

# Un modèle de propagation de spécularité dans une vidéo pour la réalité diminuée

Souheil Hadj Said<sup>1,2</sup>

Mohamed Tamaazousti<sup>1</sup>

Adrien Bartoli<sup>2</sup>

<sup>1</sup> CEA, LIST, Point Courrier 173, F-91191 Gif-sur-Yvette, France

<sup>2</sup> Institut Pascal - UMR 6602 - CNRS/UCA/CHU, Clermont-Ferrand, France

souheil.hadjsaid@cea.fr ; mohamed.tamaazousti@cea.fr ; adrien.bartoli@gmail.com

## Résumé

En Réalité Diminuée (RD), une région cible de l'image est supprimée et remplacée par des patches du reste de l'image, la région source. Le résultat est ensuite recopié aux trames suivantes du flux vidéo. Une étape importante dans cette technique consiste à mettre à jour la région cible en fonction du changement de luminosité. Ce problème complexe est récurrent lors de changements de point de vue. Nous montrons que les méthodes de RD de l'état de l'art ne résolvent pas ce problème, même dans des cas apparemment simples (source lumineuse ponctuelle et surface plane). Nous nous appuyons sur les modèles d'illumination locale pour aborder ce problème. Au sens de ces modèles, il est admis que cette variation de luminosité n'impacte que la composante spéculaire de l'image. Dans le contexte de la RD, une solution est donc de propager la spécularité dans la région cible. Nous listons un ensemble de propriétés structurales de la spécularité. Ceci nous a mené à un nouveau modèle de propagation d'une spécularité qui intègre ces propriétés. Des résultats expérimentaux sur des données synthétiques et réelles montrent que notre modèle améliore la qualité des rendus par rapport aux solutions de l'état de l'art en RD.

## Mots Clef

Réalité diminuée, spécularité, propagation, rendu, isocontour, point le plus lumineux, variation de luminosité.

## Abstract

In Diminished Reality (DR), a target region from the image is deleted and replaced with patches from the rest of the image, the source region. The result is then propagated to the next frames of the video stream. One of the important stages of this technique is to update the target region with respect to the illumination change. This is a complex and recurrent problem when the viewpoint changes. We show that the state of the art in DR fails in solving this problem, even in apparently simple scenarios. We rely on local illumination models to address this problem. According to these models, the variation in illumination only affects the specular component of the image. In the context of DR, the problem is therefore solved by a means of propagating

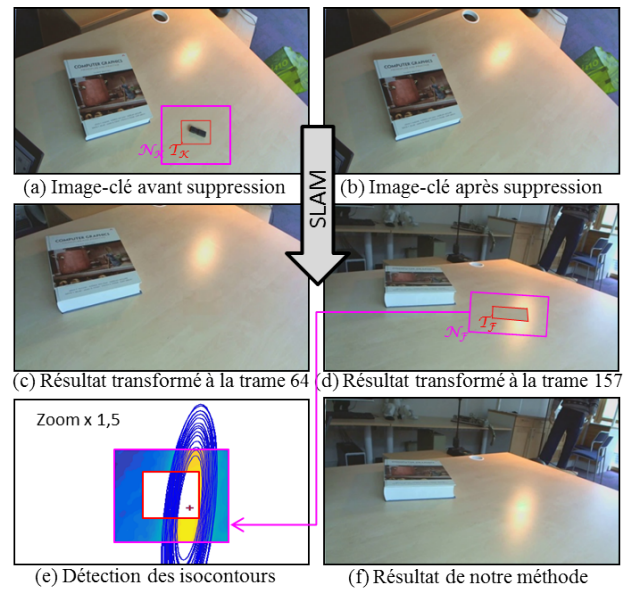


FIGURE 1 – Propagation de la spécularité dans la RD. (a) Trame 1 (Image-clé) et la région cible  $\mathcal{T}_K$  (en rouge). (b) Image-clé après suppression avec la RD. (c) Le résultat de (b) transformé à la trame 64. (d) Le résultat de (b) transformé à la trame 157 sans propagation de la spécularité. (e) Trame 157 rectifiée au plan image de la caméra. Détection et ajustement d'ellipses. (f) Résultat de notre méthode.

specularities in the target area. We list a set of structural properties of specularities. This leads us to a new model of specularity propagation that incorporates these properties. Experimental results on synthetic and real data show that our approach substantially improves the rendering quality compared to the state of the art in DR.

## Keywords

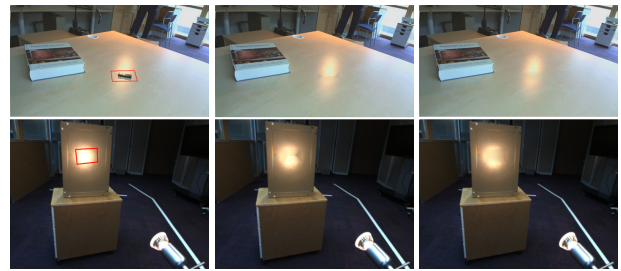
Diminished Reality, specularity, propagation, rendering, isocontour, brightest point, illumination variation.

## 1 Introduction

La Réalité Diminuée (RD) vise à supprimer un objet dans la scène à partir d'un flux vidéo en temps réel [10, 11, 21].

