A Multiple-View Geometric Model of Specularities on Non-Planar Shapes with Application to Dynamic Retexturing

Alexandre Morgand, Mohamed Tamaazousti and Adrien Bartoli

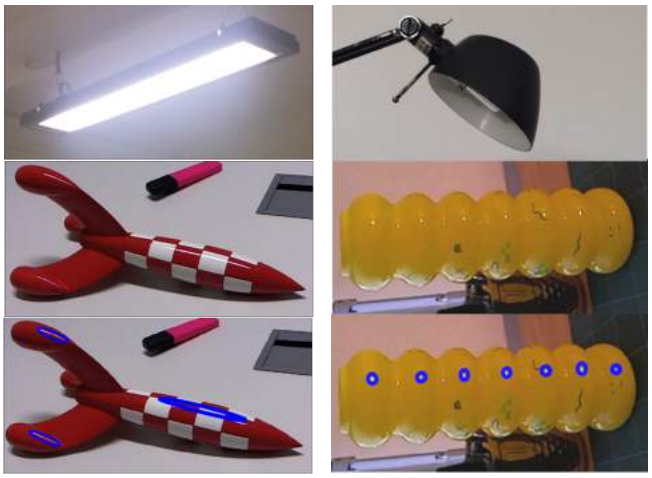
Abstract—Predicting specularities in images, given the camera pose and scene geometry from SLAM, forms a challenging and open problem. It is nonetheless essential in several applications such as retexturing. A recent geometric model called JOLIMAS partially answers this problem, under the assumptions that the specularities are elliptical and the scene is planar. JOLIMAS models a moving specularity as the image of a fixed 3D quadric. We propose dual JOLIMAS, a new model which raises the planarity assumption. It uses the fact that specularities remain elliptical on convex surfaces and that every surface can be divided in convex parts. The geometry of dual JOLIMAS then uses a 3D quadric per convex surface part and light source, and predicts the specularities by a means of virtual cameras, allowing it to cope with surface's unflatness. We assessed the efficiency and precision of dual JOLIMAS on multiple synthetic and real videos with various objects and lighting conditions. We give results of a retexturing application. Further results are presented as supplementary video material.

Index Terms—Specularity Prediction, Augmented Reality, Retexturing, Quadric, Multiple Light Sources.

1 Introduction

SLAM (Simultaneous Localization And Mapping) has matured recently. It facilitates practical industrial applications, with numerous methods publicly available, providing real-time camera localization [11, 37] and 3D reconstruction [31, 40]. Most SLAM approaches assume the scene to be Lambertian (matte surfaces), a condition which is not verified in a scene with specular materials such as metal, plastic and porcelain. Due to a mirror-like interaction between light sources and specular materials, specularities appear, adding drastic changes in the image intensity. As noted by Blake [3], specularities are used in human scene understanding to perceive the shape and type of material. In computer vision, specularities play an important role and are the subject of numerous studies [1, 9, 13, 18]. They may be discarded, but can also be used as visual cues. On the one hand, they may be used to improve camera/object localization [8, 21, 30] or 3D reconstruction [41]. On the other hand, they can drastically improve the perceptual quality in augmented and diminished reality [33, 43]. There is thus a need to predict the specularities in several applications. This is a challenging problem, even in a context where the camera pose and scene geometry are known. The visual appearance of a specularity depends on various physical properties of the scene: the geometry, the materials, the light sources and the camera. Knowing the geometry and camera pose, the visual appearance of a specularity may in some cases be equally well-explained by the light intensity and the material properties (reflectance, roughness). This ambiguity makes the problem of specularity prediction particularly challenging.

Existing methods [15, 38] do not express specularities explicitly but include them in a process of lighting condition reconstruction. These methods require one to compute numerous parameters on the materials and light sources, and do not predict the specularity on new viewpoints. More generally, existing methods fall in two categories: global illumination and light source reconstruction. However, there exists a third category of approaches where specularity prediction is cast as a multiple-view reconstruction problem. Recent works [28, 29] showed that a specularity on a planar surface can be well approximated by an ellipse under a light bulb or a fluorescent lamp illumination. A model called JOint LIght-MAterial Specularity (JOLIMAS) was proposed that abstracts the light-matter interaction and treats the problem with geometry. This model uses a fixed 3D quadric whose projection predicts the specularity's shape on existing and new viewpoints. The 3D quadric reconstruction is achieved from at least three viewpoints. We refer to this model as *primal JOLIMAS*.



(a) (b)

Figure 1: Specularity prediction on two non-planar objects. With the hypothesis of elliptical specularities on planar and convex surfaces, we reconstructed such a 3D quadric whose perspective projection fits the specularities in new viewpoints. A 3D quadric is created for each convex surface and each light source. (a) real sequence of a rocket replica (middle) with one fluorescent lamp (top). Our model is able to predict specularities (bottom) on different parts of the rocket for new viewpoints (shown as blue ellipses). In (b), we use a light bulb (top) to illuminate a yellow glass vase (middle).

We bring three main contributions to the geometric modeling of specularities. Firstly, we study the main limitation of *primal JOLIMAS*, which is its inability to work on non-planar surfaces, in section 2. Secondly, we address this limitation with *dual JOLIMAS* in section 3. This new model generalizes primal JOLIMAS to piecewise smooth surfaces while keeping its functionality on planar surfaces. The key

- *Alexandre Morgand is with CEA,LIST, Gif-sur-Yvette, France and IP-UMR 6602 - CNRS/UCA/CHU, Clermont-Ferrand, France. E-mail: alexandre.morgand@cea.fr.*
- *Mohamed Tamaazousti is with CEA, LIST. E-mail: mohamed.tamaazousti@cea.fr.*
- *Adrien Bartoli is with IP. E-mail: adrien.bartoli@gmail.com.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

idea is that specularities remain elliptical on convex surfaces. Dual JOLIMAS is based on a virtual camera representation which associates a virtual camera to each specularity with known camera pose and surface geometry. Since every surface can be divided into convex pieces, we can model and predict specularities on any piecewise smooth surface as shown in figure 1. In practice, we represent a surface by a mesh and use approximate convex decomposition, as explained in section 4. We then reconstruct a 3D quadric for each convex part and light source. We give the details of our pipeline for specularity prediction and how we compute the virtual camera representation for new viewpoints in section 5. The efficiency and precision of our method are validated on multiple synthetic and real sequences with various objects and lighting conditions in section 6. Finally, we use geometric specularity prediction to achieve dynamic retexturing, as described in section 7.

2 Related Work

Most approaches to specularity prediction explicitly model the light sources or lighting conditions for the scene and render a synthetic specularity afterwards. These methods are split in two categories: global illumination rendering and light source estimation.

Global illumination rendering. Methods in this category favor the quality of rendering by solving the rendering equation [14,16]. This equation describes the total amount of light emitted from a point $\mathbf{P}$ along a particular viewing direction, given a function for the incoming light and a BRDF. These approaches do not generally compute the physical attributes of the light sources. For instance, [15] captures a 4D light field over a small planar specular surface. By reconstructing the diffuse and specular components, it achieves a convincing rendering. However, it is unable to predict the specular component for viewpoints unknown during the initial reconstruction. Moreover, light sources with changing states (on/off) are not handled. Recently, [38] extended [15] by adding material segmentation for complex surfaces reconstructed using an RGB-D sensor, but shared the same limitations as [15]. As a consequence, [15, 38] and similar approaches such as [25] cannot predict specular reflections for new viewpoints.

