Discriminative Illumination: Per-Pixel Classification of Raw Materials Based on Optimal Projections of Spectral BRDF

Chao Liu, Student Member, IEEE, and Jinwei Gu, Member, IEEE

Abstract—Classifying raw, unpainted materials—metal, plastic, ceramic, fabric, and so on—is an important yet challenging task for computer vision. Previous works measure subsets of surface spectral reflectance as features for classification. However, acquiring the full spectral reflectance is time consuming and error-prone. In this paper, we propose to use coded illumination to directly measure discriminative features for material classification. Optimal illumination patterns—which we call “discriminative illumination”—are learned from training samples, after projecting to which the spectral reflectance of different materials are maximally separated. This projection is automatically realized by the integration of incident light for surface reflection. While a single discriminative illumination is capable of linear, two-class classification, we show that multiple discriminative illuminations can be used for nonlinear and multiclass classification. We also show theoretically that the proposed method has higher signal-to-noise ratio than previous methods due to light multiplexing. Finally, we construct an LED-based multispectral dome and use the discriminative illumination method for classifying a variety of raw materials, including metal (aluminum, alloy, steel, stainless steel, brass, and copper), plastic, ceramic, fabric, and wood. Experimental results demonstrate its effectiveness.

Index Terms—Computational illumination, appearance modeling, material classification

1 INTRODUCTION

Classifying materials—metal, plastic, ceramic, fabric, paint, and so on—has significant implications for both scientific research and industrial applications across many disciplines, such as remote sensing [25], food inspection [37], mineralogy [19], and recycling [12]. In computer vision, we primarily focus on uncoated or unpainted raw materials because we are limited to appearance related features, such as color [34], bidirectional reflectance distribution function (BRDF) [42], texture [20], [5], [39], [40], translucency, roughness [32], [38], [22], and polarization [14], [15]. Fig. 1 shows some examples of such materials.

Even for uncoated raw materials, appearance-based classification is still challenging because appearance changes with object shape, illumination, and viewing condition. Fully describing the appearance of a scene requires a 14D function, i.e., two plenoptic functions [1]. While some previous work measures subsets of this function for material classification [42], [17], due to the intrinsic high dimensionality material classification has had limited progress compared to that of object recognition.

In this paper, we focus on per-pixel classification of raw materials based on spectral BRDFs. Instead of first sparsely sampling subsets of this high-dimensional function and then performing classification, we propose to use coded illumination to directly measure discriminative features, i.e., projections of spectral BRDFs, for classification. The optimal coded illumination—which we call discriminative illumination—is learned from training samples, after projecting to which, the spectral BRDFs of different materials can be maximally separated. The projection operation is automatically realized by the integration of incident light for surface reflection in an imaging system.

While a single discriminative illumination is capable of linear, two-class classification, we show multiple discriminative illumination patterns can be used for multiclass and nonlinear classification. The proposed discriminative illumination method is more economical than conventional methods of using raw material measurements for classification in terms of the number of captured images—this enables the classification of materials that changes with time. In addition, we derive that the discriminative illumination method results in higher signal-to-noise ratio (SNR) than conventional methods thanks to light multiplexing.

In our prior work [11], we constructed an LED-based multispectral dome light (as shown in Figs. 11a and 11b) and used it as a prototype to implement the proposed discriminative illumination for classifying a variety of raw materials with flat surfaces, including metal (aluminum, alloy steel, stainless steel, cold roll and hot roll steel, brass, and copper), ceramic, plastic, fabric, and wood. Experimental results demonstrated the effectiveness of the proposed method. In this paper, we extend this work 1) to deal with raw materials with nonflat surfaces and 2) to use RGB color images for classification. Also, we analyze the SNR benefits by using the discriminative illumination. As a result, the classification rates under the discriminative light patterns are higher than...
the rates by first measuring the BRDF vector and then performing the classification.

2 RELATED WORK

Per-pixel material classification in machine vision. There are several works aiming at per-pixel material classification using various low-level appearance features, such as polarization for metal and plastics [43], [4], spectral reflectance for printed circuit board inspection [16] and near infrared reflectance [34] for wood and textiles, and 2D slice of BRDF [42] for paint classification. Unlike these methods, we design imaging systems that use learned coded illumination to directly measure discriminative features in captured images. Recently, Jehle et al. [17] proposed to select subsets of basis lights (i.e., rings and sectors) specifically for steel plate classification. In contrast, our method learns optimal weighted combinations of basis illumination for general raw material classification, which offers much more flexibility and SNR benefits due to light multiplexing. We also extend it for multiclass and nonlinear classifications, and use both the spectral and BRDF features.

Computational illumination. Our work falls in the area of computational illumination that uses coded light for efficient material and shape measurement [10], [21]. However, instead of seeking for coded light for reconstructing signals with high SNR, our goal is to find coded light with maximum discriminative ability. This is similar to the relation between EigenFaces and FisherFaces [3].

Task-specific and feature-specific imaging. Our work is also related to task-specific and feature-specific imaging [27], [28], in which the goal of such imaging systems is not to capture visually appealing images but to maximize the amount of information relevant to given tasks (in our case, material classification). An essential component in our work is the supervised learning from labeled datasets.

Statistical analysis of BRDF. In their seminal work, Matusik et al. [23] performed statistical analysis of BRDF data of many real-world materials. It showed that real-world BRDFs can be represented with 45 principal components, among which some are related to perceptual traits of materials, such as specularity and “metal-likeness.” In theory, one can project BRDFs to these principal components for material classification along these traits. Nevertheless, it is fundamentally different with our work for material classification. In [23], an unsupervised learning (e.g., PCA or LLE) is performed to analyze BRDF data, aiming to find efficient representation for dimension reduction. In our work, we perform supervised learning, aiming to find projection directions that can maximize the difference among multiple classes. Again, this is exactly analogous to the difference between the Eigenface and Fisherface approaches for face recognition [3].

