Learning Outdoor Color Classification

Roberto Manduchi

Abstract—We present an algorithm for color classification with explicit illuminant estimation and compensation. A Gaussian classifier is trained with color samples from just one training image. Then, using a simple diagonal illumination model, the illuminants in a new scene that contains some of the surface classes seen in the training image are estimated in a maximum likelihood framework using the Expectation Maximization algorithm. We also show how to impose priors on the illuminants, effectively computing a maximum a posteriori estimation. Experimental results are provided to demonstrate the performance of our classification algorithm in the case of outdoor images.

Index Terms—Color constancy, classification, expectation maximization.

1 Introduction

Recognition (or, more generally, classification) is a fundamental task in computer vision. Whereas clustering or segmentation are unsupervised processes, classification relies on prior information, in the form of physical modeling and/or of training data, to assign labels to images or image areas. This paper is concerned with the classification of outdoor scenes based on color. Color features are normally used in such different domains as robotics, image database indexing, remote sensing, tracking, and biometrics. Color vectors are generated directly by the sensor for each pixel as opposed to other features, such as texture or optical flow, which require possibly complex preprocessing. In addition, color information can be exploited at the local level, enabling simple classifiers that need not worry too much about contextual spatial information.

Color-based classification relies on the fact that the reflectance spectrum is often a discriminative feature of a surface material. Unfortunately, a camera does not take direct reflectance measurements. Even neglecting specular and other non-Lambertian components, the spectral distribution of the radiance from a surface is a function of the illuminant spectrum (or spectra) as much as of the surface reflectance. The illuminant spectrum, is in this context, a nuisance parameter, inducing undesired variability to the perceived color of a surface. Unless one is interested solely in surfaces with a highly distinctive reflectance (such as the bright targets often used in indoor robotics experiments), ambiguity and, therefore, misclassification will result when surfaces are illuminated by "unfamiliar" light.

One way to reduce the dependency on the illuminant is to use more training data and rich statistical models for the representation of the color variability. This is feasible if the space of possible illuminants is not too broad, meaning that one can hope to sample it adequately. For example, in the case of outdoor scenes (which are of interest to this work), the spectrum of the illuminant (direct sunlight or diffuse light and shade) can be well modeled by a low-dimensional linear space [15], [23]. Thus, by collecting many image samples of the surfaces of interest under all expected light conditions, one may derive the complete statistical distribution of colors within each class considered and, therefore, build a Bayesian classifier that can effectively cope with illumination variations. This approach was taken by the author and colleagues for the design and implementation of the color-based terrain typing subsystem of the eXperimental Unmanned Vehicle (XUV) DEMO III [16], which gave good classification performances.

Unfortunately, collecting and hand-labeling extensive training data sets may be difficult, time-consuming, and impractical or impossible in many real-world scenarios. If only a few labeled images are available, then a different approach must be taken. Rather than relying only on exemplars (different versions of the same surface types under different illuminants), one may build a model describing the distribution of colors in a scene as a function of illumination changes. In other words, one may try to decouple the contribution of the reflectance and of the illumination components to the color distribution within each surface class and to explicitly recover and compensate for variations of the illuminant (or illuminants) in the scene. The effect of both reflectance and illuminant on the sensed color is modeled by suitable (and simple) statistical models. The marginal contribution of reflectance and of scene geometry to the color distribution is learned by observing each class under just one "canonical" illuminant. We found that often a single training image contains enough variability (for a fixed illuminant) to provide a sufficiently rich description of the class and illuminant-conditional color distributions. Indeed, all of the experiments in this paper use a single image for training.

At the core of our algorithm is the *rendering model*, which describes the transformation of the color of a surface patch when seen under different illuminants. We use a simple diagonal rendering model [10], [7], [3] in this work. In other words, we assume that a change in illumination manifests itself as an independent rescaling of the color channels.

• *The author is with the Department of Computer Engineering, University of California, Santa Cruz, 1156 High Street, MS: SOE3, Santa Cruz, CA 95064. E-mail: manduchi@soe.ucsc.edu.*

Manuscript received 8 Oct. 2004; revised 5 Dec. 2005; accepted 25 May 2006; published online 14 Sept. 2006.
Recommended for acceptance by G. Healey.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0542-1004.

Although the diagonal model is known to be a relatively crude approximation, its use is justified in this context by two main factors. First, when compared to more complex approaches that describe color in terms of a bilinear relationship between reflectance and illuminant coefficients [25], the diagonal model allows us to drastically reduce the number and type of training images required for training. Second, the parameters of the color transformations in the scene can be computed using a simple and well-known optimization procedure (the Expectation Maximization algorithm). This is usually not possible with more complex color transformation models.

Whereas the characteristics of the reflectance of each surface class can be learned from one or more instances of the surface under the same "canonical" illuminant, modeling the distribution of illuminant-induced color transformations requires observation of the same surface type under different illumination conditions. The surface types used to learn the statistics of the transformation matrices may be different from the surfaces classes that need to be recognized. This is quite convenient: The learning procedure can be performed offline once and for all, regardless of the future choice of the surfaces to be classified.

In our algorithm, pixel classification is performed along with the estimation of which illuminant type impinges on each surface patch and of the diagonal color transformations with respect to the canonical illuminant. Note that, in practical outdoor cases, at least two illuminants are present in the same scene (direct sunlight and diffuse light or shade). Inter-reflections normally coexist in the scene as well. In line with the vast majority of the existing approaches to color classification, we neglect to model inter-reflections in our work. The diagonal transformation matrices are estimated using either a Maximum Likelihood (ML) or a Maximum A Posteriori (MAP) criterion. We provide experimental evidence that the MAP criterion is superior to the ML approach because it reduces the risk of finding unlikely solutions. Once the color transformation matrices have been found, pixels are labeled into material and illuminant type based on a MAP criterion.

Our formulation for the total scene likelihood is very similar to the one proposed by Tsin et al. [25]. However, [25] uses a more complex (and more accurate) model for color transformation, based on a linear subspace approximation. This approach requires more complex training data acquisition procedures than ours, which can use just one training image. In addition, [25] must rely on an ad hoc optimization procedure, whereas we can use the well-understood EM algorithm. Finally, our system allows one to easily incorporate prior information about the distribution of color transformation matrices. This capability is critical for dealing with scenes containing surfaces that were not part of the training set. Without some assumptions about the statistics of the transformation parameters, the system would be tempted to "explain too much" of the scene, that is, to model the color of a surface never seen before as if produced by a surface with a familiar reflectance under a very unlikely light.

This contribution is organized as follows: In Section 2, we present our algorithm for finding the illuminant-induced color transformation parameters in a new scene, based on the ML and the MAP approaches. The mathematical details of our derivation are deferred until the Appendix. Experiments showing the effectiveness of our method are presented in Section 3 on mosaics of color squares from a Macbeth chart under a fixed (and known) set of illuminants, as well as on real outdoor scenes. A preliminary version of this paper appeared in [17].

2 The Algorithm

We will start by assuming that K surface classes of interest have been selected. As a practical rule, the number of classes should represent a compromise between descriptiveness (which favors a dense taxonomy in order to better describe the perceived world) and distinctiveness (which discourages the presence of multiple classes that cannot be discriminated unambiguously). For example, in [16], the authors use three classes in the context of terrain classification for outdoor robotic navigation: green grass/foliage, dry vegetation, and soil.

We will also assume that one or more hand-labeled images are available for training, satisfying the following critical condition: *All instances of all surface classes are seen under the same "canonical" illuminant.* These labeled samples are used to learn the class-conditional color distributions under the canonical illuminant. For meaningful modeling, it is important that the training image(s) contain a significant variety of viewing conditions for each surface. Note that we use the $c = (r, g, b)^T$ color vector directly, whereas others (e.g., [25]) modeled the contribution of brightness and chromaticity independently. We have found in our experiments that separating brightness and chromaticity does not appear to bring practical advantages.