Light source estimation. We distinguish two categories of light sources: directional sources and point sources. In an outdoor context, a directional source is often assumed but can also provide interesting results indoor. For instance, [20] uses specularities to compute directional sources with a moving object observed from a fixed viewpoint. However, specularity prediction requires information on the shape, intensity and position of the light source as well as the material properties. Point light source reconstruction methods such as [6,7,10,17,42] experience the same issues. Moreover, extended sources such as fluorescent lamps are not modeled by point light sources, which limits applicability.

Geometric modeling. Considering the specularity as a geometric cue in the scene has been little studied in the literature. [4] showed that specularity movement is linked to surface curvature. In the context of known camera poses and known geometry, the specularity movement can then be predicted. However, no information is given on the shape transformation due to curvature changes. *Primal JOLIMAS* [28] showed that for planar glossy surfaces, a specularity created from a light bulb or fluorescent lamp is elliptical. By reconstructing a fixed 3D quadric from several ellipses, this method predicts the specular shape for new viewpoints by simple perspective projection of the 3D quadric. It abstracts light and material interactions, and is used for retexturing. However, it only handles planar surfaces, which drastically limits its applicability. Primal JOLIMAS uses the following hypotheses:

1. Specular reflections are elliptical on planar surfaces, as in the Phong [32] and Blinn-Phong [5] models
2. A light source is associated with a single specularity on planar surfaces
3. There is a unique fixed 3D quadric located 'under' the planar surface whose perspective projection fits the specular shape in the image

Primal JOLIMAS is closely related to the field of SfM (Structure-from-Motion) using mirror reflections such as [19, 26, 34]. In multiple-view geometry, we distinguish several cases to reconstruct a static object using perspective projection. The simplest case is when an object is directly observed in a scene. Its image corresponds to its perspective projection. For an object observed through a perfect planar mirror, its image corresponds to the perspective projection of the symmetric of the object about the mirror's normal. For a point light source, its intensity combined with the sensitivity of the sensor causes the image of the point light source to not correspond exactly to the perspective projection of the light source. In the case of a specularity, in addition to the intensity of the light source and the sensor sensitivity, the image of a point light source observed through a shiny surface (mirror-like behavior) is affected by the surface's material (reflectance and roughness). This image thus does not correspond exactly to the perspective projection of the symmetric of the light source according to the normal of the surface. This image is called a specularity. This mismatch of shape seems to be represented in the scale change of the light source. As shown in [28], this transformation in scale seems to be constant along every viewpoint for planar surfaces. [28] has proposed to match the specularity shape by projecting a fixed 3D quadric. This fails if the surface is non-planar because the 3D quadric is not fixed in space (ellipses are not epipolar consistent, as illustrated in figure 2). The problem is more complex because the reflected image of the light source (specularity) through a curved surface (mirror or specular surface) can be drastically distorted. To solve this issue, the distortion of the specularity should be included in the model.

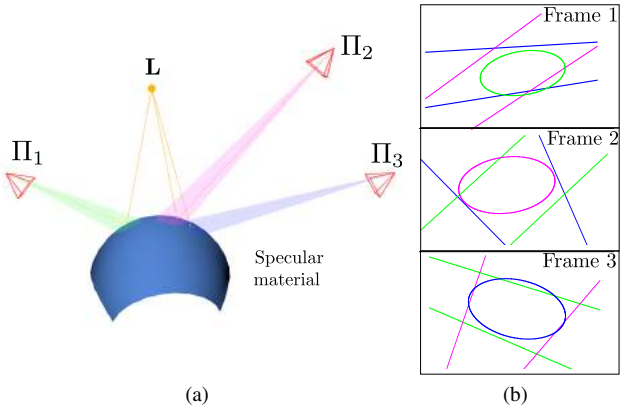


Figure 2: Epipolar geometry of the ellipses for three camera poses Π_1, Π_2 and Π_3 and a point light source $\mathbf{L}$ using primal JOLIMAS on a sphere. The scene is showed in (a) and the associated ellipses and epipolar lines in (b). The epipolar geometry is not respected. This results in an incorrect 3D quadric reconstruction and incorrect specularity prediction.

Proposed model. Our approach extends [28] to non-planar surfaces, while maintaining the elliptical specularity hypothesis even for extended light sources. We address the issues of primal JOLIMAS which fails to reconstruct a fixed 3D quadric in space on non-planar objects in our new model *dual JOLIMAS*. We use virtual cameras obtained by computing the symmetric of the real cameras about the tangent plane of particular points on the surface. As opposed to [28], our 3D quadric reconstruction is located near the real light source, which is fixed, rather than its symmetric.

3 Dual JOLIMAS

3.1 Notation and Formalization

A typical 3D scene contains a camera of pose $\Pi = \begin{bmatrix} \mathtt{R} & \mathbf{V} \end{bmatrix}$ and a light source $\mathbf{L}$. A surface point is written $\mathbf{P}$ with its normal $\hat{\mathbf{N}}(\mathbf{P})$. Defining the normalization operator $\mu(\mathbf{A}) = \frac{\mathbf{A}}{\|\mathbf{A}\|}$, we use $\hat{\mathbf{L}}$ as the

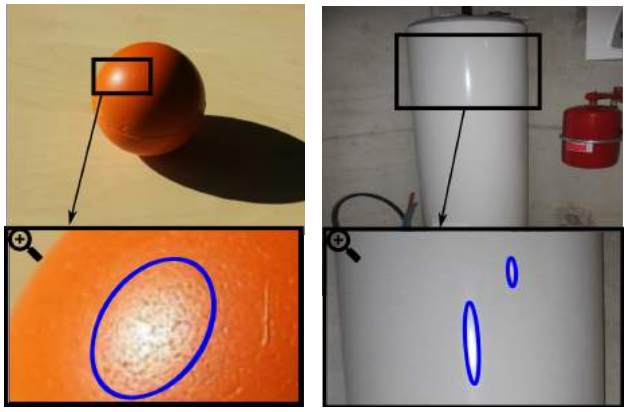
normalized vector of the incident light ray from the light source $\mathbf{L}$ such that $\hat{\mathbf{L}}(\mathbf{P}) = \mu(\mathbf{L}-\mathbf{P})$ and $\hat{\mathbf{V}}$ as the normalized vector of the reflected ray from the camera $\mathbf{V}$ such that $\hat{\mathbf{V}}(\mathbf{P}) = \mu(\mathbf{V}-\mathbf{P})$. The tangent plane $\mathbf{P}$ is defined by $\pi_{\mathbf{P}} = \begin{bmatrix} \hat{\mathbf{N}}(\mathbf{P}) \\ -d \end{bmatrix}$ with d the distance of the plane to the origin. Every virtual entity computed by symmetry is noted with $\sim$, such as the virtual camera position $\tilde{\mathbf{V}}$ and the virtual pose $\tilde{\Pi}$.