3 DISCRIMINATIVE ILLUMINATION

For a point on an opaque, unpainted surface, its material property can often be described with a spectral BRDF [30], \( f(\omega_i, \omega_o, \lambda) \), which is a 5D function describing the ratio between the incident light in the direction \( \omega_i \) and the reflected light in the direction \( \omega_o \) at the wavelength \( \lambda \). Although in principle \( f(\omega_i, \omega_o, \lambda) \) itself can be used as a feature for material classification, measuring this 5D function is time consuming and error-prone—for shiny materials, measurements close to specular reflection angles are not reliable due to limitation of the dynamic range of measurement devices [31], [29]—and thus directly using it for classification is impractical. As mentioned earlier, subsets of spectral BRDF have been used for material classification [42], [17].

3.1 Two-Class Classification

Our approach is to design imaging systems that directly measure discriminative features from spectral BRDFs for classification. Consider a canonical problem of a two-class material classification with a linear classifier:

\[
\mathbf{w}^T \mathbf{x} + b = \begin{cases} 
\geq 0 & y \in \text{Class 1}, \\
< 0 & y \in \text{Class 2}, 
\end{cases}
\]

where \( \mathbf{x} = [f(\omega_i, \omega_o, \lambda)] \) is a vector of the spectral BRDF of a point, and the projection vector \( \mathbf{w} \) and the threshold \( b \) consist of the linear classifier. The key operation here is the projection of the spectral BRDF \( \mathbf{x} \) to the direction \( \mathbf{w} \).

Instead of measuring the full spectral BRDF and then performing the projection, we can use coded illumination to directly measure the projection from reflected light. Consider illuminating the sample with multiple light sources from different angles with different spectra as shown in Fig. 11a, the measured reflected light, \( I(\omega_o) \), is...
where $L(\omega, \lambda)$ is the incident light in the direction $\omega$ at the wavelength $\lambda$, $f(\omega_i, \lambda)$ is the spectral BRDF of the sample, $S(\lambda)$ is the spectral sensitivity of the camera, and $\max(0, \cos \theta_i)$ is the visibility term.

Equation (2) shows that for a given viewing direction (i.e., fixed $\omega_i$), a given camera (i.e., fixed $S(\lambda)$), and a flat sample, the measured reflected light $I(\omega_i)$ is a dot product between the spectral BRDF, $f(\omega_i, \lambda)$, and the incident light, $L(\omega, \lambda)$. More explicitly, if we define $\hat{f}(\omega_i, \omega_j, \lambda) = f(\omega_i, \omega_j, \lambda) S(\lambda) \max(0, \cos \theta_i)$ as the spectral BRDF feature vector for the given viewing direction and the given camera, by discretizing (2) we have

$$I(\omega_i) = \mathbf{w}^T \mathbf{x},$$

where the vector $\mathbf{w} = [L(\omega, \lambda)]$ represents the coded illumination, and the vector $\mathbf{x}$ is the spectral BRDF feature vector. If the lights are narrow band, $\mathbf{x}$ is the spectral BRDF; otherwise, $\mathbf{x}$ is the projection of the spectral BRDF onto the spectra of the lights. Below, we show the derivation of (3).

Suppose the dome has $N$ clusters of light sources. Each cluster has $K$ LEDs with different spectra (in our experimental setup, $N = 25$, $K = 6$). We also assume $M$ consecutive spectral narrow bands are considered in computation (e.g., from 400 to 700 nm with 10-nm interval, $M = 31$). By discretizing (2), we have

$$I(\omega_i) = \sum_{n=1}^{N} \sum_{m=1}^{M} \hat{f}(\omega_i^{(n)}, \omega_j^{(m)}, \lambda^{(m)}) \cdot \Delta \omega_i \Delta \lambda \cdot L(\omega_i^{(n)}, \lambda^{(m)}).$$

We can write this equation in a matrix form:

$$I = \mathbf{w}^T \mathbf{x},$$

where

$$\mathbf{w} = \begin{pmatrix}
L(\omega_i^{(1)}, \lambda^{(1)}) \\
\vdots \\
L(\omega_i^{(N)}, \lambda^{(M)})
\end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix}
\hat{f}(\omega_i^{(1)}, \omega_j^{(1)}, \lambda^{(1)}) \\
\vdots \\
\hat{f}(\omega_i^{(N)}, \omega_j^{(M)}, \lambda^{(M)})
\end{pmatrix} \Delta \omega_i \Delta \lambda,$$

and both $\mathbf{w}$ and $\mathbf{x}$ are of size $NM \times 1$.

