Designing illuminant spectral power distributions for surface classification

Henryk Blasinski, Joyce Farrell, Brian Wandell
Department of Electrical Engineering, Stanford University
Stanford, CA
hblasins,joyce_farrell,wandell@stanford.edu

Abstract

There are many scientific, medical and industrial imaging applications where users have full control of the scene illumination and color reproduction is not the primary objective. For example, it is possible to co-design sensors and spectral illumination in order to classify and detect changes in biological tissues, organic and inorganic materials, and object surface properties. In this paper, we propose two different approaches to illuminant spectrum selection for surface classification. In the first approach, a supervised framework, we formulate a biconvex optimization problem where we alternate between optimizing support vector classifier weights and optimal illuminants. In the second approach, an unsupervised dimensionality reduction, we describe and apply a new sparse Principal Component Analysis (PCA) algorithm. We efficiently solve the non-convex PCA problem using a convex relaxation and Alternating Direction Method of Multipliers (ADMM). We compare the classification accuracy of a monochrome imaging sensor with optimized illuminants to the classification accuracy of conventional RGB cameras with natural broadband illumination.

1. Introduction

The spectral power distribution of the illumination source plays a fundamental role in how objects objects are imaged and analyzed by a digital camera [34]. In consumer digital photography, the spectral properties of scene illumination are unknown at capture and the main challenge in color balancing (white balancing) is estimating the spectral power of the illuminant [35].

There are many applications where color reproduction is not the primary goal of imaging systems, for example laparoscopic surgery, endoscopy, microscopy or industrial quality control [4, 17]. In these cases the spectrum of light is a design parameter that can be optimized for tasks such as detecting object features or classifying surfaces [7]. Depending on the application it may be more advantageous to adjust the spectral power distribution of the illuminant to enhance the visibility of features of interest, at the expense of accurate color rendering. Specifically, scene data captured under natural, broadband light sources and with conventional RGB cameras can be less useful in classification problems than the data acquired with monochrome cameras and highly customized, narrowband illuminants (Fig. 1). The customized set of illuminants will produce a more robust set of features that can be later leveraged by machine learning algorithms to improve classification performance.

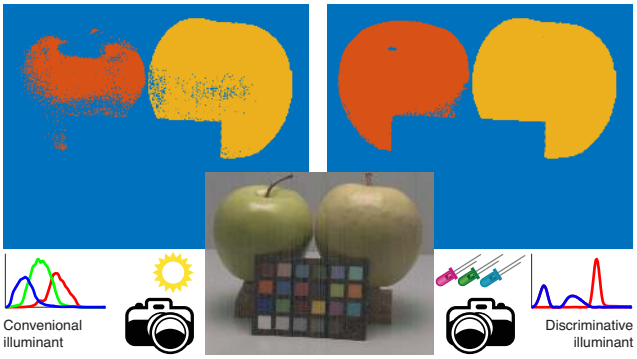


Figure 1: The spectral power distribution of the illuminant has a big impact on the data captured by an imaging system. It can be difficult for a machine learning algorithm to discriminate between a real and fake apple based on the data from an RGB camera and broadband illumination (left). In contrast a camera capturing the scene with customized lights produces better data for discrimination (right) even though the light optimal for feature discrimination or classification may not produce accurate color matching effects.

In this paper we describe and evaluate two novel algorithms to effectively and systematically search for spectral power distributions of optimal lights that improve material classification. The lights are optimized for a particular classification task, for example apple classification. We formulate the problem of illuminant spectral power distribution selection in two different frameworks: supervised and unsupervised learning. Our methods are designed to explicitly

estimate spectral power distributions, rather than to select the best illuminant from a given set. The methods are also constrained to produce solutions that are physically realizable.

In the supervised context we assume that the reflectance data is labeled and we formulate a penalty function incorporating the image formation model, classifier parameters, and the spectral power distribution of light. Such an objective function is bi-convex and locally optimal solutions can be found via alternating minimization.

In the unsupervised context we approach optimal illuminant spectral power distribution design from the dimensionality reduction perspective. Here the objective is to project high dimensional reflectance data onto a smaller number of dimensions that preserve the information necessary for classification. Optimized projection directions better capture the characteristics of the underlying data preserving more variance in the projected data set. Larger variance in the data set generally translates to better performance of the classification algorithms. The classical algorithm for dimensionality reduction is Principal Component Analysis (PCA). This algorithm computes a set of orthogonal projection directions along which the variance in the data set is the largest [12].

The space of physically realizable spectral power distributions of light is limited by certain physical constraints. First, light cannot be negative. Second, the shapes of the spectrum cannot be arbitrary, typically these are smooth functions. All these restrictions need to be taken into consideration when searching for the optimal illuminant.

In summary, our contributions include:

- Two new algorithms optimizing the spectral power distributions of illuminants for classification tasks and producing distributions that can be generated with real light sources.
- A new sparse nonnegative PCA algorithm with a single tuning parameter.
- A framework for analyzing imaging system classification performance using different number and spectra of optimized illuminants.

2. Related work

There exist few methods that solve the problem discussed in this paper. The most relevant algorithm is that of Liu and Guo [21], who used cost function minimization to select the color, position and intensity of LEDs in their capture device that enhance material classification accuracy. Unlike the methods we describe, the authors do not analyze the data in the wavelength domain and do not restrict their solution to non-negative intensities.