3.2 Principle

For a planar surface, we can see the specularity as the mirrored image of the light source. However, for curved surfaces, this mirrored image is distorted and incoherent with the actual light source position and shape. From the specularities, we cannot reconstruct a consistent 3D object located near the light source. The main idea of dual JOLIMAS is to use virtual cameras to reconstruct a 3D quadric coherent with the actual light source. These virtual cameras are computed by symmetry of the real cameras about the surface. As opposed to planar surfaces, the normal varies across curved surfaces, making the choice of the virtual camera to use ambiguous. For a given pose and specularity, which point of the surface should we choose to compute the virtual camera?

3.3 The Importance of the Brightest Point

On convex surfaces, several properties can be highlighted. Firstly, a specularity is elliptical, as seen in figure 3. Secondly, a light source is associated with a unique specularity. Finally, from a planar surface to a convex surface, a specularity is only affected by the scale of the ellipse.



(a) (b)

Figure 3: Specularities on convex surfaces such as a foam ball (a) and a water heater (b). The hypothesis of elliptical specularity still holds on these surfaces, as seen on the ellipse fitted on the close-up of (a) and (b). For each convex surface, there is a light source for each specularity (one light source in (a) and two light sources in (b)).

Several studies such as [13, 27, 39] highlights the importance of a specific point called *brightest point* which is unique within a specularity. In any direction around this point, the intensity of the specularity decreases. This point is very important in optics because for a fixed camera pose and light source position, it follows the law of reflexion:

$$\hat{\mathbf{L}}(\mathbf{P}) = 2(\hat{\mathbf{N}}(\mathbf{P})^\top \hat{\mathbf{V}}(\mathbf{P}))\hat{\mathbf{N}}(\mathbf{P}) - \hat{\mathbf{V}}(\mathbf{P}). \quad (1)$$

This implies that the brightest point is not affected by the distortion induced by the surface curvature. If we compute the symmetric of the camera about the tangent plane at this point, we obtain a virtual camera which points directly at the real light source through the specularity, as illustrated in figure 4. By computing a virtual camera for each specularity, we can thus reconstruct a fixed 3D quadric located near the real light source whose perspective projection fits the specularities' shape.

In the supplementary material, we show that the brightest point under Phong's model is unique and located at the intersection of the line containing the light source $\mathbf{L}$ and the symmetric position $\tilde{\mathbf{V}}$ of the camera for planar surfaces. For convex surfaces, this property holds if the point of the surface used to compute $\tilde{\mathbf{V}}$ is the same as the intersection point, as illustrated in figure 4. In practice, the brightest

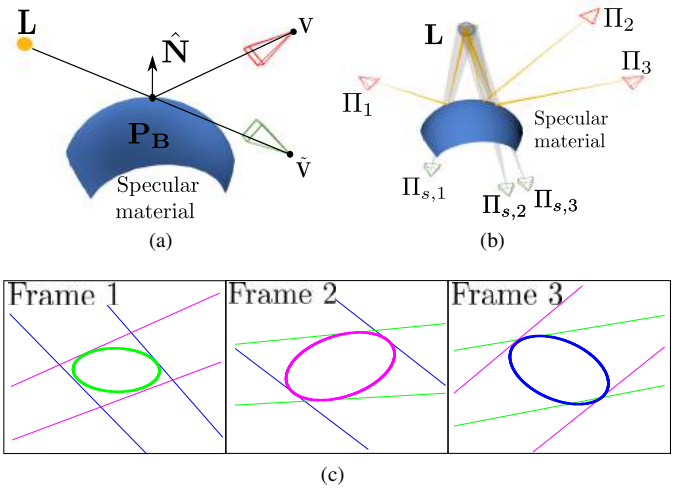


Figure 4: Dual JOLIMAS uses virtual cameras, producing epipolar consistent ellipses. In (a), we show the law of reflection on a convex surface. The brightest point $\mathbf{P_B}$ is associated to the point on the surface where the intersection of the line containing the virtual camera position $\tilde{\mathbf{V}}$ (in green) and the light source position $\mathbf{L}$ is the same as the point of the surface used to compute the virtual camera. In (b), we show dual JOLIMAS on a convex surface for three given camera poses Π_1, Π_2 and Π_3. The virtual camera poses (in green) allow us to reconstruct a 3D quadric around the light source $\mathbf{L}$. In (c), we show that our virtual representation allows the ellipses to respect the epipolar geometry.

point computation needs to be accurate because the virtual cameras can be sensitive to normal errors. We give a robust brightest point computation method from geometric priors on the surface mesh in section 5.2.

3.4 Model Computation

We compute dual JOLIMAS from the virtual cameras depending on the brightest point. As an initial guess of the brightest point $\mathbf{P_B}$, from our initial specularity detection, we take the gravity center of the specularity contour. In practice, the brightest point is not unique in the image since specularities often cause the camera to saturate. Following the symmetry used in [24], for instance, we compute a virtual camera pose $\tilde{\Pi}$ from the real camera Π as:

$$\tilde{\Pi} = \Pi \mathtt{S} \text{ with } \mathtt{S} = \begin{bmatrix} \mathtt{I}_3 - 2\hat{\mathbf{N}}(\mathbf{P})\hat{\mathbf{N}}(\mathbf{P})^\top & \mathbf{0} \\ -2d\hat{\mathbf{N}}(\mathbf{P}) & 1 \end{bmatrix}^\top. \quad (2)$$

For a single light source, the model estimation pipeline is described in figure 5. An example of our virtual representation is presented in figure 4(b). For multiple light sources, the pipeline is simply repeated.

4 Piecewise-Convex Surfaces

4.1 Principle

For each light source, we have a unique brightest point and more generally a unique specularity per convex surface. This hypothesis can be associated to a mirror behavior of convex shapes. In this mirror, the reflexion of a given object will be distorted but the image will be unique. We can represent an object's 3D model as a combination of convex surfaces. In practice, we reconstruct a 3D quadric for each convex piece and each light source.

