

Reflectance Hashing for Material Recognition

Hang Zhang¹ Kristin Dana¹ Ko Nishino²

¹Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854

²Department of Computer Science, Drexel University, Philadelphia, PA 19104

zhang.hang@rutgers.edu, kdana@ece.rutgers.edu, kon@drexel.edu

Abstract

We introduce a novel method for using reflectance to identify materials. Reflectance offers a unique signature of the material but is challenging to measure and use for recognizing materials due to its high-dimensionality. In this work, one-shot reflectance of a material surface which we refer to as a reflectance disk is capturing using a unique optical camera. The pixel coordinates of these reflectance disks correspond to the surface viewing angles. The reflectance has class-specific structure and angular gradients computed in this reflectance space reveal the material class. These reflectance disks encode discriminative information for efficient and accurate material recognition. We introduce a framework called reflectance hashing that models the reflectance disks with dictionary learning and binary hashing. We demonstrate the effectiveness of reflectance hashing for material recognition with a number of real-world materials.

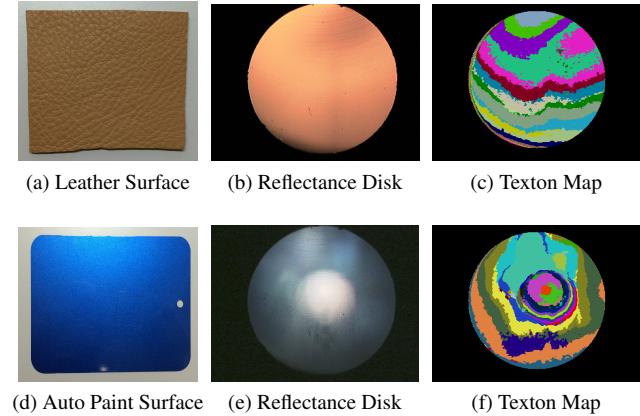


Figure 1: Reflectance disks provide a quick snapshot of the intrinsic reflectance of a surface point. Gradients of the reflectance space are captured with textons and provide a signature for material recognition.

1. Introduction

Color and geometry are not a full measure of the richness of visual appearance. Material composition of a physical surface point determines the characteristics of light interaction and the reflection to an observer. In the everyday real world there are a vast number of materials that are useful to discern including concrete, metal, plastic, velvet, satin, asphalt, carpet, tile, skin, hair, wood, and marble. A computational method for identifying these materials has important implications in developing new algorithms and new technologies for a broad set of application domains. For example, a mobile robot or autonomous automobile can use material recognition to determine whether the terrain is asphalt, grass, gravel, ice or snow in order to optimize mechanical control. An indoor mobile robot can distinguish among wood, tile, or carpet for cleaning tasks. The material composition of objects can be tagged for an e-commerce

inventory or for characterizing multi-composite 3D printed objects. The potential applications are limitless in areas such as robotics, digital architecture, human-computer interaction, intelligent vehicles, and advanced manufacturing. Furthermore, just as computer vision algorithms now use depth sensors directly from RGB-D cameras, material sensors can have foundational importance in nearly all vision algorithms including segmentation, feature matching, scene recognition, image-based rendering, context-based search, object recognition, and motion estimation. Our approach uses reflectance for material recognition because of the advantages of having an intrinsic optical signature for the surface. However, we bypass the use of a gonioreflectometer by using a novel one-shot reflectance camera based on a parabolic mirror design. The output of this camera is a *reflectance disk*, a dense sampling of the surface reflectance of the material projected into a single image as shown for two

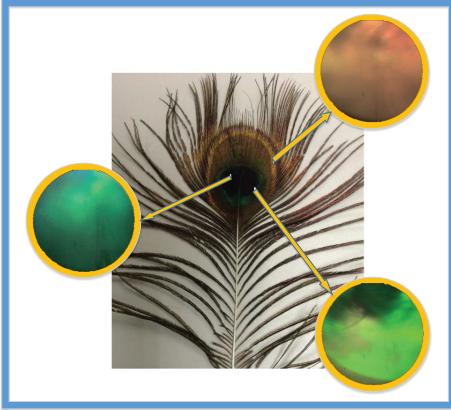


Figure 2: Reflectance disks provide a snapshot of the reflectance for spot samples on the surface. For this example of a peacock feather, iridescence causes a large variation of intensity with viewing direction as revealed in the three reflectance disks.

example surfaces in Figure 1. Each reflectance disk measures surface properties for a single point and can capture complex appearance such as the iridescence of a peacock feather as illustrated in Figure 2. We then use this convenient reflectance measurement as the discriminative characteristic for recognizing the material class. We address the issue of high dimensionality using a novel application of binary hash codes to encode reflectance information in an efficient yet discriminative representation. The key idea is to obtain sufficient sampling of the reflectance with enough discriminative power and reduce its representation size so that it can be effectively used for probing the material.

We present a database of reflectance disks comprised of twenty different diverse material classes including wood, velvet, ceramic and automotive paint with 10 spot measurements per surface and with three different surface instances per class. Measurements include three on-axis illumination angles ($-10^\circ, 0^\circ, 10^\circ$) and ten random spot measurements over the surface. Each spot measurement is a reflectance disk composed of a dense sampling of viewing angles totaling thousands of reflectance angles per disk. The database of 3600 images or reflectance disks is made publicly available. For recognition, we combine binary hash coding and texton boosting for a new framework called *reflectance hashing* for efficient and accurate recognition of materials. We compare reflectance hashing with texton boosting for the task of recognizing materials from reflectance disks.