If $K = M$ and the LEDs have narrow-band spectra, $\mathbf{x} = \mathbf{x}$ is a 2D slice of spectral BRDF sampled at $N$ locations and $M$ bands. In our work, however, because the $K$ LEDs in each cluster are broadband, the feature vector $\mathbf{x}$ is in fact the projection of the 2D slice of spectral BRDF to the $K$ basis spectra of the $K$ LEDs. Specifically, the spectrum of each of the $N$ clusters is a weighted linear combination of the $K$ basis spectra,

$$\begin{pmatrix}
L(\omega_i^{(n)}, \lambda^{(1)}) \\
\vdots \\
L(\omega_i^{(n)}, \lambda^{(M)})
\end{pmatrix} = \mathbf{B} \cdot \mathbf{l}(\omega_i^{(n)}), \quad n = 1, \ldots, N,$$

where $\mathbf{B}$ is an $M \times K$ matrix consisting of the $K$ spectra of the $K$ LEDs primaries, and $\mathbf{l}(\omega_i^{(n)})$ is the $K \times 1$ weight vector for the $n$th cluster. Thus, we can rewrite $\mathbf{w}$ as

$$\mathbf{w} = \begin{pmatrix}
\mathbf{B} \cdot \mathbf{l}(\omega_i^{(1)}) \\
\vdots \\
\mathbf{B} \cdot \mathbf{l}(\omega_i^{(N)})
\end{pmatrix}.$$

With this, we can finally simplify (4) to

$$I = \mathbf{w}^T \mathbf{x},$$

where

$$\begin{pmatrix}
f(\omega_i^{(1)}) \\
\vdots \\
f(\omega_i^{(N)})
\end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix}
f(\omega_i^{(1)}) \\
\vdots \\
f(\omega_i^{(N)})
\end{pmatrix},$$

$$\mathbf{f}(\omega_i^{(n)}) = \mathbf{B}^T \cdot \begin{pmatrix}
\hat{f}(\omega_i^{(n)}, \omega_j^{(1)}, \lambda^{(1)}) \\
\vdots \\
\hat{f}(\omega_i^{(n)}, \omega_j^{(M)}, \lambda^{(M)})
\end{pmatrix} \Delta \omega_i \Delta \lambda.$$

Note that both $\mathbf{w}$ and $\mathbf{x}$ are of size $NK \times 1$. $\mathbf{w}$ is the coded illumination, i.e., the intensities of the $N \times K$ LEDs. $\mathbf{x}$ is the spectral BRDF feature vector, which is the projection of the 2D slice of spectral BRDF to the $K$ basis spectra of the $K$ LED primaries. When $K = M$ and the LEDs are narrow band, $\mathbf{B}$ becomes an identity matrix and $\mathbf{x}$ is the 2D slice of spectral BRDF. If we turn on each of the $N \times K$ LEDs sequentially, we can directly measure $\mathbf{x}$.

As shown, $I(\omega_i)$ directly measures the projection of spectral BRDF onto coded illumination. This implies that we can learn the optimal projection vector $\mathbf{w}$ from training samples for classification, and implement the projection to $\mathbf{w}$ using coded illumination. The classification is then done according to (1).

Fig. 2a shows a schematic diagram of this idea, where the spectral BRDFs of samples are shown as points while the discriminative illumination $\mathbf{w}$ is shown as a projection vector. In the training stage, raw images of the training samples, i.e., the BRDF feature vectors $\mathbf{x}$ of the training samples, are measured. The raw images can be measured either in a straightforward way, in which each image is captured when a single LED is turned on, or with multiplexing illumination [35] to reduce the noise. Next, the discriminative illumination $\mathbf{w}$ is obtained via training by maximizing the discrimination of materials based on a variety of metrics, such as Fisher’s Linear Discriminant Analysis (LDA) or the support vector machine (SVM) with a linear kernel [6], [7]. Since $\mathbf{w}$ may have negative values, we implement it as the difference of two nonnegative vectors,$^1$ $\mathbf{w} = \mathbf{w}^+ - \mathbf{w}^-$, where $\mathbf{w}^+ = \max(0, \mathbf{w})$ and $\mathbf{w}^- = -\min(0, \mathbf{w})$.

Fig. 2b shows an example of aluminum-versus-alloy classification. As shown in Fig. 2b, these two types of metal have very similar color. Fig. 2c shows the learned classification.
Fig. 2. Discriminative illumination as a physically-based linear classifier. 
(a) A schematic diagram in which coded illumination acts as a linear classifier, after projecting to which the spectral BRDFs of different materials are maximally separated. (b) An example of aluminum-versus-alloy classification. The image is captured by one of the 150 LEDs of the dome that yields the best classification performance on training data. Its classification rate on testing data is 41 percent. (c) We train a linear kernel SVM classifier from the same training data, with the classification rate of 95 percent on the testing data. The bar graph shows the normalized learned SVM light, \( w \), where the 25 bar groups correspond to the 25 LED clusters and the six bars within each group correspond to the six LEDs. The vertical axis shows the relative brightness of each LED.

Discriminative illumination \( w \) using a linear kernel SVM—which we call SVM light. The vector \( w \) is shown as 25 vertical bar groups. Each group has six bars corresponding to the brightness of the six LEDs within an LED cluster. Figs. 2d and 2e show the two nonnegative light patterns, \( w^+ \) and \( w^- \), on the top view of the LED dome. The colors and the brightness of the nodes show the mixed spectra of the LED clusters. As shown, \( w^+ \) has mainly blue and white colors, while \( w^- \) has mainly red and orange colors. \( w^+ \) has stronger incident light from grazing angles, while \( w^- \) has stronger incident light from nearly the center of the dome. Figs. 2f and 2g are the corresponding captured images, and Fig. 2h is the difference image of Figs. 2f and 2g. We also implemented a baseline method for comparison: for each dimension of the \( x \) of the training samples, we perform a brutal-force search to find an optimal threshold by which the samples are best classified. Each dimension and its corresponding threshold are set to be a weak classifier. Then, a strong classifier is built by combining the best weak classifiers in the Adaboost [8] framework. For fair comparison, we constrain the number of weak classifiers to be the same as the number of required images in our method. This method is referred to as the raw measurement method because it uses the raw measurements directly. Fig. 2i shows the classification result for our method while the accuracy of the raw measurement method is shown under Fig. 2b. With the same number of measurements, SVM light yields much higher classification rate (95 percent) than using raw measurement for classification (41 percent). The reported accuracies in Fig. 2 and subsequent figures with classification maps are the percentages of the pixels that are correctly classified in the region-of-interest in the images.

