MVA2015 IAPR International Conference on Machine Vision Applications, May 18-22, 2015, Tokyo, JAPAN

Illuminant Classification based on Random Forest

Bozhi Liu
University of Nottingham
Wollaton Road Nottingham NG8 1BB UK
psxbl@nottingham.ac.uk

Guoping Qiu
University of Nottingham, Ningbo Campus
199 Taikang East Road, Ningbo, 315100, China
guoping.qiu@nottingham.edu.cn

Abstract

we present a novel machine learning/pattern recognition based colour constancy method. We cast colour constancy as an illumination source recognition problem, and have developed an effective and efficient random forest based classification technique for inferring the class of illumination source of an image. In an opponent colour space, we have developed a binary image representation feature that is somewhat insensitive to image contents for building the random forest classifier that infers the likely class of the illumination source of the image. The binary image feature and the tree structure of the recognition system are intrinsically efficient. We present results on colour constancy benchmark data sets and show that our new technique outperforms state of the art techniques.

1 Introduction

The vision system observe colours of objects are depended on three parameters: the light source spectral power distribution function $E(\lambda)$, where λ is the wavelength of the light source, a surface is characterized by its spectral reflectance function S(λ), and $C_k(\lambda)$ which characterizes the spectral response function of the observer, where k refers to different colour channels [1]. Then, the colour response signal ρ_k reflecting from a given surface under a given illuminant can be defined as:

$$\rho_k = \int E(\lambda)S(\lambda)C_k(\lambda)d\lambda \qquad (1)$$

Most of imaging devices usually have 3 distinct classes of sensor (k equals to 3), so that the response to light at a given pixel is defined by a triplet of responses $\rho = (\rho_1, \rho_2, \rho_3)$. And we will assume this trichromatic imaging system throughout the paper, as is commonly used in much of the previous literature [2, 3, 4].

From Equation (1), we could see the colour response ρ of a device to a given surface depends on both the reflectance properties of the surface, and the spectral power distribution of the current illuminant. When illumination changes, so too the colour response ρ. However, human vision system can ensure the perceived colour of objects remains relatively constant under varying illuminant conditions, because human vision system can recognize the object colour changing caused by the illuminant change, then recover the influence made by illuminant change. This ability to account for the change of the light source is called colour constancy which helps human identify objects [1]. But from a computer vision perspective this illumination dependence is a problem since the digital vision system can not recognize such a changing illuminant scene,this implies that using colour as a cue to help solve fundamental vision tasks such as scene segmentation, object recognition and tracking might run into problems if the scene illumination is changing. So, how to estimate the unknown illuminant is a problem we need to solve.

In this paper, under the assumption that images taken under the same illuminant should have similar colours, we present a novel learning-based technique for colour constancy. We first divide the training images into a number of classes based on the known illuminant data of the images and treat each class of images as having the same illuminant. We then build a machine learning model based on Random Forest and treat colour constancy as a problem of inferring the illuminant class of the input image. We have tested our algorithm on a number of publicly available testing databases and show that our new technique outperforms state of the art techniques.

2 Related Work

There exist many colour constancy methods in the literature, and these methods can be roughly classified into three categories. The first category is based on statistical models. These models are usually associated with some parameters determined either based on low-level statistics or the physics-based dichromatic reflection model. The most widely used algorithm in this category is the so called "grey world" algorithm [3], which is based on the assumption that the average of the surface reflectance of a typical scene is a fixed value, which is referred to as "grey". Grey-Edge is a recent version which assumes that the average of the reflectance differences in a scene is achromatic [4]. White patch method [3] uses the maximum colour value in the image as a reference "white" value instead of using mean value. As none of these assumptions could perfectly cover all illuminant conditions, some works focus on combining these different algorithms and try to select the optimal algorithm for a given image [5].

