Using Specularities for Recognition

Margarita Osadchy
NEC Labs
Princeton, NJ

David Jacobs
Dept. of Computer Science
The University of Maryland

Ravi Ramamoorthi
Dept. of Computer Science
Columbia University

Abstract

Recognition systems have generally treated specular highlights as noise. We show how to use these highlights as a positive source of information that improves recognition of shiny objects. This also enables us to recognize very challenging shiny transparent objects, such as wine glasses. Specifically, we show how to find highlights that are consistent with an hypothesized pose of an object of known 3D shape. We do this using only a qualitative description of highlight formation that is consistent with most models of specular reflection, so no specific knowledge of an object's reflectance properties is needed. We first present a method that finds highlights produced by a dominant compact light source, whose position is roughly known. We then show how to estimate the lighting automatically for objects whose reflection is part specular and part Lambertian. We demonstrate this method for two classes of objects. First, we show that specular information alone can suffice to identify objects with no Lambertian reflectance, such as transparent wine glasses. Second, we use our complete system to recognize shiny objects, such as pottery.

1 Introduction

Lighting variation has a large effect on the appearance of objects. Many recognition methods can account for this variability when objects are not shiny (eg., when they are *Lambertian*). When objects produce some specular highlight, these methods essentially treat them as noise; if the noise is minor recognition succeeds in spite of it.

We propose to use specular highlights as a source of information instead. We first focus on understanding what information specularity makes available for recognition. To isolate this key problem, we consider recognition of shiny transparent objects such as wine glasses (bottom, Figure 1). These are challenging objects to recognize, since their appearance depends almost entirely on the scene behind them. We show that given rough knowledge of lighting, approximate 3D shape and a hypothesis about the position of these objects, we can determine their identity using the shape of the specularities they produce. To our knowledge, no existing method could perform comparable identification.

Figure 1: Objects tested in the experiments. Top – objects for Lambertian plus Specular test; bottom – transparent specular objects

To do this, in Section 2 we propose a simple, qualitative reflectance model that captures the properties of existing models (eg., Torrance-Sparrow[7], Phong[23], Ward[30], ...) over any reasonable choice of parameters. Using this model, we show that we can efficiently search through a large set of candidate image highlights to find ones that are consistent with a hypothesized 3D object, provided that highlights are largely produced by a compact light source. This search also determines the sharpness of highlights (ie., the degree to which the material is mirror-like or more diffuse), so our model of specularity requires no previously known parameters.

Next, we show how to integrate knowledge about highlights with the Lambertian method of Basri and Jacobs [1]. This allows us to use highlights in recognizing objects that are part Lambertian, part specular, such as pottery (top of Figure 1). Using [1], we can determine lighting conditions from Lambertian reflectance. This allows us to remove the assumption that we have rough knowledge of lighting a priori, which we needed to identify glassware. Then, we can identify highlights consistent with a specific model, and refine the Lambertian results using knowledge of these highlights. One limitation of [1] is that it requires knowledge of

the Lambertian albedo of an object. We extend the method so that instead we can more simply use a reference image of the object, taken under unknown lighting conditions. Combining these steps, we have a measure of how well Lambertian and specular reflectance together can allow a known 3D shape to fit the image.

3D models seem important if one wants to account for specularities in recognition; image-based methods are difficult to apply since the location of specularities varies so dramatically with lighting. We can acquire 3D models using stereo or structured light systems, if shiny objects are first covered with powder or paint to reduce their shininess. This however results in noisy 3D models of shiny objects; consequently our approach is robust to noise in the model. We demonstrate this by using models built with commercial structured light systems.

We do not address the problem of pose in this paper. Since there has been so little work on recognition of specular objects, we feel that it is appropriate to focus on the photometric aspects of the problem, testing our algorithms with known pose. This approach has proven fruitful in much work on lighting that has first focused on identification with known pose, such as Basri and Jacobs[1], Georghides et al.[10], and Chen et al.[6]. Nonetheless, we feel that our approach to identification can fit into many existing approaches to solving for pose. With objects that have a Lambertian component, it is reasonable to suppose that some features can be identified and used to compute hypothetical poses (eg., the eyes and nose of a face), even though these features will not be adequate for complete recognition. We can also use an identification method as the basis for search. For example, to recognize transparent objects, we can use candidate highlights in the image to narrow the possible position of the object, then search for exact poses and objects that fit these highlights.