Fig. 11 shows another example of alloy-versus-steel classification. In addition to SVM, we also train discriminative illumination using Fisher LDA. As shown, both the Fisher light and SVM light have higher performance (93 and 96 percent) than using raw measurement for classification (72 percent).

### 3.2 SNR Benefits

In this section, we demonstrate that the discriminative illumination method results in higher SNR than the conventional methods. For simplicity of derivation, let us assume a two-class linear classification task. If we use the linear combination of the raw measurements of spectral BRDF for classification, we need to first sequentially acquire the \( M \times 1 \) measurements \( x \), and then perform linear classification based on a discriminative function \( y = w^T x + b \). Assuming there is additive noise in the imaging system, the noise \( n \) is added to the measurement \( x \), i.e., \( y = w^T (x + n) + b \). Assuming the variance of noise is \( \sigma \), the SNR of using raw measurement is \( w^T x / |w|^2 \sigma \).

For the discriminative illumination method, instead of \( M \) measurements, we at most capture two images, i.e., the images under \( w^+ \) and \( w^- \). For fair comparison, we need to make sure that the total amount of incident light used is the same as the method of using raw measurements. Since when using raw measurements the total incident light energy is \( M \), for discriminative illumination the incident light should be scaled from \( w \) to \( Mw/|w|_1 \). Therefore, the measured signal is \( Mw^T x / |w|^2 \). The additive noise is added to the measured signal:

- If read noise dominates, the SNR of the proposed method is \( Mw^T x / (\sqrt{2} \sigma |w|_1) \) because there are two captured images. The SNR gain compared to the method of using raw measurements is \( G_r = M/|w|_2 / (\sqrt{2} |w|_1) \).
- If photon noise dominates, the variance of the photon noise is \( \sqrt{M} \sigma \), and thus the SNR is \( Mw^T x / (\sqrt{M} \sigma |w|_1) \). The SNR gain compared to the method of using raw measurements is \( G_p = \sqrt{M}/|w|_2 / |w|_1 \).

Since \( |w|_1 / M \leq |w|_2 / \sqrt{M} \leq |w|_1 / \sqrt{M} \), we have

\[
\sqrt{M/2} \leq G_r \leq M/\sqrt{2} \quad \text{and} \quad 1 \leq G_p \leq \sqrt{M}. \tag{5}
\]
Alternatively, one can measure \( x \) with multiplexed illumination [35]. Multiplexed illumination requires the same number of measurements. As shown in [44], compared to single light sources, multiplexed illumination with Hadamard codes reduces the variance of read noise with a factor of \( M/4 \) but it increases the variance of photon noise with a factor of 2. Thus, if \( x \) is acquired with the multiplexed illumination scheme, the SNR gain can be estimated in a similar way and we have

\[
\sqrt{2} \leq G_r \leq \sqrt{2M}, \quad \text{and} \quad \sqrt{2} \leq G_p \leq \sqrt{2M}. \tag{6}
\]

The SNR benefit shows a considerable improvement of accuracy in classification when we are using images captured under the discriminative illumination, rather than the linear combination of the raw measurements. To evaluate the benefits of SNR by using discriminative illumination, we simulate the classification tasks for the materials in the MERL BRDF database [24]. This database includes BRDF measurements in RGB channels of 100 materials. In simulation, we assume the images are captured using a camera with a unit spectral sensitivity function and the spectra of the red, green, and blue light sources on the light dome have the same shapes as the spectral sensitivity functions of the camera used in [24]. Thus, the pixel intensity under the illumination of a single LED has a one-to-one correspondence to the (scaled) BRDF measurement in the database. The classification tasks are performed with an addition of noise of level \( r \). The noise is the sum of read noise and photon noise with zero means and variances \( \sigma_r = r \sqrt{I_{\text{all}}} \) and \( \sigma_{ph} = r \sqrt{I_{\text{all}} I_{\text{disc}}} \) respectively, where \( I_{\text{all}} \) is the average intensity of all raw measurements and \( I_{\text{disc}} \) is the average intensity of the image without noise.

Two other methods are simulated to compare with our proposed method. 1) We sequentially turn on each LED and measure the BRDF vector \( x \). After obtaining \( x \), we use the same classifier \( w^T x + b \) for classification. Both methods require \( M \) measurements, while our proposed method needs two measurements. We scale the exposures accordingly so they have the same time budget. Fig. 3 shows the classification rates for 10 tasks under different noise levels. As shown, compared with the two methods, our method has the highest performance due to the higher SNR in measurements and this performance advantage increases as the noise level increases as expected.

## 4 Extensions

In this section, we show the core idea of discriminative illumination can be extended to handle multiclass and nonlinear classification to deal with nonflat surfaces and to use images with multiple channels.

### 4.1 Multiclass Classification

Multiple discriminative illumination patterns can be used for multiclass classification, as shown in Fig. 4a. There are two common schemes to generalize binary classifiers for multiclass tasks: one-versus-all, which has \( N \) binary classifiers for \( N \) classes, and one-versus-one, which needs \( N(N - 1)/2 \) binary classifiers for \( N \) classes.