Let $p(c|k)$ denote the conditional likelihood [21] of the color c given that the color was generated by the (non-specular) reflection of the canonical illuminant by a surface of class k. This probability density is learned from the labeled training data. The total likelihood of color c given that all surfaces are lit by the canonical illuminant is thus:[1]

$$p(c) = \sum_{k=1}^{K} P_K(k) p(c|k), \tag{1}$$

where $P_K(k)$ is the prior probability of seeing a surface of type k.

In general, a scene to be classified contains a number of surfaces, some (but not all) of which belong to the set of classes used for training and are illuminated by one or more sources of light which may be different from the canonical illuminant used for training. Assume that there are L possible illuminant types in the scene. Let c be the color of the pixel that is the projection of a certain surface patch under illuminant type l (e.g., direct sunlight). We will denote by $F_l(c)$ the operator that transforms c into the color that would be seen if the surface was lit by the canonical illuminant, all other conditions being the same. Note that we are implicitly assuming that $F_l(c)$ is a function only of c. As is well-known

1. We use the symbol $P(\cdot)$ to indicate probability mass distributions and $p(\cdot)$ to indicate probability density functions.

[27], this is a simplification, since two surfaces with different reflectance spectra may produce the same color under a certain illuminant and a different color under another illuminant (metamerism). It is exactly this simplification (which is at the core of the diagonal model) that allows us to use a straightforward optimization algorithm.

If the functions $F_l(c)$ are known, one may compute the conditional likelihood $p_F(c|k,l)$ of a color c given the class type k and illuminant type l using the well-known formula for the density of a function of a random variable:

$$p_F(c|k,l) = p(F_l(c)|k)|J(F_l)|_c, \tag{2}$$

where $|J(F_l)|_c$ is the absolute value of the Jacobian of F_l at c.

2.1 Modeling Assumptions

We will begin our analysis by making the following assumptions:

1. The surface class and illuminant type at any given pixel are mutually independent random variables.
2. The surface class and illuminant type at any given pixel are independent of the surface classes and illuminant types at nearby pixels.
3. The color of any given pixel is independent of the color of nearby pixels, even when they correspond to the same surface class and illuminant type.
4. Each surface element is illuminated by just one illuminant.
5. The class-conditional color likelihood under the canonical illuminant is a Gaussian density, with mean μ_k and covariance Σ_k: $p(c|k) \in \mathcal{N}(\mu_k, \Sigma_k)$.
6. The illuminant-induced color transformation can be modeled as:[2] $F_l(c) = D_l c$, where
$$D_l = \mathrm{diag}(d_{l,1}, d_{l,2}, d_{l,3}).$$
7. The diagonal coefficient vectors $d_l = (d_{l,1}, d_{l,2}, d_{l,3})$ for different illuminants in the scene are statistically independent.

Assumption 1 is fairly well justified: The fact that a surface is under direct sunlight or in the shade should be independent of the surface type. It should be noted, however, that, in the case of "rough" surfaces (e.g., foliage), self-shading will always be present, even when the surface is under direct sunlight. Assumption 2 is not very realistic: Nearby pixels are very likely to belong to the same surface class and illuminant type. This is indeed a general issue in computer vision, by no means specific to this particular application. One could therefore resort to standard approaches to deal with spatial coherence [4], [28], [26]. In this paper, however, we neglect spatial coherence for the sake of simplicity. The "white noise" Assumption 3 is not fully realistic (nearby pixels of the same smooth surface under the same illuminant are likely to have similar color), but, in our experience, it is a rather harmless, and computationally quite convenient, hypothesis. Assumption 4 is a good approximation for outdoor scenes, where the only two illuminants (excluding inter-reflections) are direct sunlight and diffuse light (shade). Obviously, when direct sunlight is occluded, only diffuse light falls upon the surface. Otherwise, the surface receives light both from the sun and the sky dome. However, as noted experimentally by Buluswar and Draper [6], direct sunlight is usually much stronger than diffuse light and, therefore, the contribution of the latter in this case is negligible. This is not true in the case of penumbra [14], which may occur in outdoor scenes, especially in the presence of plant canopies [24]. In the case of penumbra, as well as in the case of inter-reflections, Assumption 4 does not hold, and our algorithm may give unsatisfactory results. Assumption 5 may be acceptable as a first approximation for unimodal distributions. Multimodal distributions can be accounted for by using mixture-of-Gaussians models [21]. The extension of our optimization algorithm to the case of Gaussian mixtures is trivial. Assumption 6 formalizes the diagonal color transformation model. Note that, in this case, $|J(F_l)|_{c(x)} = |\det D_l| = |d_{l,1}d_{l,2}d_{l,3}|$. If c is normally distributed with mean μ and covariance Σ, then $F_l(c) = D_l c$ is normally distributed with mean $D_l\mu$ and covariance $D_l\Sigma D_l$. Finally, Assumption 7 is a simplifying hypothesis that will be used in Section 2.5 to model the joint prior distribution of illuminant-induced color transformations. The shortcomings of this assumption are clear if one thinks about a typical outdoor scene, where the color and intensity of sunlight and of skylight are usually quite different (meaning that the induced color transformations are not really independent).

2.2 Problem Statement

Our goal is to find, given the image $\underline{c} = \{c(x)\}$, the optimal (in the Bayesian sense) assignments of surface class labels $\underline{k} = \{k(x)\}$ and illuminant type labels $\underline{l} = \{l(x)\}$ over the pixels x. Formally:

$$\begin{aligned}(\underline{k}_{opt}, \underline{l}_{opt}) &= \arg\max P(\underline{k}, \underline{l}|\underline{c}) \\ &= \arg\max p(\underline{c}|\underline{k}, \underline{l}) P(\underline{k}, \underline{l}).\end{aligned} \tag{3}$$

If the set of L illuminant-induced color transformations $F = \{F_l\}$ is known, then one easily proves that, based on the independence assumptions of the previous section, it is possible to solve the assignment problem on a pixel-by-pixel basis:

$$\begin{aligned}(k(x), l(x)) &= \arg\max P_K(k) P_L(l) p_F(c(x)|k,l) \\ &= \arg\max P_K(k) P_L(l) p(D_l c(x)|k) |\det(D_L)|,\end{aligned} \tag{4}$$

where $P_K(k)$ and $P_L(l)$ are the marginal prior distributions of surface class and illuminant type, which may be estimated from the data or, as in this work, assumed uniform.

Unfortunately, the transformation set F is unknown and, in order to find the optimal label assignments, we would need to integrate over the distribution of color transformations:

$$(\underline{k}_{opt}, \underline{l}_{opt}) = \arg\max \int_F p_F(\underline{c}|\underline{k}, \underline{l}) p(F)\ dF\ \cdot P(\underline{k}, \underline{l}). \tag{5}$$

A drawback of the expression above as an optimization criterion is that pixel-by-pixel optimization as in (4) is no longer feasible. For this reason, instead of trying to solve (5), we resort to a suboptimal strategy. Rather than integrating over the space of transformations, we find the

2. We use the following convention, inspired by the homonynous Matlab function call. $\mathrm{diag}(D)$, where D is a square matrix, is a vector formed by the main diagonal of D. $\mathrm{diag}(d)$, where d is a vector, is a diagonal matrix with main diagonal formed by d.

transformations F_{opt} that best justify the image and then plug them into (4). Formally, we solve for F_{opt} based on either an ML or a MAP criterion:

$$\begin{aligned} \text{ML criterion:} \quad & F_{opt} = \arg\max L_F(\underline{c}) \\ \text{MAP criterion:} \quad & F_{opt} = \arg max L_F(\underline{c}) p(F), \end{aligned} \tag{6}$$

where $L_F(\underline{c})$ is the total log-likelihood of the image given the transformation set F:

$$\begin{aligned} L_F &= \sum_x \log \sum_{l=1}^{L} \sum_{k=1}^{K} P_{KL}(k,l) p_F(c(x)|k,l) \\ &= \sum_x \log \sum_{l=1}^{L} \sum_{k=1}^{K} P_K(k) P_L(l) p(F_L(c(x))|k) |J(F_L)|_{c(x)} \\ &= (2\pi)^{-3/2} \sum_x \log \sum_{l=1}^{L} \sum_{k=1}^{K} P_K(k) P_L(l) \\ &\quad \cdot e^{-0.5\ (D_L c(x) - \mu_K)^T \Sigma_K^{-1} (D_L c(x) - \mu_K)} \\ &\quad \cdot \det(\Sigma_K)^{-1/2} |d_{l,1} d_{l,2} d_{l,3}|. \end{aligned} \tag{7}$$

Note that the first summation extends over all image pixels and that L and K are the number of possible illuminants and surface classes, which are assumed to be known in advance. In our experiments with outdoor scenes, we always assumed that only two illuminants (sunlight and diffuse light) would be present, hence $L = 2$.