Many other approaches involve brute force selection strategies using exhaustive search over the entire parameter space [26]. This is tractable for small problems, but the complexity grows exponentially and thus makes the method computationally impractical. One way to reduce the solution space of brute-force methods is to use genetic algorithms [6]. The methods we present involve continuous valued function minimization rather than search strategies making them more computationally tractable.

Optimal light spectral power distribution selection shares many similarities with camera responsivity design. Camera responsivities, however, are most frequently optimized for color matching tasks [14, 28, 30, 31], low light performance [8, 19] or spectral reconstruction [25, 27]. The idea of adapting camera spectral characteristics to the particular properties of the imaged scene was presented in [20] where the authors manually changed the characteristics of a liquid crystal tunable filter (LCTF) as a function of the distribution of scene radiance. In a similar effort [10] describe an adaptive spectroscopy algorithm which takes repeated measurements of the scene and adjusts the sensitivity of the spectrometer as a function of prior observations. This approach conditions the shape of optimal responsivity curves on the results of earlier measurements and thus cannot be used to derive a fixed set of spectral curves independent of specific measurement outcomes.

Dimensionality reduction techniques designed for hyperspectral data have been developed in the remote sensing community [5, 16, 18]. In these applications the weights are not constrained and are allowed to be negative and therefore wavelength distributions of those weights do not represent physically realizable light spectra.

The existing sparse PCA [1, 15, 24] or nonnegative sparse PCA [32, 36] algorithms cannot be directly used to design physically realizable illuminants. These methods do not allow nonnegativity constraints on linearly transformed variables, which is necessary to assure physical realizability of optimized illuminant spectra. Some of the methods are also impractical due to large numbers of tuning parameters that need to be adjusted. We overcome these limitations by proposing a new, flexible nonnegative sparse PCA algorithm with a single tuning parameter.

3. Image formation model

The response of a camera's photodetector $m_{j,k}$ is a linear function of the scene illuminant x_k, surface reflectance r and the spectral responsivity of the jth camera color channel c_j [34]

$$m_{j,k} = \int c_j(\lambda) r(\lambda) x_k(\lambda) d\lambda. \tag{1}$$

For most natural images, the spectral curves are smooth and slowly varying, therefore a discretization to n spectral bands simplifies modeling with little impact on the accuracy

$$m_{j,k} = \Delta\lambda c_j^T \mathbf{diag}(r) x_k = e_j^T x_k \tag{2}$$

where $e_j = \Delta\lambda(r_1 c_{j,1}, r_2 c_{j,2}, \ldots r_n c_{j,n}), e_j \in \mathbf{R}^n$ is the wavelength-wise product between the surface spectral reflectance and the camera responsivity function of the jth channel. In this work, we propose a method to choose the vector x_k, given a set of labeled or unlabeled vectors e_j so that those vectors projected onto a subspace spanned by x_k, $k \in 1, \ldots, K$ preserve the discriminative information from e in m. A classification algorithm then uses the projected data m, *i.e.* pixel intensities, to derive decision boundaries in this low dimensional subspace.

Spectral curves are often approximated using linear models

$$x = Bw, \tag{3}$$

where the columns of $B \in \mathbf{R}^{n \times n_b}$ represent model components and $w \in \mathbf{R}^{n_b}$ are the corresponding weights. Model components can be derived using two different approaches. The first approach takes advantage of the fact that spectrally smooth curves lie on a low dimensional subspace [23]. The span of this subspace is described by the columns of B, which, by definition, form an orthogonal set, *i.e.* $B^T B = I$. This modeling usually aims to simplify the problem by reducing the number of variables used to describe a light spectrum.

The second approach is to make columns of B directly represent a collection of spectral power distributions of all light sources the user controls. For example B can be created from spectral power distributions of all LEDs on offer by a manufacturer. The task now becomes selection of these spectra that are optimal. In such cases B forms a dictionary and often becomes 'fat', *i.e.* contains more columns than rows. This means that some of the dictionary entries are linear combinations of each other, and $B^T B$ is no longer full rank.

We assume that a physically realizable illuminant is non-negative and that is spanned by the columns of B. Our approach to finding the optimal light does not rely on particular properties of B, specifically B can be 'fat' and $B^T B$ non-invertible.

We also note that if one could use a large number of illuminants ($K \geq n$) the optimal strategy would be to pick monochromatic lights. In this case, acquisition of K images would correspond to capturing the full spectral characteristic of the surface, as if imaged with a hyperspectral camera. This means that in general, preferred solutions should have little overlap across spectral channels, and many zero entries.

4. Supervised framework

Many classical supervised learning algorithms attempt to find a set of hyperplanes that define boundaries between points in the feature space representing different classes. For example, to classify a set of vectors e with labels y, the parameters describing separating hyperplanes θ are found by minimizing a cost function f subject to constraints g on those parameters

$$\begin{aligned} &\text{minimize } f(\theta; e, y) \\ &\text{subject to } g(\theta; e, y) \leq 0. \end{aligned} \tag{4}$$

Often the input data e are not guaranteed to be linearly separable in the original n dimensional space. The data may be projected, using operator $\Phi(e)$, to a higher dimensional space in which such linear separability can be achieved

$$\begin{aligned} &\text{minimize } f(\theta_\Phi; \Phi(e), y) \\ &\text{subject to } g(\theta_\Phi, \Phi(e), y) \leq 0. \end{aligned} \tag{5}$$

We note that in the context of optimal illumination design the projection operator Φ naturally arises in the formulation of the image formation model, though it projects data to a subspace with fewer dimensions. Observe that the measured pixel intensities satisfy $m = X^T e = \Phi(e)$, where the columns of $X = [x_1, \ldots, x_K]$ represent the spectral power distributions of K illuminants, and e are defined as in (2). The problem is to find such linear projection directions X that facilitate class separability in the $\Phi(e)$ vector space, *i.e.* camera pixel measurement space.