In summary, our main contribution is to show that using a very general model of specular reflection, we can efficiently judge the consistency of a known 3D model and possible specularities in the image. This can be used to recognize very challenging shiny objects, or integrated with previous approaches in the recognition of objects that are partly specular and partly Lambertian. In both cases, specularity becomes a positive clue for use in recognition.

1.1 Previous Work

One approach to accounting for illumination effects in recognition has been to build representations of the set of images that a 3D object can produce under a large range of lighting conditions. This can be done by condensing a large number of training images into a low-dimensional representation (eg., Hallinan[13], Epstien et al.[8], and Murase and Nayar[19]) or by using analytic insights to generalize from a small number of training images (eg., Shashua[28], Moses[18], Belhumeur and Kriegman[2]). Basri and Jacobs[1] and Ramamoorthi and Hanrahan[25] build on these results to show how to analytically derive a 9D representation of an object's image set from a 3D model. Ramamoorthi[24] shows that a lower-dimensional 5D or 6D subspace often suffices. These results form the starting point for our Lambertian method. [2, 1] provide further discussion of the large literature on this topic.

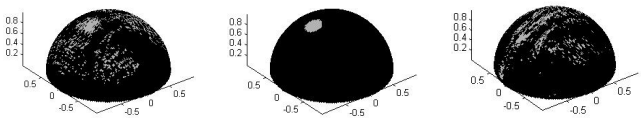


Figure 2: Mapping of the candidate image highlights onto the Gaussian sphere. Left: the candidate highlights are consistent with a hypothesized 3D object, consequently the true specularity forms a disk on a sphere. Center: the segmented specularity. Right: the candidates are not consistent with the 3D object – the points are spread everywhere and no sufficient disk can be segmented.

A second direction is to learn object appearance under different lighting with images rendered using its 3D structure (Brunelli[5], Sato et al.[26], Gremban and Ikeuchi[12]. A third approach is to use image representations that are not sensitive to lighting, such as gradient direction (Chen et al.[6]) or Gabor jets (Lades et al.[15]).

These methods have provided a great deal of insight, and have produced good experimental results for objects that are approximately Lambertian, such as faces. Few previous works have addressed the recognition of specular objects, which is the focus of our efforts. However, in Sato et al.[26] a physics-based simulator is used to predict specularities from which a set of aspects of the object is generated. For each specular aspect they prepare deformable matching templates. At runtime, an input image is first classified into candidate aspects, then the deformable templates are used to refine matching. The method proposed in Gremban et al.[12] uses multiple observations from different viewpoints to resolve an ambiguity in the recognition of specular objects. [26] gives an overview of earlier attempts at recognition of specular objects.

2 Identification Using Specularities

In this section we show how to use specularity to identify objects given a rough knowledge of lighting, 3D shape and position. We will explain this method using the example of highly specular objects, such as glass. Later we will show that when objects have Lambertian and specular reflectance, we can use Lambertian effects also, in particular automat-

ically determining the lighting direction sufficiently accurately to allow us to analyze specular highlights.

Our treatment of specularity has two stages. First we must locate pixels that are candidate highlights. This process need not be too accurate, since we refine these candidates in the second stage. Many different cues have been used to detect specularities, such as color, polarization, stereo and multiple views (see Oren and Nayar[22] for a survey). When the scene contains a dominant, compact light source, shiny objects such as glass produce highlights that are quite bright (Brelstaff and Blake[4]). For these objects it is sufficient to simply threshold the image based on intensity. This produces many false positive candidates that are rejected in the second stage of the method. We consider the problem of identifying specular candidates for less shiny objects in the next section.