The second category is Gamut Mapping algorithm [2, 6, 7]. The Gamut Mapping algorithm is based on the assumption that in real-world images, for a given illuminant, one observes only a limited number of colours. Therefore, colours form a "canonical" gamut, defined as the standard white light illuminant, which contains all possible colours can be observed under a canonical illumination. And for an input unknown gamut, an estimate of the current illuminant can be derived by mapping current pixel colour gamut to the canonical gamut. Finding such mappings is referred as finding feasible sets. There are several version of gamut mapping algorithms each differs in the way of finding the feasible sets. Gamut mapping algorithms are considered as having the best performance amongst colour constancy algorithms, however it is possible that no feasible mapping can be found that maps the input data into the canonical gamut with one single transform, if the image does not fully satisfied the diagonal model. This is one of the disadvantages of gamput mapping, also, the computation cost of gamut mapping is the highest. To simplify the algorithm and reduce the computational cost, several extensions have been proposed for gamut mapping algorithms [6, 7].

Recently, an increasing number of learning based colour constancy methods have been developed. Learning based algorithms estimate the illuminant using a model that is learned on training data. Early approaches using machine learning techniques are based on neural networks [8]. The input to the neural network consists of a large binarized chromaticity histogram of the input image, the output is two chromaticity-values of the estimated illuminant. Although this approach, when trained correctly, can deliver accurate color constancy even when only a few distinct surfaces are present, the training phase requires a large amount of training data. Another learning-based approach to illumination estimation problem is the Bayesian approach [9], bayesian approaches try to model the variability of reflectance and treat the illuminant as a random variable. The Bayes' rule computes the probability of the chromaticity of the dominant illuminant of a scene given the observed chromaticity colour. The illuminant is then estimated from the posterior distribution conditioned on the image intensity data. Earlier method in this theme attempted to do this did not outperform gamut mapping, until Finlayson [10] developed "colour by correlation", which used a nonparametric statistical model to capture the correlation between nearby pixels. This method had built a simple algorithm framework and produced very good performance close to gamut mapping methods. It was shown the learning based colour constancy methods still had scope for improvement.

3 A Novel Random Forest Approach to Colour Constancy

The basic idea of our random forest approach is to treat the problem of estimating the unknown illuminant of the input image as an illuminant classification problem. Random forest is an ensemble classifier that consists of many decision trees and outputs the class that is the mode of the classes output by individual trees [11]. Random Forest is a generic classification technique, and here we present an approach of adapting it for tackling the colour constancy problem.

3.1 Image representation

As in any pattern classification problem, an effective representation scheme for the real world data is essential. Some previous works including the first version of the gamut mapping based algorithm used all three colour channels in the RGB colour space [2]. The method in [10] used a two dimensional feature ($r = R/B, g = G/B$). This model discarded the illuminant intensity information and had reduced the problem from 3D to 2D, however some authors [5] evaluated the 2D method and showed the performance degrades slightly.

We first derive two opponent colour signals, Red-Green (RG) and Blue-Yellow (BY), from the original RGB signals according to (2):

$$RG = R - G, \quad BY = (R + G)/2 - B \tag{2}$$

We then treat each of the two opponent signals as an 8 bits per pixel gray scale image and construct pixel histograms H_{RG} and H_{BY} for the RG image and the BY image respectively. Here H_{RG} and H_{BY} are 256 dimensional vectors. $H_{RG}(i)$, $i = 0$, 1, 2,$\cdots$, 255, is the frequency of the pixels in the RG image having a quantized value of i, and same as $H_{BY}(i)$.

The two histograms are obviously dependent on the contents of the image. As our task is estimating the colours of the illuminant, our representation should be relatively content independent. In this case, we only take into account the existence of a colour rather than the number of times it appears in the image or the number of pixels having that colour. A simple way to achieve this is to binaries H_{RG} and H_{BY} to ensure that each chromaticity is counted only once. Our image representation feature is therefore two 256 dimensional binary vectors obtained as (3):

$$h_{RG}(i) = \begin{cases} 1, & H_{RG}(i) > 0 \\ 0, & otherwise \end{cases} ; h_{BY}(i) = \begin{cases} 1, & H_{BY}(i) > 0 \\ 0, & otherwise \end{cases} \tag{3}$$

And finally, we cascade the two binary vectors h_{RG} and h_{BY} into one 512 dimensional feature vector $F_{RGBY} = (h_{RG}, h_{BY})$ for the construction of our Random Forest based recognition or classification based colour constancy algorithm.

