

Modèle géométrique multi-vues de spécularité

Alexandre Morgand^{1,2}

Mohamed Tamaazousti¹

Adrien Bartoli²

¹ CEA-LIST, Point Courrier 94, Gif-sur-Yvette, F-91191, France
{alexandre.morgand, mohamed.tamaazousti}@cea.fr

² Institut Pascal - UMR 6602 - CNRS/UCA/CHU, Clermont-Ferrand, France
adrien.bartoli@gmail.com

Résumé

Prédire les spécularités dans une image, à partir d'une pose de caméra et de la géométrie de la scène connues, est un problème complexe et ouvert. Cette prédiction est néanmoins essentielle pour de nombreuses applications en réalité augmentée. Une approche récente appelée JOLIMAS répond partiellement à ce problème sous l'hypothèse que les spécularités ont une forme elliptique et que la scène est composée d'un objet plan. JOLIMAS modélise une spécularité comme étant l'image d'une quadrique fixe dans l'espace. Nous proposons JOLIMAS dual, un nouveau modèle géométrique qui adresse la limitation aux plans. Ce modèle repose sur le fait que les spécularités conservent leurs formes elliptiques sur des surfaces convexes et que chaque surface peut être divisée en plusieurs parties convexes. La géométrie de JOLIMAS dual utilise une quadrique pour chaque surface convexe et chaque source lumineuse, et prédit les spécularités en utilisant des caméras virtuelles permettant de gérer la courbure de la surface. Nous évaluons l'efficacité et la précision du JOLIMAS dual sur de multiples séquences synthétiques et réelles composées de divers types d'objets et sous différents éclairages. Des résultats supplémentaires sont présentés dans une vidéo annexe.

Mots Clef

Modèle, géométrie, spécularité, courbure, multi-vues

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 augmented reality. 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. Further results are presented as supplementary video material.

Keywords

Geometry, model, specularity, curvature, multi-view

1 Introduction

Le SLAM (Simultaneous Localization And Mapping) a récemment gagné en maturité en matière d'application industrielle pratique comme [26] qui a combiné du SLAM avec de la réalité augmentée et diminuée afin de modifier le contenu d'une pièce d'appartement et de changer le contexte lumineux dynamiquement. La plupart des méthodes de SLAM supposent que la scène est Lambertienne ce qui n'est pas toujours valable pour des scènes comportant des surfaces réfléchissantes comme du métal, plastique ou encore de la porcelaine. Pour des modèles d'illuminations comme celui de Phong [20], la partie non-Lambertienne peut être restreinte au terme spéculaire. En vision par ordinateur, les spécularités jouent un rôle important et font l'objet de nombreuses études [8, 11]. Elles sont, en général, considérées comme indésirables mais peuvent être utilisées comme des traces visuelles intéressantes. D'un côté, elles peuvent être utilisées pour améliorer la localisation de la caméra [5, 19] ou encore un processus de reconstruction 3D [24]. D'un autre côté, elles peuvent grandement améliorer la qualité de rendu des applications de réalité augmentée et diminuée. Blake *et al.* [1] montrent que les spécularités sont des éléments utilisés pour la perception humaine d'une scène (en matière de géométrie par exemple). La formation de ces éléments est dépendante de plusieurs éléments physiques de la scène comme la surface, les matériaux, les sources de lumière et la caméra. Ces nombreux paramètres rendent le problème de prédiction de spécularité complexe même si les surfaces et les paramètres de caméra sont connus.

Il existe des méthodes de prédiction de spécularité [3, 23]

modélisant la lumière en utilisant des modèles d’illuminations issus de l’infographie. Elles supposent une illumination distante et requièrent le calcul de nombreux paramètres sur les matériaux, les sources de lumière et ne fournissent pas une solution géométrique flexible pour prédire les spécularités. Ces méthodes sont classées dans deux catégories : illumination globale et reconstruction de source de lumière. Une troisième catégorie d’approches existe où le problème de prédiction de spécularité est considéré comme un problème de reconstruction multi-vues. Des travaux récents [18] ont montré qu’une spécularité sur une surface plane peut être convenablement approchée par une ellipse sous un éclairage provenant d’une ampoule ou d’un néon. Un modèle empirique appelé JOint LIght-MATerial Specularity (JOLIMAS) a été proposé pour abstraire l’interaction lumière-matière et traiter le problème de prédiction de spécularité géométriquement. Ce modèle utilise une quadrique fixe dont la projection prédit la forme des spécularités sur des points de vue existants et nouveaux. La reconstruction de quadrique est obtenue avec au moins trois points de vue. Nous nous référons à ce modèle comme étant *JOLIMAS primal*.

Dans cet article, nous mettons en avant deux contributions pour la modélisation géométrique de la forme d’une spécularité. Dans un premier temps, nous étudions la principale limitation de *JOLIMAS primal* : son incapacité à fonctionner pour des surfaces non-plans. Dans un second temps, nous traitons cette limitation par *JOLIMAS dual*. Ce nouveau modèle est une généralisation du primal aux surfaces convexes par morceaux tout en conservant sa fonctionnalité sur des surfaces planes. L’idée principale est que les spécularités conservent leur forme elliptiques sur des surfaces convexes. Notre modèle se base sur une représentation virtuelle des caméras qui associe une caméra virtuelle à chaque spécularité avec la pose de la caméra réelle et la géométrie de la surface connues. Du fait que chaque surface peut être divisée en parties convexes, nous pouvons modéliser et prédire les spécularités sur des objets plus complexes comme illustré à la figure 1. En pratique, nous représentons une surface par un maillage et utilisons un procédé de décomposition convexe approximative. Nous reconstruisons une quadrique pour chaque partie convexe et chaque source de lumière. Nous expliquons le détail de notre modèle de prédiction de spécularité et le calcul de notre représentation pour de nouveaux points de vue. Cette prédiction n’est pas triviale car les caméras virtuelles utilisées pour la prédiction requièrent le calcul du point d’intensité maximale d’une spécularité inconnue pour un nouveau point de vue. Nous détaillerons le calcul de ce point particulier par utilisation de notre quadrique reconstruite. L’efficacité et la précision de notre méthode est validée sur de nombreux exemples synthétiques et réels incluant divers types objets et divers conditions lumineuses.