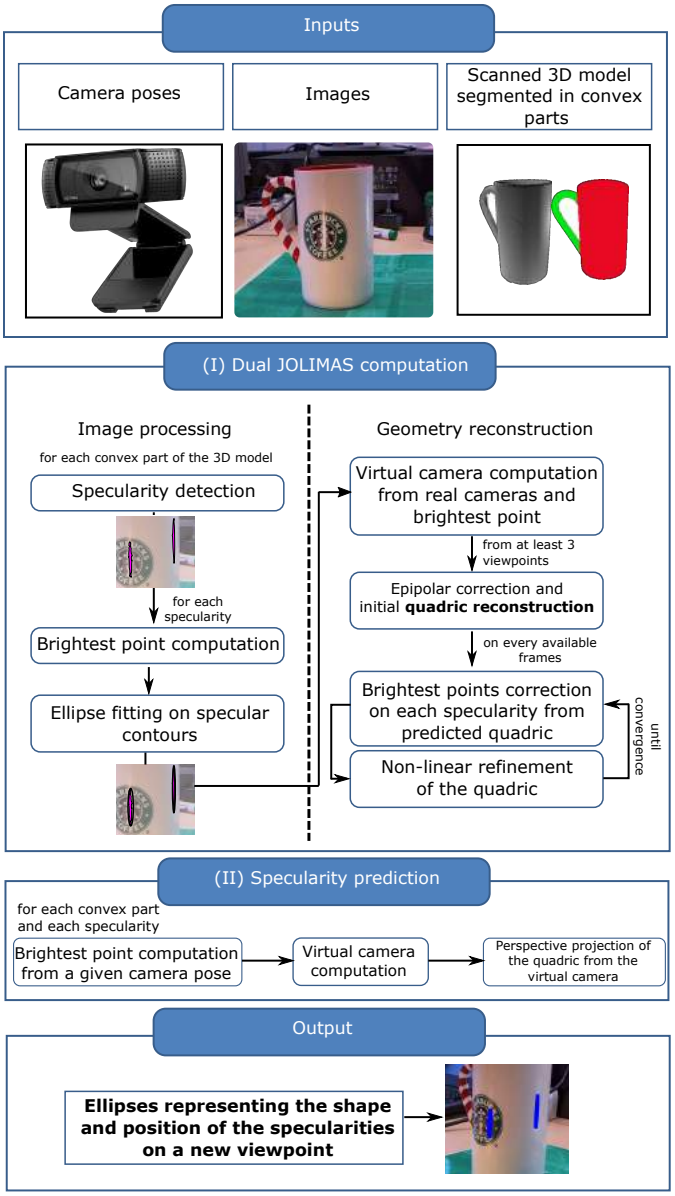


Figure 5: Pipeline of the method describing the inputs, the computation of the model and the specularity prediction process producing an ellipse. This ellipse describes the shape and the position of the specularity. This geometric representation abstracts parameters of the material (reflectance and roughness) and the light source (intensity, position, shape, color).

4.2 Approximate Convex Decomposition

According to Lien *et al.* [22], every mesh can be fairly approximated as the union of convex surfaces. Approximate convex decomposition is often used in physical engines to speed up collision detection. Our goal is to determine a convex decomposition of a surface S, which consists in partitioning it into a minimal set of convex sub-surfaces. This process aims to determine a partition of the mesh triangles with a minimal number of clusters, while ensuring that each cluster has a concavity lower than a user defined threshold. Exact convex decomposition is NP-hard and would be impractical since it may produce an overly large number of clusters. We then use an approximate convex decomposition library called V-HACD[1] to segment our mesh into several convex surfaces. Using this decomposition, our model is generalized to non-planar objects by reconstructing a 3D quadric for each convex surface. We illustrate this decomposition in figures 6(c) and 6(d). The meshes used for the decomposition are reconstructed using the HandySCAN 3D scanner from Creaform[2] as shown in figures 6(a) and 6(b).

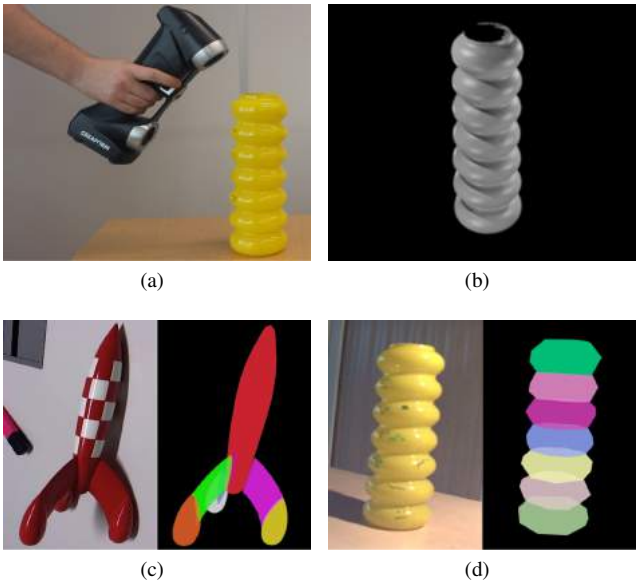


Figure 6: Object scanning and approximate convex decomposition. In (a), we use the HandySCAN 3D on a vase and show the reconstructed mesh in (b). For the two meshes scanned from real objects presented in figure 1, we show results of the approximate convex decomposition algorithm. The rocket replica mesh (a) was divided into 7 convex surfaces, so was the yellow vase . We have one specularity per convex piece and per light source, and a 3D quadric was reconstructed for each.

5 Specularity Prediction for New Viewpoints

5.1 Principle

Specularity prediction is achieved in primal JOLIMAS by projecting the resulting 3D quadric in a new viewpoint giving the ellipse fitting the new specularity. However, in dual JOLIMAS, for a new viewpoint, the position of the brightest point is required to compute the virtual camera. This virtual camera will then be used to project the reconstructed 3D quadric adequately. As an initial guess, we approximate this brightest point by considering the center of the 3D quadric as our point light source $\mathbf{L}$ position. This point is then projected according to the virtual camera pose on the input mesh.

5.2 Brightest Point Computation

According to equation (1), a surface point $\mathbf{P}$ is considered as the brightest point if $\mu(\tilde{\mathbf{V}}-\mathbf{P})$ and $\mu(\mathbf{L}-\mathbf{P})$ are collinear. This means that the surface point $\mathbf{P}$ used to compute the virtual camera must be identical to the intersection point of the line containing the virtual camera $\tilde{\mathbf{V}}$ and the light source $\mathbf{L}$ with the tangent plane $\pi_{\mathbf{P}}$ of S at $\mathbf{P}$. This property is illustrated in figure 7(a). The intersection point $\mathbf{P}_I$ is defined by:

$$\mathbf{P}_I = \tilde{\mathbf{V}} + \frac{\tilde{\mathbf{V}}^\top \hat{\mathbf{N}}(\mathbf{P}) - d}{(\mathbf{L}-\tilde{\mathbf{V}})^\top \hat{\mathbf{N}}(\mathbf{P})}(\mathbf{L}-\tilde{\mathbf{V}}) \tag{3}$$

with

$$\tilde{\mathbf{V}} = \begin{bmatrix} \mathrm{I}_3 - 2\hat{\mathbf{N}}(\mathbf{P})\hat{\mathbf{N}}(\mathbf{P})^\top & -2d\hat{\mathbf{N}}(\mathbf{P}) \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ 1 \end{bmatrix},$$

and δ the 3D euclidean distance. The brightest point computation must be done for each convex surface and each light source. Our estimation of the brightest point is described in the algorithm 1.

[1] github.com/kmammou/v-hacd

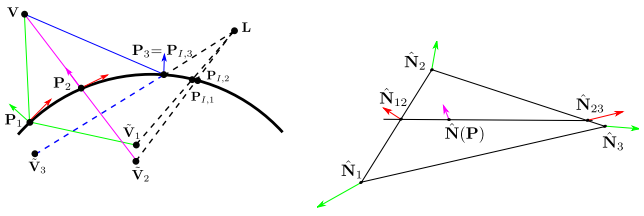
[2] www.creaform3d.com/en

Algorithm 1 Simplified brightest point computation from the camera pose, the CAD model and the light source position computer from the center of the 3D quadric.

```
procedure BRIGHTESTPOINTCOMPUTATION(Pose, Model, L)
    BP ← initial_BP
    do
        VirtualPose ← compute_virtual_pose(Pose, BP)
        P ← model_intersection(L, VirtualPose, Model)
        search_direction ← (P−BP) / ||P−BP||
        BP ← BP + search_direction
    while BP ≠ P
return BP
```

5.3 Normal Interpolation

To compute coherent virtual cameras, an accurate normal information is important. On a mesh, this information is only provided by the edges and the faces. To find the most accurate brightest point, we need to retrieve the normal variation accurately within the different mesh polygons. To answer this issue, we use a normal interpolation process between the different mesh polygons which is similar to the normal interpolation of Phong using [32]. We start from a given point $\mathbf{P}$ within a mesh polygon of unknown normal $\hat{\mathbf{N}}(\mathbf{P})$. After computing the normals $\hat{\mathbf{N}}_i$ for every vertices i of the triangle, we draw a line including $\mathbf{P}$ in a chosen direction (the $\mathbf{y}$ axis for instance) which will intersect several edges of the polygon. For these intersection points, we compute their normal by linear interpolation. The final normal $\hat{\mathbf{N}}(\mathbf{P})$ is computed from the linear interpolation of the intersection points' normals along the line, as illustrated in figure 7(b).