L'approche standard consiste à remplacer un objet par une texture synthétisée à partir d'éléments du reste de la scène. Cette technique peut être utilisée dans de nombreuses applications. Par exemple, il est possible de remeubler virtuellement l'intérieur d'une pièce en ajoutant ou supprimant des meubles [18], dans des applications mobiles de Réalité Augmentée (RA), les panneaux de signalisation peuvent être retirés pour la simulation d'un paysage et les marqueurs RA peuvent être masqués pour obtenir une fusion plus réaliste entre les objets virtuels et le monde réel [12, 14]. En RD, la qualité de l'expérience utilisateur dépend fortement de la qualité du rendu. Pour l'images-clé, la complétion d'image [2] permet de remplacer de manière cohérente la région cible. Pour les images suivantes, nous devons mettre à jour cette région en recalant l'image mais aussi tout en respectant la variation de la luminosité. Cette étape est cruciale afin d'obtenir une qualité de rendu photo-réaliste. En effet, le problème de la variation de la luminosité peut apparaître lors du changement de point de vue. C'est pourquoi toute approche de RD doit avoir une étape pour résoudre ce problème. Parmi les approches de l'état de l'art [10, 11, 14, 18], seulement deux proposent, explicitement, une méthode pour mettre à jour la luminosité [10, 11]. Herling *et al.* calculent les variations en intensité sur les points de la bordure de la région cible. Ils utilisent une fonction continue pour propager ces variations à l'intérieur. Kawai *et al.* proposent une approche par patches. Les variations des patches du voisinage sont utilisés comme des contraintes de limites dans une équation de minimisation qui sert à calculer les variations des patches de la région cible. Ces deux travaux proposent des techniques d'interpolation heuristiques qui suggèrent uniquement la continuité de la variation de la luminosité à travers une spécularité. Ainsi, nous montrons sur des exemples réels en figure 2 que cette hypothèse est insuffisante en terme de qualité des rendus. Comme on le voit sur ces exemples, les méthodes de l'état de l'art génèrent des artefacts, même dans des scénarios apparemment simples (une surface plane et une source de lumière ponctuelle).

Dans cet article, nous soulignons que cette variation est due principalement à la composante spéculaire dans l'image. Il est d'ailleurs admis dans les modèles d'illumination locale comme Phong [17] et Blinn-Phong [5] que seule la composante spéculaire dépend de l'angle de vue. Le problème revient donc à propager la spécularité à l'intérieur de la région cible. Il a également été démontré que la spécularité joue un rôle important dans la perception de la scène par le cerveau humain [4]. Ainsi, nous analysons dans un premier temps la structure de la spécularité, et en nous basant sur nos observations, nous proposons un modèle qui prend en compte de manière explicite ses propriétés structurales. Dans le cas simple où l'état de l'art échoue (surface plane et source de lumière ponctuelle), notre approche améliore considérablement la qualité des rendus. Notre méthode d'estimation utilise une interpolation par les fonctions plaques minces ("Thin Plate Splines" ou TPS) qui est



(a) Les images origi- (b) Méthode de Her- (c) Méthode de Ka-  
nales ling *et al.* [10] wai *et al.* [11]

FIGURE 2 – Illustration des limites des méthodes précédentes. (a) L'image originale et la zone cible (en rouge). (b) Résultats de la méthode de Herling *et al.* [10]. (c) Résultats de la méthode de Kawai *et al.* [11].

une représentation de fonction lisse. Ensuite, nous contrainsons cette fonction par deux contraintes principales qui sont l'ellipticité des isocontours d'intensité lumineuse et l'existence d'un unique point le plus lumineux.

La section 2 décrit les principales propriétés structurales d'une spécularité. La section 3 présente le problème que nous cherchons à résoudre. La section 4 présente le modèle et l'algorithme proposés. Enfin, la section 5 est une discussion de nos résultats expérimentaux.

## 2 Les propriétés structurales de la spécularité

En observant des images de reflets spéculaires sur des surfaces planes en présence d'une source lumineuse ponctuelle telles que celles présentées dans la figure 3, nous avons établi certaines propriétés structurales de la spécularité :

- **Propriété 1 : continuité.** La variation de la luminosité est lisse et donc continue.
- **Propriété 2 : point le plus lumineux.** La spécularité a un seul point le plus lumineux, situé approximativement en son centre.
- **Propriété 3 : ellipticité.** Les isocontours d'intensité d'une spécularité sont de forme elliptique.
- **Propriété 4 : monotonie.** Plus on s'éloigne du point le plus lumineux, plus l'intensité est basse. Cela implique que les isocontours ne se croisent pas.

Certaines de ces propriétés ont été théoriquement et expérimentalement vérifiées sur des modèles de synthèse d'image (notamment Phong [17] et Blinn-Phong [5]) dans [16]. Cependant, elles n'ont pas été considérées dans la propagation des spécularités en RD dans les méthodes existantes [10, 11]. Nous les exploitons afin d'améliorer le rendu de la spécularité dans la RD.

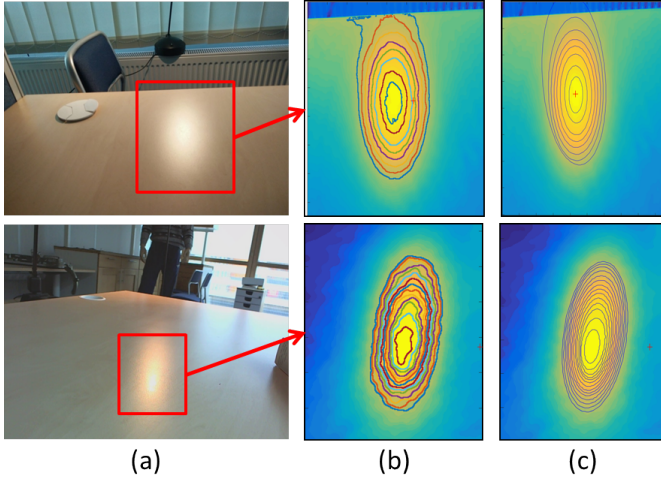


FIGURE 3 – Dans (a) les images montrent des reflets spéculaires sur une surface plane. Dans (b), nous montrons la carte de luminosité correspondante autour de la spécularité pour chaque image et les isocontours de l'intensité. Dans (c), nous montrons les ellipses ajustées à ces isocontours.