3.2 Constructing the illuminant recognition random forest

In building a random tree, we randomly select a dimension of the feature vector of the input image and then based on the value of that chosen dimension (which is either 0 or 1 in our case) distribute the image to either the right child or the left child of a binary tree node. For example, if dimension i of F_{RGBY} has been selected as the testing dimension, the training images will be split according to Equation (4):

$$\begin{cases} \textit{Image goes to the left child,} & F_{RGBY}(i) = 0 \\ \textit{Image goes to the right child,} & F_{RGBY}(i) = 1 \end{cases} \tag{4}$$

After each split, the input images at the current node I_n will be divided into two groups: I_l and I_r; which are then further split until it has reached the leaf nodes.

For a typical decision tree, we should generate multiple hypothesized tests and pick the best split dimension and threshold value as the split function of each tree node. At each non-leaf node, N different dimensions in the feature vector F_{RGBY} are randomly selected and then we perform N different splits according to (4). For each split, we calculate the information gain to evaluate the merit of the split. The score function (information gain) is calculated as (5):

$$Score(split) = -\frac{|I_l|}{|I_n|}E(I_l) - \frac{|I_r|}{|I_n|}E(I_r) \quad (5)$$

where $E(I)$ is the Shannon entropy of the class distributions in the set of samples I. $|I_l|$ is the number of samples contained in the left child and $|I_r|$ is the number in the right. I_n is the set of training sample in node n. After compared the scores of the splits, the split with the best score will be selected and the dimension i related with this split will be recorded at the current node as the split dimension.

The decision tree generation process will be stopped when certain tree construction criteria such as maximum tree depth have been met. Each leaf node is then associated with an illuminant class distribution histogram, $H(k)$, $k = 1, 2, \cdots, K$, where K is the totally number of possible illuminant classes, and $H(k)$ records the probability of the input image belongs to the k-th illuminant when it falls onto that leaf node.

3.3 Illuminant estimation based on random forest

For a test input image, after extracted the h_{RG} and h_{BY} features, we let the image features go through all the trees. When the image reaches a leaf node of the m-th tree, we save the illuminant class distribution histogram of the leaf node $H_m(k)$, $k = 1, 2, \cdots, K$, where K is the totally number of possible illuminant classes. Suppose we have M trees in the random forest, we sum all M illuminant class histograms together:

$$H(k) = \sum_{m=1}^{M} H_m(k) \quad (6)$$

The l-th illuminant is estimated as the input image's illuminant if $H(l) \geq H(k)$ for all k.

3.4 Recover the colour under canonical illuminant

After labeled the l-th illuminant as the estimate illuminant for the input image, we could simply use diagonal model of illumination change to recover the image colour under a canonical illuminant. The diagonal model is colour channel independent, for example, the colour signal values of a white patch taken under a white, canonical illuminant is (R_c, G_c, B_c), and the response under an unknown illuminant is (R_e, G_e, B_e),then the mapping from the unknown illuminant to canonical illuminant could be achieved by scaling the three channels by $(R_c/R_e, G_c/G_e, B_c/B_e)$, and the same scaling works for other non-white patches.

4 Algorithms Evaluation

We evaluate our algorithm's performances on two public benchmark datasets: The first dataset consists of 321 images of constructed scenes taken under 11 different illuminant sources in the Lab SFU [12]. For each illuminant source the ground truth is known, so the class number in the SFU dataset is 11. The second database [13] is a large database contains 11346 images of several indoor or outdoor scenes. The images were actually frames in a video; each image has a grey ball at the bottom right corner to calculate the ground truth of the illuminant. The database contains 15 different scenes and we regard each scene as having one class of illuminant. In each class, 80% of the image is used for training and the other 20% is for test. The original images and the ground truth are under sRGB colour space. Note that all the results are averaged over 50 different random trials (each with 80% different training and 20 % testing samples).

