge (n=2)	5.0°
Two (from [47])	White-Patch	6.7°
	Grey-World	6.4°
	Grey-Edge (n=1)	5.6°
	Grey-Edge (n=2)	5.1°
One	**Exemplar-Based**	**5.1°**
Two		**3.8°**
Multi		**4.3°**

Table 3 shows the median of per-pixel angular error for this dataset assuming a single illuminant and using the White-Patch, Grey-World and Grey-Edge methods; as well as assuming two illuminants with these algorithms with the method proposed in [47], compared to our three experiments as outlined above. For these images with multiple illuminants, almost all methods show significant improvement when including the knowledge that there are two distinct light sources in the image. Although the quality of the images is quite low and indeed this may affect our texture features in our exemplar-based surface model, our proposed method works well assuming a single illuminant is the goal. Moreover using the surface estimates (called Multi in Table 3) and our proposed method for multiple illuminants, with a 2-illuminant assumption, the performance of illumination estimation improves respectively by 14% and 25%. As already mentioned, knowing the number of illuminants makes our final estimate more robust and resistant to incorrect surface matching; therefore the exemplar-based method performs better for this dataset when we use the knowledge that the are exactly two illuminants in the scene. Figure 10 displays images from this dataset, showing ground truth as well as calculated back-projection images using the three mentioned approaches of assuming a single illuminant, two illuminants, and illumination estimation for each surface separately.

6 Conclusion

In this paper we present a completely new line of approach to the colour constancy problem, which we call Exemplar-Based colour constancy. We use both texture features and weakly colour constant three-channel colour values in order to find the nearest neighbour surfaces from training data for each surface, and then we estimate the illuminant for each surface based on histogram matching of each surface to its candidate nearest neighbour surfaces from training data. The final step

Fig. 10. Five example images from an outdoor image dataset with two distinct illuminants [47]. For each example, from left to right: original image, the ground-truth back projection (chromaticity of pixelwise illumination) for that image and its estimated value using exemplar based colour constancy assuming single illuminant, two illuminants and illumination estimation for each surface separately.

is integrating these estimates into a unique illuminant estimation for the whole image. The proposed method has the advantage of overcoming the difficulty of multi-illuminant situations, which is not possible for most current methods. We show that the proposed method performs very well for four standard datasets commonly used in colour constancy tests compared to current colour constancy algorithms.

We also extend our proposed method to overcome the problem of multiple illuminants in the scene by clustering all estimates correspond to nearest neighbour surfaces. The proposed method is shown to work well for an image set of outdoor images with two distinct light sources.

In future, we can apply more complex methods of integrating estimated surface illuminants into a unique illumination estimate. We may also utilize spatial information in order to improve the estimation as well as spatial smoothness in the colour corrected output image. As well, we should construct a dataset for multi-illuminant colour constancy in order to evaluate Exemplar-Based Colour Constancy for images with more than one light source colour, in that this is a uniquely challenging scenario which could produce real benefits in image understanding.

Already, the method provides a real, substantive improvement over current methods, and further tests using inter-dataset calculations are called for to see whether we could generate a standardized and much faster pipeline for colour constancy.

REFERENCES

[1] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 24, pp. 509–522, 2002.

[2] ——, "Shape context: A new descriptor for shape matching and object recognition," in *Neural Inf. Proc. Sys.*, 2000, pp. 831–837.

[3] G. Shakhnarovich, P. Viola, and T. Darrell, "Fast pose estimation with parameter sensitive hashing," in *Int. Conf. on Comp. Vis.*, 2003, pp. 750–757.

[4] C. Liu, J. Yuen, and A. Torralba, "Nonparametric scene parsing: Label transfer via dense scene alignment," in *Comp. Vis. and Patt. Rec.*, 2009, pp. 1972–1979.

[5] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba, "Sun database: large-scale scene recognition from abbey to zoo," in *Comp. Vis. and Patt. Rec.*, 2010, pp. 3485–3492.

[6] T. Malisiewicz, A. Gupta, and A. Efros, "Ensemble of exemplar-SVMs for object detection and beyond," in *Int. Conf. on Comp. Vis.*, 2011, pp. 89–96.

[7] A. Torralba, R. Fergus, and W. Freeman, "80 million tiny images: a large data set for nonparametric object and scene recognition," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 30, no. 11, pp. 1958–1970, 2008.

[8] D. Aha, D. Kibler, and M. Albert, "Instance-based learning algorithms," *Machine learning*, vol. 6, no. 1, pp. 37–66, 1991.

[9] M. Olkkonen, T. Hansen, and K. Gegenfurtner, "Color appearance of familiar objects: effects of object shape, texture, and illumination changes," *Journal of Vision*, vol. 8, no. 5, pp. 13.1–13.16, 2008.

[10] J. Granzier and K. Gegenfurtner, "Effects of memory colour on colour constancy for unknown coloured objects," *i-Perception*, vol. 3, no. 3, pp. 190–215, 2012.