2. Related work

Prior methods for material recognition use two distinct approaches. One approach assesses material identity using reflectance as an intrinsic property of the sur-

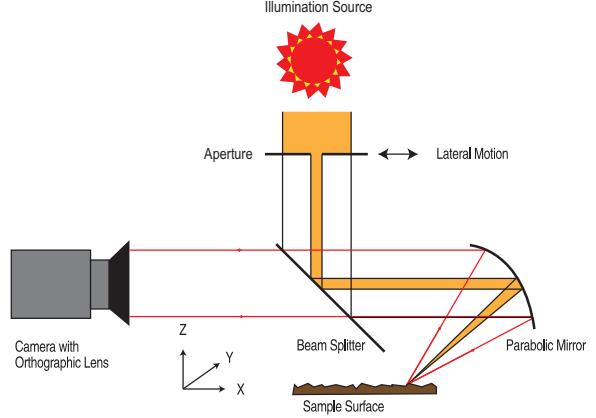


Figure 3: Schematic of the mirror-based camera. Reflectance disks are obtained by viewing a single point under multiple viewing directions using a concave parabolic mirror viewed by a telecentric lens.

face [5, 13, 17, 19, 26]. Another main approach identifies material labels using the appearance of the surface within the real world scene [21, 23, 34]. Using reflectance instead of scene appearance has the advantage that reflectance is a direct measurement of the material characteristics, instead of its phenomenological appearance [1]. Reflectance is mostly unique to the material, whereas the appearance is the convoluted end result of the interaction of all the intrinsic and extrinsic factors and thus is more difficult to decipher.

A challenge in using reflectance for material recognition is that measurements are typically difficult and cumbersome. For example, most methods require knowledge of the scene like geometry [26, 27] or illumination [30, 31]. Other methods require lab-based measurements of the BRDF (bidirectional reflectance distribution function) or BRDF slices such as light domes for illumination patterns [22]. Acquiring full BRDF requires gonioreflectometers that are comprised of multiple cameras and light sources covering the hemisphere of possible directions using geodesic domes or robotics as in [8, 9, 20, 28, 29, 45, 46]. Surface appearance in terms of relief texture is captured by the GelSight camera [15] using three-color photometric stereo by imprinting the surface geometry on an elastomer surface with Lambertian skin. With this device very fine geometric texture can be recovered but surface appearance is lost. An additional challenge of using reflectance for measurements is the high dimensionality, reflectance as a function of illumination and viewing direction especially if densely sampled can lead to thousands or more samples per surface point.

Specialized cameras have been developed in prior research to obtain reflectance measurements of surfaces efficiently. The mirror-based camera illustrated in Figure 8 is one such a device [6, 7]. This camera views an image

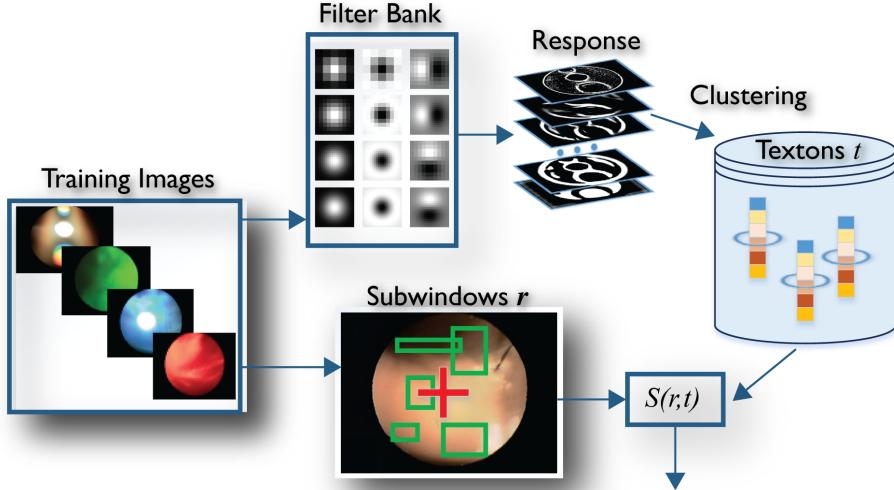


Figure 4: Gradients on the Reflectance Disk Extracted with Textons: Images are filtered by a filter bank comprised of Gaussian, Laplacian and oriented gradient filters. The 24×1 responses vectors are clustered to form visual words or textons. A random subsampling of rectangular subregions of the reflectance space defines regions of interest r . The feature of interest is $S(r, t)$, the count of pixels identified as texton t in region r .

of a single surface point by using a concave parabolic mirror with the focus coincident with the target point. In this manner, the camera records a multiview image of a surface point where each pixel records a different angle. Multiple views from mirror-based optics is also achieved with a kaleidoscope-like configuration [14]. However, the viewing angles captured in this device are discrete and sparse. Because our approach relies uses gradients in angular space a dense sampling of reflectance is needed and we use the parabolic mirror-based camera as a reflectance sensor.

In prior work, recognition of standard scene images is typically accomplished using image features that capture the spatial variation of image intensity. For example, image filtering with a set of filter banks followed by grouping the outputs into characteristic textons has been used for image texture [4, 19, 40, 42], object recognition [35, 36] and image segmentation [18]. Similarly, we encode the discriminative optical characteristics of materials captured in the reflectance disks with a texton-based representation. Texton methods for the general problem of scene and image recognition have been improved significantly by incorporating boosting as in *Texton-boost* [36] and *Joint-boost* [38]. We incorporate the utility of boosting to identify a discriminative and compact representation.