### 3 Contexte et énoncé du problème

#### 3.1 Notation

Les scalaires sont en italique (par exemple  $x$ ), les vecteurs en gras (par exemple  $\mathbf{v}$ ) et les matrices en sans-serif (par exemple  $M$ ). Les éléments d'un vecteur sont écrits comme dans  $\mathbf{a}^\top = (a_1 \ a_2 \ a_3)$  où  $^\top$  est la transposée des vecteurs et des matrices. Les coordonnées d'un point de l'image sont écrites avec un vecteur  $\mathbf{p}^\top = (x \ y)$ . Un domaine d'image est écrit en police calligraphique en majuscule (par exemple  $\mathcal{R}$ ). Un groupe de points est écrit avec des lettres italiques en majuscule (par exemple  $B$ ) et le nombre de points dans un groupe par  $|B|$ . Les fonctions sont écrites en lettres grecques droites (par exemple  $\psi$ ) ou minuscules latines en italique.

#### 3.2 Énoncé du problème

**Contexte.** Nous présentons deux techniques principales utilisées en RD. Tout d'abord, nous expliquons la technique de complétion d'image, aussi appelée "inpainting". Une image  $\chi$  peut être mathématiquement définie comme :

$$\chi : \begin{cases} \mathcal{O} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^n \\ \mathbf{p} \rightarrow \chi(\mathbf{p}), \end{cases} \quad (1)$$

où  $\mathbf{p}$  représente un vecteur indiquant les coordonnées spatiales d'un pixel. Dans le cas d'une image couleur, chaque pixel porte trois composantes de couleur ( $n = 3$ ) définies dans l'espace de couleur  $RVB$ . Alors,  $\chi$  peut être écrit comme  $\chi^\top = (\chi^R \ \chi^G \ \chi^B)$ . Le terme "inpainting" a été introduit par Bertalmio *et al.* [3]. Depuis, de nombreux articles ont proposé des techniques d'inpainting en temps réel. Elles peuvent être classées en deux groupes : *diffusion-based inpainting* [3, 1] et *example-based inpainting* [2, 7]. En général, dans le problème d'inpainting,

l'image  $\chi$  (*i.e.* chaque canal de l'espace couleur) subit une opération de dégradation qui retire une partie de l'image. En conséquence, la définition générique du domaine  $\mathcal{O}$  de l'image originale  $\chi$  peut être vu comme composée de deux parties  $\mathcal{O} = \mathcal{S} \cup \mathcal{T}$ ,  $\mathcal{S}$  étant la partie connue de  $\chi$  (la région source) et  $\mathcal{T}$  la partie inconnue de  $\chi$  qu'on cherche à estimer (la région cible). Le but de l'inpainting est d'estimer les composantes de couleur des pixels  $\mathbf{p}$  situés dans la région cible  $\mathcal{T}$ , à partir des pixels situés dans  $\mathcal{S}$ , afin de reconstruire finalement l'image "inpaintée" que nous désignons par  $\hat{\chi}$ . L'objectif en termes de qualité est que la région retrouvée paraisse naturelle à l'œil humain. La deuxième technique utilisée est le SLAM, "Simultaneous Localization and Mapping". On désigne un point 3D par  $\mathbf{x} \in \mathbb{R}^3$ , la rotation de la caméra par  $R \in \mathbb{SO}(3)$  et sa translation par  $\mathbf{t} \in \mathbb{R}^3$ . Pour chaque trame  $f$ , le SLAM détermine les coefficients de  $R$  et de  $\mathbf{t}$  qui projettent de manière cohérente le point 3D  $\mathbf{x}$  sur le plan image de la caméra. Nous utilisons la technique SLAM de [20] qui permet de résoudre ce problème en temps réel.

**La RD comme est un problème de complétion spatio-temporel.** Nous considérons la RD comme un problème de complétion spatio-temporel. Pour l'image-clé, la cohérence spatiale est assurée par la technique d'inpainting. Pour les trames suivantes de la vidéo, le SLAM propage le résultat obtenu à l'image-clé tout en assurant la cohérence temporelle dans la vidéo. Cependant, une spécularité peut apparaître autour de  $\mathcal{T}$ , ce qui provoque des variations de luminosité. Ainsi, la structure spatiale du résultat d'inpainting doit être correctement mise à jour. Cette condition est essentielle pour obtenir un résultat de rendu réaliste. Nous nous référons à l'image-clé par  $\chi_K$  et l'image actuelle par  $\chi_F$ . Nous appliquons sur  $\chi_K$  une technique d'inpainting similaire à PatchMatch [2] pour reconstruire la région cible  $\mathcal{T}_K$ . Ensuite, nous utilisons le SLAM pour transformer les trames suivantes au plan de l'image-clé. Pour chaque pixel  $\mathbf{p}_K$  dans l'image-clé, nous définissons un pixel correspondant dans l'image actuelle  $\mathbf{p}_F = \eta(\mathbf{p}_K)$ ,  $\eta$  étant une homographie. Nous définissons la région voisine  $\mathcal{N}_K \subset \mathcal{S}$  centrée autour de  $\mathcal{T}_K$  avec la largeur  $w_{\mathcal{N}} = z w_{\mathcal{T}}$  et de hauteur  $h_{\mathcal{N}} = z h_{\mathcal{T}}$ .  $z$  dépend à la fois de la taille de la spécularité ainsi que la taille du plan observé dans l'image. Nous nous proposons de propager la variation de la luminosité du voisinage  $\mathcal{N}_F$  vers l'intérieur de  $\mathcal{T}_F$ . Pour chaque pixel  $\mathbf{p}_F$  dans  $\mathcal{N}_F$ , le coefficient de variation de luminosité est défini comme suit :

$$v_{\mathbf{p}_K} = \frac{\chi_F(\mathbf{p}_F)}{\chi_K(\mathbf{p}_K)}. \quad (2)$$

Pour chaque trame  $f$ , connaissant les coefficients de variation d'intensité des pixels dans  $\mathcal{N}_F$ , nous estimons la fonction  $\psi_F$  qui retourne la variation d'illumination pour tous les pixels dans  $\mathcal{T}_F \cup \mathcal{N}_F$  et ainsi, nous mettons à jour leurs intensités :

$$\hat{\chi}_F(\mathbf{p}_F) = \psi_F(\mathbf{p}_K) \hat{\chi}_K(\mathbf{p}_K), \quad \forall \mathbf{p}_F \in \mathcal{T}_F \cup \mathcal{N}_F. \quad (3)$$