Fig. 4 shows an example of three-class classification, fabric-versus-ceramic-versus-plastic. Samples from these three classes have similar colors, as shown in Fig. 4a. If we select the four most distinctive raw measurements from the 150 measured images for classification and combined them (as described in Section 3.1), the best classification rate is 62 percent, as shown in Fig. 4g. To use discriminative illumination for this task, we implement the one-versus-all method, and thus train three discriminative illuminations with linear kernel SVM classifiers, \( w_1, w_2, w_3 \), as shown in Fig. 4c. Since these three discriminative illuminations have negative values, we implement them as four nonnegative illumination patterns, \( w^- = -\min(0, w_1, w_2, w_3) \), and \( w_+^k = w_k - w^- \), \( k = 1, 2, 3 \), as shown in Fig. 4d. Fig. 4e shows one of the four captured images under discriminative illumination, and Fig. 4f shows the classification result. With the

\[
\begin{align*}
I_{\text{Id}} &= \text{the average intensity of the image without noise.} \\
I_{\text{all}} &= \text{the average intensity of all raw measurements} \\
I_{\text{disc}} &= \text{the average intensity of the image without noise.}
\end{align*}
\]

2. These factors change if we consider sensor saturation. In this paper, we assume no saturation limit in sensors.
same number of measurements, our methods achieve 94 percent classification rate.

4.2 Cascade Classifier: From Linear to Nonlinear

Multiple discriminative illumination patterns can also be constructed as an ensemble classifier for nonlinear classification [33], [8], [9], such as boosting, bagging, random subspace, and cascade classification.

In this section, we focus on training a cascade classifier to solve the detection problem (i.e., one-class classification) [18], where the goal is to distinguish one class of samples from all other possible samples. Often the number of samples of the positive class (i.e., the target class) is much smaller than that of the negative class (i.e., all nontarget classes). As shown in [41], subsets of negative samples and all positive samples are used to train a classifier for each stage of a cascade classifier. We adjust the threshold of each stage to meet a given false negative rate while minimizing the overall false positive rate, as shown in Fig. 5a. Fig. 5b shows a toy example of detecting red + (target class) from blue circles. As shown in Fig. 5c, for this nonlinear classification task, a four-stage cascade classifier (where each stage is linear classification) is sufficient.

We use the following method to train a cascade classifier:

- **Input.** Positive sample set $P^+$, negative sample set $P^-$, and false negative rate $e^f$.
- **Step 0.** Initially set all negative samples as misclassified: $Q^- = P^-$.  
- **Step 1.** Randomly select a subset of misclassified negative samples $\hat{P}^-$ from $Q^-$ and make sure $\text{size}(\hat{P}^-) = \min(\text{size}(P^+), \text{size}(Q^-))$.  
- **Step 2.** Train a classifier based on $P^+$ and $\hat{P}^-$, and adjust the threshold so that the false negative rate of this stage $e^- \leq e^f$.  
- **Step 3.** Classify all negative samples in $P^-$ using the classifiers of the current and all previous stages. Set $Q^-$ as the misclassified samples in $P^-$.  
- **Step 4.** If the maximum stages have been trained, or $Q^-$ is empty, break and finish the training. Otherwise, go back to **Step 1**.

For material classification, we choose the task of detecting aluminum as an example and perform classification between aluminum and three other materials (steel, ceramic, and plastic), as shown in Fig. 6. This task has practical implications in recycling because efficiently finding and
recycling scrap aluminum from all other waste materials produces significant cost savings over the production of new aluminum [36], [13]. As shown in Fig. 6, we train a four-stage cascade classifier. Fig. 6a shows one of the raw measurements (under LED #137). Figs. 6b, 6c, 6d, and 6e show the corresponding SVM light $w$ (shown as a bar graph), the classification result (shown as a binary image), and the false negative rate and false positive rate for each stage. With four stages of discriminative illumination, we can achieve 4.2 percent false negative rate and 0.07 percent false positive rate. In comparison, if we train a single linear light for this problem with the same false negative rate, the false positive rate is 6 percent.

### 4.3 Nonlinear SVM Kernels

In addition to ensemble classifiers, another way for nonlinear classification with discriminative illumination is to employ classifiers, where only a dot product operation is needed. For example, consider a two-class classification problem with an SVM classifier:

$$\sum_{i=1}^{N} \alpha_i k(x, w_i) + b = \begin{cases} \geq 0 & y \in \text{Class 1}, \\ < 0 & y \in \text{Class 2}, \end{cases}$$

where the kernel function $k(x, w)$ is either the polynomial kernel

$$k(x, w_i) = (\gamma w_i^T x + r)^q$$

or the sigmoid kernel

$$k(x, w_i) = \tanh(\gamma w_i^T x + r),$$

and $N$ is the number of support vectors. In both cases, the kernel functions, although nonlinear, depend only on $w^T x$—projections of $x$. Therefore, these projections of $x$ can also be measured directly with coded illumination. Note, we need as many projection vectors as support vectors.

Table 1 shows the comparison of the linear SVM classifier and SVMs with nonlinear kernels. As shown, the classification accuracy increases with both nonlinear kernels (especially for the polynomial kernel). On the other hand, both nonlinear SVMs have many support vectors, which means we need to capture more coded images. The SNR analysis in Section 3.2 is also complicated for the nonlinear case. Nevertheless, nonlinear classifiers may still serve as a useful option for some tasks where classification accuracy is critical. If necessary, tradeoffs between accuracy and measurement cost can also be made by simplifying nonlinear classifiers to piecewise linear classifiers.

### 4.4 Material Classification for Nonflat Surface

In real-world material classification tasks, the surface normals of materials of interest are usually not pointing directly to the camera, as we have assumed before, due to the orientation of the surface or the fact that the surface is not flat. So it is necessary to consider those cases.