Following a common practice in the literature, we calculate the angular error according to equation (7) from [1] as the performance indicator:

$$Error_{angular} = cos^{-1}((T \cdot E)|T|^{-1}|E|^{-1}) \quad (7)$$

where T is the ground truth illuminant value and E is the estimated value by the colour constancy algorithms. The error is calculated by the degree distances between the two colour vectors.

4.1 Colour features evaluation

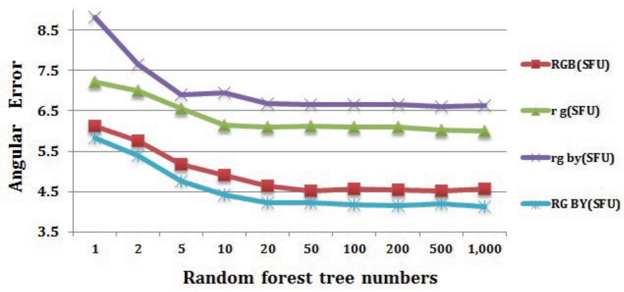

Figure 1. colour features performance.

In addition to using the RG and BY colour signals to derive the image representation features as described in Section 3, we have also tested other colour signals for deriving the binary histogram of equation (3) for building the random forests, including, the original R, G and B signals to build a 3x256 dimensional binary feature vector; $r = R/(R+G+B)$ and $g = R/(R+G+B)$ to build a 2x256 dimensional binary feature; and rg and by to build 2x256 dimensional binary feature, where $r = R/(R+G+B)$, $g = G/(R+G+B)$, and $b = B/(R+G+B)$, $rg = r - g$, $by = ((r+g)/2 - b$.

The results are shown in Figure 1. As the number of trees increases, the performances are getting better. RG and BY feature always have a better result. r and g, rg and by combinations performed worse than the original R, G and B signals. One possible reason is that the illuminant intensity may still have useful information and RG and BY colour signals contain illuminant information.

4.2 Comparison of colour constancy algorithms

The second experiment is the evaluation of different algorithms for the two datasets with three ground truths (The gray ball dataset's ground truth values have been modified under a linear colour space since the original dataset is captured under a non-linear colour space [14]). The compared algorithms included are classic algorithms like the grey world [3], the max-RGB algorithms [3], and some leading algorithms like the grey edge [4], and the gamut mapping algorithms with different gamut weighting methods [7, 15]. All the results for these algorithms are from the website www.colorconstancy.com created by [14], an evaluation platform for colour constancy. The tree number in the forest is 100 (We have varied the number of the trees and results are similar).

Table 1. Algorithms Comparison.

Algorithms	Lab SFU	Grey Ball (Original)	Grey Ball (Linear)
Grey world[3]	9.8	7.9	13
Max-RGB[3]	9.1	6.8	12.7
Grey edge 2nd order[4]	5.2	6.1	10.6
pixel-based Gamut mapping[7]	3.7	7.1	11.8
Edge-based Gamut mapping[15]	3.9	6.8	12.8
Intersection-based Gamut mapping[15]	3.6	6.9	11.8
Random forest	4.2	5.6	8.2

Table 1 shows the results of the random forest algorithm have close performance to gamut mapping methods in Lab SFU dataset and the lowest angular errors of all the algorithms for Grey-ball datasets for both ground truths. As Grey edge and Gamut mapping also need to set some parameters in different datasets (Minkowski-norm ρ is varied between 1 and 15, and the smoothing value σ is varied between 1 and 12), the results are from the website www.colorconstancy.com which are the best results of the algorithms.Figure 2 shows a visual example of colour correction results of an image based on illuminants estimated by different algorithms. The image is selected from Grey ball dataset and it is seen that the visual result of the random forest method is also better with lower angular errors.