2.3 ML Estimation

The set of $3L$ color transformation coefficients $\{d_{l,m}\}$ that maximize the total image log-likelihood $L_F(\underline{c})$ in (7) can be found using Expectation Maximization (EM) [19]. EM is an iterative algorithm that reestimates the model parameters in such a way that the log-likelihood of the observables is increased (or remains the same) at each iteration. It is shown in the Appendix that each iteration is comprised of L independent estimations, one per illuminant type. For the lth illuminant type, one needs to compute:

$$\{d_{l,m}\} = \arg\max Q_L \tag{8}$$

with:

$$Q_L = u_l \sum_{\bar{m}=1}^{3} \log |d_{l,\bar{m}}| - 0.5\ d_l^T G_l d_l + H_l^T d_l, \tag{9}$$

where $d_l = (d_{l,1}, d_{l,2}, d_{l,3})^T$. The scalar u_l, the 3×3 matrix G_l and the 3×1 vector H_l are defined in the Appendix and are recomputed at each iteration. Maximization of (9) can be achieved by setting to zero the partial derivatives of Q_l with respect to the variables $d_{l,m}$. This yields the following system of quadratic equations for $m = 1, 2, 3$:

$$\sum_s G_{l,m,s} d_{l,s} d_{l,m} - H_{l,m} d_{l,m} - u_L = 0. \tag{10}$$

Unfortunately, this system cannot be solved for d_l in closed form. An iterative solution is possible based on the following observation. For each illuminant type l, if two variables (say, $d_{l,1}$ and $d_{l,2}$) are kept fixed, then Q_l can be maximized with respect to the third variable (in this case, $d_{l,3}$), since this is equivalent to solving a single quadratic equation (in this case, over $d_{l,3}$) in (10). Hence, Q_l can be maximized using the Direction Set Method [20]. For each illuminant type l, one iterates maximization of Q_l over each variable $(d_{l,1}, d_{l,2}, d_{l,3})$, in turn, until a predetermined convergence criterion is reached. Note that the complexity of this procedure is independent of the number of pixels in the image.

Finally, note that the EM algorithm converges to a local minimum of $L_F(\underline{c})$ and the result may thus depend on the choice of the starting point for the iterations.

2.4 Outliers

Any classification algorithm should account for unexpected situations that were not considered during training by recognizing outlier points which can be assigned to a "none of the above" class. A popular strategy treats outliers as an additional class for which a prior probability and a least informative conditional likelihood are defined [13]. For example, if one knows that the observables can only occupy a bounded region in the measurement space (a realistic assumption in the case of color features), one may allocate a uniform outlier distribution over such a region. In addition, we marked as "outlier" any saturated or undersaturated pixel, that is, such that its color has one or more components equal to 255 or equal to 0.

In our work, we defined an outlier surface class with an associated prior probability (defined by the user) and uniform conditional likelihood over the (r, g, b) color cube. Note that the outlier class contributes to the parameter estimation only indirectly, in the sense that (2) does not apply to it. In other words, color transformations do not change the conditional likelihood given the outlier class.

The use of an outlier class is highly recommended in the implementation of our algorithm. Parameter estimation is performed based on the currently estimated posterior class probabilities (23) at each pixel, limited to the set of nonoutlier classes. Without an outlier class, a pixel from a surface that was not considered for training would not only be erroneously classified, but it would also contribute incorrectly to the estimation of the illumination parameters. Experimentally, we found that the outlier class is particularly useful during the first iterations of the algorithm, steering the search away from obviously incorrect solutions.

2.5 MAP Solution

A problem with the ML approach is that the algorithm, when presented with scenes that contain surfaces with very different colors from the ones used for training, may arbitrarily "invent" unusual illuminants in order to maximize the scene likelihood. The variety of direct sunlight and diffuse illuminants (excluding inter-reflections) in an outdoor scene occupy only a small portion of the space of all possible illuminants [15], [6]. This knowledge may be used to constrain the search space in our algorithm. We propose a simple Gaussian model for the distribution of the diagonal transformation coefficients and learn the model parameters (mean, μ_d, and covariance, Σ_d) from a number of training images taken under many different illumination conditions. A Macbeth color chart, photographed at different times of the day, can be used for this purpose. For each ordered pair of pictures taken of the chart under

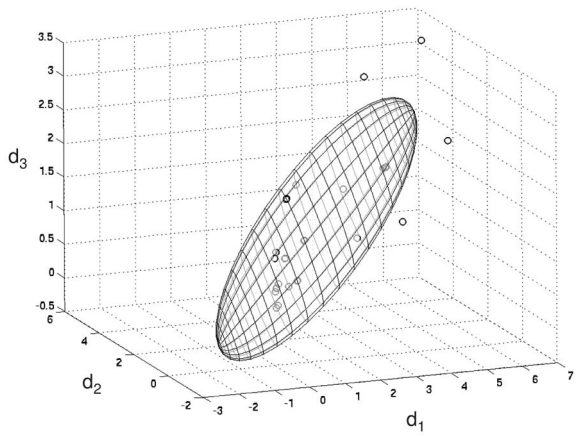

Fig. 1. The ensemble of diagonal matrices (represented as points in 3D space) used to estimate the prior distribution of the diagonal color transformations. The ellipsoid represents an equilevel surface of the best-fitting Gaussian.

different illuminants, one diagonal transformation matrix can be computed by least-squares regression. The hope is that the ensemble of all such matrices can adequately model the distribution of diagonal color transformations. Note that this operation requires no prior color camera calibration, except for compensating for the camera's nonlinearities [12].

Fig. 1 shows a set of diagonal matrices $\{D\}$ (represented as 3D vectors $d = \mathrm{diag}(D)$) corresponding to six daylight illuminants (overall, $6 \cdot 5 = 30$ transformations), together with an equilevel surface of the best fitting Gaussian density $p_d(d)$. It can be seen from Fig. 1 that the data is not normally distributed and that a Gaussian density can only provide a rough approximation to the actual distribution (a mixture of Gaussians would be a better fit). However, this choice is justified by algorithmic convenience since the presence of a Gaussian prior implies only a simple modification of the ML algorithm, as discussed below. Note from Fig. 1 that the data is mostly concentrated in a 2D subspace (the eigenvalues of the covariance matrix are: (6.201, 0.266, 0.006)). This is not surprising if one considers that the variability of daylight is fairly well modeled by a parabola or even a line in chromaticity space [15], [6].

The prior distribution of color transformation matrices can be plugged into the same EM machinery used for ML optimization, by suitably modifying the function to be maximized at each iteration. More precisely, as discussed in [19], imposing a prior distribution on the parameters $\{D_l\}$ translates into adding the term $\log p_D(\{D_l\})$, representing the prior probability of the ensemble of diagonal coefficients $\{D_l\}$, to the function $Q(\{D_l\}, \{D_l^0\})$ in the Appendix before the maximization step (24). Based on Assumption 7 of Section 2.1, we may write:

$$\log p_D(\{D_L\}) = \sum_{l=1}^{L} \log p_D(d_L), \tag{11}$$

where $D_l = \mathrm{diag}(d_{l,1}, d_{l,2}, d_{l,3})$. For what concerns the algorithm of Section 2.3, it is necessary to modify the definition of Q_l in (9) as follows:

$$\begin{aligned} \bar{Q}_L = u_l \sum_{\bar{m}=1}^{3} \log |d_{l,\bar{m}}| - 0.5\ d_l^T (G_l + \Sigma_D^{-1}) d_l \\ + (H_l^T + \mu_d^T \Sigma_d^{-1}) d_l. \end{aligned} \tag{12}$$

Fortunately, even with this new formulation of the optimization criterion, the Direction Set algorithm can still be used for maximization of $\bar{Q}_l$.