(a) (b)

Figure 7: Brightest point computation and normal interpolation. In (a), our brightest point computation algorithm is illustrated. From a rough initialization $\mathbf{P}_1$, we compute the virtual camera $\tilde{\mathbf{V}}_1$ and the line containing $\tilde{\mathbf{V}}_1$ and $\mathbf{L}$. Since the intersection $\mathbf{P}_{I,1}$ between the line and the surface S is not equal to $\mathbf{P}_1$, we iterate to the next point on the surface in the direction of the intersection point. This process stops when the distance between $\mathbf{P}$ and $\mathbf{P}_I$ is minimal. In order to retrieve the normal information on any point of the surface, we compute a precise normal using normal interpolation in (b). In this example, from a polygon with three vertex normals, the normal $\hat{\mathbf{N}}(\mathbf{P})$ of $\mathbf{P}$ is computed by first interpolating the normals of the horizontal line containing $\mathbf{P}$ with the polygon edges giving $\hat{\mathbf{N}}_{12}$ and $\hat{\mathbf{N}}_{23}$. From linear interpolation of these two normals, $\hat{\mathbf{N}}(\mathbf{P})$ is computed.

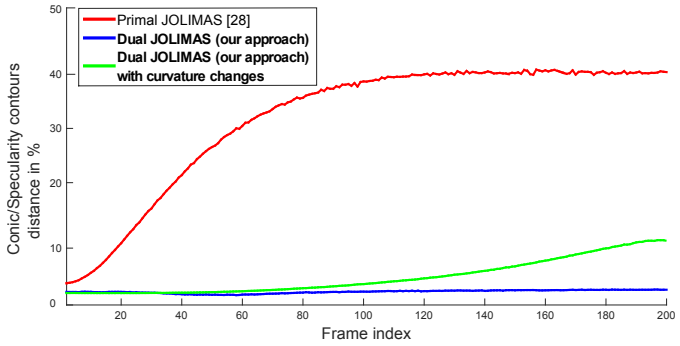
6 Experimental Results

We have tested specularity prediction on synthetic and real data to evaluate the improvement over previous work, the sensitivity of the model to curvature changes and the robustness of our method to a noisy brightest point. Specularity prediction is evaluated using the 2D distance between the specularity contour and the predicted ellipse, as defined in [35]. We compute a prediction error as a percentage by computing the mean error per contour point compared to the size of the frame.

6.1 Comparison with Primal JOLIMAS

We use a sequence of 300 images illustrated in figure 8(a). A planar surface is gradually curved into a cylinder to assess the functionality of our dual model on both planar and non-planar surfaces. For each image, we reconstruct a 3D quadric from 6 virtual viewpoints and compute the 2D distance between the predicted specularity and the detected specularity. Our method has a mean error of 1% as opposed to 33% per ellipse for [28].

(a)



(b)

Figure 8: Comparison between primal JOLIMAS and the proposed dual JOLIMAS. In this synthetic sequence, the curvature is increasing at each frame as shown in (a). Using 6 camera poses, both the primal and dual model are reconstructed and compared by comparing the 2D distance between specularity prediction and the contours of the detected specularity. Our dual approach outperforms [28] with an average error of 1% as opposed to 33% error on average for primal JOLIMAS. We also tested the specularity prediction of our dual model by reconstructing a 3D quadric on the planar surface (first frame) and use it for the remaining curvature changes. Even if the error increases, the specularity prediction remain reliable (7% on average).

6.2 Sensitivity to Curvature Changes

Since a specularity varies with geometry, and more particularly curvature, is it difficult to predict an accurate specularity for a new viewpoint with a different curvature. Using the previous sequence, we have tested our sensitivity to curvature changes by reconstructing a 3D quadric for the first case (planar) and predict the specularity for every frame. Even if the error increases, the information provided remain coherent with an average error of 7%, as seen in figure 8.

6.3 Specularity Prediction on Real Data

For real sequences, qualitative results are conducted on the specularity prediction by comparing the predicted ellipse with the specularity contours on each surface presenting specularities and with the specularity prediction of primal JOLIMAS.

Context. We choose different objects with various curvatures and materials under different lighting conditions. For the sequences rocket replica, mug and yellow vase, the images have a resolution of 1280×720. The camera poses are calculated with the constrained SLAM [36] adapted to curved objects [23]. This SLAM uses the mesh reconstructed with the HandySCAN 3D scanner which was computed beforehand. For the billiard pool sequence, the images have a resolution of 640×480. For this sequence, the camera poses are retrieved using Agisoft PhotoScan[3]. The specularity detection is provided by [27] as a gray

[3] www.agisoft.com

scale image. We only use 6 images to reconstruct the 3D quadric for each sequence.

All the results are illustrated in figure 1 and 9. A summary of the results comparing our dual JOLIMAS to the primal one is given in table 1.

Rocket replica. In this sequence of 1410 frames, a metal rocket replica is illuminated by a fluorescent lamp. As seen in figure 6(c), our rocket mesh is divided in 7 convex pieces. However only 3 convex pieces are used since the remaining one did not present specular reflections. We achieve a precision of 2.1%, as opposed to a precision of 32.8% for the primal model.

Mug. For this sequence of 300 frames, a porcelain mug is illuminated by two light sources: a spot light and a fluorescent lamp different from the rocket sequence. We are able to predict the two specularities accurately with a prediction error of 1.3% since the curvature of the object of interest is constant. The primal JOLIMAS predicts the specularities with the appropriate shape but with the wrong positions, with an error on average of 18.3%.

Yellow Vase. In this sequence of 1900 frames, a yellow vase made of glass is illuminated by a spot light. This sequence is particularly challenging due to the symmetry of the objects making the camera pose estimation process prone to errors. In addition, the specularities on the vase are very small, making the refinement of the 3D quadric estimation difficult. However, we achieved a good precision with a specularity prediction error per ellipse of 1.9% as opposed to a precision of 35.2% for the primal model. This object was divided in 7 convex pieces as shown in 6(d).

Billiard pool. In this sequence of 415 frames, three billiard balls are illuminated by two light sources: a desktop lamp and a ceiling lightbulb. This sequence is also challenging due to the small size of the specularities. Dual JOLIMAS achieves a precision of 1.1% as opposed to 51.4% for the primal model.

Table 1: Empirical validation of our model and its ability to predict specularities in images for the 3 real sequences presented in figure 9. For varied curvatures and light sources, our method provides substantially accurate results in comparison with primal JOLIMAS.