Next, we determine which of these candidates are consistent with a known 3D object. We do this using a simple, qualitative model of specular reflection that captures the properties of many standard models. Let $\hat{v}$ be a vector giving the direction from the object to the viewer. Let $\hat{l}$ give the direction from the object to the center of a compact light source. We will assume that the object is reasonably distant from the viewer and the lighting, so that $\hat{v}$, $\hat{l}$ are constant. This assumption will introduce some noise, to which our algorithm will be robust. Next we define $\hat{n}_p$ to be the surface normal coplanar with $\hat{l}$ and $\hat{v}$, and halfway between them. That is, $\hat{n_p}$ is the unit vector in the direction of $\hat{l} + \hat{v}$. Our model first assumes that specular reflections are most intense at surface normals $\hat{n}_p$. We also assume that the specular reflection produced by a surface normal, $\hat{n}$, that is close to $\hat{n}_p$, only depends on the angle between $\hat{n}$ and $\hat{n}_p$. That is, the intensity of specular reflectance is rotationally symmetric about $\hat{n}_p$. Finally, we assume that reflectance will be a monotonically decreasing function of $\hat{n_p} \cdot \hat{n}$. This model holds strictly for a number of models of specularity, such as Phong [23] and Ward[30]. For other models, such as Torrance-Sparrow[7] and Beckmann-Spizzichino [3] it is true under the conditions listed in Ikeuchi and Sato[14], which usually hold. This model is helpful because it implies that if we threshold specularities based on intensity, the surface normals that produce specular points will form a disk on the Gaussian sphere.

Therefore, we proceed by selecting candidate specularities consistent with such a disk. We map each candidate specularity to the point on the sphere having the same surface normal (Figure 2 left). There we find a plane that separates the specular pixels from the other pixels with a minimal number of misclassifications. The plane will intersect the sphere in a disk. The plane $(\hat{w}, b)$, defined by $\hat{w} \cdot v - b = 0$ is found by minimizing (1)

$$\min_{\hat{w},b} \left\{ \sum_i \left(sign(\hat{w} \cdot n_i - b) - y(n_i) \right) \right\} \qquad (1)$$

$$y(n_i) = \begin{cases} +1 & n_i \text{ is specular} \\ -1 & \text{otherwise} \end{cases}$$

where $\{n_i\}$ are the surface normals. Finding a good linear separator is a well-studied problem. Since our problem is low dimensional, we solve it simply. We choose $\hat{w}$ from $\{n_i \mid y(n_i) = 1\}$, and for every $\hat{w}$ we search for an optimal b. We label all image pixels with normals in this disk as highlight points (Figure 2 center).

The thresholded, binary image may also have false positives that approximately map to a disk on the sphere. To reject these false positives we add a constraint that the specular disk should be consistent with the light source direction, which is roughly known. Since viewing direction is known, and $\hat{n_p}$ is at the center of the specular disk, the disk implies the lighting direction. In our experiments, it worked well to require that this implied direction lie within 10 degrees of the roughly known light source. In sum, this algorithm finds the maximal set of specular points consistent with our knowledge of the object's shape and of the lighting. We can use the degree to which an object can explain possible specular points in the image as a measure of how well it fits the image. In Section 4.1 we use this information to identify objects having no Lambertian reflectance.

3 Recognition of Hybrid Objects

In this section we show how to incorporate knowledge about highlights into a more general scheme for recognition of objects that have diffuse and specular components. In this method, we do not assume prior knowledge of lighting conditions. We also do not assume knowledge of how Lambertian and how specular the object is. It seems impossible to recognize a Lambertian object without some knowledge of its Lambertian albedo, since an object with unknown albedo might have any image painted on it. We first present our algorithm assuming we know the Lambertian albedo. Since this can be tedious to measure for shiny objects, we will later present a method that only requires a reference image of the object, taken in known pose but with unknown lighting.

We will describe each step of the algorithm in a separate subsection. Overall, the algorithm proceeds as follows:

- Find an approximation to the lighting using a purely Lambertian model. This also gives us an estimate of the luminance (light reaching the camera) that is due to purely Lambertian reflectance.

- Use this estimate of luminance to locate candidates for specular highlights.
- Find specularities using the method in Section 2.
- Re-estimate the lighting and Lambertian luminance excluding specularities.
- (Optional) At this stage, since we have separated Lambertian and specular reflectance in the image, if we desire we can recover the intensity profile of the specularity, and the Lambertian albedo within highlights.
- Finally, we compare the image to the specular and lambertian luminance produced by the resulting model and recovered lighting.

3.1 Recovering Lighting and Lambertian Appearance

In the first step of our algorithm, we use a Lambertian model to recover information about lighting in the scene. To do this, we apply the model proposed by Basri and Jacobs[1] and Ramamoorthi and Hanrahan[25]. We describe this for completeness, and because we will extend it later.