3 Experiments

3.1 Macbeth Color Chart Experiments

In order to provide a quantitative evaluation of our algorithm's performance, we first experimented using mosaics formed by Macbeth color squares under different illuminant types. We generated colors using the spectral curves for reflectances, illuminants, and sensor responses in the database[3] described in [2]. Since this work is mainly concerned with outdoor images, we only used the six illuminants in the database that emulate daylight.

We picked five color squares from the chart (No. 2, 22, 9, 17, 12) as representative of some hypothetical classes of interest. The first two and the last three patches were purposely chosen to be have rather similar colors in order to make the test more challenging. These patches, which are used for system training, are shown on top of Fig. 2a, as seen under a "canonical" illuminant (*Illuminant 1*, Solux 4100 K°). Also shown in Fig. 2a are the same surfaces seen under *Illuminant 2* (Solux 3500 K°) and *Illuminant 3* (Solux 4700 K°). In the bottom of Fig. 2a, we show four "distractor" patches (No. 4, 6, 11, 19) seen under Illuminant 2 and 3. Test queries were formed by picking random numbers (n_1 and n_2) of randomly chosen training patches seen under Illuminant 2 and 3, respectively, with $3 \leq n_1 + n_2 \leq 9$, along with n_d distractors ($0 \leq n_d \leq 3$) seen under Illuminant 2 and n_d distractors seen under Illuminant 3. An example of query formation is shown in Fig. 2a. Overall, a query contains $n_1 + n_2 + 2n_d$ patches. The goal of the test was to estimate two diagonal rendering matrices ($\bar{D}_{2\to1}$, $\bar{D}_{3\to1}$), mapping the colors under Illuminant 2 and 3, respectively, into colors under Illuminant 1, as well as to classify each patch in the query. The number of correct matches for nondistractor colors at each test was recorded, together with the number of nondistractor colors in each query. The ratio between the cumulative number of correct matches and the cumulative number of nondistractor squares in the test mosaics over all tests provides an indication of the probability of correct match. The estimated diagonal rendering matrices $\bar{D}_{2\to1}$, $\bar{D}_{3\to1}$ for each test were compared to the "ground truth" diagonal matrices:

$$D_{2\to1} = \mathrm{diag}(1.43, 1.67, 1.84),\ D_{3\to1} = \mathrm{diag}(1.51, 1.35, 1.07) \tag{13}$$

obtained by least square regression. For each query, normalized errors were computed as:

3. www.cs.sfu.ca/~colour/data/.

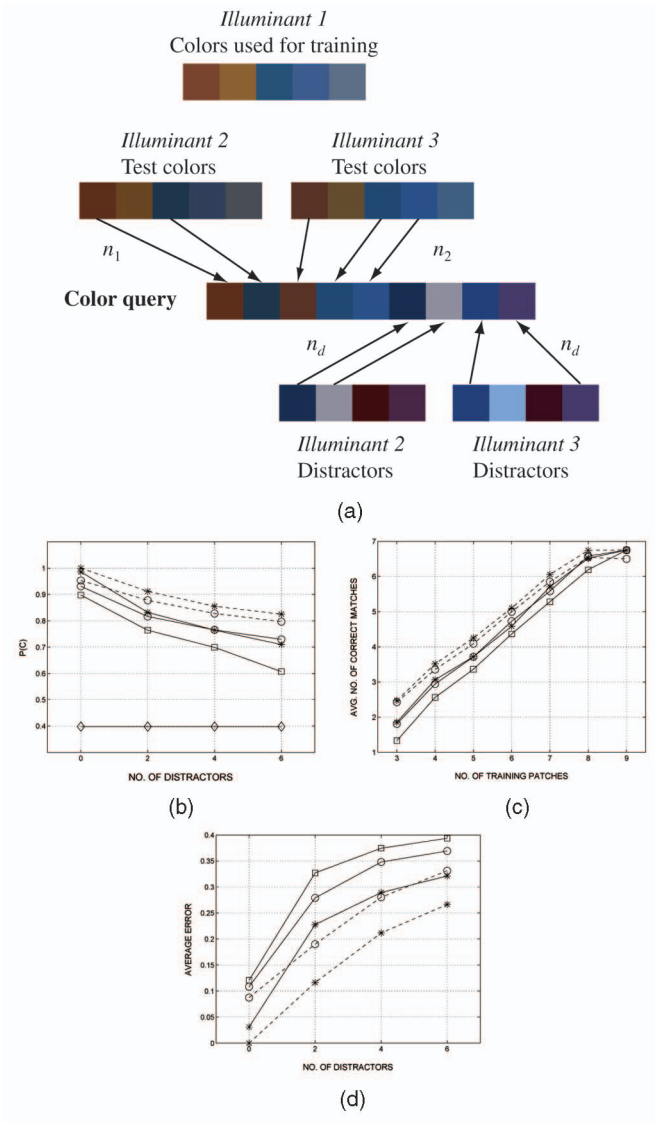

Fig. 2. (a) An example of query formation for the experiments of Section 3.1. (b) The probability of correct classification $P(C)$ as a function of the overall number of distractors for the following cases: without color compensation (diamonds), with ML-based compensation (squares), with MAP-based compensation based on six illuminants (circles, solid line) and three illuminants (circles, dashed line), exhaustive search based on six illuminants (stars, solid line) and three illuminants (stars, dashed line). (c) The average number of correct matches as a function of the overall number $(n_1 + n_2)$ of "training patches" in the query for a fixed overall number $(2n_d = 4)$ of distractors. (d) The average normalized parameter estimation error (14) as a function of the overall number of distractors.

$$e_D = 0.5 \cdot \frac{\|\mathrm{diag}(D_{2\to1}) - \mathrm{diag}(\bar{D}_{2\to1})\|}{\|\mathrm{diag}(D_{2\to1})\|} + 0.5 \cdot \frac{\|\mathrm{diag}(D_{3\to1}) - \mathrm{diag}(\bar{D}_{3\to1})\|}{\|\mathrm{diag}(D_{3\to1})\|}. \quad (14)$$

The errors in (14) were then averaged over all tests.

For each value of n_d between 0 and 3, we ran 1,000 tests, each time creating a new query color mosaic, and estimated the probability of correct match, $P(C)$, as well as the average of the parameter estimation errors, e_D. We tested with both the ML and MAP algorithms for parameter estimation, setting the number L of illuminants to 2. Since only one color sample per class and per illuminant was available during training, in this experiment, we artificially set the covariance matrices in (7) to $\Sigma_k = \sigma^2 I_3$, where I_3 is the 3×3 identity matrix and σ was $1/10$ of the maximum value measured in any color channel. In all experiments, the initial values for the two triplets of diagonal coefficients were as follows:

$$d_{1,1} = d_{1,2} = d_{1,3} = 0.5\,,\ d_{2,1} = d_{2,2} = d_{2,3} = 1.5. \quad (15)$$

When running MAP parameter estimation, we considered two choices for the prior density of the diagonal matrices. In the first case, we considered all six available illuminants in the data set, while, in the second case, we only considered the three illuminants used when creating the query mosaics. This test is meant to assess the importance of prior knowledge of the distribution of expected illuminants in the scene. Clearly, we expected better results in the second case, where the prior is concentrated around the actual values used in the experiment.

In addition, we considered a direct search algorithm to maximize the log-likelihood function in (7). Since, in this experiment, the set of illuminants is known in advance, one may compute a priori all of the diagonal rendering matrices mapping colors under one illuminant into another illuminant. For N different illuminants, $N(N-1)(N-2)/2$ tests need to be performed for exhaustive search. Similarly to the MAP case, we experimented with $N = 6$ and $N = 3$ illuminants. The expectation was that exhaustive search should perform at least as well as EM. The reason why it may not give the optimal result (i.e., it may not select the matrices $D_{2\to1}$ and $D_{3\to1}$ in (13) even though they are in the pool of candidates) is the limited number of training patches in the query mosaic along with the presence of distractors.