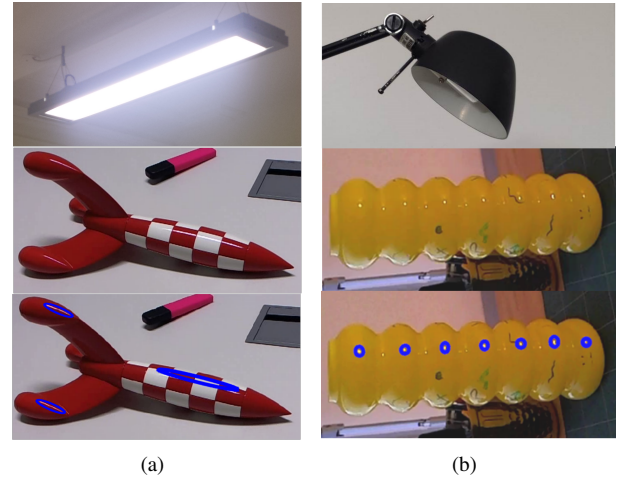


FIGURE 1 – Sous l’hypothèse que les spécularités ont une forme elliptique sur les surfaces convexes, nous reconstruisons une quadrique dont la projection perspective correspond aux spécularités pour de nouveaux points de vue. Une quadrique est créée pour chaque surface convexe et chaque source de lumière. (a) présente une séquence réelle d’une réplique de fusée (au milieu) éclairée par un néon (en haut). Dans (b), nous utilisons une ampoule (en haut) pour illuminer un vase en verre (au milieu). Notre modèle a la capacité de prédire les spécularités (en bas) sur les différentes parties de la fusée et du vase pour de nouveaux points de vue (ellipses bleues).

2 État de l’art

Rendu utilisant l’illumination globale. Les méthodes dans cette catégorie privilégie la qualité de rendu en résolvant l’équation de rendu [10]. Cette équation décrit la quantité de lumière émise au point P à partir d’un point de vue selon une fonction de la lumière incidente et une BRDF (Bidirectional Reflectance Distribution Function). Ces approches ne calculent pas en général les attributs physiques des sources de lumière. [9] capture un champ lumineux 4D à partir d’une petite surface spéculaire plane. En reconstruisant les composantes spéculaires et diffuses, cette méthode obtient un rendu convaincant en réalité augmentée. Cependant, cette méthode n’est pas capable de prédire la composante spéculaire pour des points de vue non utilisés dans la reconstruction initiale. De plus, les sources de lumière avec état variable (allumées ou éteintes) ne sont pas gérées. [23] est une extension de [9] en ajoutant une segmentation de matériaux pour des surfaces complexes et reconstruites à l’aide d’une caméra RGB-D. Cette méthode partage les mêmes limitations que [9]. En conséquence, [9, 23] et d’autres approches similaires comme [16] ne peuvent pas prédire les spécularités pour de nouveaux points de vue.

Estimation de source lumineuse. Nous distinguons deux catégories de sources de lumière : sources directionnelles et sources ponctuelles. Dans un contexte extérieur,

une source directionnelle est souvent considérée. Ce type de lumière peut également fournir des résultats intéressants dans un contexte intérieur : [12] utilise les spécularités pour calculer des sources directionnelles avec un objet en mouvement à partir d'un point de vue fixe. Cependant, la prédiction de spécularité requiert des informations sur la forme, l'intensité et la position de la source de lumière ainsi que sur les propriétés des matériaux. [4, 7, 25] partagent les mêmes problèmes. Dans [3], les auteurs reconstruisent une source ponctuelle en utilisant le terme diffus sur des surfaces Lambertiennes à partir d'une caméra RGB-D. Les méthodes utilisant des sources ponctuelles ne peuvent pas modéliser des sources étendues comme des néons ce qui limite leurs champs d'application.

Modélisation géométrique. Récemment, [18] a proposé une nouvelle façon de modéliser l'interaction lumière/matière avec JOLIMAS primal. En reconstruisant une quadrique à partir d'une spécularité sur des surfaces planes, cette méthode prédit la forme spéculaire pour de nouveaux points de vue en utilisant une projection perspective de cette quadrique. Cette méthode abstrait les interactions lumière/matière et est utilisée efficacement pour du retexturing. Cependant, elle gère uniquement des surfaces planes ce qui limite grandement son champ d'application. JOLIMAS primal se base sur les hypothèses suivantes :

1. Les spécularités ont une forme elliptique sur les surfaces planes comme dans les modèles de Phong [20] et Blinn-Phong [2]
2. Une source de lumière est associée à une spécularité sur une surface plane
3. Il y a une quadrique fixe unique localisée 'sous' la surface plane dont la projection perspective correspond à la forme de la spécularité dans l'image. La projection est réalisée à partir des caméras réelles

JOLIMAS primal utilise des surfaces planes. Le symétrique par rapport à la surface est utilisé pour comparer la reconstruction avec la position de la source de lumière mais n'est pas utilisé pour la reconstruction de quadrique. Utiliser cette symétrie serait trivial puisque la normale d'un plan est uniforme. Pour une surface non-plan, cette propriété n'est plus valable : la normale sur la surface. Ceci met en échec JOLIMAS primal puisqu'il n'y a pas de quadrique dans le repère caméra dont la reprojction correspondrait aux spécularités dans l'image. Comme montré dans la figure 2 pour une sphère synthétique utilisant trois points de vue et leurs trois spécularités respectives dans l'image, la géométrie épipolaire n'est pas respectée. En effet, pour chaque image, nous pouvons voir que les lignes épipolaires provenant des autres points de vue ne sont pas tangentes à l'ellipse correspondant à la spécularité. Notre modèle adresse la question de l'existence d'une quadrique fixe dont la reprojction correspondrait aux spécularités dans les images respectives pour des surfaces non-plans.