	Prediction error (2D distance in %)	
Sequence	Primal [28]	Dual (proposed)
Rocket replica	32.8	2.1
Mug	18.3	1.3
Vase	35.2	1.9
Billiard balls	51.4	1.1

7 Application in Dynamic Retexturing

In augmented reality, it is sometimes useful to switch a previous real texture by a synthetic one in order to display/highlight a new information or cover an unwanted surface. This process is called retexturing. In most cases, retexturing must be achieved seamlessly according to the lighting conditions (intensity, shadows) and the geometry of the scene. To improve the realism, we propose a dynamic retexturing application using our specularity prediction to model the specularities on the new textures. We improve the retexturing of [28] by adding a diffuse term which is essential to represent self-shadows and intensity variations which often appeared on non-planar surfaces. To further improve the realism, we also compute the intensity and the color of the specularity by learning them on every frame used for the quadric reconstruction.

Diffuse term computation We consider the input texture as the ambient term commonly used in Phong's model [32]. To add a diffuse term, we use the centers of the reconstructed 3D quadrics as light source position $\mathbf{L}$ and compute the diffuse term for each surface point $\mathbf{P}$ as:

$$I_d = \sum_{i=1}^{k} \left(\hat{\mathbf{L}}_i(\mathbf{P}) \cdot \hat{\mathbf{N}}(\mathbf{P}) \right), \quad (4)$$

with I_d the diffuse image, k the number of light sources and i the index of the light source used.

Specularity learning and blending. For each frame used for the 3D quadric reconstruction, we use the predicted ellipse and the specularity detection to learn the specularity pattern. By matching the shape of the different specularities, we retrieve the intensity evolution and the color of the specularity. To ensure an adequate blending, we interpolate the color of the boundary of the specularities with the diffuse term around it. The full pipeline of the retexturing is illustrated in figure 10.

As illustrated in figure 11, our retexturing application allows us to synthesize the specularity's intensity and color and to create coherent self-shadows in real-time.

8 Computation Time

In table 2, we present the computation time for each step of our method including the specularity prediction, the brightest point computation, the 3D quadric reconstruction, the specularity prediction and the retexturing. We compute the results on an Intel i7 processor with a 2.70 GHz frequency on the real sequences presented in figure 9. The 3D quadric reconstruction takes 76.2 milliseconds on average and is performed when a minimum of 6 keyframes is obtained. The 3D quadric is refined for every new keyframe but takes this time 15 milliseconds on average because only a few iterations is needed as opposed to the initial refinement. The other steps are performed for each frame. Our method is running at approximately 27 frames per second making it well adapted to a real-time context. Our analysis was conducted without any GPU implementation and parallelization.

Table 2: Computation time of JOLIMAS prediction and retexturing analysis. Our method is running at approximately 27 frames per second making our application suitable for real-time.

JOLIMAS's reconstruction step	Computation time (ms)
Specularity detection	12.1
Brightest point correction	8.3
3D quadric reconstruction	76.2
Specularity prediction	2.4
Retexturing	15.4

9 Limits of the Approach

The proposed dual JOLIMAS significantly improves on JOLIMAS [28] but has limitations and failure cases, which we now describe. The main limitation of dual JOLIMAS is the fact that it reconstructs one quadric per convex surface and light source. This may be seen as an incoherence in the model and may be computationally inefficient for complex objects with many convex pieces. It would be more coherent and efficient to have one quadric per light source and for all convex surface pieces. In the example of the vase shown in figure 1, we could use the seven specularities to reconstruct the 3D quadric associated with the light source from a single image. Strictly speaking, dual JOLIMAS does however not allow this type of reconstruction because it does not handle curvature changes. In terms of the images, it requires the observed specularity ellipses to fulfill the epipolar geometry. In practice however, we have seen that mild variations of the surface's curvature can be handled well by dual JOLIMAS for smooth surfaces. Note however that for stronger changes, turning a convex into a concave surface, we have that the specularity's shape changes topology and can clearly not be modeled anyhow by the projection of a quadric. An example is shown in figure 12(b). A failure case of dual JOLIMAS is when a convex surface piece is made of materials with significantly different reflectance properties. This will cause the scale of the best fit specularity ellipse to change across the materials, and thus negatively impact the quadric based model of dual JOLIMAS. An example is shown in figure 12(b).