Fig. 2b shows the probability of correct classification $P(C)$ for the different cases considered as a function of the total number of distractors $2n_d$, together with the results in the case of a nearest neighbor classifier without color compensation. The performance of the latter is obviously independent of the number of distractor. The classifier based on (4) using the optimal rendering matrices $D_{2\to1}$ and $D_{3\to1}$ in (13) gives perfect classification ($P(C) = 1$, not shown in the graph), regardless of the number of distractors. Fig. 2c shows the average number of correctly classified patches as a function of the number of "training patches" (i.e., patches that were used during training, albeit under a different illuminant) in the query for a fixed number (four) of distractors. Finally, Fig. 2c shows the average normalized error (14) as a function of the number of distractors.

A number of observations can be drawn from these experiments. First, the presence of distractors impairs the system performances, as expected. Second, ML parameter estimation provides a substantial improvement with respect to the noncompensated case. Third, prior knowledge of the distribution of rendering matrices under the MAP framework of Section 2.3 can further improve the classifiers' performances. Better prior distribution modeling leads to higher classification rates. Finally, direct search in the discrete space of possible solutions performs better than MAP optimization, although the difference in terms of correct classification rate $P(C)$ is not very large (the difference of estimation errors is more pronounced). It should be noted that, in this particular case, the optimal (in

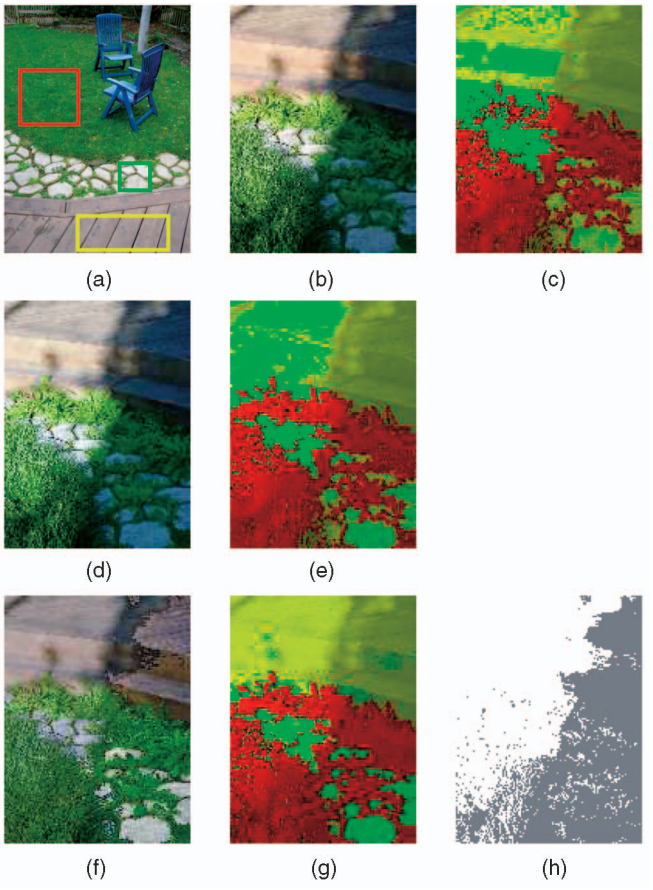

Fig. 3. Color classification experiments. (a) Training image (regions actually used for training are contoured by a line whose color is representative of the surface class). (b) Test image. (c) Surface classification without illuminant compensation (red: grass, green: cobblestones, yellow: wood). $P(C) = 0.77$. (d) and (e) Color-compensated image and surface classification using the gray-world method. $P(C) = 0.72$. (f), (g), and (h) Color-compensated image, surface classification, and illuminant classification with ML-based parameter estimation. $P(C) = 0.94$.

least squares terms) solutions were part of the pool of candidates, which is not necessarily the case in general. We defer further comments about this point until Section 4.

3.2 Experiments with Real-World Scenes

We present three sets of experiments of color classification of real scenes in Figs. 3, 4, and 5. For the images of Fig. 3 and of Fig. 5, we used a Sony DSC-S75 camera with automatic exposure, taking care to compensate for the camera's gamma function. In all three cases, we set the number of illuminants L to two, with equal prior probability. Likewise, the class prior probabilities $P_K(k)$ of the nonoutlier classes were identical. The prior probability of the outlier class was kept constant for all tests within each experiment set (when the outlier class was actually used). Initial values for the triplets of coefficients were as in (15). For comparison, we also included a case with color normalization based on the well-known gray-world model. In this case, we found the diagonal transformation matrix D_{gw} such that the average color of the transformed image was identical to the average color of the training image. Then, classification was performed assuming $L = 1$, and solving for (4) with $D_1 = D_{gw}$.

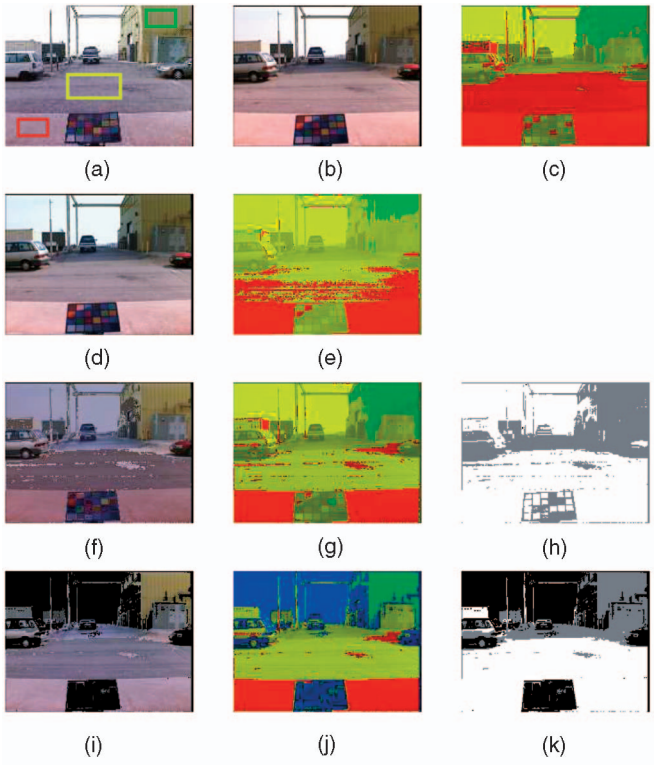

Fig. 4. Color classification experiments. (a) Training image (regions actually used for training are contoured by a line whose color is representative of the surface class). (b) Test image. (c) Surface classification without illuminant compensation (red: concrete, yellow: asphalt, green: painted wall). $P(C) = 0.51$. (d) and (e) Color-compensated image and surface classification using the gray-world method. $P(C) = 0.71$. (f), (g), and (h) Color-compensated image, surface classification, and illuminant classification with ML-based color compensation (without outlier class). $P(C) = 0.93$. (i), (j), and (k) Using an outlier class. Outliers are marked in blue in the classification image and in black in the color-compensated and illuminant classification image. $P(C) = 0.94$.

In order to provide a quantitative assessment of the classifier's performance using the different proposed compensation methods, we manually labeled the test images for the considered classes and built the confusion matrices CM after each classification. The (i, j) entry of CM represents the number of pixels that were hand-labeled as class i and classified as class j. Our estimate of the correct classification rate is the following:

$$P(C) = \frac{\sum_i CM_{i,i}}{\sum_i \sum_j CM_{i,j}}. \tag{16}$$

Note that pixels belonging to a hand-labeled area but classified as outliers are automatically considered as misclassified.