For a point on a nonflat surface, as shown in [42], the local coordinates around the point (where the surface normal is the north pole) and the global world coordinates are related by a rotation corresponding to the surface normal. Thus, by tilting a flat sample plate at multiple angles, we can obtain multiple rotational variants of the spectral BRDF feature vectors. To handle material classification without explicitly recovering the surface normal, one solution is to augment the training dataset with these

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The Performance of SVM Classifiers with Different Kernels and the Amount of Support Vectors for Each Class for Nonlinear Kernels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>Alloy vs. Steel</td>
</tr>
<tr>
<td></td>
<td>87.5%</td>
</tr>
<tr>
<td>Polynomial</td>
<td>100%(36, 28)</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>87.0%(100, 100)</td>
</tr>
</tbody>
</table>
rotational variants of the spectral BRDF feature vector. The learned discriminative illumination can then tolerate the variation of surface normals for material classification to some degree. This solution can only deal with mild surface normal variation because the augmentation may complicate the classification boundary as surface normal variation increases, and thus decrease classification performance. With large surface normal variation, it is recommended to explicitly recover surface normals or use multiple cameras.

Fig. 7 shows a toy example that demonstrates the feasibility of the proposed method to deal with surface normal variation for material classification. We create a sample with random surface normals (within \(\pm 10\) degrees variation in its tilt angle), as shown in Fig. 7c. The sample consists of two BRDFs. Figs. 7a and 7b show the renderings of the two BRDFs, and Fig. 7d shows the distribution of the two BRDFs on the sample. With a conventional point light source, the appearance of the sample is shown in Fig. 7e. Because of the random surface normal, it is difficult to separate the two BRDFs. Under the learned discriminative illumination, as shown in Fig. 7f, the two BRDFs can be separated more accurately. Certainly, this approach will work within some range of surface normal variation, depending on the complexity of the BRDFs.

To evaluate the method discussed above, we trained a set of SVM light patterns taking into consideration the surface normal variations in which the zenith angle of the surface normal, \(\phi_n\), ranges from 0 to 360 degrees. In this simulation, we use the BRDF measurements from the MERL BRDF database [24] as it has a rich collection of materials with dense BRDF measurements. The 100 samples are classified manually into five categories: paint, plastic, phenolic, fabric, and metal. The accuracies for SVM light patterns trained with and without considering the surface normal variation are shown in Table 2. For most tasks, the SVM light patterns trained using nonflat samples are more robust to surface normal variation. Note that for the tasks “paint versus plastic” and “paint versus metal,” this is not the case. This is because the points of different classes in the BRDF space before and after varying the surface normals are far enough so that these changes have little influence on the classification performance.

Figs. 8 and 9 show two experimental results on material classification with surface normal variation. In Fig. 8, we classify varnished and unvarnished paints. The patches are painted with touches of brushes so that there are (subtle) surface normal variations. The classification rate improves after the training set is augmented to include the nonflat paints, as shown in Figs. 8c and 8e. Fig. 9 is a more

<table>
<thead>
<tr>
<th>Task</th>
<th>PA v PH</th>
<th>PA v PL</th>
<th>PA v ME</th>
<th>PA v FA</th>
<th>PH v PL</th>
<th>PH v ME</th>
<th>PH v FA</th>
<th>PL v ME</th>
<th>PL v FA</th>
<th>ME v FA</th>
</tr>
</thead>
<tbody>
<tr>
<td>without</td>
<td>0.808</td>
<td>0.852</td>
<td>0.796</td>
<td>0.928</td>
<td>0.713</td>
<td>0.849</td>
<td>0.782</td>
<td>0.865</td>
<td>0.945</td>
<td>0.858</td>
</tr>
<tr>
<td>with</td>
<td>0.853</td>
<td>0.849</td>
<td>0.797</td>
<td>0.972</td>
<td>0.780</td>
<td>0.932</td>
<td>0.954</td>
<td>0.999</td>
<td>0.950</td>
<td>0.992</td>
</tr>
</tbody>
</table>

Higher accuracies are printed in bolder face.

Fig. 7. Simulation results of extending discriminative illumination for material classification with unknown surface normals. (a) and (b) Renderings of two BRDFs under natural lighting. (c) A sample with random surface normal (\(\pm 10\) degrees variation in the tilt angle, color coded). (d) The distribution of the two BRDFs on the sample. (e) Simulated image under a point light, with which it is difficult to separate the two BRDFs. (f) Simulated image under a discriminative illumination, with which we can separate the two BRDFs more accurately.

Fig. 8. Captured images and classification of nonflat diffuse (first row) and varnished (second row) paints. (a) Test samples. (b) Image under SVM light, without augmented training sets. (c) Classification from (b). The accuracy is 94 percent. (d) Image under SVM light, with augmented training sets. (e) Classification from (d). The accuracy is 97 percent.

Fig. 9. Captured images and classification of dented aluminum (first row) and steel (second row) plates. (a) Test samples, with region of interest marked in red. (b) Images under SVM light, without augmented training sets. (c) Classification from (b). The accuracy is 51 percent. (d) Image under SVM light, with augmented training sets. (e) Classification from (d). The accuracy is 61 percent.

Fig. 8. Captured images and classification of nonflat diffuse (first row) and varnished (second row) paints. (a) Test samples. (b) Image under SVM light, without augmented training sets. (c) Classification from (b). The accuracy is 94 percent. (d) Image under SVM light, with augmented training sets. (e) Classification from (d). The accuracy is 97 percent.