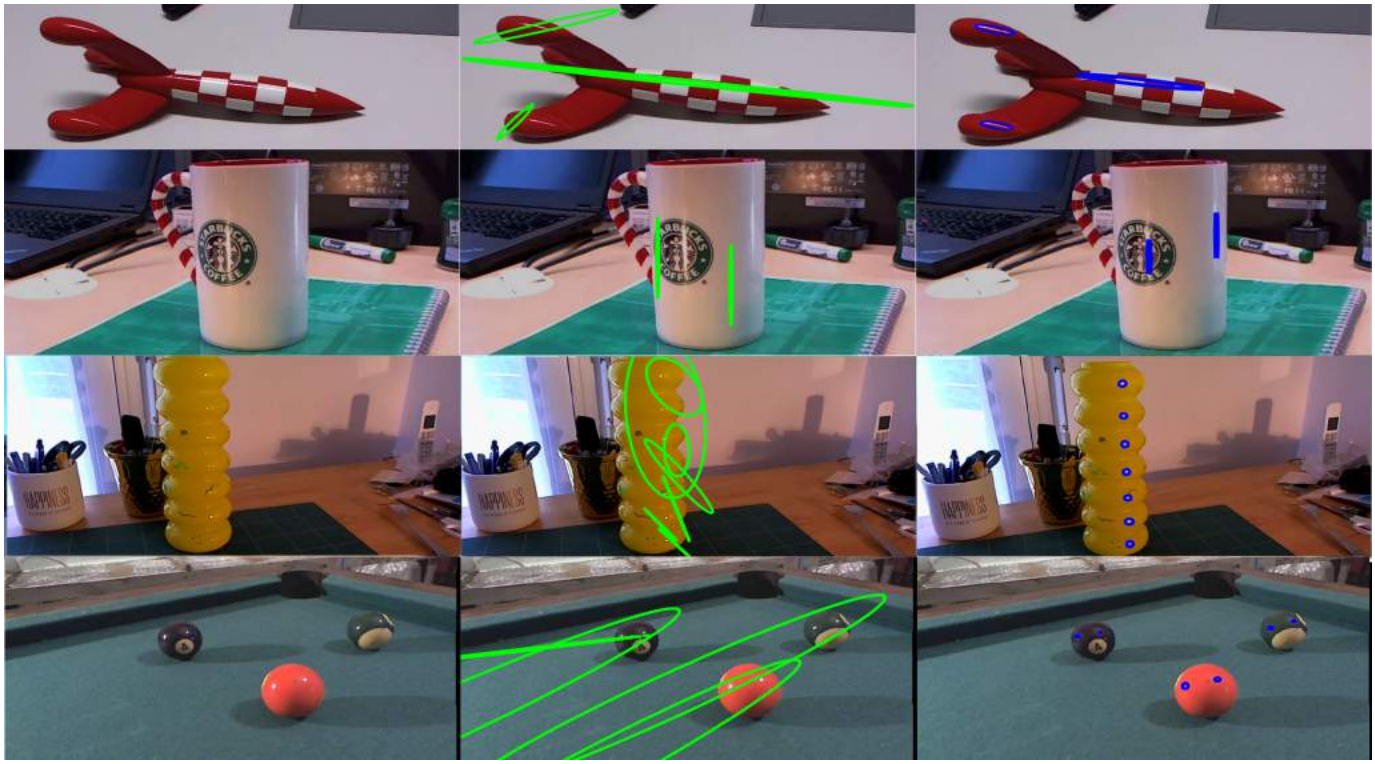


Figure 9: Specularity prediction on real sequences with various light sources and curvatures. In the first row, a rocket replica is illuminated by a fluorescent lamp. We predict 3 specularities accurately by using dual JOLIMAS (right) as opposed to primal JOLIMAS (middle) which fails in terms of position and shape of the predicted specularities. In the second row, a mug is illuminated by two light sources a desk and a fluorescent one, which are used to test the capability of our model for multiple light sources varied in shape and intensity. For this object, our result is particularly accurate due to the constant curvature of the object. Primal JOLIMAS predicts the correct scale of the specularities with wrong position. In the third row, we use a vase illuminated by a desk lamp. This object is composed of 7 convex part giving seven specularities. In the fourth row, we use three billiard balls illuminated by a desk lamp and a ceiling lightbulb. These last two sequences are particularly challenging due to the camera pose errors caused by the difficulty of our localization method to handle symmetric objects and the limitations of the specularity detection to handle very small specularities.

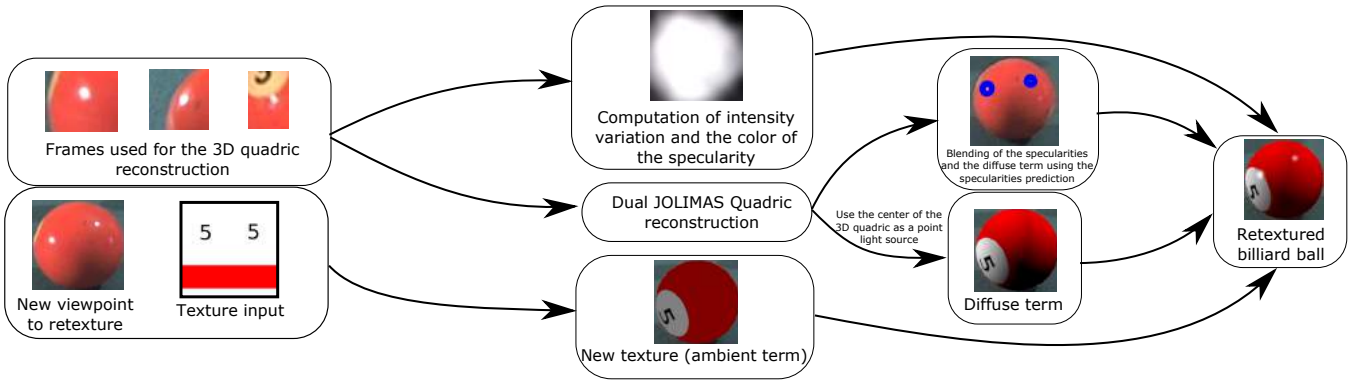


Figure 10: Pipeline of the retexturing process. The method is divided in three main phases: the specularity reconstruction from frames used for the 3D quadric reconstruction (top part), the diffuse term computation (bottom part) and the blending process between the specularity reconstruction, alignment with the specularity prediction and the diffuse term. The specularity is reconstructed in terms of intensity variation and color empirically from the predicted conics and the specularity detection image. To compute the diffuse term, we consider the center of the 3D quadrics reconstructed as point light sources. We then compute the diffuse component according to [32]. The specularity and the diffuse term are merged to create our retextured object. The addition of the diffuse term and blending of the specularities greatly improve the rendering.

10 Discussion

Our future work will be to study the impact of roughness and curvature changes on specularities. When reconstructing the 3D quadric, we need to include slight changes of curvature and roughness because these elements impact the specularity's shape. To handle curvature changes, it would be interesting to find an analytic solution linking the curvature to the specularity's shape close to what [4] achieved. Moreover, we will need to study the materials on the specularity's shape. In our model, the reconstruction is still limited to a specific material because the specularity's shape is strongly influenced by the material type. A first step would be to use material segmentation methods such as [2, 12] to indicate to the algorithm a change in the specularity

Figure 11: Retexturing example on 3 sequences: the billiard pool, the mug and the yellow vase. We can see that the specularities and the shadows are coherent to the lighting context of the input image.

(a) (b)

Figure 12: Hard cases for JOLIMAS dual. In (a), we show how the specularity is affected by the surface material in terms of scale even for similar curvature. In this context, the ellipses associated to the specularities would not respect the epipolar geometry. In (b), we illustrate a fire extinguisher presenting a consequent curvature change that will affect the specularity shape. In that case, we will not be able to correctly predict the specularity in terms of scale.

shape. A perspective would be to train an algorithm from a database of specularities on different materials with various light source shapes and intensities to best predict the appropriate changes to apply to the specularities to match the material of the surface. Our intensity function uses the keyframe images to reconstruct the specularity appearance but does not include the roughness. It could be interesting to compare our specularity prediction with the segmentation of the specularity's contours to learn a roughness pattern (normal map or bump map) and include it in the specularity prediction and retexturing steps.

11 Conclusion

To address the need for a specularity prediction process for augmented reality and computer vision applications, we have presented a new multiple-view geometric model of the specular shape called dual JOLIMAS. It extends primal JOLIMAS [28] under the same hypothesis of elliptical specularity on convex surfaces. The main idea of the model is to represent a specularity as a reflected image of the light source on a mirror-like surface. This virtual light source can be modeled by a 3D quadric whose perspective projection fits the specularity's shape for existing and new viewpoints according to the camera pose. The previous approach was limited to planar surfaces. For non-planar surfaces, the ellipses fitted to the specularities are not epipolar consistent, which defeats primal JOLIMAS. We used a virtual representation, computed from a specific point in the specularity: the brightest point, which follows the reflexion law of optics and which is not affected by the distortion induced by the surface curvature. However, the computation of this brightest point is not trivial due to the presence of noise in the camera poses and in the normal of the meshes. For these reasons, we implemented a robust brightest point estimation method following optics rules. To ensure the genericity of our model to advanced meshes, we used approximate convex decomposition to divide a mesh in convex pieces. To predict every specularity in terms of shape, we reconstruct a 3D quadric per convex piece and per light source. The predicted specularity (as an ellipse) is obtained by projecting the reconstructed 3D quadric according to the virtual camera pose of the new viewpoint. Our specularity prediction was evaluated in terms of robustness and precision on synthetic and real sequences for multiple light sources and objects. Among the possible applications of this prediction, we presented a retexturing application to predict the intensity and color of the specularity.

Acknowledgments

This research has received funding from the EU's FP7 through the ERC research grant 307483 FLEXABLE.

References

[1] A. Artusi, F. Banterle, and D. Chetverikov. A survey of specularity removal methods. *Computer Graphics Forum*, 30(8):2208–2230, 2011.

[2] S. Bell, P. Upchurch, N. Snavely, and K. Bala. Material recognition in the wild with the materials in context database. In *Computer Vision and Pattern Recognition*, CVPR, 2015.

[3] A. Blake. Does the brain know the physics of specular reflection? *Nature*, 343(6254):165–168, 1990.

[4] A. Blake and G. Brelstaff. Geometry from specularities. In *International Conference on Computer Vision*, ICCV, 1988.