The goal of efficient recognition of images for large scale search has led to numerous methods for binary hash codes [2, 16, 24, 25, 44]. A common approach in recent work [16, 33, 37, 47] is learning binary codes where the Hamming distance approximates the kernel distance between appear-

ance descriptors. This method is used when the kernel distance is known to give good performance but is expensive to compute and store. Inspired by the work in image search, we develop *reflectance hashing* with binary codes that allow matching of measured reflectance disks with those in the labeled training set. The binary codes are learned using the match metric obtained from the texton-based descriptor we present in Section 3. We use a suite of state-of-the-art hashing techniques applied to our measured reflectance disk dataset to evaluate recognition performance in terms of speed and recognition accuracy.

3. Methods

3.1. Reflectance Disk Measurements

We use the *texture camera* approach introduced in [6, 7, 41, 43] for measuring reflectance of surface points. The measurement camera mainly consists of a parabolic mirror, a CCD camera, a movable aperture and a beam splitter. The imaging components and their arrangement are illustrated in Figure 3. A parabolic mirror section is fixed so that its focus is at the surface point to be measured. The illumination source is a collimated beam of light parallel to the global plane of the surface and passing through a movable aperture. The angle of the incident ray reaching the surface is determined by the intersection point with the mirror. Therefore, the illumination direction is determined by the aperture position. Light reflected from the surface point is reflected by the mirror to a set of parallel rays directed

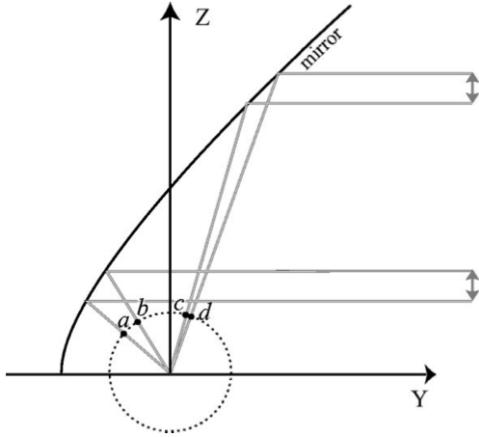


Figure 5: Gradients on the reflectance disk are computed using image filtering. However, the angular resolution varies over the surface as shown above. The angle at ab is different from cd . Therefore differences (or filtering) computed with a uniform spacing on the reflectance disk have a non-uniform mapping to angles.

through a beam splitter to the camera. Each pixel of this camera image corresponds to a viewing direction of the surface point. Therefore, the recorded image is an observation of a single point on the surface but from a dense sampling of viewing directions. Multiple illumination directions can be measured by planar motions of the illumination aperture. A key advantage of this approach is the potential for a handheld reflectance sensor where the mirror and light source are attached to a handheld camera.

We refer to the camera’s output images as *reflectance disks* since they depict circular regions from on-axis projections of a paraboloid. Examples of these reflectance disks are shown in Figure 1 for two surfaces. These examples show the variation of the surface reflectance with viewing angle. Reflectance disks have an intensity variation that corresponds to angular change with the viewing direction instead of spatial change on the surface. However, the reflectance disks can be interpreted as images and filtered to find characteristic gradients. This approach of using *angular gradients in reflectance space* is a novel contribution of our work. Our goal is to represent the measurements in an efficient and meaningful way that supports material recognition.

3.2. Textons on the Reflectance Disk

We find characteristic patterns of intensity in a reflectance disk by computing *spatial change* in the reflectance disk via image filtering. Since spatial coordinates of the reflectance disk map to angular coordinates of reflectance, filtering is a convenient way to compute angular

gradients of the reflectance. A common method of quantifying intensity patterns is using *textons*. Textons are computed by first filtering the image with a set of gradient filters at different orientations and scales. We employ the same filter banks as in [19], comprised of Gaussian low-pass filters at four scales, Laplacian filters at four scales and eight gradient filters at different orientations and scales for a total of 24 filters (ranging in size from 7×7 to 25×25). The resulting 24×1 responses at each pixel are clustered using K -means clustering and a dictionary of visual words called textons is created. The underlying assumption is that the local intensity variation can be captured with the finite set of characteristic filter responses that are the centers of each of the K clusters. Each pixel in the reflectance disk is assigned a texton label and the resulting label image is called a texton map (as illustrated in Row 2 of Figure 6).

We use textons as a reflectance feature to provide a dense, per-pixel description of reflectance variation. Methods to detect key-points are useful for scene images, but reflectance disks do not typically have specific key-points of interest.

For texton computation, we use gradient filters that approximate the derivative operator with a discrete spatial difference filter. This discrete approximation of the gradient computes change over a non-infinitesimal distance, typically the difference of neighboring pixels. The distance between these neighboring pixels is constant over the reflectance disk, i.e. the same gradient filter is applied over the entire disk image. However, because of the mapping from spatial coordinates to angular coordinates by the parabolic mirror, the angular distance is not constant. This situation is illustrated in Figure 5 where the cone angle at the surface is shown for two different locations on the parabolic mirror. A constant distance d on the projected disk leads to two different cone angles. Consequently, the resolution for the angular gradient varies as a function of disk coordinates.