This model approximates the set of images an object produces under varying illumination by a 9D linear subspace that is computed from the 3D model. The dimensions of this subspace are low-degree polynomial functions that are spherical harmonics (MacRobert[17]) of the surface normals of the object in a specific position, scaled by albedo, ie.:

$$b_{nm}(x,y) = \lambda(x,y) h_{nm}(\theta(x,y), \phi(x,y)), \qquad (2)$$

where b_{nm} are basis functions in the image for ($0 \leq n \leq 2$, $-n \leq m \leq n$), each of which describes how the object will look under a different, low-frequency lighting condition. $\lambda(x,y)$ is the albedo at pixel (x,y) and h_{nm} is the spherical harmonic evaluated at the surface normal (θ, ϕ) corresponding to pixel (x,y). For example, b_{00} represents the appearance of the object when lighting intensity is constant from all directions. b_{10} describes the object under light that is brightest from above, and falls off in intensity with the cosine of its angle to the z axis. [1, 25] show that nine basis images accurately approximate all images of an object.

Given an image I, [1] seek a vector a that minimizes $\|Ba - I\|$. B denotes the basis images, arranged as a $p \times 9$ matrix, where p is the number of points in the image. Every column of B contains one harmonic image b_{nm}, as per equation 2. The vector Ba corresponds to a rendered image that can be produced by the model with low frequency lighting and that best fits I.

After solving this linear minimization problem, we can derive the low frequency components of the lighting from the coefficients a (see [1] for details). As [25] point out, it is not possible to use a Lambertian object to accurately determine the high frequency components of lighting. In Section 3.3 we discuss how to use this low frequency information to determine the direction of a dominant, compact light source.

3.2 Finding Specular Candidates

Very shiny objects, like glass, produce bright specularities that we can find by simple thresholding. Lambertian plus specular objects may produce highlights that are less bright. Brighter pixels can arise from Lambertian shading, or from light albedos. To account for this, we consider $I_{diff} = I - Ba$, the difference between the real image and the image we render using Lambertian reflectance. This is the intensity that we cannot explain with a Lambertian model. Where this difference is due to specular highlights, the rendered image will be dimmer than the true image. We therefore threshold I_{diff}, designating the brightest 5% of the pixels as candidates for specular highlights. We set the threshold to get many false positives and fewer false negatives, since we don't want to miss the specular highlight. The next phase of the algorithm is capable of rejecting candidates that do not correspond to specular highlights.

3.3 Finding the Highlight

Using these candidate highlights, we can find the ones most consistent with an hypothesized model using the method in Section 2. The only variation for partly Lambertian objects is that we now insist that the lighting direction that we find be consistent with the computed, low-frequency lighting, rather than a known lighting direction. We estimate the light source direction with the vector v that minimizes $\|E - Nv\|$, where N is a $p \times 3$ matrix of the model's surface normals, and E is a rendering of these normals under low-frequency lighting.

3.4 Re-computing the Lambertian Part

In this step, we simply recompute the lighting (as described in Section 3.1 or in Section 3.7, Equation 3), excluding the pixels we have identified as specular. This improves our estimate of luminance from Lambertian reflectance.

3.5 Computing the Specular Profile

For the case of glossy objects, we can then recover the intensity profile of the specular highlight. If we allow this profile to be arbitrary, then we could perfectly fit the image by setting the profile to be I_{diff}. Alternately, we can

constrain the specular intensities at a surface normal to be a monotonically decreasing function of the angle between that normal and the normal at the specular peak. In practice we find that we can only recover the specular profile when our models are fairly accurate.

In experiments reported in this paper, we fit the specular profile in a simple way. We divide I_{diff} into concentric rings so that each ring contains pixels with surface normals at roughly the same angle relative to the specular peak. Then we choose the value of each ring in the specular profile to be the mean of the corresponding values in I_{diff}. In our experiments, these values happen to obey the monotonicity constraint.

In general, this need not be the case. However, we can efficiently find the optimal specular profile that does obey the monotonicity constraint. To explain this, we index the pixels in I_{diff} as $I_{diff,i}$, arranging them in a 1D list. Let J_i denote the value assigned in the specular profile to the point $I_{diff,i}$. We order the pixels so that if $i < j$ then the monotonicity constraint requires $J_i \leq J_j$. Then, our goal is to find $J_1 ... J_n$ that minimizes $\Sigma(I_{diff,i} - J_i)^2$ subject to the constraint that $J_i \leq J_{i+1}$.