Modèle proposé. Notre modèle étend [18] aux surfaces non-plans tout en reprenant l'approximation de spéculari-

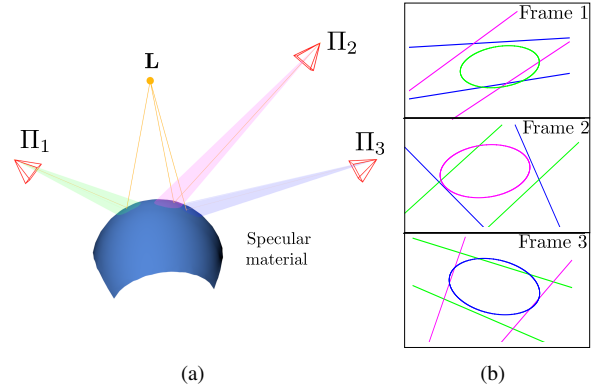


FIGURE 2 – Lignes épipolaires des ellipses pour trois poses de caméra données Π_1 , Π_2 et Π_3 et pour une source de lumière L utilisant JOLIMAS primal sur une sphère. La scène est illustrée dans (a) et les ellipses associées à leurs droites épipolaires dans (b). Dans la forme primale, la géométrie épipolaire n'est pas respectée ce qui donne une mauvaise reconstruction de quadrique et une mauvaise prédiction de spécularité.

tés de forme elliptique même pour des sources étendues. Nous proposons JOLIMAS dual qui utilise des caméras virtuelles obtenues par symétrie des caméras réelles selon le plan tangent à un point particulier de la surface à l'opposé de JOLIMAS primal [18], qui utilise des caméras réelles. Notre approche améliore la reconstruction en terme de généricité avec un nombre de points de vue nécessaire inférieur à la version primale grâce à une meilleure répartition des caméras virtuelles. De nouvelles caméras virtuelles peuvent être ajoutées par l'utilisation de spécularités sur des objets similaires (courbure et matériaux) à partir de la même source de lumière.

3 JOLIMAS dual

3.1 Principe

Une spécularité est le résultat d'une interaction de la lumière avec une surface agissant de façon similaire à un miroir. Pour une surface plane, nous pouvons interpréter une spécularité comme étant une image miroir de la source de lumière. Cependant, pour des surfaces courbes, cette image miroir est distordue et incohérente avec la vraie position et forme de la source de lumière. A partir des spécularités, nous ne pouvons pas reconstruire un objet 3D cohérent localisé autour de la source de lumière dans l'état. L'idée principale de JOLIMAS dual est d'utiliser des caméras virtuelles pour reconstruire la quadrique de façon cohérente avec la source de lumière. Ces caméras virtuelles sont calculées par symétrie de la vraie caméra avec la surface. Contrairement aux surfaces planes, la normale varie le long de la surface courbe ce qui rend le choix de la caméra virtuelle ambigu. Pour une pose de caméra et une spécularité donnée, quel point de la surface devons-nous choisir pour calculer cette caméra virtuelle ?

3.2 Le point d'intensité maximale

Nous avons observé empiriquement que cette approximation que l'hypothèse de la forme elliptique et de l'unicité d'une spécularité est valide (voir section 6) pour des surfaces convexes. [8, 18] ont étudié l'évolution de l'intensité d'une spécularité autour du point d'intensité maximale. Ce point de la surface avec l'intensité la plus intense est très importante en optique car pour une pose de caméra et une position de source lumineuse données, il suit la loi de la réflexion. Nous définissons l'opérateur de normalisation μ tel que $\mu(\mathbf{A}) = \frac{\mathbf{A}}{\|\mathbf{A}\|}$ avec \mathbf{P} un point de la surface, $\hat{\mathbf{N}}(\mathbf{P})$ le vecteur normalisé représentant la normale de la surface au point \mathbf{P} , $\hat{\mathbf{L}}$ le vecteur normalisé du rayon de lumière incident provenant de la source de lumière \mathbf{L} tel que $\hat{\mathbf{L}}(\mathbf{P}) = \mu(\mathbf{L} - \mathbf{P})$ et $\hat{\mathbf{V}}$ le vecteur normalisé du rayon réfléchi au point de surface jusqu'à la position de caméra \mathbf{V} tel que $\hat{\mathbf{V}}(\mathbf{P}) = \mu(\mathbf{V} - \mathbf{P})$, la loi de la réflexion est décrite par :

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

Selon [18], le point le plus intense sous le modèle de Phong est localisé à l'intersection de la surface S avec la ligne passant par la source de lumière \mathbf{L} et le symétrique de la position de la caméra $\tilde{\mathbf{V}}$ pour des surfaces planes. Ce résultat suit la loi de la réflexion. Pour des surfaces convexes, cette propriété est toujours valable si le point de la surface utilisé pour calculer le symétrique de la caméra $\tilde{\mathbf{V}}$ coïncide avec le point d'intersection comme illustré sur la figure 3. En pratique, le calcul du point le plus intense se doit d'être le plus précis possible car les caméras virtuelles sont sensibles aux erreurs de normale. Nous présentons une méthode de calcul de ce point le plus intense à partir d'aprioris géométriques sur le maillage dans la section 5.2.