The equation of the parabolic mirror surface is given by,

$$y + F = \frac{z^2 + x^2}{4F}, \quad (1)$$

where the y -axis is the optical axis of the mirror aligned with the camera optical axis, the $x - z$ plane is aligned with the disk-shaped projection of the mirror surface on camera’s image plane, and F is the mirror focal length (12.7 mm). Define the cone angle α as the set of angles subtended by arc ab or cd in Figure 5, where d is taken as 2 pixel units. From this equation, the variation of the cone angle α as a function of the vertical spatial coordinate z on the reflectance disk is derived in [6] as

$$\alpha(z) = \arctan \frac{z - 1}{\frac{(z-1)^2}{4F} - F} - \arctan \frac{z + 1}{\frac{(z+1)^2}{4F} - F}. \quad (2)$$

Since the geometry of the mirror is known, the spatial filters can be adaptive to ensure a uniform cone angle over the

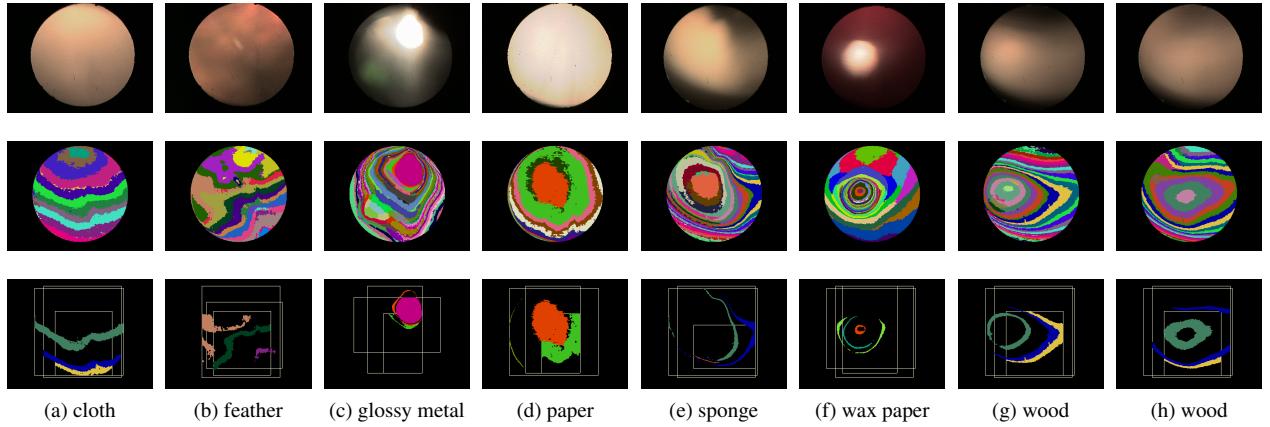


Figure 6: Visualizing the reflectance-layout filters. The first row shows the reflectance disks for different materials. The second row shows the corresponding texton maps. The third row shows three of the dominant region-texton (r, t) combinations chosen for each class.

disk. However, the spatially invariant gradient filter provides a discriminative representation and uniformity of angular resolution is not required.

Reflectance-Layout Filter Textons are a local measure of image structure, and texton histograms are used to represent a region. The histogram representation has the advantage of invariance to shifts of structures. Reflectance disks show the structure of the reflectance field, but often the distribution within that structure is the characteristic property. Histogram representations have the drawback that crucial spatial information is lost. We follow the texton boost framework [35, 36, 38] to encode the characteristic spatial structure of the reflectance disk by defining a set of randomly shaped rectangular regions over the disk. For a given texton t and region r , $S(r, t)$ is the count of pixels labelled with texton t that are located within region r as illustrated in Figure 4. We use this response as a key feature for describing the reflectance disks. This response is referred to as the texture-layout filter response when applied to ordinary images [36]. For material recognition using reflectance disks, we use the term *reflectance-layout filter* to indicate that the representation captures the angular layout of reflectance gradients. For computing reflectance-layout filter responses, the number of particular textons in a region, the texton count is calculated using a soft-weighting for the 8 nearest cluster centers. Soft-weighted textons as in [18, 36, 39] allows each pixel to be associated with 8 nearest clusters instead of a single texton label. This method alleviates the problem of two similar responses assigned to different texton labels.

Each reflectance layout filter characterizes a specific region in reflectance space where the change of reflectance as

a function of viewing angle is a descriptive feature for the material. There are many possible combinations of (r, t) as illustrated in Figure 4 and we want to choose the ones that are most discriminative. The boosting approach in [35, 38] is used to identify the most discriminative (r, t) combinations. After boosting, the typical numbers of reflectance layout filters is approximately 700.

3.3. Texton Boosting

The algorithm for *Joint Boosting* [38] provides feature selection where the feature is the reflectance-layout filter $S(r, t)$ specified by a particular (r, t) pair that indicates the number of times texton t occurs in region r . The approach builds a strong learner iteratively by adding the strongest weak learner which is based on the dominant feature at each iteration, allowing the weak learner to help classify a subset of classes. At each step m , the method maximally reduces the weighted squared error by choosing the strongest weak learner $h_m(I_i, c)$, updating the strong learner as: $H(I_i, c) := H(I_i, c) + h_m(I_i, c)$, where I_i is the input image and c is the class. The weighted square error is given by

$$J_{wse} = \sum_{c=1}^C \sum_{i=1}^N w_i^c (z_i^c - h_m(I_i, c))^2, \quad (3)$$

where C is number of classes, and N is the number of training samples, $w_i^c = e^{-z_i^c H(I_i, c)}$ defines the weights for class c and sample i that emphasize the incorrectly classified samples for that iteration, z_i^c is the membership label (± 1) for class c . The optimal solution that gives the selected learner at step m has the form

$$h_m(I_i, c) = \begin{cases} a\delta(S(r, t) > \theta) + b & , \text{if } c \in T_m \\ k^c & , \text{otherwise} \end{cases} \quad (4)$$

with the parameters (a, b, k^c, θ) given by the minimization of Equation 3, $\delta(\cdot)$ is an indicator function (0 or 1), $S(r, t)$ is the count or texton t in region r , T_m is the class subset that share this feature. For the classes sharing the feature ($c_i \in T_m$), the weak learner gives $h(I_i, c) \in \{a + b, b\}$ determined by the comparison of reflectance-layout filter response $S(r, t)$ with the threshold θ . For the classes not sharing the feature, there is a constant k^c different for each class c that avoids the asymmetry caused by the number of positive and negative samples for each class.