To do this, we first make an observation about the optimal solution. Suppose $J_1 ... J_n$ is optimal, and a set of consecutive values are equal, but different from adjacent ones, ie., $J_{k-1} \neq J_k = J_{k+1} = ... = J_l \neq J_{l+1}$. Then it must be the case that $J_k = \Sigma_{i=k}^{l} \frac{I_{diff,i}}{1+k-l}$. That is, a constant set of values in the specular profile must equal the mean of the values in I_{diff} to which they correspond. Otherwise, we could reduce the difference between I_{diff} and the specular profile by adjusting J_k to be closer to the mean. Given this, we can also note that given an optimal solution for the first k values of the specular profile, an optimal solution for the entire profile depends only on the value of J_k, and on the number of previous values that are identical. In light of these observations, it is straight-forward do construct a dynamic program to find an optimal solution.

3.6 Comparison

At this point, we have computed the lighting, luminance due to this lighting and Lambertian reflectance, and the position and intensity caused by specular highlights that are consistent with the lighting and object geometry. We identify an object by doing this for all models and measuring the residual error between the query image and the rendered models. The one that best fits the image results in the minimal SSD. For correct objects, we can fit the Lambertian and specular parts of the image they produce; for incorrect models incorrect specularity computation leads to extra errors, plus a failure to account for true specularities.

3.7 Handling Unknown Lambertian Albedos

When Lambertian albedo is unknown we can use a reference image of the object, aligned with the 3D model, taken under unknown lighting conditions to separately estimate illumination and albedo. Let I_1 and I_2 be the reference image and the query image. Let E represent the irradiance, the image that would be produced by the illumination if the object had uniform, white albedo. The image is the pointwise product of E and the albedo. Since E is dominated by low frequency components of lighting, we estimate it by accounting for as much of the image as possible with low frequency variation in lighting. Similar ideas are used in homomorphic filtering (Gonzalez and Woods[11]) methods and the Retinex theory of lightness constancy (Land and McCann[16]). Specifically, if I is the image (a column vector of p elements) and H is a $p \times 9$ matrix of the nine spherical harmonic basis vectors h_{nm} on the image, then by solving the least squares problem

$$\min_a \|Ha - I\|, \tag{3}$$

we find the coefficients a, which approximate the incident irradiance or illumination effects,

$$E \approx Ha. \tag{4}$$

This approximation will be exact when the albedo contains no low frequency components, except for a zero order, DC component. In practice, we find that we obtain good results for a variety of real objects in which this is not strictly true.

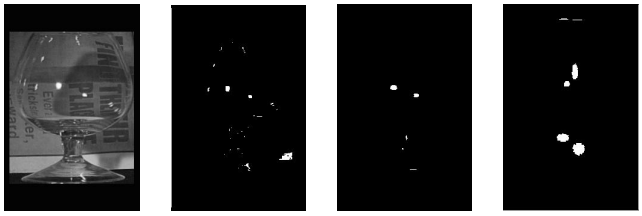
We estimate the low-frequency incident irradiance or illumination effects E_1 and E_2 separately for both images, as per equations 3 and 4. At each pixel, for both images we now know the effects of illumination. We pick an albedo λ to scale these intensities to minimize the sum of square error in both images.

$$\min_\lambda \|I_1 - \lambda E_1\| + \|I_2 - \lambda E_2\| \tag{5}$$

For pixel j, this is done by choosing λ_j so that:

$$\lambda_j = \frac{E_{1,j} I_{1,j} + E_{2,j} I_{2,j}}{(E_{1,j})^2 + (E_{2,j})^2}. \tag{6}$$

Now we have an initial approximation of albedo and irradiance functions in both images. One could use an iterative process where the current estimation of albedo is used for re-estimation of irradiance. However, we found that this iterative process didn't improve the initial estimates. We emphasize that we neither require nor recover accurate parametric models, since we only use them to tell whether the images might be consistent with the 3D model, not to render new images. In our empirical tests, we have seen that the initial estimates suffice for the purposes of recognition.



(a) (b) (c) (d)

Figure 3: Recognition of transparent specular objects. (a)-query image, (b)-specular candidates, (c)-specular highlights determined using the correct model, (d)-specular highlights determined using the incorrect model, which do not match the image.