### 3.3 Les fonctions TPS

Nous présentons brièvement la paramétrisation du TPS. En entrée, nous considérons un ensemble de  $l$  centres  $(\mathbf{c}_k, u_k)$  où  $\mathbf{c}_k \in \mathbb{R}^2$  contient les coordonnées d'un centre et  $u_k \in \mathbb{R}$  est sa valeur cible correspondante. Nous définissons la matrice de coordonnées des centres  $\mathbf{C} = (\mathbf{c}_1 \cdots \mathbf{c}_l)$  et le vecteur cible des centres  $\mathbf{u} = (u_1 \cdots u_l)$ . Le TPS est une fonction lisse de  $\mathbb{R}^2$  vers  $\mathbb{R}$  entraînée par ces centres et donnée pour tout point  $\mathbf{p}$  par :

$$\phi_{\text{tps}}(\mathbf{p}; \mathbf{u}) = \mathbf{I}_{\mathbf{p}}^{\top} \mathbf{E}_{\lambda} \mathbf{u}, \forall \mathbf{p} \in \mathbb{R}^2, \quad (4)$$

où  $\mathbf{I}_{\mathbf{p}}^{\top} = \left( \rho(d^2(\mathbf{c}_1, \mathbf{p})) \cdots \rho(d^2(\mathbf{c}_l, \mathbf{p})) \right)$  avec  $\rho(d) = d \log(d)$  le noyau TPS pour la distance euclidienne.  $\mathbf{E}_{\lambda}$  est la matrice de paramétrisation, et intègre un poids de régularisation interne  $\lambda \in \mathbb{R}^+$  [6, 8].

En pratique, nous avons des centres arbitraires avec des valeurs cibles inconnues  $\mathbf{u}$ . Alors, étant donné  $m$  points de données  $(\mathbf{q}_i, v_i)$ , nous estimons les valeurs cibles en minimisant l'erreur suivante :

$$\min_{\mathbf{u}} \sum_{i=1}^m (\phi_{\text{tps}}(\mathbf{q}_i; \mathbf{u}) - v_i)^2. \quad (5)$$

Il s'agit d'un problème des moindres carrés linéaires, que nous résolvons avec une matrice pseudo-inverse.

## 4 Modèle et méthode proposés

### 4.1 Description

La variation de luminosité peut être considérée comme l'évolution d'une carte d'altitude, comme le montre la figure 4. La base de cette carte est représentée par les coordonnées des pixels de l'image et la hauteur par la valeur de variation. La modélisation de la variation de luminosité par une fonction lisse permet de respecter la propriété 1. Notons que cette propriété a été exploitée par Kim *et al.* [13] pour segmenter les réflexions spéculaires en une seule image. Cette fonction est altérée par la suite afin de

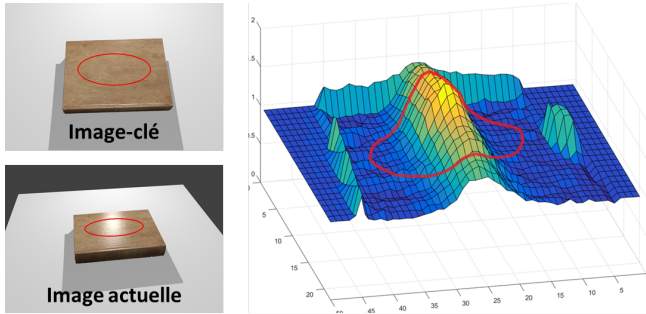
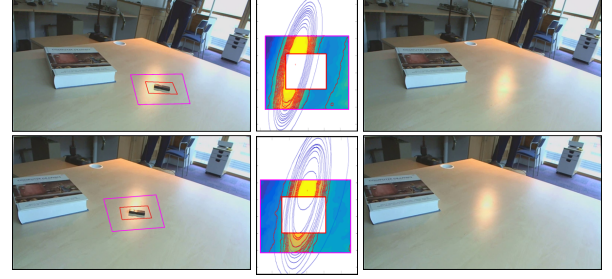


FIGURE 4 – La variation de luminosité dans le plan image entre l'image-clé  $K$  et la trame courante  $F$  est vue comme une carte d'altitude. Dans cet exemple, une spécularité traverse la région cible  $\mathcal{T}_F$ . Ceci démontre la propriété de continuité des reflets spéculaires.



(a) Les images origi- (b) Ajustement (c) Résultats après ajustement d'ellipses

FIGURE 5 – Estimation des isocontours sur la carte de luminosité. En (b), nous représentons en rouge les points utilisés pour estimer les ellipses et en bleu l'estimation finale. Les ellipses interpolées peuvent s'étendre à l'intérieur de la région cible et ceci nous permet de propager la specularité.

respecter quelques contraintes. En outre, nous considérons que les isocontours d'intensité lumineuse ont une forme elliptique et qu'il y a un unique point le plus lumineux.

### 4.2 Estimation

**La fonction lisse.** Nous utilisons la TPS pour représenter la fonction  $\psi_F$  qui retourne la variation de luminosité pour tous les pixels dans  $\mathcal{N}_F \cup \mathcal{T}_F$  :

$$\psi_F : \begin{cases} \mathcal{O} \subset \mathbb{R}^2 \rightarrow \mathbb{R} \\ \mathbf{p} \rightarrow \begin{cases} v_{\mathbf{p}_K} = \frac{\chi_F(\mathbf{p}_F)}{\chi_K(\mathbf{p}_K)}, \text{ si } \mathbf{p}_F \in \mathcal{N}_F \\ \phi_{\text{tps}}(\mathbf{p}_K; \mathbf{u}), \text{ if } \mathbf{p}_F = \eta_F(\mathbf{p}_K) \in \mathcal{T}_F \end{cases} \end{cases} \quad (6)$$