3.3 Estimation du modèle

Nous calculons JOLIMAS dual à partir de la caméra virtuelle estimée à partir du point le plus intense pour chaque image sur la surface. Nous utilisons une estimation initiale du point du plus intense \mathbf{P}_B , à partir d'une détection de spécularité calculée par [17] initiale en prenant le centre de gravité des contours de la spécularité. En pratique, ce point le plus intense n'est pas unique dans l'image car la spécularité provoque souvent une saturation de l'image. Une pose de caméra virtuelle $\tilde{\Pi}$ est calculée en suivant une formule de [15] telle que :

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

avec Π la pose de la caméra réelle et le plan tangent $\pi = \left[\begin{array}{c} \hat{\mathbf{N}}(\mathbf{P}) \\ d \end{array} \right]$ au point \mathbf{P} . Pour une source de lumière unique, le pipeline de l'estimation de notre modèle est :

- Détection de spécularité calculant simultanément le point le plus intense pour chaque spécularité

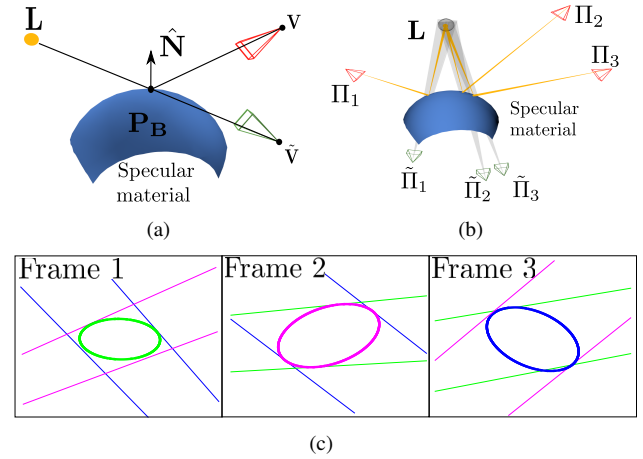


FIGURE 3 – Loi de la réflexion sur une surface convexe (a). Le point le plus intense \mathbf{P}_B correspond au point de la surface où l'intersection de la ligne contenant la position de la caméra virtuelle $\tilde{\mathbf{V}}$ (en vert) et la source de lumière \mathbf{L} est confondue avec le point de la surface utilisé pour calculer la caméra virtuelle. Dans (b) nous présentons JOLIMAS dual sur une surface convexe pour 3 poses de caméra données Π_1, Π_2 et Π_3 . Les poses de caméra (en vert) $\tilde{\Pi}_1, \tilde{\Pi}_2$ et $\tilde{\Pi}_3$ permettent de reconstruire une quadrique proche de la source de lumière réelle. Dans (c), nous montrons que notre représentation virtuelle de caméra permet d'obtenir des ellipses respectant la géométrie épipolaire.

- Estimation des caméras virtuelles initiales par symétrie des caméras réelles avec la surface pour chaque point d'intensité maximale
- Correction épipolaire à partir de [21] et reconstruction de la quadrique à partir de [6] pour au moins 3 points de vue
- Correction des points d'intensité maximale de chaque spécularité à partir de la reconstruction de quadrique (décrite dans 5.2)
- Raffinement non-linéaire à partir de tous les points de vue disponibles

Un exemple de notre représentation de caméra virtuelle est présenté à la figure 3(b). Ce pipeline est répété pour chaque source de lumière.

4 Surfaces convexes par morceau

4.1 Principe

Pour chaque source de lumière, nous avons un point d'intensité maximale unique et plus généralement une spécularité unique par surface convexe. Pour améliorer la généralité de notre modèle dual, nous décomposons la scène en une combinaison de plusieurs morceaux convexes. Nous reconstruisons une quadrique pour chaque partie convexe et pour chaque source de lumière.

4.2 Décomposition convexe approximative

La décomposition convexe approximative est souvent utilisée dans les moteurs physiques pour accélérer la détection de collisions. Notre objectif est de déterminer une décomposition convexe d'une surface S , qui consiste à la partitionner en un ensemble minimal de sous-surfaces convexes. Ce procédé a pour but de déterminer la partition des triangles du maillage avec un nombre d'amas minimal tout en assurant que chaque amas a une concavité plus petite qu'un seuil prédéfini. Une décomposition convexe exacte est un problème *NP-complet* et ne serait pas utilisable du fait d'une fragmentation trop importante des amas. Dans notre cas, nous avons besoin d'une répartition suffisante des points de vue (au moins 3) pour reconstruire notre modèle ce qui requiert une surface suffisamment large. Nous utilisons ainsi une bibliothèque de décomposition convexe approximative appelée V-HACD¹ pour segmenter notre maillage en plusieurs surfaces convexes. Selon Lien *et al.* [13], chaque maillage peut être approximé correctement comme union de surfaces convexes. En utilisant cette décomposition, notre modèle peut être généralisé aux objets non-plans en reconstruisant une quadrique pour chaque surface convexe. Nous illustrons cette décomposition sur la figure 4. Notre maillage est reconstruit à partir du scanner HandySCAN 3D de l'entreprise Creaform².