This method chooses a dominant weak learner $h_m(I_i, c)$ at each iteration to minimize the error function. For each selected learner, finding the class subset T that maximally reduces the error on weighted training set is expensive. Instead of searching all possible $2^C - 1$ combinations, a greedy search heuristic [38] is used to reduce the complexity to $O(C^2)$. We first find the best feature for each single class, and pick the class that has the best error reduction. Then we add the class that has best joint error reduction until we go through all the classes. Finally, we select the set that provides the largest overall error reduction. It is also time consuming to go through all the features at each iteration, so only a fraction $\tau \ll 1$ reflectance-layout filters are examined, randomly chosen in each iteration. The bottom row of Figure 6 shows the selected dominant features (reflectance-layout filters) for a number of different reflectance disks

3.4. Reflectance Hashing

Boosting allows the selection of learners $h_m(I_i, c)$ that specify (r, t) (region, texton) pairs of interest. The selected learners can simply be viewed as the input feature vector for a basic nearest neighbor classification. Nearest neighbors in high dimensions is problematic, but the computational tools of binary hash codes enables efficient and accurate representation. We use the reflectance layout filters directly in a concatenated feature vector with $700 \times c$ component where each component is a region-texon (r, t) pair that is known to be useful in material recognition. There is likely to be redundancy in the (r, t) combinations so the number of components in the feature vector may be reduced accordingly. This high dimensional feature vector is used directly for fast computation by employing binary coding. This efficiency is highly desirable for the embedded hardware implementation.

We use a suite of state-of-the-art approaches for forming binary code words including circulant binary embedding (CBE) [49], bilinear embedding [10], iterative quantization (ITQ) [12], angular quantization-based binary codes(AQBC) [11], spectral hashing (SH) [48], locality sensitive hash(LSH) [3], and locality-sensitive binary codes from shift-invariant kernels (SKLSH) [32]. We refer to this direct encoding and use of the reflectance layout filters as reflectance hashing.

4. Experiments

Reflectance Disk Database To evaluate the performance of our reflectance hashing framework, a database of 3600 reflectance disks is collected consisting of the following 20 surface classes: cardboard, CD, cloth, feather, textured rubber, glossy ceramic, glossy metal, leather, linen, mattel metal, automotive painting, non-ferrous metal, paper, plastic, rock, rubber, sponge, velvet, glossy paper and wood. The measurement camera shown in Figure 8 is arranged as in Figure 3 with a video camera, telecentric lens, light source, beam splitter and parabolic mirror. The database includes 3 instances per class, i.e. three different example surfaces per class, 10 spot samples per surface with 3 illumination directions $(-10^\circ, 0, 10^\circ)$ where 0° is frontal illumination aligned with the surface normal. Additionally, two exposure settings are collected for each reflectance disk. Therefore, the dataset consists of 180 reflectance disks per class, for a total dataset of 3600 reflectance disks each of size 320×240 . This reflectance dataset is made publicly available.¹

Material Recognition We compare the performance of texton boosting with reflectance hashing for material classification. We compared the classification result with a training set of N random selected images for both methods, and test the classification accuracy on a test set of $3600 - N$ random selected images. We vary N from 360 (10%) to 1500. For the large set of training images, both methods perform well; however, the recognition rate of joint boosting decreases significantly, when decreasing the size of training set. Reflectance hashing gives a more stable recognition rate as a function of the training set size as shown in Figure 9.

Figure 7 shows the confusion matrix from $N = 360$ classification with joint boosting with 700 iterations in the training stage, and mean recognition rate of 84.38%. The parameter settings were $K = 512$ textons and 200 subwindows or regions, with the random feature selection fraction $\tau = 0.01$. Figure 10 shows the confusion matrix obtained by our method of reflectance hashing, with the features selected in 700 iterations, and using iterative quantization [12] as the binary embedding method. The overall recognition rate is 92.3%, and several individual class recognition rates are significantly higher than the boosting method of Figure 7. For both methods, 5-fold cross validation was performed.

We also make a baseline comparison with texton histogram classification where no boosting is done and no region subwindows are used. The histogram is computed over the entire reflectance disk. The recognition rate with 1500 training images is 79.53% with a max rate of 100% (plastic, glossy paper) and a minimum of 45.6% (sponge). The recognition rate with 360 training images is 41.94% with a