Nous considérons une grille uniformément distribuée  $\mathcal{G}_{\mathcal{F}} \subset \mathcal{N}_{\mathcal{F}} \cup \mathcal{T}_{\mathcal{F}}$ . On fixe la grille pour avoir  $r^2 = l$  nœuds. En utilisant le paramétrage de la TPS présenté dans la section 3.3, nous considérons les nœuds de la grille comme les centres  $\mathbf{c}_k$  et les variations d'intensité des pixels entre l'image-clé et la trame courante comme valeur cible  $u_k$ . Les points dans  $\mathcal{N}_F$  sont considérés comme les points de données utilisés pour estimer  $\mathbf{u}$ . En utilisant la méthode d'estimation de la section 3.3, on obtient la fonction TPS  $\phi_{\text{tps}}$ . Le poids de régularisation interne est fixé à  $\lambda = 10^{-1}$ ,  $l = 100$  et  $m$  dépend du nombre de pixels que nous avons dans la région voisine. En termes de calcul, cette méthode nécessite un ajustement de moindres carrés linéaires à chaque trame pour minimiser les critères dans (5). Sachant que la matrice  $\mathbf{E}_{\lambda}$  est constante, résoudre  $\psi_F$  revient à minimiser (5) avec une simple multiplication entre une matrice connue et le vecteur mesuré de variations  $\mathbf{v}^{\top} = (v_1 \dots v_m)$ . Dans l'espace de couleur RVB, nous avons besoin d'estimer séparément trois rapports de variation d'intensité pour chaque pixel  $\mathbf{p}_F$ ,  $\mathbf{E}_{\lambda}$  étant la même pour les trois canaux de couleur. La TPS a déjà été utilisée dans [19] pour modéliser un changement de luminosité dans le contexte du recalage d'image.

**Détection des isocontours.** Pour évaluer la variation de la luminosité, nous convertissons l'image vers l'espace couleur *Lab* et considérons seulement le canal *L*



(Lightness). Un algorithme de détection des reflets spéculaires [15, 13] a été exploité pour vérifier si une spécularité traverse la région cible. On applique le filtre de Wiener [22] pour segmenter les niveaux de luminosité et réduire le bruit généré par la rugosité des surfaces planes. Ensuite, nous utilisons un histogramme de quantification pour segmenter l'image en niveaux d'intensité lumineuse. Le résultat est la carte de luminosité. Nous définissons  $s$  isocontours de reflets spéculaires et estimons leurs points correspondants.

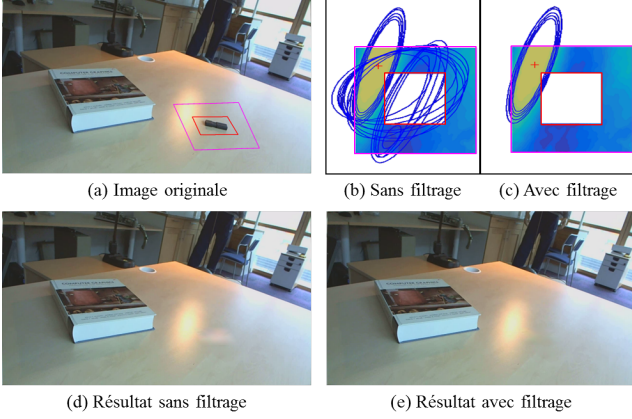


FIGURE 6 – Comparaison des résultats d'ajustement des ellipses sur la même trame de la vidéo sans et avec filtrage. Dans (b), avant filtrage, de nombreuses fausses estimations d'ellipses sont observées. Cela génère des artefacts visuels sur le résultat final (d). Dans (c), nous supprimons les fausses estimations. En (e), le résultat final de notre modèle. Notons que la carte de luminosité en (b) et (c) est rectifiée au plan image de l'image-clé et agrandie pour une meilleure visualisation.

**Ajustement d'ellipses.** En considérant la carte de luminosité obtenue dans la région voisine  $\mathcal{N}_F$ , comme le montre la figure 5, les points des isocontours qui sont en dehors de  $\mathcal{T}_F$  peuvent être extrapolés en ellipses s'étendant dans la région cible. Notre but ici est d'estimer l'extension de ces isocontours en supposant qu'ils ont une forme elliptique. Ceci équivaut à minimiser le critère suivant :

$$\min_E \sum_{j=1}^r \left( h - \chi^L(E(j)) \right)^2, \quad (7)$$

où  $\chi^L$  retourne la valeur du  $L$  (Lightness) dans l'espace couleur  $Lab$  d'un pixel et  $h$  l'intensité de l'isocontour. Nous utilisons l'algorithme de Fitzgibbon *et al* [9] pour ajuster la meilleure ellipse à l'isocontour. Ceci permet non seulement de trouver tous les points de l'isocontour qui se trouvent dans la région cible, mais aussi la position du centre de l'ellipse. Pour un isocontour  $C$ , le rapport de variation de luminosité pour un canal de couleur  $z$  est calculé à partir des points dans  $C \cap \mathcal{N}_F$  :

$$v_{\mathbf{p}}^z = \frac{\sum_{\mathbf{q} \in C \cap \mathcal{N}_F} v_{\mathbf{q}}^z}{|C \cap \mathcal{N}_F|}, \forall \mathbf{p} \in C \cap \mathcal{T}_F. \quad (8)$$

avec  $C$  l'ensemble des points de l'isocontour réel et  $v_{\mathbf{p}}^z$  le rapport de la variation de luminosité dans le pixel  $\mathbf{p}$ . Par référence à l'estimation de la TPS, nous ajoutons les points des ellipses qui se trouvent à l'intérieur de la région cible comme points de données. Leurs valeurs cibles respectives sont les moyennes des rapports de variation d'intensité des points connus des isocontours réels. Nous considérons, également, le centre de l'ellipse la plus petite comme étant le point le plus lumineux. Sa valeur cible est l'intensité lumineuse maximale dans l'image.