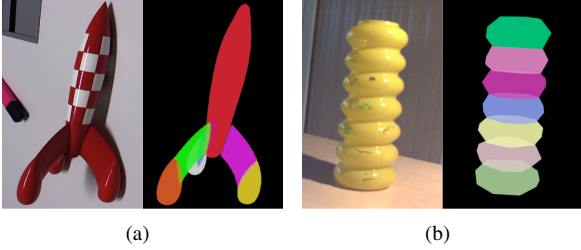


FIGURE 4 – Décomposition convexe approximative pour deux maillages présentés dans la figure 1. Notre réplique de fusée (a) est divisée en 7 morceaux convexes comme pour le vase jaune dans (b). Nous avons une spécularité par morceau convexe et par source de lumière, et nous reconstruisons une quadrique virtuelle à chaque fois.

5 Prédiction de spécularité pour de nouveaux points de vue

5.1 Principe

La prédiction de spécularité est obtenue dans JOLIMAS primal en projetant la quadrique résultante pour de nouveaux points de vue avec une pose donnée. Cependant, dans JOLIMAS dual, pour de nouveaux points de vue, la position du point le plus intense est requise pour calculer la caméra virtuelle. Cette caméra virtuelle sera utilisée par la

suite pour projeter la quadrique reconstruite de façon adéquate. Pour l'initialisation, nous utilisons une approximation du point le plus intense en partant du centre de la quadrique considéré comme étant la source ponctuelle \mathbf{L} , et projetons ce point par rapport à la pose de la caméra réelle sur le maillage en entrée.

5.2 Calcul du point le plus intense sur un maillage convexe par morceaux

Le point le plus intense est localisé à l'intersection de la ligne contenant la caméra virtuelle $\tilde{\mathbf{V}}$ et la source de lumière \mathbf{L} avec le plan tangent $\pi_{\mathbf{P}}$ d'une surface S au point \mathbf{P} tel que $\pi_{\mathbf{P}} = \begin{bmatrix} \hat{\mathbf{N}}(\mathbf{P}) \\ d \end{bmatrix}$. Ce point d'intersection \mathbf{P}_I est défini par :

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

À l'opposé des surfaces planes, cette ligne doit intersecter la surface S au même point utilisé que celui utilisé pour le calcul de la caméra virtuelle. Pour trouver ce point, nous résolvons :

$$\arg \min_{\mathbf{P} \in S} \delta(\mathbf{P}, \tilde{\mathbf{V}} + \frac{\tilde{\mathbf{V}}^\top \hat{\mathbf{N}}(\mathbf{P}) - d}{(\mathbf{L} - \tilde{\mathbf{V}})^\top \hat{\mathbf{N}}(\mathbf{P})} (\mathbf{L} - \tilde{\mathbf{V}})), \quad (4)$$

avec

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

et δ la distance euclidienne 3D. Le calcul du point le plus intense doit être réalisé pour chaque surface convexe et chaque source de lumière. En raison de la convexité du problème, nous résolvons ce problème de minimisation par descente de gradient. Ce procédé est illustré à la figure 5(a).

5.3 Interpolation de normale

Le calcul de la caméra virtuelle est très sensible aux variations de normale. Ainsi une interpolation de normale est utilisée entre les différents polygones du maillage. Cette procédure est similaire à l'interpolation utilisée par Phong [20]. Nous commençons à partir d'un point \mathbf{P} de normale inconnue $\hat{\mathbf{N}}(\mathbf{P})$ sur un polygone du maillage. Après avoir calculé les normales $\hat{\mathbf{N}}_i$ pour chaque sommet i du triangle, nous traçons une ligne incluant \mathbf{P} dans une direction donnée (l'axe y par exemple) qui intersecte plusieurs arrêtes du polygone. Pour ces points d'intersection, nous calculons leurs normales par interpolation linéaire. La normale finale $\hat{\mathbf{N}}(\mathbf{P})$ est calculée à partir de l'interpolation linéaire des normales des points d'intersection le long de la ligne, comme illustré à la figure 5(b). Cette procédure nous permet d'obtenir une normale plus précise.

6 Résultats expérimentaux

Dans cette section, la prédiction de spécularité est testée sur des données synthétiques et réelles pour évaluer les améliorations par rapport aux travaux antérieurs, la sensibilité

1. github.com/kmammou/v-hacd
2. www.creaform3d.com/en

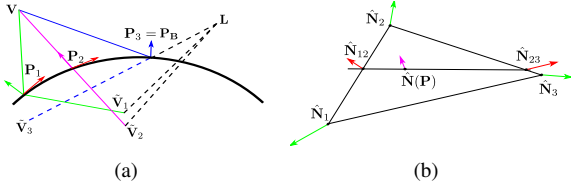


FIGURE 5 – Calcul du point le plus intense et interpolation de normale. Dans (a), notre algorithme pour calculer le point le plus intense est illustré. A partir d’une initialisation P_1 , nous calculons la position de la caméra virtuelle \tilde{V}_1 (symétrique de V) par rapport au plan tangent à la surface en P_1 et la ligne contenant \tilde{V}_1 et L . Puisque l’intersection entre la ligne et la surface S n’est pas égale à P_1 , nous itérons à la face adjacente dans la direction du point d’intersection. Ce procédé s’arrête quand nous convergeons sur une même face du maillage. Pour améliorer l’estimation sur la face restante, nous calculons une normale plus précise en utilisant une interpolation de normale dans (b). Dans cet exemple, à partir d’un polygone comportant trois sommets avec normales associées, la normale $\hat{N}(P)$ en P est calculée en interpolant les normales sur la ligne horizontale contenant P avec les arrêtes du polygone donnant \hat{N}_{12} et \hat{N}_{23} .

du modèle aux changements de courbure et la robustesse de notre méthode à un point d’intensité maximal bruité. La prédiction de spécularité est évaluée en utilisant la distance 2D entre les contours de la spécularité et l’ellipse prédite comme définie dans [22]. Cette erreur de prédiction est décrite en pixels pour les séquences de synthèse et correspond à la somme des distances entre chaque point de contour (spécularité) avec les point associés sur l’ellipse. Cette erreur est calculée pour chaque ellipse. Dans les séquences réelles, nous calculons l’erreur de prédiction en pourcentage en calculant la moyenne d’erreur par point de contour comparée à la taille de l’image.