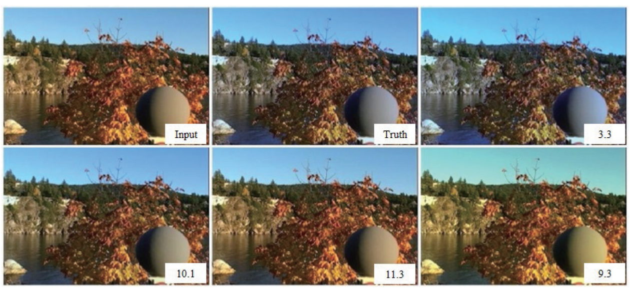

Figure 2. Colour constancy colour correction results of grey world, max RGB, grey edge and random forest. Top row, left: original, middle: corrected with ground truth data, right: random forest. Bottom row, left: grey world, middle: max RGB, right: grey edge.

5 Concluding Remarks

In this paper, we have developed a new colour constancy algorithm. Based on the assumption that images taken under the same illuminant should have similar colours, treating the unknown image illuminant estimation task as an illuminant classification problem, we developed a new binary colour feature to represent the colours of an image and used random forest method as a classifier to estimate the illuminate class of the image. Compared with other algorithms on two benchmark datasets, random forest is more efficient and can outperform other state of the art algorithms. However, as a classification method, random forest can only provide a few discrete illuminant class estimations based on how the datasets have been classified. In these two test datasets, the SFU dataset has 11 know illuminant classes. But for the grey ball dataset, or other datasets with real world images, how to divide the dataset into an appropriate number of classes is an important issue for further study.

References

[1] S. D. Hordley, "Scene illuminant estimation: Past, present, and future," *Color Research & Application*, vol. 31, no. 4, pp. 303–314, 2006.

[2] D. Forsyth, A novel approach to color constancy, in Proc. ICCV, Tampa, FL, USA, 1988, pp. 918.

[3] E. Provenzi, C. Gatta, M. Fierro, and A. Rizzi, "A Spatially Variant White-Patch and Gray-World Method for Color Image Enhancement Driven by Local Contrast," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 1757–1770, 2008.

[4] J. V. D. Weijer, T. Gevers, and A. Gijsenij, "Edge-Based Color Constancy," *IEEE Transactions on Image Processing*, vol. 16, pp. 2207–2214, 2007.

[5] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Improving color constancy using indoor–outdoor image classification," *IEEE Transactions on Image Processing*, vol. 17, no. 12, pp. 2381–2392, 2008.

[6] G. D. Finlayson and S. D. Hordley, "Improving gamut mapping color constancy," *IEEE Transactions on Image Processing*, vol. 9, pp. 1774–1783, 2000.

[7] A. Gijsenij and T. Gevers, "Generalized Gamut Mapping using Image Derivative Structures for Color Constancy," *International Journal of Computer Vision*, vol. 86, pp. 127–139, 2010.

[8] V. C. Cardei, B. Funt, and K. Barnard, "Estimating the scene illumination chromaticity by using a neural network," *JOSA A*, vol. 19, no. 12, pp. 2374–2386, 2002.

[9] P. V. Gehler, C. Rother, A. Blake, T. P. Minka, and T. Sharp, "Bayesian color constancy revisited," in *Computer Vision and Pattern Recognition*, pp. 1–8, 2008.

[10] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Colour by Correlation: A Simple, Unifying Approach to Colour Constancy," in *ICCV*, pp. 835–842, 1999.

[11] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.

[12] K. Barnard, L. Martin, B. Funt, and A. Coath, A data set for color research, Color Research and Application, vol. 27, no. 3, pp. 147151, 2002.

[13] F. Ciurea and B. V. Funt, "A Large Image Database for Color Constancy Research," in *Color Imaging Conference*, pp. 160–164, 2003.

[14] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational Color Constancy: Survey and Experiments," *IEEE Transactions on Image Processing*, vol. 20, pp. 2475–2489, 2011.

[15] A. Gijsenij, T. Gevers, and J. van de Weijer, "Improving Color Constancy by Photometric Edge Weighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 918–929, 2012.