**Filtrage des fausses estimations.** Des fausses estimations peuvent survenir quand le nombre de points de l'isocontour est trop petit ou quand une grande portion de l'isocontour est à l'intérieur de la région cible. Pour assurer une estimation robuste des isocontours, nous aurons à respecter la condition suivante : *chaque ellipse de niveau d'intensité  $h_e$  doit être totalement à l'intérieur d'une ellipse de niveau d'intensité  $h < h_e$* . Ceci revient à définir le coefficient de confiance  $\gamma(e)$  :

$$\gamma(e) = \sum_{j=1, j \neq e}^s \delta(j, e) \quad (9)$$

où :

$$\delta(j, e) = \begin{cases} 1(E_j \subset E_e) & \text{si } h_j > h_e \\ 1(E_e \subset E_j) & \text{si } h_j < h_e \end{cases} \quad (10)$$

Après l'initialisation, nous calculons les coefficients de confiance des différents contours estimés et nous conservons le plus grand ensemble d'ellipses cohérentes (avec les valeurs de confiance maximales). Nous nous référons à cette opération de filtrage dans la figure 6. Nous raffinons ensuite les ellipses extraites en minimisant l'équation (7).

## 5 Résultats et discussion

Nous considérons une vidéo synthétisée (figure 7) et trois vidéos réelles (figures 8, 9 et 10). Dans les deux premières vidéos, aucun objet indésirable n'est placé dans la région cible afin de comparer la spécularité rendue à la vérité terrain. Dans les deux autres, nous nous positionnons dans le cas pratique de la RD. La méthode de Kawai *et al.* [11] fonctionne bien avec des variations de luminosité faibles, en particulier lorsque le point le plus lumineux est en dehors de la région cible. Cependant, elle génère des artefacts lorsqu'une spécularité entre dans la région cible et avec des variations d'intensité élevées. La méthode d'Herling *et al.* [10] est très dépendante de la forme de la frontière entre la région cible et la région source. Par conséquent, le résultat de rendu de spécularité n'est pas convaincant visuellement. Pour l'ensemble des résultats expérimentaux, notre méthode réussit à imiter l'apparence réelle des reflets spéculaires. La qualité visuelle de la spécularité rendue par notre méthode est ainsi la meilleure. Ceci est dû au fait que l'interpolation des coefficients de variation de luminosité dans les approches précédentes ne respecte pas la structure de la spécularité. Notre méthode ainsi que celles de l'état

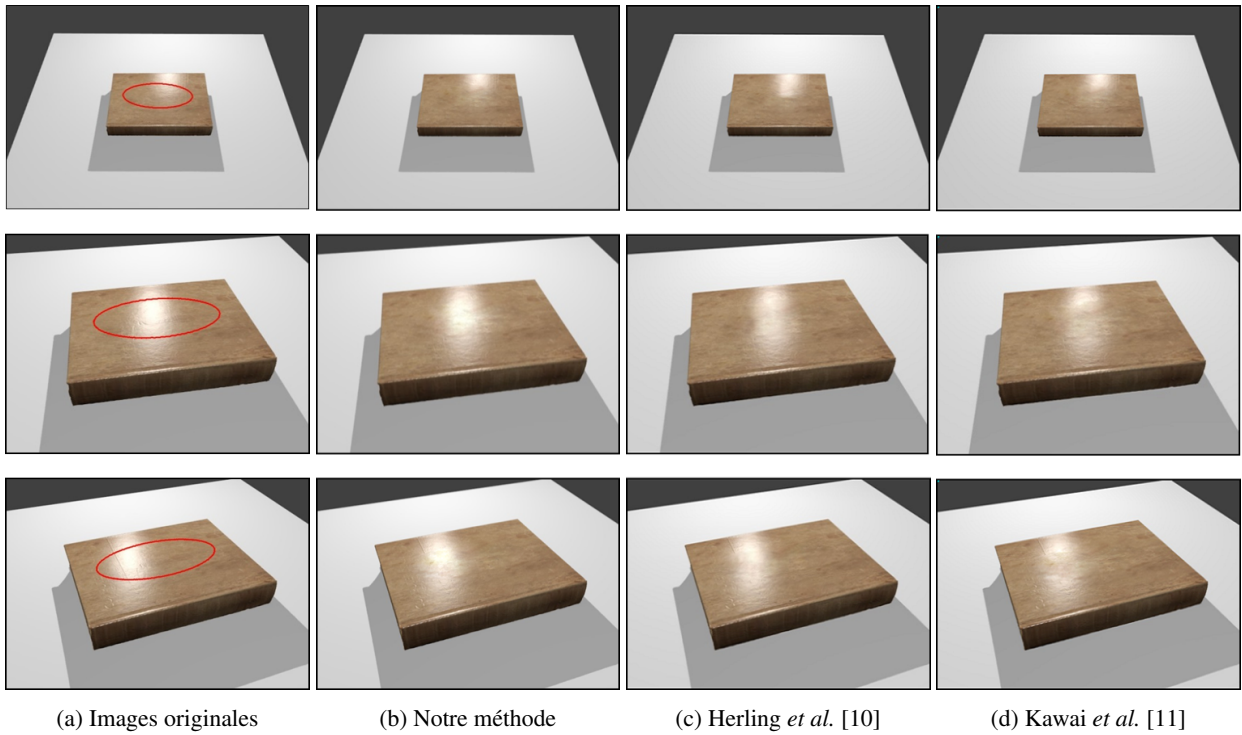
de l'art ne considèrent pas les surfaces rugueuses et de géométrie complexe. Dans ces cas, la propriété de continuité n'est en fait plus valide. Aussi, un cas limite est lorsque une spécularité rentre complètement à l'intérieur de la région cible. Pour résoudre ça, il faut un modèle de prédiction de spécularité comme celui présenté dans [16].

## 6 Conclusion

Cet article aborde le problème de variation de la luminosité lors du changement de point de vue dans le contexte de la Réalité Diminuée. Nous soulignons que c'est un problème difficile et qu'il n'a pas été proprement résolu par les méthodes précédentes. Nous notons, également, que ce problème peut être reformulé en un problème de propagation de spécularité. Ainsi, à partir d'observations empiriques, nous proposons une liste des propriétés structurales d'une spécularité. Nous proposons ensuite un nouveau modèle qui intègre ces propriétés pour estimer la variation de la luminosité. Les résultats rapportés dans ce document montrent la pertinence de cette approche par rapport aux travaux précédents, même sous des conditions pratiques apparemment simples. La cohérence spatio-temporelle du résultat de rendu final s'en trouve améliorée. Notre modèle instaure une base pour résoudre ce problème. Elle peut être étendue pour des conditions plus générales, et plus complexes. Dans nos travaux futurs, nous prévoyons introduire d'autres contraintes induites par la forme de la spécularité, le matériau et la rugosité de la surface.