6.1 Comparaison avec JOLIMAS primal

Nous utilisons une séquence comportant 200 images illustrée à la figure 6(a). Dans cette séquence, une surface plane est courbée progressivement afin d’évaluer la fonctionnalité de notre modèle dual sur des surfaces planes et non-plans. Pour chaque image, nous reconstruisons une quadrique à partir de 6 points de vue et nous calculons la distance 2D entre la spécularité prédite et la spécularité détectée. Notre méthode a une erreur moyenne de 11 pixels en comparaison à [18] avec 150 pixels d’erreur.

6.2 Sensibilité aux changements de courbure

La forme d’une spécularité varie en fonction des changements géométriques et plus particulièrement les changements de courbure. Il est difficile de prédire de façon précise pour un nouveau point de vue avec un changement de courbure de la scène. En utilisant notre séquence précédente, nous avons testé la sensibilité de notre méthode aux

changements de courbure en reconstruisant une quadrique à partir de la première image (surface plane) et en prédisant la spécularité pour toutes les images. Même si l’erreur augmente, l’information fournit par la prédiction reste cohérente avec une erreur relativement faible comme illustré à la figure 6.

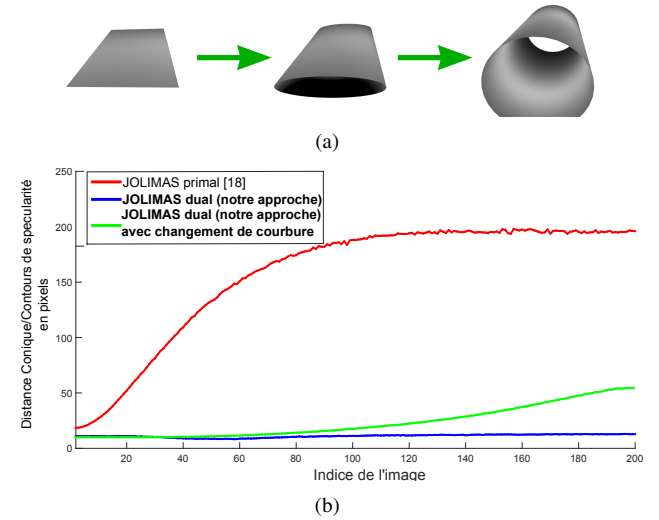


FIGURE 6 – Comparaison entre JOLIMAS primal et notre modèle JOLIMAS dual. Dans cette séquence de synthèse, la courbure augmente progressivement à chaque image comme illustré dans (a). En utilisant 6 poses de caméra, les deux modèles primal et dual sont reconstruits et comparés en utilisant la distance 2D entre la prédiction de spécularité et les contours de la spécularité détectée. Notre approche duale réalise de meilleurs résultats que [18] avec une erreur moyenne de 11 pixels par ellipse contrairement à une erreur de 150 pixels par ellipse en moyenne pour JOLIMAS primal. Nous testons également la prédiction de spécularité de notre modèle dual en reconstruisant une quadrique sur une surface plane (première image) et en l’utilisant pour les changements de courbure restants. Même si l’erreur croît pour chaque image, la prédiction de spécularité reste fiable avec une erreur relativement faible (maximum 50 pixels).

6.3 Prédiction de spécularité sur des données réelles

Pour les séquences réelles, des expériences qualitatives sont réalisées sur la prédiction de spécularité en comparant l’ellipse prédite avec les contours de la spécularité associée pour chaque surface présentant des spécularités et avec la prédiction de spécularité de JOLIMAS primal. Nous choisissons différents objets avec des courbures et matériaux variés sous différentes conditions de lumière. Chaque image des séquences a une taille de 1280×720 . La pose de la caméra est calculée à partir d’un SLAM contraint modèle de Loesch *et al.* [14]. L’ensemble des résultats est illustré sur la figure 7. Un résumé des résultats de notre modèle dual comparé au modèle primal donné par

le tableau 1.

Réplique de fusée. Dans cette séquence comportant 1410 images, une réplique de fusée en métal est illuminée par un néon. Comme illustré à la figure 4(a), notre maillage de la fusée est divisé en 7 morceaux convexes. Cependant, seulement 3 morceaux convexes sont utilisés car les pièces restantes ne présentent pas de spécularité. Nous atteignons une précision de 2.1% au niveau de l’erreur de prédiction en utilisant la distance contours/ellipse 2D de [22].

Mug. Pour cette séquence comportant 300 images, un mug en porcelaine est illuminée par 2 sources de lumière : une lampe de bureau et un néon différent de celui utilisé pour la séquence de la fusée. Notre méthode est capable de prédire les deux spécularités précisément avec une erreur de prédiction de 1.3% puisque la courbure de l’objet d’intérêt est constant. JOLIMAS primal est capable de prédire les spécularités avec la taille appropriée mais avec une mauvaise position.

Vase jaune. Dans cette séquence de 1900 images, un vase jaune en verre est illuminé par une lampe de bureau. Cette séquence est particulièrement difficile due à la symétrie de l’objet rendant l’estimation de la pose de caméra propice aux erreurs. De plus, les spécularités sur le vase ont une petite taille rendant le raffinement de la quadrique difficile. Cependant, nous arrivons à obtenir de bons résultats avec une erreur de prédiction de l’ordre de 1.9%. Cet objet est divisé en 7 morceaux convexes comme illustré sur la figure 4(b).