To illustrate the approach, consider a classifier where f and g correspond to a multiclass, one-vs-all, Support Vector Classifier (SVC) [13]. The linear decision boundaries p_t for each of the t classes are given by a solution to the convex optimization problem

$$\begin{aligned} &\text{minimize } \sum_t \|p_t\|_2^2 + C \sum_{i, t \neq y_i} \xi_i^t \\ &\text{subject to } p_{y_i}^T \Phi(e_i) + b_{y_i} \geq p_t^T \Phi(e_i) + b_t + 2 - \xi_i^t \\ &\qquad \xi_i^t \geq 0, \quad i = 1, \ldots, L, \quad t \in \{1, \ldots, T\} \setminus y_i, \end{aligned} \tag{6}$$

where $\{y_i, e_i\}$ represent the label and the reflectance-camera responsivity product of the ith data point respectively with the size of the training set given by L. The scalar C controls how many data points are allowed to be misclassified during training.

In order to find the optimal illuminant spectral power distributions we re-formulate the SVC problem by replacing the projection $\Phi(e_i) = (BW)^T e_i$ and adding constraints on the illuminant power distributions BW assuring physical realizability. The resulting biconvex optimization problem

$$\begin{aligned} &\text{minimize } \sum_k \|p_k\|_2^2 + C \sum_{i, k \neq y_i} \xi_i^k + \alpha z(BW) \\ &\text{subject to } p_{y_i}^T (BW)^T e_i + b_{y_i} \geq p_k^T (BW)^T e_i + b_k + 2 - \xi_i^k \\ &\qquad \xi_i^k \geq 0, \quad i = 1, \ldots, L, \quad k \in \{1, \ldots, K\} \setminus y_i, \\ &\qquad BW \geq 0, \end{aligned} \tag{7}$$

can be solved through alternating minimization over p, b, ξ and b, ξ, W. The matrix W describes the illuminant spectra in terms of the linear basis weights (3). Note that the cost function has been augmented with a function z, controlled with a tuning parameter α, penalizing certain types of solutions. We promote sparsity with an l_1 penalty

$$z(X) = \|X\|_1 = \sum_{i,j} |x_{i,j}| \tag{8}$$

where the l_1 norm is generalized to matrices and computes the sum of the absolute values of all matrix entries.

5. Unsupervised framework

Principal Component Analysis (PCA) is a standard technique used in machine learning to compactly represent a data set. The PCA algorithm finds a set of projection vectors along which the original data has the largest variance. These projections are given by the principal eigenvectors of the data set covariance matrix. In the context of illuminant spectrum selection, we note that the measured pixel intensity represents the projection of the reflectance-camera responsivity product onto a vector describing the spectral distribution of the illuminant.

For convenience, we express PCA as an iterative algorithm. Let $\Sigma \in \mathbf{S}_+^n$ represent the positive semi-definite sample covariance matrix of vectors e from (2). The ith optimal PCA direction is a solution to the problem

$$\begin{array}{ll} \text{maximize} & x^T \Sigma^{(i-1)} x \\ \text{subject to} & x^T x = 1, \end{array} \tag{9}$$

where $\Sigma^{(i-1)}$ is the covariance matrix estimate from the previous step. Before proceeding to the next iteration Σ needs to be recomputed to account for the removal of the solution dimension (*i.e.* deflated).

The iterative PCA algorithm can be modified to produce nonnegative and sparse directions by introducing additional constraints into the optimization problem (9)

$$\begin{array}{ll} \text{maximize} & x^T \Sigma^{(i-1)} x \\ \text{subject to} & x^T x = 1 \\ & \mathbf{card}(x) \leq \delta \\ & x \geq 0, \end{array} \tag{10}$$

where δ is a sparsity parameter and $\mathbf{card}(x)$ specifies the number of nonzero entries in x. Equation (10) is nonconvex due to the incorporation of sparsity and nonnegativity constraints. Hence, eigen decomposition is no longer a valid method for finding a solution.

5.1. Convex relaxation

A useful approach to solving nonconvex problems is to perform an approximation with a similar convex problem *i.e.* relaxation. For example, using the PCA relaxations described in [9, 33], the original nonconvex problem can be approximated with

$$\begin{array}{ll} \text{maximize} & \mathbf{tr}(\Sigma X) - \alpha \|X\|_1 \\ \text{subject to} & 0 \preceq X \preceq I, \\ & \mathbf{tr}(X) = 1 \\ & X \geq 0 \end{array} \tag{11}$$

(we drop the superscript $(i-1)$ for clarity). The function $\mathbf{tr}(X)$ is the trace operator, *i.e.* the sum of the entries of X along the diagonal, and α controls the sparsity enforcing penalty.

The first constraint forces X to be positive, semi-definite and $X - I$ to be negative, semi-definite. The last constraint restricts the solution to a set of matrices with nonnegative entries. The optimal projection direction is given by the principal eigenvector of a solution $X^\star = x^\star x^{\star T}$. Note that the matrix entry-wise inequality enforces all the entries of $x^\star$ to have the same sign, therefore if $x^\star$ is a solution so is $-x^\star$.