Our first experiment is shown in Fig. 3. Fig. 3a shows the training image. The conditional color distributions for three classes of interest (grass, marked in red, cobblestones, marked in green, wood, marked in yellow) were estimated using the pixels within the rectangles visible in Fig. 3a. Fig. 3b shows the test image, containing surfaces belonging to all three classes under mixed illumination (sunlight and shade). The classification without illuminant compensation is shown in Fig. 3c. Large portions of the wooden patio and of the cobblestones were misclassified $(P(C) = 0.77)$. The compensated image

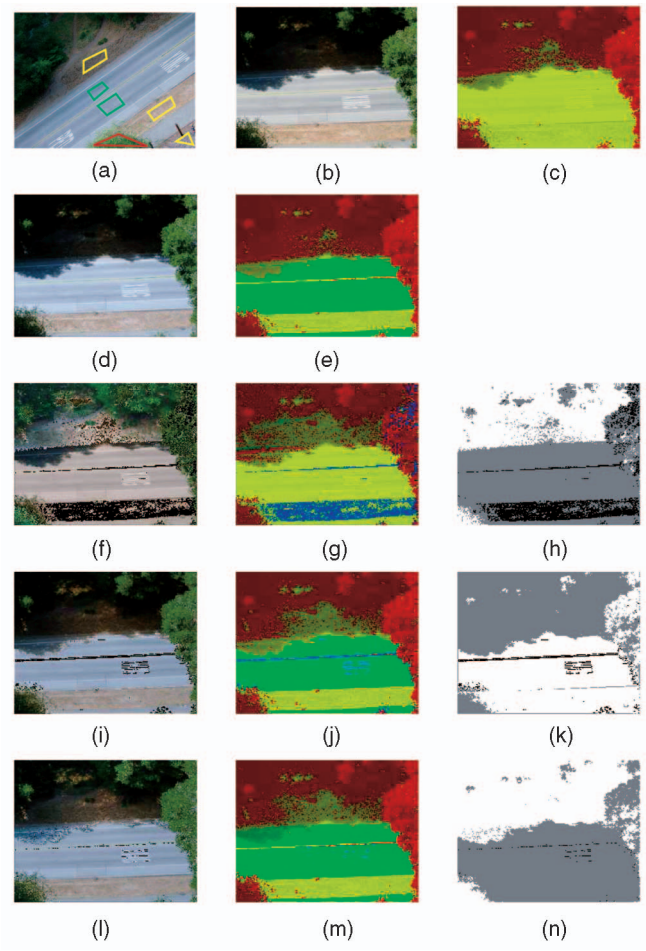

Fig. 5. Color classification experiments. (a) Training image (regions actually used for training are contoured by a line whose color is representative of the surface class). (b) Test image. (c) Surface classification without illuminant compensation (red: foliage, green: asphalt, yellow: soil, blue: outlier). $P(C) = 0.43$. (d) and (e) Color-compensated image and surface classification using the gray-world method. $P(C) = 0.86$. (f), (g), and (h) Color-compensated image, surface classification, and illuminant classification with ML-based color compensation. Outliers are marked in black in the color-compensated and illuminant classification image. $P(C) = 0.36$. (i), (j), and (k) Using the MAP procedure. $P(C) = 0.89$. (l), (m), and (n) Hand-selecting the transformation coefficients. $P(C) = 0.94$.

using the gray-world method is shown in Fig. 3d, with the corresponding classification in Fig. 3e. This type of color compensation actually impaired performances in this case $(P(C) = 0.72)$, probably because the color distributions in the training and test image were rather different. Color correction and classification after ML-based illuminant compensation (without an outlier class) are shown in Figs. 3f and 3g. Note that most of the image has been correctly classified $(P(C) = 0.94)$. Note also how color compensation "casts light" over the areas in the shade, except for the region in the penumbra. As noted in Section 2.1, penumbra is not well modeled by our assumptions and, therefore, color is not correctly compensated in this case. The assignment of illuminant types is shown in Fig. 3h.

Figs. 4a and 4b show two images of the same scene under very different illuminants. Gaussian color models for three surface classes (concrete, marked in red, asphalt marked in yellow, painted wall, marked in green) were trained over the rectangular areas shown in Fig. 4a. Note that several other surface types are present in the image. The result of classification without illuminant compensation is shown in Fig. 4c. Due to the similar color of concrete and asphalt, and to the change of illuminant, this classifier mistakenly labels the asphalt region as concrete $(P(C) = 0.51)$. Gray-world color compensation (Figs. 4d and 4e) improves the correct classification rate $(P(C) = 0.71)$. After ML-based color compensation without the outlier class (Figs. 4f, 4g, and 4h), almost all of the scene is correctly classified $(P(C) = 0.93)$. However, color compensation is not very good since a large portion of the sky (which was not used during training) is visible. The sky is classified as asphalt under sunlight, creating an undesired bias in the image log-likelihood (7). When the outlier class is introduced (Figs. 4i, 4j, and 4k), the sky is classified as outlier and no longer affects the estimation of the color compensation matrices, thus enabling more accurate color compensation and a small increase in the correct classification rate $(P(C) = 0.94)$. For more classification results of the same scene, taken at different times throughout the day, please see www.soe.ucsc.edu/~manduchi/theColorTestMovie.mov.

Our third experiment, shown in Fig. 5, involves pictures of a road taken from a tall building at different times of the day. In this case, the surface classes of interest are: asphalt (marked in green), soil (marked in yellow), and foliage (marked in red). As in the previous figure, outliers are marked in blue. The training image (Fig. 5a) was taken in the morning in overcast conditions; the test image (Fig. 5b) was taken in the afternoon with clear sky. The result of classification without illuminant compensation is shown in Fig. 5c. The correct classification rate is very low $(P(C) = 0.43)$, as expected given the color shift between the training and the test image. Gray-world color compensation (Figs. 5d and 5e) in this case performs very well, increasing the correct classification rate to $P(C) = 0.86$. Classification using ML illuminant compensation (Figs. 5f, 5g, and 5h) produces disappointing results, even with an outlier class $(P(C) = 0.36)$. The situation improves considerably using MAP color compensation (Fig. 5i, 5j, 5k, and 5l), leading to $P(C) = 0.89$. This example highlights the importance of constraining the search space of diagonal transformations based on their prior distribution. Note, however, that even though most of the image is correctly classified, some areas of error remain. In particular, the region of shadow on the asphalt is incorrectly classified as vegetation. To verify whether this remaining misclassfication is simply due to a suboptimal choice of the diagonal coefficients or to the inability of the diagonal model to adequately compensate for the color change, we hand-picked values for the diagonal coefficients in an empirical trial-and-error procedure in order to maximize the correct classification rate. The result is shown in Figs. 5l, 5m, and 5n. Now, the shadow has been (correctly) classified as asphalt, proving that, at least for this image, the diagonal model was powerful enough for classification purposes. It should be noted, however, that the color of surfaces in the shade was not perfectly recovered, probably due to the presence of penumbra under the tree canopy. The fact that our technique only found a suboptimal solution for this rather difficult case is probably due to the several simplifying assumptions underpinning the algorithm and discussed in Section 2.1.

4 Discussion and Conclusions

We have presented a method for outdoor color classification that explicitly estimates the parameters of the illuminant-induced color transformations in the scene, determines which illuminant impinges on each surface element, and classifies surfaces into different types in a predefined taxonomy. We proposed two criteria (ML and MAP) for estimating the color transformation parameters. The ML approach requires no previous color calibration nor any knowledge of the statistics of illumination. The MAP criterion uses a simple procedure, whereby pictures of a color target are taken under different illumination conditions. Differently from previous approaches, our method uses the Expectation Maximization algorithm for parameter estimation. Our experiments show the effectiveness of our technique for color-based classification of outdoor images when only one image is used for training.

A competing approach to EM optimization is exhaustive search in a discrete space of possible color transformations. As shown in Section 3.1, this strategy gives very good results (when the actual solution is within the search space) at a complexity of $\mathcal{O}(N^3)$, where N is the number of different illuminants considered when building the search space. For example, for the estimation of the prior density of parameters (as in Section 2.5) in the experiments of Section 3.2, we used 39 pictures of a color target during different times of the day and under different weather conditions. Exhaustive search would require more than 25,000 tests, whereas EM converges in less than 50 iterations on average. Of course, one could reduce the size of the search space by some sort of subsampling, although this brings in the risk of removing the optimal solution from the pool of candidates. In conclusion, with respect to exhaustive search, our approach to MAP seems to offer a good trade-off between computational complexity and accuracy.