## Références

- [1] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. *IEEE transactions on image processing*, 10(8) :1200–1211, 2001.
- [2] C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch : A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics (TOG)*, 28(3) :24, 2009.
- [3] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *SIGGRAPH*, pages 417–424. ACM, 2000.
- [4] A. Blake and H. Bülthoff. Does the brain know the physics of specular reflection? *Nature*, 343(6254) :165–168, 1990.
- [5] J. F. Blinn. Models of light reflection for computer synthesized pictures. In *SIGGRAPH*, pages 192–198. ACM, 1977.
- [6] F. L. Bookstein et al. Principal warps : Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence (PAMI)*, 11(6) :567–585, 1989.
- [7] A. Criminisi, P. Perez, and K. Toyama. Object removal by exemplar-based inpainting. In *CVPR*, pages 714–721. IEEE, 2003.
- [8] J. Duchon. Splines minimizing rotation-invariant semi-norms in sobolev spaces. In *Constructive Theory of Functions of Several Variables*, pages 85–100. Springer, 1977.
- [9] A. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least square fitting of ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 21(5) :476–480, 1999.
- [10] J. Herling and W. Broll. High-quality real-time video inpainting with pixmix. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, page 1, 2014.
- [11] N. Kawai, T. Sato, and N. Yokoya. Diminished reality based on image inpainting considering background geometry. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 2015.
- [12] N. Kawai, M. Yamasaki, T. Sato, and N. Yokoya. Ar marker hiding based on image inpainting and reflection of illumination changes. In *ISMAR*, pages 293–294. IEEE, 2012.
- [13] H. Kim, H. Jin, S. Hadap, and I. Kweon. Specular reflection separation using dark channel prior. In *CVPR*, pages 1460–1467, 2013.
- [14] O. Korkalo, M. Aittala, and S. Siltanen. Light-weight marker hiding for augmented reality. In *ISMAR*, pages 247–248. IEEE, 2010.
- [15] A. Morgand and M. Tamaazousti. Generic and real-time detection of specular reflections in images. In *VISAPP*, pages 274–282. IEEE, 2014.
- [16] A. Morgand, M. Tamaazousti, and A. Bartoli. An empirical model for specular prediction with application to dynamic retexturing. In *ISMAR*. IEEE, 2016.
- [17] B. T. Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6) :311–317, 1975.
- [18] S. Siltanen. Diminished reality for augmented reality interior design. *The Visual Computer*, pages 1–16, 2015.
- [19] G. Silveira and E. Malis. Unified direct visual tracking of rigid and deformable surfaces under generic illumination changes in grayscale and color images. *International journal of computer vision*, 89(1) :84–105, 2010.
- [20] M. Tamaazousti, V. Gay-Bellile, S. Collette, S. Bourgeois, and M. Dhome. Nonlinear refinement of structure from motion reconstruction by taking advantage of a partial knowledge of the environment. In *CVPR*, pages 3073–3080. IEEE, 2011.
- [21] P. Tiefenbacher, M. Sirch, and G. Rigoll. Mono camera multi-view diminished reality. In *WACV*, pages 1–8. IEEE, 2016.
- [22] N. Wiener. *Extrapolation, interpolation, and smoothing of stationary time series*, volume 2. MIT press Cambridge, MA, 1949.



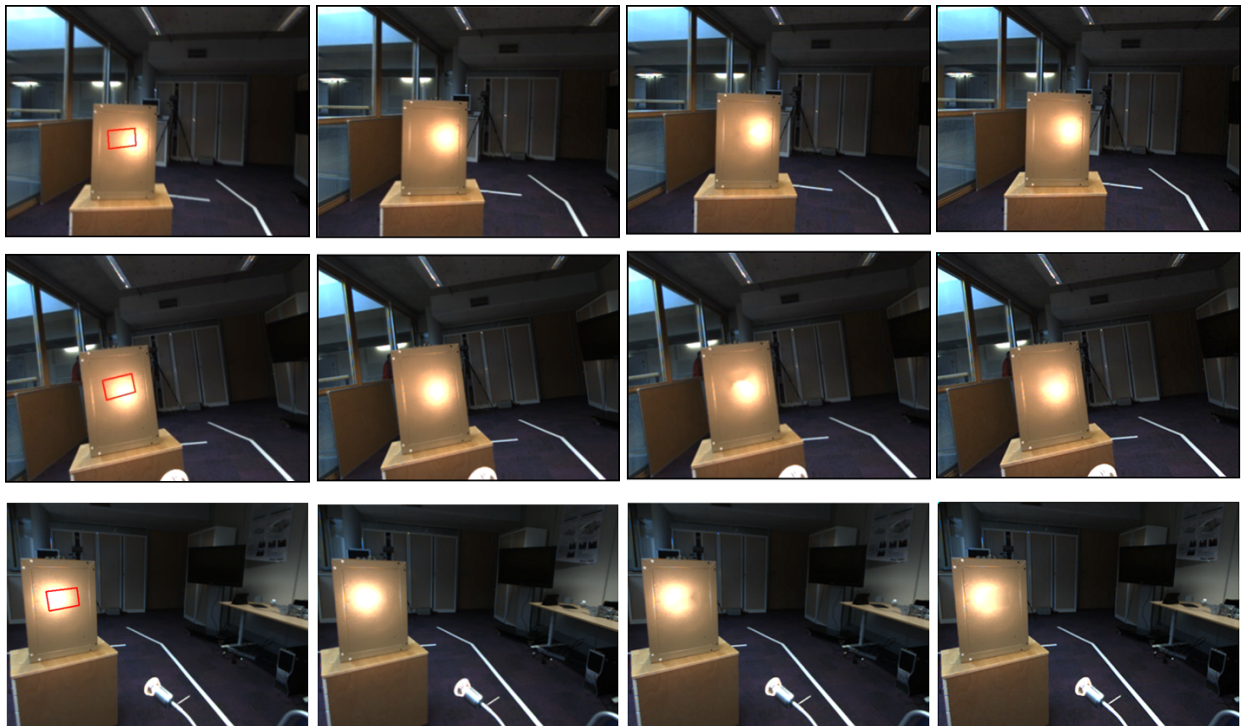
(a) Images originales

(b) Notre méthode

(c) Herling *et al.* [10]