In order to express the PCA problem in terms of the basis weights, $X = xx^T = Bw(Bw)^T = BWB^T$ is substituted into (11) yielding

$$\begin{array}{ll} \text{maximize} & \mathbf{tr}(\Sigma BWB^T) - \alpha \|BWB^T\|_1 \\ \text{subject to} & 0 \preceq BWB^T \preceq I, \\ & \mathbf{tr}(BWB^T) = 1 \\ & BWB^T \geq 0, \end{array} \tag{12}$$

where $W = ww^T$.

The above optimization problem can be efficiently solved using the Alternating Direction Method of Multipliers (ADMM) [3, 33]. The computational steps of this approach are summarized in Algorithm 1, derivation details are available in the Supplemental Material. This algorithm requires the user to provide the desired accuracy ϵ and a learning rate ρ. Following the recommendation of [33] we used an update heuristic proposed in [3]. It may happen that the function computing the largest eigenvector of $Y_1^{(t+1)}$ will return a vector with all nonpositive entries. In this case the sign of such a solution should be reversed.

5.2. Matrix deflation

The iterative approach to PCA makes it necessary to deflate the current sample covariance matrix estimate $\Sigma^{(i-1)}$ before proceeding to the next iteration. The commonly used Hotelling's deflation scheme may not preserve the semidefiniteness of $\Sigma^{(i)}$ if the projection direction x_i is not an eigenvector of $\Sigma^{(i-1)}$ [29]. To avoid such issues we used a generalized deflation algorithm proposed in [22]. Algorithm 2 summarizes this deflation method and outlines the steps required to compute r sparse nonnegative PCA directions for a particular data set.

Algorithm 1 Single nonnegative sparse PCA direction.

function FINDDIRECTION(Σ,B, α, ϵ, ρ)
$Y_1^{(0)} \leftarrow 0, U_1^{(0)} \leftarrow 0, Y_2^{(0)} \leftarrow 0, U_2^{(0)} \leftarrow 0$
repeat
$W^{(t+1)} \leftarrow \arg\min \left(\|BWB^T - Y_1^{(t)} + U_1^{(t)}\|_F^2 + \|BWB^T - Y_2^{(t)} + U_2^{(t)}\|_F^2 \right)$
$Y_1^{(t+1)} \leftarrow \mathcal{P}_{\mathcal{F}}(BW^{(t+1)}B^T + U_1^{(t)} + \Sigma/\rho), \quad Y_2^{(t+1)} \leftarrow \mathcal{H}_{\alpha/\rho}(BW^{(t+1)}B^T + U_2^{(t)})$
$U_1^{(t+1)} \leftarrow U_1^{(t)} + W^{(t+1)} - Y_1^{(t+1)}, \quad U_2^{(t+1)} \leftarrow U_2^{(t)} + BW^{(t+1)}B^T - Y_2^{(t+1)}$
$t \leftarrow t + 1$
until $\max\{\|BW^{(t+1)}B^T - Y_1^{(t+1)}\|_F^2, \|BW^{(t+1)}B^T - Y_2^{(t+1)}\|_F^2, \rho^2\|Y_1^{(t+1)} - Y_1^{(t)}\|_F^2\}, \rho^2\|Y_2^{(t+1)} - Y_2^{(t)}\|_F^2\} \leq \epsilon$
return Eigenvector of $Y_1^{(t+1)}$ corresponding to the largest eigenvalue.
end function

Algorithm 2 Nonnegative sparse PCA

Require: $\Sigma \in \mathbf{S}_+^n, \alpha \geq 0, \epsilon \geq 0, \rho \geq 0, r \in \mathbf{N}, B$
$Q^{(0)} \leftarrow I, \Sigma^{(0)} \leftarrow \Sigma$
for $t = 1, \ldots, r$ **do**
$x_t \leftarrow$ FINDDIRECTION($\Sigma^{(t-1)}$,B, α, ϵ, ρ)
$q_t \leftarrow Q^{(t-1)}x_t$
$\Sigma^{(t)} \leftarrow (I - q_t q_t^T)\Sigma^{(t-1)}(I - q_t q_t^T)$
$Q^{(t)} \leftarrow Q^{(t-1)}(I - q_t q_t^T)$
$x_t \leftarrow x_t / \|x_t\|$
end for
return $x_1, \ldots, x_r$

6. Experiments

To evaluate the proposed approaches we conducted a set of experiments where we captured images of known test targets under different illumination conditions and used standard machine learning algorithms to classify captured data. We implemented all computations in Matlab, and our code repository is available online[1].

First, we measured the performance of classification algorithms on data captured with conventional RGB cameras and under a single broadband illuminant. This condition represents the typical imaging scenario where a color camera with three types of color pixels is used to capture a single image in natural lighting conditions (*i.e.* $J = 3$, $K = 1$). Overall, we evaluate the performance of 34 cameras with different spectral properties (see Supplemental Material for the list of evaluated camera models). For simplicity we will refer to this mode of data capture as *conventional camera*.

Next, we captured the data using a monochrome camera and optimized illuminants. To make comparisons between conventional and optimized systems meaningful we projected spectral measurements onto a three-dimensional space. To obtain this representation with our optimized setup we captured three successive frames with a monochrome camera, each under different, optimized light (*i.e.* $J = 1$, $K = 3$) derived with either supervised or unsupervised approach. We refer to this capture mode as *optimal camera*.