4 Experiments

4.1 Purely Specular Objects

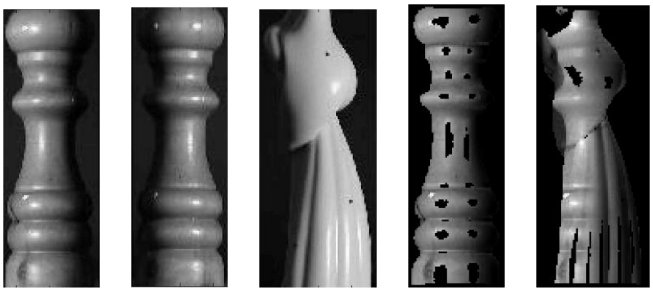
We use the method described in Section 2 to identify the transparent objects shown in Figure 1,bottom. Given a query image, I, we produce a binary image I_{bin} by thresholding I to find the brightest pixels, which we consider to be candidates for specular highlights. Since the object is transparent, two normals help produce every pixel in the query image: one for the front surface and one for the back, and either of the surfaces may produce specularity. Consequently, we map each candidate highlight to two points on the sphere. As a result, specular highlights map to a disk in the correct location, and also in another place on the sphere. The disk found is then mapped back and the resulting binary image $I^{'}_{bin}$ is matched to image candidates I_{bin} by computing the overlap between them:

$$overlapp = \frac{size(I_{bin} \wedge I^{'}_{bin})}{size(I_{bin} \vee I^{'}_{bin})} \quad (7)$$

The best match corresponds to the biggest overlap. If the image and the model belong to the same object, then the two methods will produce similar segmentations, otherwise the mapping method that uses the wrong model will detect erroneous specularity that doesn't match the segmentation produced by thresholding.

Figure 3 shows an example in which we attempt to explain the specular highlight in the query image (Figure 3a). The binary image in Figure 3b shows the candidates for specular highlights, the brightest image pixels. Figure 3c presents the correctly computed specularity ($I^{'}_{bin}$) using the disk model and the corresponding geometry. Figure 3d shows the results of the algorithm using the wrong model. The specularity in Figure 3d is inconsistent with the specular highlights in the query image.

We have tested the glass algorithm on nine objects. These form groups with similar shapes (Figure 1 bottom).



(a) (b) (c) (d) (e)

Figure 4: Recognition of glossy objects. From left to right: query image, reference image of correct model, reference image of wrong model, rendering using a correct model (with specularities cut out), rendering using a wrong model.

The test set is small because we see this as a preliminary result that shows the feasibility of the approach for such a challenging task. The shape and position of specular highlights alone probably does not provide enough information to distinguish among a large number of very similar shapes. The 3D models were computed from 2D images assuming rotational symmetry. The query images were taken under side illumination. All nine objects were recognized correctly using only consistency between the geometry and the specular highlights in a query image.

4.2 Lambertian plus Specular Objects

We have tested the algorithm for Lambertian plus specular surfaces on a database of 21 objects made from shiny ceramic and wood (Figure 1 top). Range images of these objects were obtained using a "Cyberware Model 3030 Rapid 3D Scanner". These models were somewhat noisy and inaccurate due the difficulty of scanning shiny objects. Four intensity images of each object were taken under different illuminations in a range of approximately 70 degrees from the frontal. The specular highlights were significant in all images. We have registered the brightness images with the range images using three feature points, which were enough since depth rotation wasn't present in the data. Since we didn't measure the albedos of the objects, we used a reference image to approximate the albedos as described in Section 3.7. We have performed two tests: one with the reference taken under frontal illumination and queries taken under 3 different side illuminations, the other with a side illumination of about 70 degrees as a reference image and three others as query images (a total of 63 query images in each experiment).

Figure 4 illustrates the algorithm on a specific example. If query image (Figure 4 a), reference image (Figure 4 b) and 3D structure belong to same object, the parameters are consistent and the rendered image (Figure 4 d) is similar to

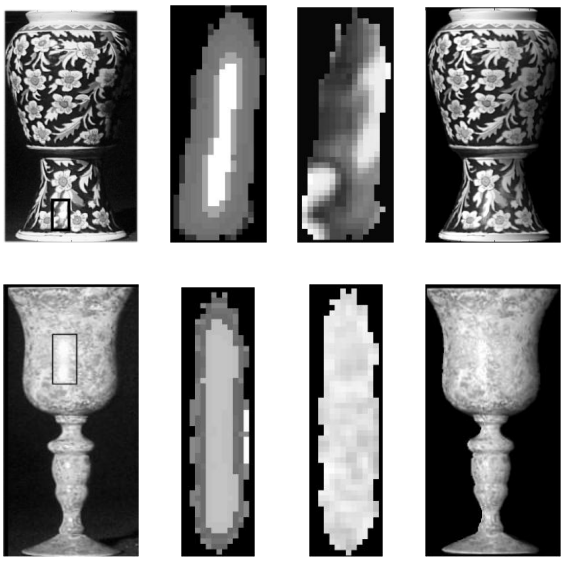


Figure 5: Nonuniform albedo objects with accurate models. From left to right column: query images, specular profiles, albedos inside specular highlights, rendered model. We zoomed in on the albedos and specular profiles at the areas marked by rectangles in the query images.