Fig. 9. Captured images and classification of dented aluminum (first row) and steel (second row) plates. (a) Test samples, with region of interest marked in red. (b) Images under SVM light, without augmented training sets. (c) Classification from (b). The accuracy is 51 percent. (d) Image under SVM light, with augmented training sets. (e) Classification from (d). The accuracy is 61 percent.
challenging task where we aim to classify dented aluminum and stainless steel plates. These two materials have very similar appearance and are both highly specular (which means the appearance changes dramatically with surface normal variation). The surface normal variation in this example is also much larger than Fig. 8. As shown in Figs. 9c and 9e, by adding samples oriented in four different directions into the training set, we can achieve a higher classification rate. We notice the result still has a lot to improve in this case—one potential solution is to explicitly recover surface normals.

4.5 Using Images with Multiple Channels

The proposed method can be directly extended to work with images with multiple channels, such as images taken with an RGB camera or multiple cameras. More specifically, with an RGB camera as the sensor, (3) becomes

\[ I(\omega_w) = X^T w, \]

where \( I \) is a column vector of pixel values in RGB channels, \( X \) is a \( M \)-by-3 matrix in which each column is the BRDF feature vector in one color channel, and \( w \) is the optimal projection vector. The optimal \( w \) that minimizes the within-class distance and maximizes the between-class distance of projected spectral BRDF feature vectors could be estimated using the same methodology as in the Fisher LDA. Instead of projecting the feature vectors onto a one-dimensional space, we project the feature vectors onto an \( N \)-dimensional space, where \( N = 3 \) for RGB cameras.

More specifically, let us define the BRDF vectors of class 1 and class 2 as \( X_{k,i} \) (\( i = 1, 2, 3, \ldots, N_k \) and \( k = 1, 2 \)), and the mean BRDF vectors of the two classes as \( \bar{X}_k \). Since we are considering the BRDF vectors in three channels, both \( X_{k,i} \) and \( \bar{X}_k \) are \( M \)-by-3 matrices, where \( M \) is the number of BRDF vectors in each channel. We estimate the optimal projection vector \( w \) such that the sum of within-class distances of projections, \( D_w \), is minimized,

\[
D_w = \sum_{k=1}^{2} N_k \sum_{i=1}^{N_k} (X_{k,i}^T w - \bar{X}_k^T w)^T (X_{k,i}^T w - \bar{X}_k^T w) = \sum_{k=1}^{2} \sum_{i=1}^{N_k} w^T (X_{k,i} - \bar{X}_k) (X_{k,i}^T - \bar{X}_k^T) w = w^T S_b w,
\]

while the between-class distance, \( D_b \), is maximized,

\[
D_b = (\bar{X}_1^T - \bar{X}_2^T) (\bar{X}_1^T - \bar{X}_2^T) w = w^T S_b w = w^T S_b w. \]

Following the same methodology of LDA, we maximize the ratio \( J = D_b / D_w \). We solve \( w \) by taking the derivative of \( J \) with respect to \( w \) and setting it to zero:

\[
\frac{\partial J}{\partial w} = \frac{\partial}{\partial w} \left( \frac{w^T S_b w}{w^T S_w w} \right) = 0. \]

Combining (11), (12), and (13), we have

\[ S_b w = \frac{w^T S_b w (S_w w)}{w^T S_w w} = \lambda S_w w, \]

which is a generalized eigenvalue problem. We can find the largest ratio \( J \) by estimating the largest eigenvalue. The optimal projection vector is the corresponding eigenvector.

The result for a multiclass classification task using the color features is shown in Fig. 10. As shown, there is a large improvement in classification rate by using images with multiple channels. This method can also be applied if we have multiple cameras mounted on the dome.

5 EXPERIMENT SETUP AND RESULTS

5.1 Experiment Setup

As shown in Figs. 11a and 11b, we construct an LED-based multispectral dome for material classification. The hemispherical geodesic dome is 1 meter in diameter with 70 nodes, 25 of which are mounted with LED clusters. Each LED cluster has six color LEDs—blue, green, amber, yellow, red, and orange—plus a white LED in the center, as shown in Fig. 12a. The six colors are chosen to cover the visible spectrum. UV and near infrared LEDs can also be used for certain material classification tasks. A thin diffuser is used to uniformly mix the colors within an LED cluster. Each LED cluster is controlled by an Arduino Duemilanove (with ATmega328) board that can provide six 8-bit pulse width modulation (PWM) outputs. At the center of the dome, we have a Lumenera Lu165 monochromatic CCD camera. Fig. 12b shows a top view of the dome, with labels of the 25 LED clusters. Both the LEDs and the camera have been geometrically and radiometrically calibrated beforehand. We also perform flat fielding to ensure the uniformity of incident illumination on the sample mounted at the center of the dome.

In the application of discriminative illumination for material classification, it is important to know how much contribution the color of light source or the available light source directions makes to the performance of the classification system. As shown in Fig. 13, we perform
Simulations for three classification tasks: alloy-versus-steel, aluminum-versus-alloy, and fabric-versus-ceramic-versus-plastic. For each sample, we measure \( \frac{25}{\binom{6}{2}} = 150 \) images corresponding to the 150 LEDs in the dome. Thus, each point on a sample plate has a \( \frac{150}{1} \) feature vector for classification. The left column of Fig. 13 shows the classification rates if we only use the color features for classification, i.e., we turn on the LED of the same color in all the LED clusters. For each number of color bands, we randomly select five possible combinations of light sources, then average the classification rates. The curves show that as the number of color bands increases, the performance increases but soon reaches the limit. The middle column of Fig. 13 shows the classification rates if we only use the angular distribution of reflectance (i.e., BRDF) for classification, i.e., we use only white LEDs and disable all other color LEDs. Similarly, as the number of LED clusters increases, the performance increases, while it saturates around 25. Finally, if we use both color and BRDF information, as shown, the classification rate is at most 72 percent if we use one of the 150 LEDs. With the learned discriminative illumination (Fisher light or SVM light), we can achieve much higher classification rates (93 or 96 percent). In addition, we show in Section 3.2 there is also an SNR benefit of using discriminative illumination. A real-time demo is given in the supplementary video, which can be found in the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPAMI.2013.110.
shown in the right column of Fig. 13, the classification performance will further improve. These plots show that 1) color and BRDF are complimentary to each other for material classification, and 2) six colors, 25 LED clusters, and 150 LEDs are necessary and sufficient because the performance improvement (versus the number of colors, LED clusters and LEDs) saturates around these values.