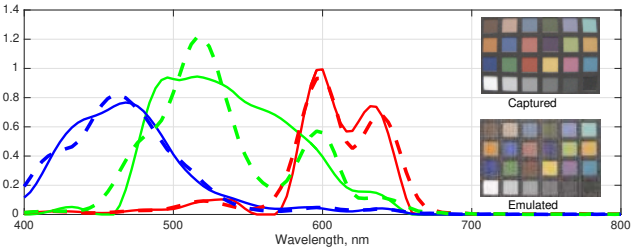


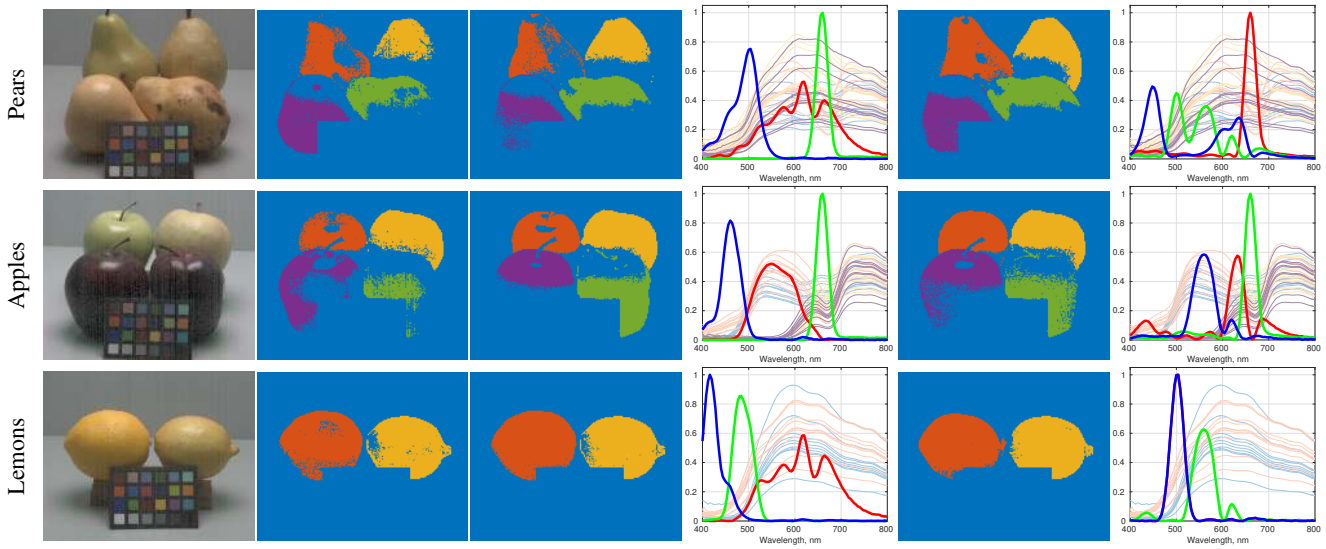
Figure 2: A comparison between spectral properties of a GoPro Hero 5 RGB camera (solid lines) and their approximation with narrowband lights (dashed lines). For a given scene, pixel values produced by an RGB camera (inset, captured) can be reproduced with a mochorome camera and appropriately adjusted narrowband lights that match the spectral responsivities of the camera (inset, emulated).

We investigated a simple classification task; assigning labels to image pixels based on raw image data. We originally designed these algorithms for use with biological multispectral data, however for presentation clarity we chose much simpler targets requiring no domain specific knowledge. In our tests we used genuine and visually similar artificial fruit pairs: apples, pears and lemons. Our goal was to discriminate between different objects. For this purpose we constructed three test scenes. Two of them contained two pairs of differently colored fruit of the same type; red and green apples (*Apples*) and yellow and green pears (*Pears*). The *Lemons* scene contained a single pair of objects; lemons.

Scenes were assembled in a Thouslite LEDCube[2] light booth. The LEDCube contained three broadband and eight narrowband LEDs whose intensities could be independently

[1] https://github.com/hblasins/optIll

[2] http://www.thouslite.com/show.asp?id=16



(a) RGB Image | (b) RGB classification | (c) Unsupervised classification | (d) Unsupervised spectra | (e) Supervised classification | (f) Supervised spectra

Figure 3: Pixel classification accuracy is increased when the scene is illuminated with optimal lights. Columns present the RGB rendering of the scene (a), pixel classification maps for conventional illumination (b), and optimal lights derived with unsupervised (c) and supervised (e) approaches. Color in (b, c, e) encodes pixels where the SVM classifier assigned correct labels, all errors are not represented. The spectral power distributions of the optimal illuminants (d, f, thick lines) overlap with areas of increased variability in the surface spectral reflectance (d, f, thin lines).

adjusted thus changing the overall spectrum of the illuminant. The spectral power distributions of the individual LEDs formed columns in the model matrix B (3).

We measured the surface spectral reflectance of all targets as well as spectral power distributions of the LEDs in the 400 to 800nm range at 4nm increments using a SpectraScan PR715 spectrophotometer. We calibrated the LEDCube LED spectra by taking measurements of the light radiance reflected from a reference white, Spectralon test target. Similarly, we estimated the surface spectral reflectance by illuminating the surface of fruits with broadband tungsten light, taking 10 radiance measurements followed by one measurement of a Spectralon sample, and by computing the ratio between the two at every wavelength [2].

Given a set of reflectance curves we used the supervised and unsupervised approaches to generate optimal illuminants for a particular setting of algorithm tuning parameters. We then programmed those spectra into the LEDCube and captured images of the so illuminated target with a PointGrey FL3-U3-13Y3M-C, 1.3MP monochrome camera with a Schneider Optics Xenoplan 1.4/23mm lens. The aperture was set to f#/11 in order to limit depth of field effects.