(d) Kawai *et al.* [11]

FIGURE 7 – Résultats des différentes méthodes sur les trames 64, 71 et 130 de la vidéo 1.



(a) Images originales

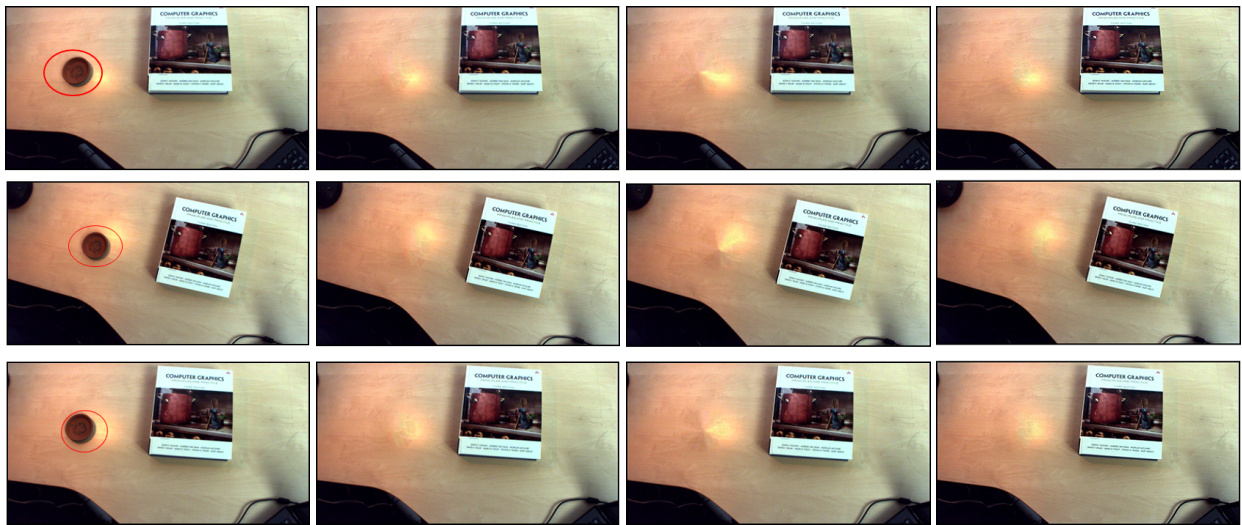
(b) Notre méthode

(c) Herling *et al.* [10]

(d) Kawai *et al.* [11]

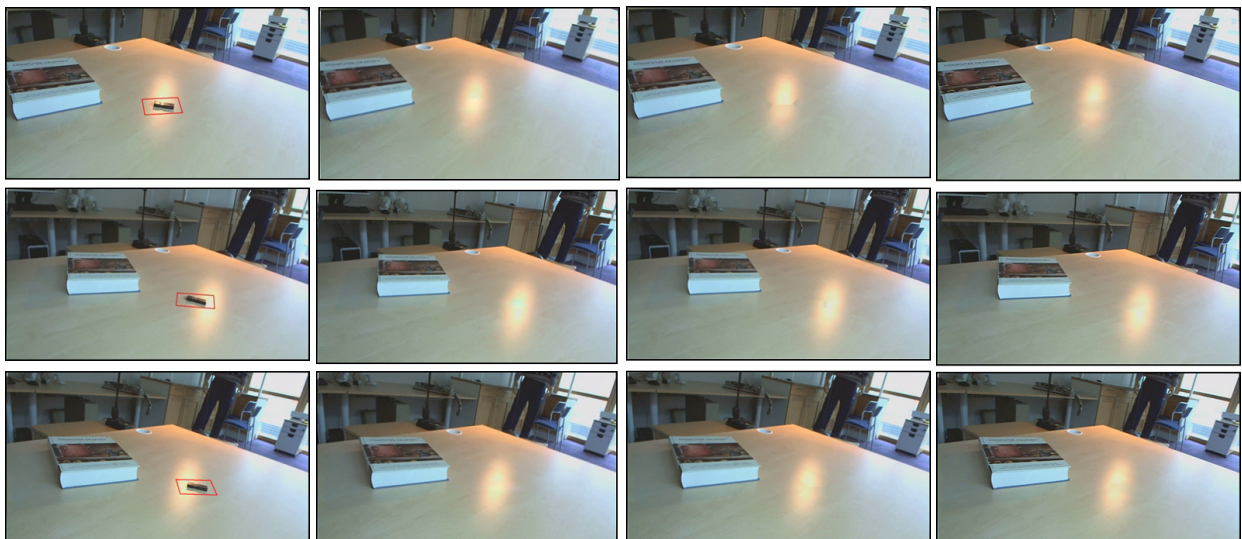
FIGURE 8 – Résultats des différentes méthodes sur les trames 367, 541 et 1227 de la vidéo 2.





(a) Images originales      (b) Notre méthode      (c) Herling *et al.* [10]      (d) Kawai *et al.* [11]

FIGURE 9 – Résultats des différentes méthodes sur les trames 168, 202 et 328 de la vidéo 3.



(a) Les images originales      (b) Notre méthode      (c) Herling *et al.* [10]      (d) Kawai *et al.* [11]

FIGURE 10 – Résultats des différentes méthodes sur les trames 415, 477 et 554 de la vidéo 4.