Other color constancy algorithms based on statistical models have been proposed. In general, the idea is to use prior knowledge of the statistics of the parameters involved to solve an otherwise ill-posed problem. For example, Brainard and Freeman [5] define a 3-3 model, with reflectance and illumination coefficients following truncated Gaussian distributions. Then, they show how to determine the illuminant of a Mondrian synthetic scene with eight randomly selected color patches, using MAP as well as a so-called MLM (*Maximum Local Mass*) criteria, comparing their results with other mainstream nonstatistical color constancy algorithms. They maintained that MLM and the gray world model outperform other approaches, while MAP fails to estimate the correct magnitude of the illuminant coefficients. The main problem with the implementation of this algorithm is that the optimal MAP or MLM solution is obtained via exhaustive search in a $3(N_s + 1)$-dimensional parameter space, where N_s is the number of surfaces in the scene. Note that a single illuminant is considered in the algorithm.

A similar problem is addressed by Forsyth et al. [9]. They propose a 6-6 model with an additional specular component (described by Phong's model) under a point source at unknown position. The number of patches, placed in a rectangular grid in the considered Mondrian image, is also supposed unknown but less than 8. The prior distributions of the reflectance and illumination parameters are set to be constant within the "realizability" area (i.e., such that the resulting illumination spectra are positive and the resulting reflectance spectra are positive and less than 1) and exponentially decaying outside the constraints set. The locations of patch edges are modeled by a Poisson distribution. The authors show how to sample a set of parameters from their posterior distribution using MCMC. Although the results shown are very good, it is not clear how to apply this technique to non-Mondrian images. In addition, MCMC sampling can be extremely slow to "burn in."

The "color by correlation" method [8], [1] is, in many aspects, similar to our technique in that it uses a representation of the likelihood of the image colors given a certain illuminant. However, there are some important differences between the two approaches. First, color by correlation uses a nonparametric representation of the color likelihood (based on histograms), while our method represents color distributions by parametric models (Gaussians). A problem with histograms is that the choice of the bin size is critical. In addition, histograms do not scale well in multiple dimensions [21], [1]. Parametric models, on the converse, can represent well only specific families of distributions. Second, in order to "generalize" from the set of observed colors, color by correlation takes the convex hull of the chromaticities [8] (or filters the histograms [1]). Our parametric model (which produces inherently smooth densities) naturally generalizes from the observed samples. Third, when computing the conditional likelihood of the colors of a given image, color by correlation only counts each color once (even if several pixels in the image have the same color). According to [8], this is done in order to avoid that large image areas with the same color dominate in the computation. However, this also means that an "outlier" color will be given the same weight as any legal color. We believe that maximizing the actual image likelihood (under the hypotheses of Section 2.1) is a more natural and robust optimization criterion. Fourth, in order to build the correlation matrix, all expected surfaces must be seen by all possible illuminants. This training phase may be costly and impractical. In our method, the distribution of color transformations (which is independent of the actual surfaces) is learned once and for all. Then, only a small sample of the surfaces of interest need to be observed under just one illuminant. Finally, search over the illuminant space is performed exhaustively by the color by correlation algorithm, whereas our method uses an iterative fix point procedure (EM). Although exhaustive search is certainly feasible if only chromaticities are considered (as in [8]), it becomes more problematic if the illuminant intensity is also considered [1].

Our method (as well as the one of [25]) explicitly models the case of two (or more) illuminants in the scene, whereas the majority of algorithms for color constancy assume a

single illuminant. An outstanding exception is the approach of [3], which implements an initial segmentation to separate areas with possibly different illumination, and uses variation of illumination to constrain the solution.

Finally, we should comment on the fact that our method performs color compensation and classification in the (r, g, b) space, rather than in chromaticity space. Previous work with varying illumination or scene geometry often projects colors onto some chromaticity space (e.g., [11], [18]). Our experience is that neglecting image brightness is often counterproductive for recognition (see also [1], [22]). Tsin et al. [25] model and estimate the brightness at each pixel explicitly, whereas, in our case, the brightness information is embedded in the (r, g, b) representation. Although color representations other than (r, g, b) could be used in our algorithm (provided that the diagonal rendering model is still supported, at least approximately), our experiments seem to indicate that modeling distributions in (r, g, b) space is good enough for our purpose.

APPENDIX

ALGORITHMIC DETAILS

In this appendix, we show how the total log–likelihood $L_F(\underline{c})$ in (7) can be maximized over the parameters $\{d_{l,m}\}$ using Expectation Maximization (EM). As is customary with EM, we introduce the "hidden" variables $z_{l,k}(x)$ representing the (unknown) label assignments: $z_{k,l}(x) = 1$ if the pixel x is assigned to the illuminant type l and surface class k; $z_{k,l}(x) = 0$ otherwise. We will denote the set of $z_{k,l}(x)$ over the image by $\underline{z}$.

The EM algorithm starts from arbitrary values for the diagonal matrices $\{D_l^0\}$, and iterates over the following two steps:

1. Compute:

$$\begin{aligned} Q(\{D_l\},\{D_l^0\}) &= E_{\{D_l^0\}}[\log p_{\{D_l\}}(\underline{c},\underline{z})|\underline{c}] \\ &= E_{\{D_l^0\}}[\log p_{\{D_l\}}(\underline{c}|\underline{z})|\underline{c}] \\ &\quad + E_{\{D_l^0\}}\big[\log p_{\{D_l\}}(\underline{z})|\underline{c}\big], \end{aligned} \tag{17}$$

   where $E_{\{D_l^0\}}[\cdot|\underline{c}]$ represents expectation over $P_{\{D_l^0\}}(\underline{z}|\underline{c})$ and $P_{\{D_l\}}(\underline{z}|\underline{c})$ is the posterior probability of the label assignment given the image $\underline{c}$ and the diagonal transformation matrices $\{D_l\}$.
2. Replace $\{D_l^0\}$ with arg max $Q\big(\{D_l\},\{D_l^0\}\big)$.

Given the assumed independence of label assignments and the particular form chosen for variables $z_{k,l}(x)$, we can write:

$$\log p_{\{D_l\}}(\underline{c}|\underline{z}) = \sum_x \sum_{k=1}^{K} \sum_{l=1}^{L} z_{k,l} \log p(D_l c(x)|k)| \det D_l|, \tag{18}$$

where, given that the conditional likelihood are assumed to be Gaussian:

$$\begin{aligned} &\log p(D_l c(x)|k) = \\ &-1.5\ \log(2\pi) - 0.5 \log|\det \Sigma_k| \\ &-0.5\ (D_l c(x) - \mu_k)\Sigma_k^{-1}(D_l c(x) - \mu_k). \end{aligned} \tag{19}$$

Also,

$$\log P_{\{D_l\}}(\underline{z}) = \sum_x \sum_{k=1}^{K} \sum_{l=1}^{L} z_{k,l}(\log P_K(k) + \log P_L(l)). \tag{20}$$