Next, we manually segmented 100×100 (200×200 for binary classification) pixel regions of interest (ROI) within each fruit sample and used this set for classification. The classifiers used single pixel intensity data. The data was split into training (70%) and test (30%) sets using the stratified approach, *i.e.* preserving the distribution of class labels [11]. For fairness of comparisons we preserve data separation across all classifiers and conditions. In our evaluations we used five standard machine learning classifiers: Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Linear Discriminant Analysis (DA), Decision Trees (Tree) and Naive Bayes (NB).

Finally, we repeated the procedure; optimal illuminant computation followed by image capture and classifier cross validation, for different settings of illuminant selection algorithms tuning parameters (refer to source code for details). In all cases training data set was used for training and parameter selection only, and all performance numbers we report were calculated using the test set.

Rather than using physical RGB cameras in our evaluations, we instead emulated their spectral properties with narrowband LEDs from LEDCube and the PointGrey monochrome camera. We approximated the effective spectral responsivity (*i.e.* the wavelength-wise product between spectral responsivity and the illuminant) of red, green and blue camera channels with an appropriate weighted sum of the LEDCube LEDs, accounting for the monochrome sensor quantum efficiency (see Supplemental Material for details). Figure 2 shows an example comparison between the actual responsivity curves and their approximations using

Table 1: Single pixel per cent classification accuracy for the *Pears* data (four-way classification).

Camera	Classifier				
	SVM	KNN	DA	Tree	NB
RGB (worst)	92.3	92.7	91.2	90.7	66.0
RGB (avg.)	94.6	95.2	93.6	94.1	67.8
RGB (best)	98.4	98.4	97.9	97.8	74.2
Ours (Unsup.)	97.2	97.9	96.9	97.8	**82.8**
Ours (Sup.)	**99.9**	**99.9**	**98.9**	**99.8**	76.6

Table 2: Single pixel per cent classification accuracy for the *Apples* data (four-way classification).

Camera	Classifier				
	SVM	KNN	DA	Tree	NB
RGB (worst)	75.2	78.4	74.9	73.6	66.0
RGB (avg.)	84.6	82.3	84.1	84.2	68.7
RGB (best)	85.6	83.3	85.3	85.8	72.7
Ours (Unsup.)	87.1	89.4	86.4	89.2	**83.7**
Ours (Sup.)	**91.5**	**92.6**	**91.2**	**91.2**	79.5

Table 3: Single pixel per cent classification accuracy for the *Lemons* data (binary classification).

Camera	Classifier				
	SVM	KNN	DA	Tree	NB
RGB (worst)	92.2	96.4	90.0	85.9	58.8
RGB (avg.)	97.5	99.0	96.3	96.6	62.0
RGB (best)	99.9	**100.0**	99.8	**99.9**	**75.6**
Ours (Unsup.)	99.8	99.8	99.6	98.3	63.1
Ours (Sup.)	**100.0**	**100.0**	**99.9**	98.8	70.9

LEDs. Insets in this figure present an image of a Macbeth chart captured with a conventional RGB camera and compares it to a view of the same chart rendered with the data from the proposed emulation approach.

There are several advantages to using the emulation approach. First, the geometry of the setup remains fixed and any differences between various test conditions arise only from illumination rather than image alignment or resampling between different cameras. Second, the same sensor is used throughout all experiments and so the noise properties remain fixed. Finally, the resolution and optics remain constant throughout all experiments.

6.1. Classification performance

Figure 3 presents the results of pixel classification outcomes for three test scenes captured under daylight and optimized illuminants. Classification maps were created by applying classifiers trained on the training data to all pixels in the image. For clarity we show and label only correctly identified pixels. In general, maps obtained for the optimized illumination cases correctly classify larger number of pixels. We also observed that those maps have more contiguous object segmentation and less 'salt and pepper' classification noise. However, the optimal illuminant selection strategies do not always increase the number of correctly classified pixels. They can also make some pixels easier to classify at the expense of others. Take the *Apple* scene as an example. Under optimized illuminants the increased accuracy in recognizing the bottom right apple (green label) is offset by erroneously classifying parts of the bottom left apple (purple label) that had been correctly classified by a conventional camera.

We also present the optimal illuminant spectra derived for each condition (Fig. 3d and 3f). The curves derived by supervised and unsupervised approaches differ, but they share a common characteristic. In all cases light energy is concentrated at wavelengths with significant amount of variability in the surface reflectance data.

The quantitative performance of the proposed approaches is presented in Tables 1, 2 and 3 which summarize the pixel based classification accuracy achieved with the optimal and conventional cameras. For every classification algorithm, we report the average accuracy computed over 34 RGB cameras as well as the accuracy of the best and worst performing model for a particular algorithm. The model of the camera achieving the highest classification accuracy varied across different scenes and classifiers. This metric represents the performance of illuminant spectral power distributions found using a brute force search algorithm and serves as a baseline for comparisons with our methods.

We experimented with illuminating our scenes with the emission spectra of black body radiators at different temperatures; 2000, 4000, 6500, and 10000K. These spectra offer good approximations to light distributions occurring in natural environments. For a particular classification algorithm, performance differences between these illuminants were minor (see Supplemental Material for details). The black body emission at 6500K closely resembles the standard D65 illuminant (daylight), and in general provided the best, or close second best classification accuracy. Furthermore, many conventional RGB cameras are designed to faithfully reproduce colors specifically under this illuminant. For these reasons we use this condition as reference in numerical comparisons.