[11] S. Hordley, "Scene illuminant estimation: past, present, and future," *Color Research and App.*, vol. 31, no. 4, pp. 303–314, 2006.

[12] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: survey and experiments," *IEEE Trans. on Im. Proc.*, vol. 20, pp. 2475–2489, 2011.

[13] E. Land, "The retinex theory of color vision," *Scientific American*, vol. 237, no. 6, pp. 108–128, Dec. 1977.

[14] G. Buchsbaum, "A spatial processor model for object colour perception," *J. Franklin Inst.*, vol. 310, pp. 1–26, 1980.

[15] J. van de Weijer and T. Gevers, "Color constancy based on the grey-edge hypothesis," in *Int. Conf. on Image Proc.*, 2005.

[16] G. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Color Im. Conf.*, 2004, pp. 37–41.

[17] D. Forsyth, "A novel approach to color constancy," in *Int. Conf. on Comp. Vis.*, 1988, pp. 9–18.

[18] A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy beyond bags of pixels," in *Comp. Vis. and Patt. Rec.*, 2008.

[19] H. Vaezi Joze and M. Drew, "Improved machine learning for image category recognition by local color constancy," in *Proceedings of IEEE Int. Conf. on Image Proc.*, 2010.

[20] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Trans. on Im. Proc.*, vol. 16, no. 9, pp. 2207–2214, 2007.
[21] R. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," *J. Opt. Soc. Am. A*, vol. 21, no. 3, pp. 321–334, 2004.
[22] M. Drew, H. Vaezi Joze, and G. Finlayson, "Specularity, the zeta-image, and information-theoretic illuminant estimation," in *Eur. Conf. on Comp. Vis. Wr. on Color and Photometry in Comp. Vis.*, 2012.
[23] G. Finlayson, S. Hordley, and P. Hubel, "Colour by correlation: A simple, unifying approach to colour constancy," in *Int. Conf. on Comp. Vis.*, 1999, pp. 835–842.
[24] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *Int. J. of Comp. Vis.*, vol. 86, no. 2-3, pp. 127–139, 2008.
[25] H. Vaezi Joze and M. Drew, "White patch gamut mapping colour constancy," in *Proceedings of IEEE Int. Conf. on Image Proc.*, 2012.
[26] C. Rosenberg, T. Minka, and A. Ladsariya, "Bayesian color constancy with non-Gaussian models," in *Neural Inf. Proc. Sys.*, 2003.
[27] P. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, "Bayesian color constancy revisited," in *Comp. Vis. and Patt. Rec.*, 2008.
[28] V. Cardei, B. Funt, and K. Barnard, "Estimating the scene illumination chromaticity using a neural network," *J. Opt. Soc. Am. A*, vol. 19, no. 12, pp. 2374–2386, 2002.
[29] W. Xiong and B. Funt, "Estimating illumination chromaticity via support vector regression," *J. Im. Sci. Tech.*, vol. 50, no. 4, pp. 341–348, 2006.
[30] A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy with spatio-spectral statistics," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 34, no. 8, pp. 1509–1519, 2012.
[31] V. Cardei and B. Funt, "Committee-based color constancy," in *Color Im. Conf.*, 1999, pp. 311–313.
[32] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 33, no. 4, pp. 687–698, 2011.
[33] M. Wu, J. Sun, J. Zhou, and G. Xue, "Color constancy based on texture pyramid matching and regularized local regression," *J. Opt. Soc. Am. A*, vol. 27, no. 10, pp. 2097–2105, 2010.
[34] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Automatic color constancy algorithm selection and combination," *Patt. Rec.*, vol. 43, no. 3, pp. 695–705, 2010.
[35] B. Li, D. Xu, and C. Lang, "Colour constancy based on texture similarity for natural images," *Coloration Tech.*, vol. 125, no. 6, pp. 328–333, 2009.
[36] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Improving color constancy using indoor-outdoor image classification," *IEEE Trans. on Im. Proc.*, vol. 17, no. 12, pp. 2381–2392, 2008.
[37] J. van de Weijer, C. Schmid, and J. Verbeek, "Using high-level visual information for color constancy," in *Int. Conf. on Comp. Vis.*, 2007.
[38] E. Rahtu, J. Nikkanen, J. Kannala, L. Lepisto, and J. Heikkila, "Applying visual object categorization and memory colors for automatic color constancy," in *Int. Conf. on Im. Anal. and Proc.*, 2009, pp. 873–882.
[39] G. Finlayson, B. Funt, and K. Barnard, "Color constancy under varying illumination," in *IEEE Int. Conf. on Comp. Vis.*, 1995.
[40] M. Tsukada and Y. Ohta, "An approach to color constancy using multiple images," in *Int. Conf. on Comp. Vis.*, 1990.
[41] M. Ebner, "Color constancy using local color shifts," in *Eur. Conf. on Comp. Vis.*, 2004, pp. 276–287.