¹Available at <http://www.ece.rutgers.edu/vision>

True class	Inferred class	cardboard	CD	woven fabric	feather	textured rubber	glossy ceramic	glossy metal	leather	linen	matte metal	automotive paint	non-ferrous metal	paper	plastic	rock	rubber	sponge	velvet	glossy paper	wood
		96.1	0.1																		3.8
	CD	99.1	0.1					0.2	0.1			0.3	0.1								
woven fabric			74.1	0.2					5.5	0.8	0.1		0.7		8.3			1.3	6.9		2.0
feather	0.1			92.9					1.9		0.2							3.1			1.7
textured rubber		0.1		72.8	0.4				0.9	0.6		3.4		1.2	17.6	0.5	2.5				
glossy ceramic					5.0	84.1	2.2				5.1		1.5			0.1					2.0
glossy metal						1.0	98.7	0.2			0.1										
leather	8.2		1.1					83.7	1.2			0.2				0.7		0.1			4.8
linen	0.6		5.9	0.2	1.5	1.3	2.4	3.5	61.9	2.4	3.1	0.5	0.5	0.5	12.3		0.4	0.3	0.2	0.2	2.3
matte metal						3.6	0.2				94.8	0.2	0.9				0.2				
automotive paint			0.4		2.0	2.8	1.0		3.3	0.4	81.7	3.6			0.6	0.8	1.1	2.4			
non-ferrous metal					0.3		0.8	0.5	1.9	1.5	2.3	85.3	0.3	0.3	1.1		0.6				5.2
paper	0.1			0.1	3.9					2.0		89.8	1.1		1.3		1.7				
plastic											0.1		99.4	0.1		0.1		0.2			
rock	0.2		10.5		2.9			4.2	4.0		0.1	0.1		66.8	1.4	4.2	3.9			1.6	
rubber						1.1				0.1		2.0		0.1	96.0	0.1	0.5				
sponge			23	1.7	43					0.7					12.0	0.2	66.1	9.8		2.7	
velvet			1.9	3.2	2.2					0.2	0.4		1.1		0.7	6.4	4.8	78.6	0.3	0.1	
glossy paper			2.3	0.2		0.1		0.5					0.2					96.4	0.2		
wood	9.1		1.1	0.1				5.7	2.0		0.5		1.5		3.8		1.4			74.9	

Figure 7: Joint Boosting Classification. Confusion matrix with percentages row-normalized. Overall accuracy is 84.38%.

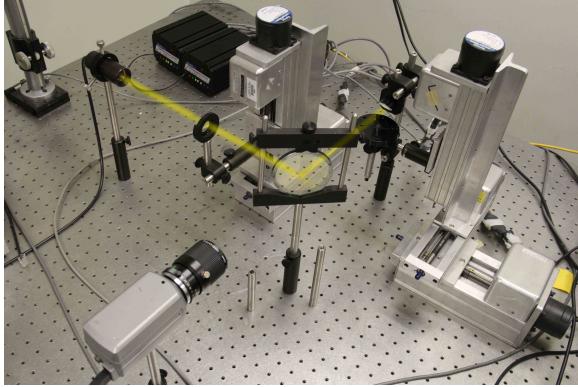


Figure 8: Prototype of the camera for reflectance disk capture. The components follow the arrangement in Figure 3.

max rate of 89.2% (cardboard) and a minimum of 18.3% (linen).

From the empirical results, we see that reflectance hashing provides reliable recognition even for a small training set. This result has important implications for real time material recognition since the approach may support online training with compact training sets.

Evaluation of Binary Codes We compare the performance of recent binary code methods for use in reflectance hashing. These codes are used to represent the response from the reflectance-layout filters as described in Section 3.4. We test the following binary embedding methods: randomized (CBE-rand) version of circulant binary embedding (CBE) [49], randomized (BITQ-rand) version of bilinear embedding [10], iterative quantization (ITQ) [12], angular quantization-based binary codes(AQBC) [11],

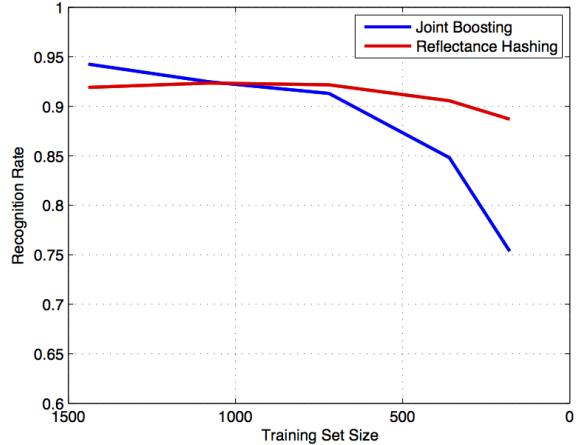


Figure 9: Recognition rate as a function of training set size for both boosting and reflectance hashing. Notice the performance of reflectance hashing is high even when the training set number is low.

spectral hashing (SH) [48], locality-sensitive binary codes from shift-invariant kernels (SKLSH) [32], and a baseline method locality sensitive hash (LSH) [3]. CBE and BITQ also have learned versions, but these are more appropriate for higher dimensional data and longer bit code. Recognition is accomplished using a nearest neighbor search with a Hamming distance metric and ten nearest neighbors. From Figure 11, we see the binary embedding recognition rate reaches around 90%, when using 128 or 256 bit binary codes. The method of ITQ gives the best results for this material recognition task.

True class	Inferred class	cardboard	CD	woven fabric	feather	textured rubber	glossy ceramic	glossy metal	leather	linen	matte metal	automotive paint	non-ferrous metal	paper	plastic	rock	rubber	sponge	velvet	glossy paper	wood
cardboard	99.6	0.1						0.2													
CD	99.8	0.2																			
woven fabric		89.9	0.2					3.3		0.2			0.1	4.5	0.4	1.0					0.4
feather	0.2		96.1					1.7		0.1		0.1					1.0	0.4	0.3		
textured rubber		0.1	82.1	0.2					0.9	0.8	0.2	4.3	1.9	8.0	0.8	0.7					
glossy ceramic			3.0	95.7	1.0				0.4												
glossy metal					99.6	0.2	0.1														
leather	0.9		0.2			97.3	0.2					0.1	1.0								0.2
linen	0.1		4.5	0.5	1.2	0.2	2.9	3.8	66.7	2.0	3.7	2.5	1.1	0.7	6.2	0.7	0.7	2.3			
matte metal					0.6	0.1	0.1			98.4	0.2	0.5									
automotive paint									0.1	0.3	99.0	0.6									
non-ferrous metal	0.1						1.2	0.6	1.2	0.4	2.5	91.6	0.3	0.4	0.6					1.1	
paper					0.5					2.2		95.2	0.6	0.4		1.1					
plastic												99.7	0.3								
rock		4.8		0.2			2.0	1.7				84.0	0.4	1.8	2.0			3.1			
rubber				0.1								1.5		98.1	0.3						
sponge		2.0	2.6	2.2								4.3	0.4	82.3	4.4			1.8			
velvet		1.2	0.1	2.6	0.1		0.8		0.1		0.7		0.6	4.3	3.4	85.6	0.3			98.2	0.2
glossy paper		0.8			0.6							0.1									
wood	7.4		0.5	0.3			5.3				0.7		0.5		0.1					85.2	

Figure 10: Reflectance Hashing Classification. Confusion matrix with percentages row-normalized. Overall accuracy is 92.3%.