The imaging system using illuminant spectral power distributions derived with our supervised method consistently outperforms conventional RGB cameras and broadband illuminants for the SVM, KNN, DA and Tree classification algorithms. It comes second only when the Tree classifier is applied to the *Lemons* scene. These gains are due to direct

incorporation of the search for optimal spectra into the classification objective function, and are achieved despite non-convexity of the problem and local optimality of the solution. Interestingly, the supervised selection method is typically outperformed by a combination of the Naive Bayes classifier with the unsupervised approach.

Our unsupervised selection method is better than the average conventional RGB camera, and its accuracy is typically on par with that of the best RGB camera. Note that this selection strategy works best together with a Naive Bayes classifier. The unsupervised approach is a variant of PCA, which provides a set of uncorrelated features. In our case we use a constrained version; it also decorrelates data, but to a lesser extent. The Naive Bayes classifier implicitly assumes feature conditional independence, which means that the unsupervised approach provides the type of data the classifier expects.

6.2. Number of illuminants

We investigated the relationship between classification accuracy and the number of different illuminants under which a scene is captured. Optimization over the illuminant spectrum offers another system design parameter. For example, a suitable illuminant could eliminate the need for a color filter array thus increasing the effective sensor resolution.

Increasing the number of optimal illuminants from 1 to 10 increases the classification performance, irrespective of the selection method used. Figure 4 presents the classification accuracy in the *Apples* set as a function of the number of optimal illuminants. The supervised selection algorithm maintains a small advantage over the unsupervised approach across all conditions. The classification accuracy asymptotes at about 95%, ten percentage points above the accuracy of the best conventional RGB camera. However, at about 3–4 illuminants the curves plateau; using more illuminants offers modest gains in performance that may not be worth pursuing given other limitations such as acquisition time. Note that the classification accuracy of an optimal two-illuminant system, *i.e.* representing pixel data with two numbers, is comparable to that of the best conventional RGB camera and a broadband illuminant, which represent pixel data with three numbers. Both systems achieve about 86% classification accuracy. This reduction in data dimensionality can translate to smaller data storage requirements or fewer processor cycles necessary for computation.

7. Conclusions

We presented two new approaches to selecting spectral power distributions of illuminants used during the image capture process that increase the accuracy of pixel classification algorithms applied to captured data. Our methods produce physically realizable distributions that can be generated using off-the-shelf LEDs.

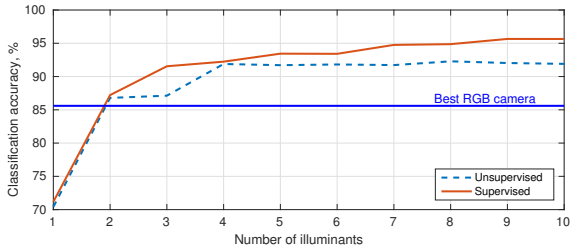


Figure 4: Pixel classification accuracy increases when a scene (*Apples*) is captured under more optimized illuminants, with supervised selection outperforming the unsupervised algorithm. Using just two optimally chosen lights allows to achieve the performance level of a conventional, three channel RGB system.

We evaluated the optimal lights produced by our selection methods in a simple surface classification task through a series of laboratory experiments using real targets, illuminants, and cameras. In all cases classifier performance on pixel data obtained under optimized lights is greater than the performance of the same algorithm on data captured with a conventional RGB camera and broadband illuminants. We showed that the supervised selection method achieves the best performance, but it requires a set of labeled training reflectance spectra. In comparison the unsupervised method is a close second in terms of accuracy, but can work with unlabeled object spectral reflectance curves.

Both algorithms are useful in analyzing the impact of additional imaging channels on classification performance. This is an important consideration because, we have shown, under the constraints of physical realizability having more channels may provide only modest performance gains that may be deemed not worth pursuing given the increased system complexity and other design trade-offs such as speed, resolution or computation.

Finally, the methodology we present in this work can be extended to other spectral selection problems in imaging system design. For example, the same algorithms can be applied to selecting the responsivity functions of the color filter array or could be used to select transmissive filters for biological image analysis.

Acknowledgements

The authors would like to thank Gordon Wetzstein from Stanford University for helpful comments and discussions during the preparation of the manuscript. The LEDCube image booth was a gift from Thouslite.

References

[1] A. Amini and M. Wainwright. High-dimensional analysis of semidefinite relaxations for sparse principal components. In *IEEE International Symposium on Information Theory ISIT*, pages 2454–2458, 2008. 2

[2] H. Blasinski, J. Caves, J. Farrell, B. Wandelt, and P. Wang. Multispectral imaging of tissue ablation. In *IEEE International Symposium on Biomedical Imaging, ISBI*, pages 360–363, April 2016. 6

[3] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011. 4

[4] J. Cadeddu, R. Fernandez, M. Desai, R. Bergs, C. Tracy, S.-J. Tang, P. Rao, M. Desai, and D. Scott. Novel magnetically guided intra-abdominal camera to facilitate laparoendoscopic single-site surgery: initial human experience. *Surgical endoscopy*, 23(8):1894–1899, 2009. 1

[5] C. Chang, Q. Du, T. Sun, and M. Althouse. A joint band prioritization and band decorrelation approach to band selection for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(6):2631–2641, Nov 1999. 2

[6] C. Chi, H. Yoo, and M. Ben-Ezra. Multispectral imaging by optimized wide band illumination. *International Journal of Computer Vision*, 86(2–3):140–151, 2010. 2