	frontal(%)	side(%)
Lambertian+Specular	100	95
Lambertian (Section 3.7)	98	94
9D subspace	76	79
Gradient direction	43	32
Correlation	52	30

Table 1: Pottery Recognition

the query image. However when the query image (Figure 4 c) and 3D structure belong to a different model, then the inconsistency shows up in all parameters: lighting, specularity and albedo, and the resulting rendering contains a mixture of two objects with incorrect lighting (Figure 4 e). The reason the result looks slightly like the correct object is that the computed albedos are a compromise between the reference image and the query image. Figure 5 shows two examples of objects with accurate models where we not only locate the highlights, but also recover the specular profile and the albedo inside the highlight.

We have compared the performance of the Lambertian plus specular (L+S) approach to different Lambertian methods that ignore specularities. First we tested the method for estimation of lighting and albedos described in Section 3.7 as a separate algorithm for recognition of Lambertian objects. The second test was a variation of the 9D linear subspace algorithm proposed by [1] in which we treat the reference image as giving the albedos of the object, texture mapping them onto its surface. We also compare to the direction of gradient method, which is proposed in Chen et al.[6] as an illumination insensitive image comparison. Finally, we compare to correlation. This is known to perform poorly under lighting changes, but it provides a baseline result. For these last two methods, no 3D information is used, and the query image is directly compared to the reference image. In all the experiments our algorithm was superior to other techniques, showing a 98% overall recognition rate. Table 1 shows the results of these tests. As apparent from the results, the difference in performance between L+S and the Lambertian method is very small. However, the performance of these methods is close to perfect, which leaves no room for improvement. We hope that on a larger database the difference will be larger, since the L+S algorithm is specifically designed to handle specular highlights. We conjecture that the gradient direction performed even worse than correlation because the objects are very smooth with specular highlights as the only features.

We've tested our method on glossy objects with a small albedo variation to demonstrate that we could handle objects that previous algorithms find especially difficult. However, some of the objects have nonuniform albedo patches and the algorithm works well for those objects too. Figure 5 shows that our approach can handle nonuniform albedo objects even when the specular highlight is not the brightest part of the image (Figure 5, top row).

5 Conclusions

This paper discusses how to use specular highlights as a positive clue for recognition of shiny objects. We have proposed a simple, qualitative reflectance model that captures the properties of existing models over any reasonable choice of parameters. Using this model,we can efficiently judge the consistency of a known 3D structure and possible specularities in the image. We have demonstrated that this information can be used to successfully identify very challenging objects with no Lambertian reflectance, such as glassware. We have also shown how to integrate knowledge about highlights with Lambertian methods, so that we can use them in recognizing objects that are part Lambertian, part specular. Experiments with glossy pottery show that using specular highlights as a source of information improves recognition.

6 Acknowledgments

The authors thank Technion for providing 3D acquisition equipment. We are also grateful to Daniel Osadchy for his help in database acquisition and Ronen Basri for fruitful discussions.

References

[1] R. Basri and D. Jacobs. "Lambertian reflectance and linear subspaces". In *Proc. of ICCV*, vol. II, pp. 283–390, 2001.

[2] P. N. Belhumeur and D. J.Kriegman,"What is the set of images of an object under all possible lighting conditions?" *IJCV*, 28(3): 245–260, 1998.

[3] P. Beckmann and A. Spizzochino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, New York: Pergamon, 1963.

[4] G.Brelstaff and A. Blake. "Detecting specular reflections using Lambertian constraints", Proc. of ICCV, pp. 297–302, 1988.

[5] R. Brunelli. "Estimation of Pose and Illuminant Direction for Face Processing", *Image and Vision Computing*, 15, pp. 741–748, 1997.

[6] H.F. Chen, P.N. Belhumeur, and D. W. Jacobs. "In search of illumination invariants". *IEEE Conf. on CVPR*, pp.254–261, June 2000.