It is easy to see that

$$\begin{aligned} &E_{\{D_l^0\}}\big[\log p_{\{D_l\}}(\underline{c}|\underline{z})|\underline{c}\big] = \\ &\sum_x \sum_{k=1}^{K} \sum_{l=1}^{L} \log p(D_l c(x)|k)| \det D_l| P_{\{D_l^0\}}(k,l|c(x)) \end{aligned} \tag{21}$$

and that

$$\begin{aligned} &E_{\{D_l^0\}}\big[\log P_{\{D_l\}}(\underline{z})|\underline{c}\big] = \\ &\sum_x \sum_{k=1}^{K} \sum_{l=1}^{L} (\log P_K(k) + \log P_L(l)) P_{\{D_l^0\}}(k,l|c(x)), \end{aligned} \tag{22}$$

where $P_{\{D_l^0\}}(k,l|c(x)) = E_{\{D_l^0\}}[z_{k,l}(x)]$ is the posterior probability of surface class k and illumination type l given the observation $c(x)$ and color transformation matrices $\{D_l^0\}$. Using Bayes' rule, we can compute $P_{\{D_l^0\}}(k,l|c(x))$ as by

$$\begin{aligned} &P_{\{D_l^0\}}(k,l|c(x)) = \\ &\frac{P_K(k) P_L(l) p(D_l^0 c(x)|k)| \det D_l^0|}{\sum_{\bar{k}=1}^{K} \sum_{\bar{l}=1}^{L} P_K(\bar{k}) P_L(\bar{l}) p(D_{\bar{l}}^0 c(x)|\bar{k})| \det D_{\bar{l}}^0|}. \end{aligned} \tag{23}$$

Remembering that $|\det D_l| = |d_{l,1} d_{l,2} d_{l,3}|$, one sees that

$$\max_{\{d_{l,m}\}} Q\big(\{D_l\},\{D_l^0\}\big) = \tag{24}$$

$$= \max_{\{d_{l,m}\}} \sum_{l=1}^{L} \left[u_l \sum_{m=1}^{3} \log|d_{l,m}| - 0.5\ d_l^T G_l d_l + H_l^T d_l \right],$$

where $d_l = (d_{l,1}, d_{l,2}, d_{l,3})^T$ and

$$u_l = \sum_x \sum_{k=1}^{K} P_{\{D_l^0\}}(k,l|c(x)), \tag{25}$$

$$G_{l,m,s} = \sum_x \sum_{k=1}^{K} S_{k,m,s} c_m(x) c_s(x) P_{\{D_l^0\}}(k,l|c(x)), \tag{26}$$

$$H_{l,m} = \sum_x \sum_{k=1}^{K} \sum_{s=1}^{3} S_{k,m,s} c_s(x) \mu_{k,s} P_{\{D_l^0\}}(k,l|c(x)), \tag{27}$$

where $S_k = \Sigma_k^{-1}$. Note that the terms in the summation over l in (24) can be maximized independently, meaning that the L sets of diagonal transformations can be computed independently at each iterations.

ACKNOWLEDGMENTS

This material is based on work supported in part by the US Defense Advanced Research Projects Agency through subcontract 1235249 from JPL and in part by the US National Science Foundation through contract IIS-0222943. The author is grateful to the anonymous reviewers whose comments and suggestion helped improve the original manuscript.

REFERENCES

[1] K. Barnard, L. Martin, and B. Funt, "Colour by Correlation in a Three Dimensional Colour Space," *Proc. Sixth European Conf. Computer Visualiztion (ECCV '00),* pp. 375-389, 2000.

[2] K. Barnard, L. Martin, B. Funt, and A. Coath, "A Data Set for Colour Research," *Color Research and Application,* vol. 27, no. 3, pp. 147-151, 2002.

[3] K. Barnard, G. Finlayson, and B. Funt, "Color Constancy for Scenes with Varying Illumination," *Computer Visualization and Image Understanding,* vol. 65, no. 2, pp. 311-321, Feb. 1997.

[4] J. Besag, "On the Statistical Analysis of Dirty Pictures," *J. Royal Statistical Soc. B.,* vol. 48, pp. 259-302, 1986.

[5] D.H. Brainard and W.T. Freeman, "Bayesian Color Constancy," *J. Optical Soc. Am., A,* vol. 14, pp. 1393-1411, July 1997.

[6] S. Buluswar and S.D. Draper, "Color Models for Outdoor Machine Vision," *Computer Visualiztion and Image Understanding,* vol. 85, no. 2, pp. 71-99, 2002.

[7] G.D. Finlayson, M. Drew, and B. Funt, "Color Constancy: Generalized Diagonal Transform Suffices," *J. Optical Soc. Am., A,* vol. 11, no. 11, pp. 3011-3019, Nov. 1994.

[8] G.D. Finlayson, S.D. Hordley, and P.M. Hubel, "Color by Correlation: A Simple, Unifying Framework for Color Constancy," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 11, pp. 1209-1221, Nov. 2001.

[9] D.A. Forsyth, J. Haddon, and S. Ioffe, "The Joy of Sampling," *Int'l J. Computer Visualization,* vol. 41, pp. 109-134, 2001.

[10] D. Forsyth, "A Novel Approach for Color Constancy," *Int'l J. Computer Visualization,* vol. 5, pp. 5-36, 1990.

[11] G. Healey, "Segmenting Images Using Normalized Colors," *IEEE Trans. Systems, Man, and Cybernetics,* vol. 22, no. 1, pp. 64-73, Jan. 1992.

[12] M.D. Grossberg and S.K. Nayar, "What Is the Space of Camera Response Functions?" *Proc. IEEE Computer Vision and Pattern Recognition,* 2003.

[13] E.T. Jaynes, *Probability Theory: The Logic of Science.* Cambridge Univ. Press, 2003.

[14] C. Jiang and M.O. Ward, "Shadow Identification," *Proc. IEEE Computer Vision and Pattern Recognition,* 1992.

[15] D. Judd, D. MacAdam, and G. Wyszecki, "Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature," *J. Optical Soc. Am.,* vol. 54, no. 8, pp. 1031-1040, 1964.

[16] R. Manduchi, A. Castano, A. Talukder, and L. Matthies, "Obstacle Detection and Terrain Classification for Autonomous Off-Road Navigation," *Autonomous Robots,* vol. 18, pp. 81-102, 2005.

[17] R. Manduchi, "Learning Outdoor Color Classification from Just One Training Image," *Proc. European Conf. Computer Vision,* 2004.

[18] J. Matas, R. Marik, and J. Kittler, "Illumination Invariant Colour Recognition," *Proc. Fifth British Machine Vision Conf.,* 1994.

[19] G.J. McLachlan and T. Krisnan, *The EM Algorithm and Extensions.* Wiley, 1997.

[20] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing.* Cambridge Univ. Press, 1986.

[21] B.D. Ripley, *Pattern Recognition and Neural Networks.* Cambridge Univ. Press, 1996.

[22] G. Schaefer, "How Useful Are Colour Invariants for Image Retrieval?" *Proc Second Int'l Conf. Computer Vision and Graphics,* 2004.

[23] D.A. Slater and G. Healey, "What Is the Spectral Dimensionality of Illumination Functions in Outdoor Scenes?" *Proc. IEEE Computer Vision and Pattern Recognition,* pp. 105-110, 1998.

[24] W.K. Smith, A.K. Knapp, and W.A. Reiners, "Penumbral Effects on Sunlight Penetration in Plant Communities," *Ecology,* vol. 70, no. 6, pp. 1603-1609, Dec. 1989.

[25] Y. Tsin, R.T. Collins, V. Ramesh, and T. Kanade, "Bayesian Color Constancy for Outdoor Object Recognition," *Proc. IEEE Computer Vision and Pattern Recognition,* Dec. 2001.

[26] Y. Weiss and E.H. Adelson, "A Unified Mixture Framework for Motion Segmentation: Incorporating Spatial Coherence and Estimating the Number of Models," *Proc. IEEE Computer Vision and Pattern Recognition,* pp. 321-326, 1996.

[27] G. Wyszecki and W.S. Styles, *Color Science: Concepts and Methods, Quantitative Data and Formulae.* Wiley, 1982.

[28] J. Zhang, J.W. Modestino, and D.A. Langan, "Maximum-Likelihood Parameter Estimation for Unsupervised Stochastic Model-Based Image Segmentation," *IEEE Trans. Image Processing,* vol. 3, no. 4, pp. 404-420, July 1994.

Roberto Manduchi received the doctorate degree in electrical engineering from the University of Padova, Italy, in 1993. He is an associate professor of computer engineering at the University of California, Santa Cruz. Prior to joining UC Santa Cruz in 2001, he worked at Apple Computer, Inc., and at the Jet Propulsion Laboratory.