We use the multispectral dome as a prototype to implement discriminative illumination for classifying a variety of raw, unpainted materials. As shown in Fig. 1, we collect a database of five material classes: metal, plastic, ceramic, fabric, and wood. For metal, we also have six subclasses: alloy (#4130 steel), aluminum, steel (hot roll and cold roll), stainless steel, brass, and copper. In total, there are 100 samples. Each sample is a nearly flat plate of size $4 \times 4$ inches. We use the dome and measure the spatially varying spectral BRDF data (i.e., 150 images) for each sample plate. This database is available at http://compimg1.cis.rit.edu/data/metal.

### 5.2 Results

We randomly choose half of the sample plates from each material category as the training data, and use the other half as the testing data. Figs. 2, 4, 11, and 6 show the classification results for several tasks. Table 3 summarizes classification rates for several other tasks. We evaluate a variety of linear classifiers and find in general Fisher light (i.e., LDA) and SVM light (i.e., SVM with a linear kernel) have better performance. For comparison, we also show the classification results of using the same number of raw measurements. For example, for a three-class classification task, the discriminative illumination method (i.e., Fisher light and SVM light) needs four images, and for using raw measurements we select the best four from the 150 images for classification (as described in Section 3.1). The experimental results show that the discriminative illumination method in general has higher performance.

We also try discriminative illumination for a challenging but desirable task for recycling aluminum scrap—classifying aluminum by alloy family. Depending on the alloying elements, aluminum alloys include 2000 series (alloyed with copper), 5000 series (alloyed with magnesium), 6000 series (alloyed with magnesium and silicon), 7000 series (alloyed with zinc), and so on. Current approaches are mainly based on laser-induced breakdown spectroscopy [2], which is expensive. Here, we use the discriminative illumination to classify four types of aluminum alloys, i.e., #2024, #5052, #6061, and #7075. As shown in Fig. 14, for this challenging task the discriminative illumination yields reasonably good results (for Fisher light, 71 percent accuracy and for SVM light, 73 percent accuracy). In comparison, using raw measurements we can only achieve 37 percent classification rate. This example shows the effectiveness of the proposed method for raw material classification.

### 6 Conclusions and Discussions

We propose a novel approach of using coded illumination for classifying raw materials based on projections of
s spectral BRDFs. Optimal illumination patterns are learned from training samples, which directly measure discriminative features for classification. This approach is more efficient than using raw measurements and also has high SNR due to illumination multiplexing. We construct an LED-based multispectral dome as a prototype to implement this approach for classifying a variety of raw materials. We show even for some challenging tasks, the discriminative illumination approach can achieve high classification rates. There are several limitations in our current approach that we plan to address in our future work.

Large surface normal variation. So far, for cases of nonflat surfaces, we augment the training dataset with variants of spectral BRDF feature vectors by tilting flat sample plates at different angles. As discussed, this method would decrease the distinctiveness between different classes. In the future, we want to tackle this issue by either increasing the dimensionality of the space in which the spectral BRDFs are represented (i.e., introduce more spectral bands and directions of light sources/sensors), or using other features that are more robust to the orientations (e.g., recovered parameters of BRDF model, texture).

Texture. In addition to color and BRDF, another important appearance feature is texture. In this work, we focus on per-pixel material classification and use only spectral BRDFs. In the future, we plan to look into texture features for material classification.

Inter-reflection and subsurface scattering. Global illumination such as inter-reflection and subsurface scattering can be separated from direct illumination by modulating the discriminative light patterns with high-frequency patterns [26]. These global components capture surface roughness (in the case of inter-reflection) and translucency (in the case of subsurface scattering), and can also be used as additional features for material classification.

Going beyond the outer layer of materials. Using surface spectral BRDFs, we can only classify raw, unpainted materials. For painted or coated materials, we need to extract features that goes beyond the outer layer, e.g., by using X-rays, or rely on other sensory inputs such as sound, density, hardness, smell, and so on. Nevertheless, the proposed methodology of finding optimal projection vectors via training while implementing the projection with computational illumination or imaging may still apply.

ACKNOWLEDGMENTS
The authors thank Profs. Oliver Cossairt, Shree K. Nayar, and Gabrielle Gaustad for the helpful discussions, as well as suggestions from anonymous reviewers. This work was supported by a grant from NYSP2I and US National Science Foundation (NSF) IIS-1257163.

REFERENCES


Chao Liu received the BS degree from the University of Electronic Science and Technology of China (UESTC) in 2010. Since 2010, he has been a PhD candidate under the supervision of Jinwei Gu at the Rochester Institute of Technology (R.I.T). He is a student member of the IEEE.

Jinwei Gu received the bachelor’s and master’s degrees from Tsinghua University, China, in 2002 and 2005, respectively, and the PhD degree from Columbia University in May 2010. He is currently an assistant professor at the Munsell Color Science Laboratory, Center for Imaging Science, Rochester Institute of Technology. His research interests include computer vision and computer graphics. He is a member of the IEEE.

For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.