[5] J. F. Blinn. Models of light reflection for computer synthesized pictures. In *Special Interest Group on Computer Graphics and Interactive Techniques*, SIGGRAPH, 1977.

[6] B. J. Boom, S. Orts-Escolano, X. X. Ning, S. McDonagh, P. Sandilands, and R. B. Fisher. Point light source estimation based on scenes recorded by a rgb-d camera. In *British Machine Vision Conference*, BMVC, 2013.

[7] P.-E. Buteau and H. Saito. Retrieving lights positions using plane segmentation with diffuse illumination reinforced with specular component. In *International Symposium on Mixed and Augmented Reality*, ISMAR, 2015.

[8] J. Y. Chang, R. Raskar, and A. Agrawal. 3d pose estimation and segmentation using specular cues. In *Computer Vision and Pattern Recognition*, CVPR, 2009.

[9] A. DelPozo and S. Savarese. Detecting specular surfaces on natural images. In *Computer Vision and Pattern Recognition*, CVPR, 2007.

[10] F. Einabadi and O. Grau. Discrete light source estimation from light probes for photorealistic rendering. In *British Machine Vision Conference*, BMVC, 2015.

[11] J. Engel, V. Koltun, and D. Cremers. Direct sparse odometry. *arXiv preprint arXiv:1607.02565*, 2016.

[12] S. Georgoulis, K. Rematas, T. Ritschel, M. Fritz, L. Van Gool, and T. Tuytelaars. Delight-net: Decomposing reflectance maps into specular materials and natural illumination. *arXiv preprint arXiv:1603.08240*, 2016.

[13] S. H. I. K. Hyeongwoo Kim, Hailin Jin. Specular reflection separation using dark channel prior. CVPR, 2013.

[14] D. S. Immel, M. F. Cohen, and D. P. Greenberg. A radiosity method for non-diffuse environments. In *Special Interest Group on Computer Graphics and Interactive Techniques*, SIGGRAPH, 1986.

[15] J. Jachnik, R. A. Newcombe, and A. J. Davison. Real-time surface light-field capture for augmentation of planar specular. In *International Symposium on Mixed and Augmented Reality*, ISMAR, 2012.

[16] J. T. Kajiya. The rendering equation. In *Special Interest Group on Computer Graphics and Interactive Techniques*, SIGGRAPH, 1986.

[17] M. Kanbara and N. Yokoya. Real-time estimation of light source environment for photorealistic augmented reality. In *International Conference on Pattern Recognition*, ICPR, 2004.

[18] G. J. Klinker, S. A. Shafer, and T. Kanade. A physical approach to color image understanding. *International Journal of Computer Vision*, 4(1):7–38, 1990.

[19] S. Kuthirummal and S. K. Nayar. Flexible mirror imaging. In *International Conference on Computer Vision*, ICCV, 2007.

[20] P. Lagger and P. Fua. Using specularities to recover multiple light sources in the presence of texture. In *International Conference on Pattern Recognition*, ICPR, 2006.

[21] P. Lagger, M. Salzmann, V. Lepetit, and P. Fua. 3d pose refinement from reflections. In *Computer Vision and Pattern Recognition*, CVPR, 2008.

[22] J.-M. Lien and N. M. Amato. Approximate convex decomposition of polyhedra. In *Solid and Physical Modeling*, SPM, 2007.

[23] A. Loesch, S. Bourgeois, V. Gay-Bellile, and M. Dhome. Generic edgelet-based tracking of 3d objects in real-time. In *Intelligent Robots and Systems*, IROS, 2015.

[24] G. Long, L. Kneip, X. Li, X. Zhang, and Q. Yu. Simplified mirror-based camera pose computation via rotation averaging. In *Computer Vision and Pattern Recognition*, CVPR, 2015.

[25] M. Meilland, C. Barat, and A. Comport. 3d high dynamic range dense visual slam and its application to real-time object re-lighting. In *International Symposium on Mixed and Augmented Reality*, ISMAR, 2013.

[26] H. Mitsumoto, S. Tamura, K. Okazaki, N. Kajimi, and Y. Fukui. 3-d reconstruction using mirror images based on a plane symmetry recovering method. *Pattern Analysis and Machine Intelligence*, 14(9):941–946, 1992.

[27] A. Morgand and M. Tamaazousti. Generic and real-time detection of specular reflections in images. In *8th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, VISAPP, 2014.

[28] A. Morgand, M. Tamaazousti, and A. Bartoli. An empirical model for specularity prediction with application to dynamic retexturing. In *International Symposium of Mixed and Augmented Reality*, ISMAR, 2016.

[29] A. Morgand, M. Tamaazousti, and A. Bartoli. A geometric model for specularity prediction on planar surfaces with multiple light sources. In *Transactions on Visualization and Computer Graphics*, TVCG, 2017.

[30] A. Netz and M. Osadchy. Recognition using specular highlights. *Pattern Analysis and Machine Intelligence*, 35(3):639–652, 2013.

[31] R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. Molyneaux, S. Hodges, D. Kim, and A. Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *ACM symposium on User interface software and technology,*, ISMAR, 2011.

[32] B. T. Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975.

[33] S. H. Said, M. Tamaazousti, and A. Bartoli. Image-based models for specularity propagation in diminished reality. In *Transactions on Visualization and Computer Graphics*, TVCG, 2017.

[34] S. Savarese, M. Chen, and P. Perona. Local shape from mirror reflections. *International Journal of Computer Vision*, 64(1):31–67, 2005.

[35] P. Sturm and P. Gargallo. Conic fitting using the geometric distance. In *Asian Conference on Computer Vision*, ACCV. 2007.

[36] M. Tamaazousti, V. Gay-Bellile, S. Naudet Collette, S. Bourgeois, and M. Dhome. Nonlinear refinement of structure from motion reconstruction by taking advantage of a partial knowledge of the environment. In *Computer Vision and Pattern Recognition*, CVPR, 2011.

[37] W. Tan, H. Liu, Z. Dong, G. Zhang, and H. Bao. Robust monocular slam in dynamic environments. In *International Symposium on Mixed and Augmented Reality*, ISMAR, 2013.

[38] R.-T. Thomas, K. Denis, P. Jinwoo, and S. Dieter. Instant mixed reality lighting from casual scanning. In *International Symposium of Mixed and Augmented Reality*, ISMAR, 2016.

[39] Q. Tian and J. J. Clark. Real-time specularity detection using unnormalized wiener entropy. In *Computer and Robot Vision*, CRV, 2013.

[40] T. Whelan, S. Leutenegger, R. F. Salas-Moreno, B. Glocker, and A. J. Davison. Elasticfusion: Dense slam without a pose graph. In *Robotics: science and systems*, vol. 11, 2015.

[41] T. Whelan, R. F. Salas-Moreno, B. Glocker, A. J. Davison, and S. Leutenegger. Elasticfusion: Real-time dense slam and light source estimation. *The International Journal of Robotics Research*, 2016.

[42] K.-Y. K. Wong, D. Schnieders, and S. Li. Recovering light directions and camera poses from a single sphere. In *European Conference on Computer Vision*, ECCV. 2008.

[43] E. Zhang, M. F. Cohen, and B. Curless. Emptying, refurnishing, and relighting indoor spaces. In *Special Interest Group on Computer Graphics and Interactive Techniques*, SIGGRAPH Asia, 2016.