TABLE 1 – Validation empirique de notre modèle et sa capacité à prédire les spécularités dans les images pour 3 séquences réelles présentées à la figure 7. La distance 2D de [22] est utilisée pour calculer l’erreur de prédiction entre les contours de la spécularité et l’ellipse prédite. Cette erreur de prédiction correspond au ratio entre la moyenne de la distance point/ellipse et la taille de l’image. Pour des courbures et sources de lumière variées, notre méthode fournit des résultats précis en termes de prédiction de spécularité en comparaison de JOLIMAS primal.

Séquence	Erreur de prédiction (distance 2D en %)	
	Primal [18]	Dual (notre méthode)
Fusée	32.8	2.1
Mug	18.3	1.3
Vase	35.2	1.9

7 Discussion et Conclusion

Nous avons présenté un nouveau modèle géométrique multi-vues pour la forme d’une spécularité appelé JOLIMAS dual. Ce modèle est capable de prédire la forme d’une spécularité sur des surfaces planes mais également sur des surfaces convexes. Il étend JOLIMAS primal [18] sous l’hypothèse que les spécularités conservent une forme elliptique sur une surface convexe. Nous avons prouvé que

JOLIMAS primal est limité aux surfaces planes car il ne prend pas en compte les normales de la surface (constantes pour une surface plane). Nous savons que la forme d’une spécularité est fortement influencée par la géométrie de la surface sous-jacente. Nous soulignons l’importance d’utiliser un point spécifique appelée le point d’intensité maximale qui suit la loi de la réflexion en optique. A partir de ce point, nous calculons des caméras virtuelles orientées vers la source de lumière réelle et dont les ellipses associées aux spécularités respectent la géométrie épipolaire. Cependant, le calcul de ce point n’est pas trivial due à la présence de bruit dans les poses de caméra et dans les normales des maillages. Pour toutes ces raisons, nous avons implémenté une méthode d’estimation du point le plus intense respectant les règles de l’optique. Pour s’assurer de la généralité de notre modèle, nous utilisons une décomposition convexe approximative pour diviser le maillage en morceaux convexes. Notre modèle reconstruit une quadrique à partir des ellipses calculées à partir des spécularités en utilisant une représentation par caméras virtuelles. La projection perspective de cette quadrique dans les caméras virtuelles correspond aux spécularités dans les images. Nous avons testé notre modèle dans des contextes variés avec plusieurs sources de lumière et plusieurs objets. Afin de prédire chaque spécularité, nous reconstruisons une quadrique pour chaque morceau convexe et chaque source de lumière. Notre prédiction de spécularité a été évaluée en termes de robustesse et de précision sur des séquences synthétiques et réelles.

Nos travaux futurs s’intéresseront à l’étude de l’impact des changements de rugosité et de courbure sur les spécularités. Pendant la reconstruction de la quadrique, il serait judicieux d’inclure les changements de courbure et de rugosité car ces deux éléments impactent fortement la forme de la spécularité. Pour gérer ces changements de courbures, il serait intéressant de trouver une solution analytique reliant la courbure de la surface à la forme de la spécularité, proche des travaux de [1].

Références

- [1] Andrew Blake and Gavin Brelstaff. Geometry from specularities. In *International Conference on Computer Vision, ICCV, 1988*.
- [2] James F Blinn. Models of light reflection for computer synthesized pictures. In *Special Interest Group on Computer Graphics and Interactive Techniques, SIGGRAPH, 1977*.
- [3] Bastiaan J Boom, Sergio Orts-Escolano, Xin X Ning, Steven McDonagh, Peter Sandilands, and Robert B. Fisher. Point light source estimation based on scenes recorded by a rgb-d camera. In *British Machine Vision Conference, BMVC, 2013*.
- [4] Paul-Emile Buteau and Hideo Saito. Retrieving lights positions using plane segmentation with diffuse illumination reinforced with specular component. In *International Symposium on Mixed and Augmented Reality, ISMAR, 2015*.
- [5] Ju Yong Chang, Ramesh Raskar, and Amit Agrawal. 3d pose estimation and segmentation using specular cues. In *Computer Vision and Pattern Recognition, CVPR, 2009*.
- [6] Geoffrey Cross and Andrew Zisserman. Quadric reconstruction from dual-space geometry. In *International Conference on Computer Vision, ICCV, 1998*.