[42] M. Bleier, C. Riess, S. Beigpour, E. Eibenberger, E. Angelopoulou, T. Troger, and A. Kaup, "Color constancy and non-uniform illumination: Can existing algorithms work?" in *Int. Conf. on Comp. Vis. Wrkshp on Color and Photometry in Comp. Vis.*, 2011, pp. 774–781.
[43] J. Kuang and W. Xiong, "Color constancy for multi-illuminants high-dynamic-range scenes," in *Color Im. Conf.*, 2008.
[44] G. Finlayson, M. Drew, and C. Lu, "Intrinsic images by entropy minimization," in *Eur. Conf. on Comp. Vis.*, 2004, pp. 582–595.
[45] G. Finlayson, S. Hordley, C. Lu, and M. Drew, "On the removal of shadows from images," *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. 28, pp. 59–68, 2006.
[46] M. Drew and H. Vaezi Joze, "Sharpening from shadows: Sensor transforms for removing shadows using a single image," in *Proceedings of the 17th IS&T/SID Color and Imaging Conference*, 2009.
[47] A. Gijsenij, R. Lu, and T. Gevers, "Color constancy for multiple light sources," *IEEE Trans. on Im. Proc.*, vol. 21, no. 2, pp. 697–707, 2012.
[48] S. Hordley and G. Finlayson, "Reevaluation of color constancy algorithm performance," *J. Opt. Soc. Am. A*, vol. 23, no. 5, pp. 1008–1020, 2006.
[49] A. Gijsenij, T. Gevers, and M. Lucassen, "A perceptual analysis of distance measures for color constancy algorithms," *J. Opt. Soc. Am. A*, vol. 26, no. 10, pp. 2243–2256, 2009.
[50] K. Barnard, B. Funt, and V. Cardei, "A comparison of computational color constancy algorithms, part one: theory and experiments with synthetic data," *IEEE Trans. on Im. Proc.*, vol. 11, no. 9, pp. 972–984, 2002.
[51] M. Mosny and B. Funt, "Reducing worst-case illumination estimates for better automatic white balance," in *Color Im. Conf.*, 2012.
[52] H. Vaezi Joze and M. Drew, "Exemplar-based colour constancy," in *Brit. Mach. Vis. Conf.*, 2012, pp. 26.1–26.12.
[53] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. 24, pp. 603–619, 2002.
[54] M. Varma and A. Zisserman, "Classifying images of materials: achieving viewpoint and illumination independence," in *Eur. Conf. on Comp. Vis.*, 2002, pp. III:255–271.
[55] ——, "A statistical approach to texture classification from single images," *Int. J. of Comp. Vis.*, vol. 62, no. 1-2, pp. 61–81, 2005.
[56] M. Drew, J. Wei, and Z. Li, "Illumination–invariant color object recognition via compressed chromaticity histograms of color–channel–normalized images," in *Int. Conf. on Comp. Vis.*, 1998.
[57] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *Int. J. of Comp. Vis.*, vol. 43, no. 1, pp. 29–44, 2001.
[58] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. of Comp. Vis.*, vol. 60, pp. 91–110, 2004.
[59] L. Shi and B. Funt, "Re-processed version of the gehler color constancy dataset of 568 images," 2010, http://www.cs.sfu.ca/∼colour/data.
[60] G. Finlayson, M. Drew, and B. Funt, "Color constancy: diagonal transforms suffice," in *Int. Conf. on Comp. Vis.*, 1993, pp. 164–171.
[61] A. Gijsenij, "Color constancy: research website on illuminant estimation," http://staff.science.uva.nl/∼gijsenij/ colorconstancy/index.html.
[62] F. Ciurea and B. Funt, "A large image database for color constancy research," in *Color Im. Conf.*, 2003, pp. 160–164.
[63] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand, "Light mixture estimation for spatially varying white balance," *ACM Trans. on Graphics*, vol. 27, no. 3, pp. 70:1–70:7, 2008.
[64] I. Boyadzhiev, K. Bala, S. Paris, and F. Durand, "User-guided white balance for mixed lighting conditions," *ACM Trans. on Graphics*, vol. 31, no. 6, pp. 200:1–200:10, 2012.

Hamid Reza Vaezi Joze received the B.Sc. and M.Sc. degrees in computer engineering from Sharif University of Technology, Iran, in 2006 and 2008, respectively. Since 2009, he has been a PhD student at School of Computing Science at Simon Fraser University in Vancouver. His research interests include computer vision focusing on colour and illumination. He has published over 15 journal and conference papers and is the holder of a patent application in image processing.

Mark S. Drew is a Professor in the School of Computing Science at Simon Fraser University in Vancouver, Canada. His background education is in Engineering Science, Mathematics, and Physics. His interests lie in the fields of image processing, colour, computer vision, computer graphics, multimedia, and visualization. He has published over 150 refereed papers and is the holder of six patents and patent applications in computer vision, image processing, and image reconstruction and accenting.