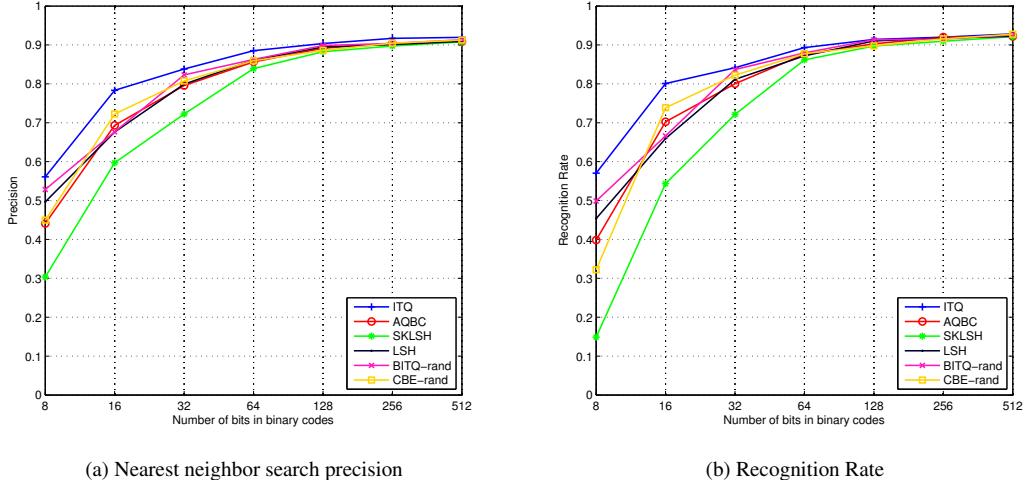


Figure 11: Precision and recognition rate as a function of the number of bits for binary embedding with 10 nearest neighbors.

5. Summary and Conclusion

We have introduced a novel framework for measurement and recognition of material classes. The approach encodes high dimensional reflectance with a compact binary code. We have compared several existing binary code methods to choose the most appropriate for this material recognition approach. The coding supports discrimination among the classes and can be realized in embedded vision implementations. Our method of reflectance hashing is compared to two popular baseline texton-based methods, boosting and texton histograms for material recognition. The results show

excellent performance even for small training sets and provide a novel method based on reflectance for fast sensing and recognition of real-world materials.

Acknowledgement

The authors would like to thank Felix Yu from lab of Shih-Fu Chang, Columbia University and Sanjiv Kumar from Google Research for supplying binary embedding codes. This work was supported by the National Science Foundation award IIS-1421134.