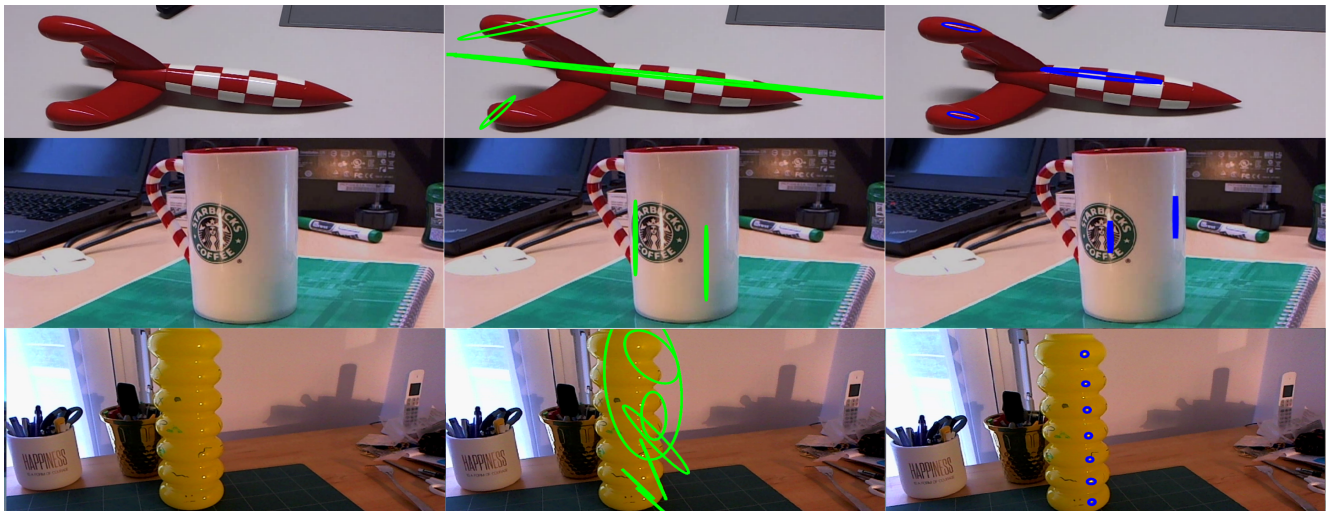


FIGURE 7 – Prédiction de spécularité sur des séquences réelles avec différentes conditions de lumière et pour différentes courbures. Sur l’image supérieure, sur une réplique de fusée illuminée par un néon, nous prédisons 3 spécularités précisément en utilisant JOLIMAS dual à droite contrairement à JOLIMAS primal (au centre) qui échoue en terme de position et de forme de spécularité prédite. Dans l’image du milieu, un mug, illuminé par une lampe de bureau et une lampe néon, est utilisé pour tester la capacité de notre modèle à prédire plusieurs sources de lumière de forme et d’intensité différentes. Pour cet objet, notre résultat est particulièrement précis due à la constance de la courbure de l’objet. JOLIMAS primal parvient à prédire correctement l’échelle des spécularités mais avec des mauvaises positions. En bas de la figure, nous utilisons un vase illuminé par une lampe de bureau. L’objet comporte 7 parties convexes, ce qui donne 7 spécularités. Cet objet est particulièrement difficile à cause de sa symétrie et de la petite taille des spécularités qui apparaissent sur celui-ci.

- [7] Farshad Einabadi and Oliver Grau. Discrete light source estimation from light probes for photorealistic rendering. In *British Machine Vision Conference, BMVC*, 2015.
- [8] Sunil Hadap Inso Kweon Hyeongwoo Kim, Hailin Jin. Specular reflection separation using dark channel prior. *CVPR*, 2013.
- [9] Jan Jachnik, Richard A. Newcombe, and Andrew J. Davison. Real-time surface light-field capture for augmentation of planar specular. In *International Symposium on Mixed and Augmented Reality, ISMAR*, 2012.
- [10] James T Kajiya. The rendering equation. In *Special Interest Group on Computer Graphics and Interactive Techniques, SIGGRAPH*, 1986.
- [11] Gudrun J Klinker, Steven A Shafer, and Takeo Kanade. A physical approach to color image understanding. *International Journal of Computer Vision*, 4(1) :7–38, 1990.
- [12] Pascal Lagerer and Pascal Fua. Using specularities to recover multiple light sources in the presence of texture. In *International Conference on Pattern Recognition, ICPR*, 2006.
- [13] Jyh-Ming Lien and Nancy M Amato. Approximate convex decomposition of polyhedra. In *Solid and Physical Modeling, SPM*, 2007.
- [14] Angélique Loesch, Steve Bourgeois, Vincent Gay-Bellile, and Michel Dhôme. Generic edgelet-based tracking of 3d objects in real-time. In *Intelligent Robots and Systems, IROS*, 2015.
- [15] Gucan Long, Laurent Kneip, Xin Li, Xiaohu Zhang, and Qifeng Yu. Simplified mirror-based camera pose computation via rotation averaging. In *Computer Vision and Pattern Recognition, CVPR*, 2015.
- [16] Maxime Meilland, Christian Barat, and Andrew 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.
- [17] Alexandre Morgand and Mohamed 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.
- [18] Alexandre Morgand, Mohamed Tamaazousti, and Adrien Bartoli. An empirical model for specular prediction with application to dynamic retexturing. In *International Symposium of Mixed and Augmented Reality, ISMAR*, 2016.
- [19] Aaron Netz and Margarita Osadchy. Recognition using specular highlights. *Pattern Analysis and Machine Intelligence*, 35(3) :639–652, 2013.
- [20] Bui Tuong Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6) :311–317, 1975.
- [21] Leo Reyes and Eduardo Bayro-Corrochano. The projective reconstruction of points, lines, quadrics, plane conics and degenerate quadrics using uncalibrated cameras. *Image and Vision Computing*, 23(8) :693–706, 2005.
- [22] Peter Sturm and Pau Gargallo. Conic fitting using the geometric distance. In *Asian Conference on Computer Vision, ACCV*, 2007.
- [23] Richter-Trummer Thomas, Kalkofen Denis, Park Jinwoo, and Schmalstieg Dieter. Instant mixed reality lighting from casual scanning. In *International Symposium of Mixed and Augmented Reality, ISMAR*, 2016.
- [24] Thomas Whelan, Renato F Salas-Moreno, Ben Glocker, Andrew J Davison, and Stefan Leutenegger. Elasticfusion : Real-time dense slam and light source estimation. *The International Journal of Robotics Research*, 2016.
- [25] Kwan-Yee K Wong, Dirk Schnieders, and Shuda Li. Recovering light directions and camera poses from a single sphere. In *European Conference on Computer Vision, ECCV*, 2008.
- [26] Edward Zhang, Michael F Cohen, and Brian Curless. Emptying, refurbishing, and relighting indoor spaces. In *Special Interest Group on Computer Graphics and Interactive Techniques, SIGGRAPH Asia*, 2016.