[7] J. Choi, S. Park, J. Cho, and E. Yoon. A 3.4-μ w object-adaptive CMOS image sensor with embedded feature extraction algorithm for motion-triggered object-of-interest imaging. *IEEE Journal of Solid-State Circuits*, 49(1):289–300, Jan 2014. 1

[8] L. Condat. A new color filter array with optimal properties for noiseless and noisy color image acquisition. *IEEE Transactions on Image Processing*, 20(8):2200–2210, Aug 2011. 2

[9] A. d'Aspremont, L. El Ghaoui, M. Jordan, and G. Lanckriet. A direct formulation for sparse PCA using semidefinite programming. In *Neural Information Processing Systems*, volume 16, pages 41–48, 2004. 4

[10] D. Dinakarababu, D. Golish, and M. Gehm. Adaptive feature specific spectroscopy for rapid chemical identification. *Optics express*, 19(5):4595–4610, 2011. 2

[11] G. Forman and M. Scholz. Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *ACM SIGKDD Explorations Newsletter*, 12(1):49–57, 2010. 6

[12] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer, 2009. 2

[13] C.-W. Hsu and C.-J. Lin. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2):415–425, 2002. 3

[14] F. Imai, S. Quan, M. Rosen, and R. Berns. Digital camera filter design for colorimetric and spectral accuracy. In *Proceedings of the 3rd International Conference on Multispectral Color Science*, pages 23–26, 2001. 2

[15] M. Journee, Y. Nesterov, P. Richtarik, and R. Sepulchre. Generalized power method for sparse principal component analysis. *The Journal of Machine Learning Research*, 11:517–553, 2010. 2

[16] N. Keshava. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1552–1565, July 2004. 2

[17] S. Kukkonen, H. Kalviainen, and J. Parkkinen. Color features for quality control in ceramic tile industry. *Optical Engineering*, 40(2):170–177, 2001. 1

[18] S. Kumar, J. Ghosh, and M. Crawford. Best-bases feature extraction algorithms for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7):1368–1379, Jul 2001. 2

[19] H. Kuniba and R. S. Berns. Spectral sensitivity optimization of color image sensors considering photon shot noise. *Journal of Electronic Imaging*, 18(2):023002–023002–14, 2009. 2

[20] A. Lin and F. Imai. Efficient spectral imaging based on imaging systems with scene adaptation using tunable color pixels. In *Color and Imaging Conference*, volume 2011, pages 332–338. Society for Imaging Science and Technology, 2011. 2

[21] C. Liu and J. Gu. Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral brdf. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1):86–98, 2014. 2

[22] L. Mackey. Deflation methods for sparse PCA. In *Neural Information Processing Systems*, volume 21, pages 1017–1024, 2008. 4

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

[24] B. Moghaddam, Y. Weiss, and S. Avidan. Generalized spectral bounds for sparse LDA. In *Proceedings of the 23rd international conference on Machine learning*, pages 641–648. ACM, 2006. 2

[25] Y. Monno, T. Kitao, M. Tanaka, and M. Okutomi. Optimal spectral sensitivity functions for a single-camera one-shot multispectral imaging system. In *IEEE International Conference on Image Processing, ICIP*, pages 2137–2140, 2012. 2

[26] J. Park, M. Lee, M. D. G., and S. Nayar. Multispectral imaging using multiplexed illumination. In *IEEE International Conference on Computer Vision, ICCV*, pages 1–8, Oct 2007. 2

[27] M. Parmar, S. Lansel, and J. Farrell. An LED-based lighting system for acquiring multispectral scenes. In *Proc. SPIE*, volume 8299, pages 82990P–82990P–8, 2012. 2

[28] M. Parmar and S. Reeves. Selection of optimal spectral sensitivity functions for color filter arrays. *IEEE Transactions on Image Processing*, 19(12):3190–3203, Dec 2010. 2

[29] Y. Saad. Projection and deflation method for partial pole assignment in linear state feedback. *IEEE Transactions on Automatic Control*, 33(3):290–297, 1988. 4

[30] Z. Sadeghipoor, Y. Lu, and S. Susstrunk. Optimum spectral sensitivity functions for single sensor color imaging. In *Proc. SPIE*, volume 8299, pages 829904–829904–14, 2012. 2

[31] H.-L. Shen, H.-G. Zhang, J. Xin, and S.-J. Shao. Optimal selection of representative colors for spectral reflectance reconstruction in a multispectral imaging system. *Applied Optics*, 47(13):2494–2502, May 2008. 2

[32] C. Sigg and J. Buhmann. Expectation-maximization for sparse and non-negative PCA. In *Proceedings of the 25th International Conference on Machine Learning*, pages 960–967, 2008. 2

[33] V. Vu, J. Cho, J. Lei, and K. Rohe. Fantope projection and selection: A near-optimal convex relaxation of sparse PCA. In *Advances in Neural Information Processing Systems*, pages 2670–2678, 2013. 4

[34] B. Wandell. *Foundations of Vision.* Sinauer Associates, 1995. 1, 2

[35] F. Xiao, J. E. Farrell, J. M. DiCarlo, and B. A. Wandell. Preferred color spaces for white balancing. In *Electronic Imaging 2003*, pages 342–350. International Society for Optics and Photonics, 2003. 1

[36] R. Zass and A. Shashua. Nonnegative sparse PCA. *Advances in Neural Information Processing Systems*, 19:1561, 2007. 2