[7] R. L. Cook and K.E. Torrance. "A reflectance model for computer graphics," *ACM Tranc. Graphics*, 1(1):7–24, January 1982.

[8] R. Epstein, P. Hallinan, A. Yuille. "5 ± 2 eigenimages suffice: an empirical investigation of low-dimensional lighting models", *IEEE Workshop on Physics-Based Vision*: 108-116, 1995.

[9] A.S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Generative Models for Recognition under variable pose and illumination". *IEEE Trans. on PAMI*.23(6):643–660, 2001

[10] A.S. Georghiades, D. J. Kriegman, and P. N. Belhumeur, "Illumination cones for recognition under variable lighting: Faces", Proc. *IEEE Conf. on CVPR.*, 1998.

[11] R.C. Gonzalez and R.E. Woods. "Digital Image Processing". Addison-Wesley, Reading, MA, 1992.

[12] K.D. Gremban and K. Ikeuchi. "Planning Multiple Observations for Specular Object Recognition", *IEEE Conf. on Robotics and Automation*, Vol. 2, May, 1993, pp. 599–604.

[13] P. Hallinan. "A low-dimensional representation of human faces for arbitrary lighting conditions", *IEEE Conf. on CVPR*: 995–999, 1994.

[14] K. Ikeuchi and K. Sato. "Determining reflectance properties of an object using range and brightness images", *IEEE Tran. PAMI*, Vol. 13, No. 11, November, 1991, pp. 1139–1153.

[15] M. Lades, J. C. Vortbrggen, J. Buhmann J. Lange, C. von der Malsburg, R. P. Wrtz, and W. Konen. "Distortion Invariant Object Recognition in the Dynamic Link Architecture", *IEEE Transactions on Computers*, 42:300–311, 1993.

[16] E. Land and J. McCann. "Lightness and retinex theory", *Journal of the Optical Society of America*,61(1):1–11, 1971.

[17] T. MacRobert. *Spherical harmonics; an elementary treatise on harmonic functions, with applications, Dover Publications*, 1948.

[18] Y. Moses. *Face recognition: generalization to novel images*, Ph.D. Thesis, Weizmann Institute of Science, 1993.

[19] H. Murase and S. K. Nayar. "Visual learning and recognition of 3d objects from appearance". *IJCV*, 14(1):5–24, 1995.

[20] S.K. Nayar, K. Ikeuchi, T. Kanade. "Surface Reflection: Physical and Geometrical Perspectives", *IEEE Tran. PAMI*,13(7):611–634,1991.

[21] K. Nishino, Z. Zhang, and K. Ikeuchi. "Determining Reflectance Parameters and Illumination Distribution from a Sparse Set of Images for View-dependent Image Synthesis", *ICCV01*, I:599-606, 2001.

[22] M. Oren, S.K. Nayar. "A theory of specular surface geometry", *IJCV*, 24(2):105–124, 1997.

[23] B. Phong, "Illumination for computer generated images", *Comm. ACM* 18, 6(June 1975) 311–317.

[24] R. Ramamoorthi, "Analytic PCA construction for theoretical analysis of lighting variability in images of a lambertian object", *IEEE Tran. PAMI*, 24(10):1322–1333, 2002.

[25] R. Ramamoorthi and P. Hanrahan. "On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object", *JOSA A*, 18(10):2448–2459, 2001.

[26] K. Sato, K. Ikeuchi, and T. Kanade."Model-Based Recognition of Specular Objects Using Sensor Models", *Proc. IEEE Workshop on Directions in Automated CAD-Based Vision*, June, 1991, pp. 2–10.

[27] Y. Sato, M.D. Wheeler, and K. Ikeuchi. "Object Shape and Reflectance Modeling from Observation", *SIGGRAPH 97* pp. 379–388, 1997.

[28] A.Shashua, *Geometry and Photometry in 3D Visual Recognition*, PhD thesis, MIT, 1992.

[29] P. Viola and W. Wells, "Alignment by maximization of mutual information," *Int. J. of Comp. Vis.*, 24(2):137-154, 1997.

[30] G. J. Ward. "Measuring and modeling anisotropic reflection," *SIGGRAPH 92*, 26(2):265–272, 1992. G. J. Ward. "Measuring and modeling anisotropic reflection," *SIGGRAPH 92*, 26(2):265–272, 1992.