References

- [1] E. H. Adelson. On seeing stuff: the perception of materials by humans and machines. In *Photonics West 2001-Electronic Imaging*, pages 1–12. International Society for Optics and Photonics, 2001. 2
- [2] A. Bergamo, L. Torresani, and A. Fitzgibbon. Picodes: Learning a compact code for novel-category recognition. *Advances in Neural Information Processing Systems 24*, pages 2088–2096, 2011. 3
- [3] M. S. Charikar. Similarity estimation techniques from rounding algorithms. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 380–388. ACM, 2002. 6, 7
- [4] O. G. Cula and K. J. Dana. Recognition methods for 3d textured surfaces. *Proceedings of SPIE Conference on Human Vision and Electronic Imaging VI*, 4299:209–220, January 2001. 3
- [5] O. G. Cula and K. J. Dana. 3D texture recognition using bidirectional feature histograms. *International Journal of Computer Vision*, 59(1):33–60, August 2004. 2
- [6] K. Dana and J. Wang. Device for convenient measurement of spatially varying bidirectional reflectance. *Journal of the Optical Society of America A*, 21:pp. 1–12, January 2004. 2, 3, 4
- [7] K. J. Dana. Brdf/btf measurement device. *International Conference on Computer Vision*, 2:460–6, July 2001. 2, 3
- [8] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real world surfaces. *ACM Trans. Gr.*, 18(1):1–34, January 1999. 2
- [9] S.-C. Foo. A gonioreflectometer for measuring the bidirectional reflectance of material for use in illumination computation. Master’s thesis, Cornell University, 1996. 2
- [10] Y. Gong, S. Kumar, H. A. Rowley, and S. Lazebnik. Learning binary codes for high-dimensional data using bilinear projections. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 484–491, 2013. 6, 7
- [11] Y. Gong, S. Kumar, V. Verma, and S. Lazebnik. Angular quantization-based binary codes for fast similarity search. In *Advances in Neural Information Processing Systems*, pages 1196–1204, 2012. 6, 7
- [12] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(12):2916–2929, 2013. 6, 7
- [13] J. Gu and C. Liu. Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral brdf. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 797–804. IEEE, 2012. 2
- [14] J. Y. Han and K. Perlin. Measuring bidirectional texture reflectance with a kaleidoscope. *ACM Trans. Graph.*, 22(3):741–748, July 2003. 3
- [15] M. Johnson and E. Adelson. Retrographic sensing for the measurement of surface texture and shape. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1070–1077, 2009. 2
- [16] B. Kulis and T. Darrell. Learning to hash with binary reconstructive embeddings. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1042–1050, 2009. 3
- [17] R. Kumar, M. Jones, and T. Marks. Morphable reflectance fields for enhancing face recognition. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2606–2613, 2010. 2
- [18] L. Ladicky, J. Shi, and M. Pollefeys. Pulling things out of perspective. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 89–96, June 2014. 3, 5
- [19] T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textures. *International Journal of Computer Vision*, 43(1):29–44, 2001. 2, 3, 4
- [20] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH ’96*, pages 31–42, New York, NY, USA, 1996. ACM. 2
- [21] W. Li and M. Fritz. Recognizing materials from virtual examples. In *Computer Vision–ECCV 2012*, pages 345–358. Springer, 2012. 2
- [22] C. Liu and J. Gu. Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral brdf. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(1):86–98, 2014. 2
- [23] C. Liu, L. Sharan, E. Adelson, and R. Rosenthal. Exploring features in a bayesian framework for material recognition. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 239–246, 2010. 2
- [24] W. Liu, J. Wang, and S. fu Chang. Hashing with graphs. In *International Conference on Machine Learning (ICML)*, 2011. 3
- [25] W. Liu, J. Wang, R. Ji, Y.-G. Jiang, and S.-F. Chang. Supervised hashing with kernels. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2074–2081, 2012. 3
- [26] S. Lombardi and K. Nishino. Reflectance and Natural Illumination from A Single Image. In *European Conference on Computer Vision*, volume VI, pages 582–595, 2012. 2
- [27] S. Lombardi and K. Nishino. Single Image Multimaterial Estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 238–245, 2012. 2
- [28] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A Data-Driven Reflectance Model. *ACM Trans. on Graphics*, 22(3):759–769, 2003. 2
- [29] W. Matusik, H. Pfister, M. Brand, and L. McMillan. Efficient Isotropic BRDF Measurement. In *Proc. of Eurographics Symposium on Rendering*, 2003. 2
- [30] G. Oxholm and K. Nishino. Shape and reflectance from natural illumination. *European Conference on Computer Vision*, pages 528–541, 2012. 2
- [31] G. Oxholm and K. Nishino. Multiview shape and reflectance from natural illumination. *IEEE Conference on Computer Vision and Pattern Recognition*, 2014. 2
- [32] M. Raginsky and S. Lazebnik. Locality-sensitive binary codes from shift-invariant kernels. In *Advances in neural information processing systems*, pages 1509–1517, 2009. 6, 7

- [33] M. Raginsky and S. Lazebnik. Locality-sensitive binary codes from shift-invariant kernels. *Advances in Neural Information Processing Systems (NIPS)*, pages pp. 1509–1517, 2010. 3
- [34] L. Sharan, C. Liu, R. Rosenholtz, and E. Adelson. Recognizing materials using perceptually inspired features. *International Journal of Computer Vision*, 103(3):348–371, 2013. 2
- [35] J. Shotton, J. Winn, C. Rother, and A. Criminisi. Textonboost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation. In *Proceedings of the 9th European Conference on Computer Vision - Volume Part I, ECCV'06*, pages 1–15, Berlin, Heidelberg, 2006. Springer-Verlag. 3, 5
- [36] J. Shotton, J. Winn, C. Rother, and A. Criminisi. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *Int. J. Comput. Vision*, 81(1):2–23, Jan. 2009. 3, 5
- [37] A. Torralba, R. Fergus, and Y. Weiss. Small codes and large image databases for recognition. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, 2008. 3
- [38] A. Torralba, K. P. Murphy, and W. T. Freeman. Sharing visual features for multiclass and multiview object detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(5):854–869, 2007. 3, 5, 6
- [39] J. C. van Gemert, J.-M. Geusebroek, C. J. Veenman, and A. W. M. Smeulders. Kernel codebooks for scene categorization. In *ECCV 2008, PART III. LNCS*, pages 696–709. Springer, 2008. 5
- [40] M. Varma and A. Zisserman. A statistical approach to material classification using image patch exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):2032–2047, 2009. 3
- [41] J. Wang and K. J. Dana. A novel approach for texture shape recovery. *ICCV International Conference on Computer Vision*, pages 1374–1380, October 2003. 3
- [42] J. Wang and K. J. Dana. Hybrid textons: Modeling surfaces with reflectance and geometry. *IEEE CVPR Conference on Computer Vision and Pattern Recognition*, pages 372–378, 2004. 3
- [43] J. Wang and K. J. Dana. Relief texture from specularities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3):446–457, 2006. 3
- [44] J. Wang and S. fu Chang. Sequential projection learning for hashing with compact codes, 2010. 3
- [45] O. Wang, P. Gunawardane, S. Scher, and J. Davis. Material classification using brdf slices. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 2805–2811, 2009. 2
- [46] G. Ward. Measuring and modeling anisotropic reflection. In *ACM SIGGRAPH 92*, pages 265–272, 1992. 2
- [47] Y. Weiss, A. Torralba, and R. Fergus. Spectral hashing. *Advances in Neural Information Processing Systems (NIPS)*, pages 1753–1760, 2008. 3
- [48] Y. Weiss, A. Torralba, and R. Fergus. Spectral hashing. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1753–1760. Curran Associates, Inc., 2009. 6, 7
- [49] F. X. Yu, S. Kumar, Y. Gong, and S.-F. Chang. Circulant binary embedding. In *International Conference on Machine Learning*, 2014. 6, 7