

Images intrinsèques de scènes en extérieur à partir de multiples vues

Pierre-Yves Laffont, Adrien Bousseau et George Drettakis
REVES / INRIA Sophia-Antipolis

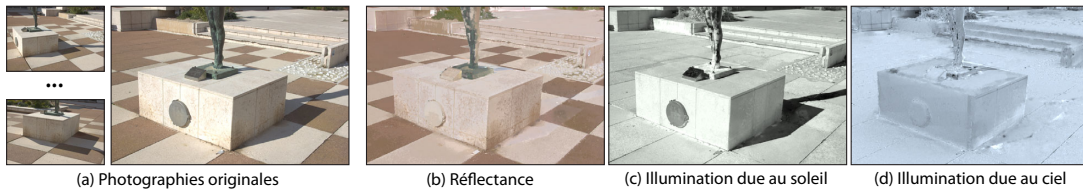


Figure 1: A partir de plusieurs vues de la scène (a), notre méthode décompose automatiquement les photographies en trois composantes intrinsèques : la réflectance (b), l'illumination due au soleil (c) et l'illumination due au ciel (d). Chaque composante intrinsèque peut alors être éditée de manière indépendante pour des manipulations d'image avancées.

Résumé

Nous présentons une méthode capable de décomposer les photographies d'une scène extérieure en trois composantes intrinsèques — la réflectance, l'illumination due au soleil et l'illumination due au ciel (figure 1). L'extraction d'images intrinsèques à partir de photographies est un problème difficile, généralement résolu en utilisant des méthodes de propagation guidée par l'image nécessitant de multiples indications par l'utilisateur. De récentes méthodes en vision par ordinateur permettent l'acquisition simple mais approximative d'informations géométriques sur une scène, à l'aide de plusieurs photographies selon des points de vue différents. Nous développons un nouvel algorithme qui nous permet d'exploiter cette information bruitée et peu fiable pour automatiser et améliorer les algorithmes d'estimation d'images intrinsèques par propagation. En particulier, nous développons une nouvelle approche afin d'estimer les ombres portées dans l'image, en peaufinant une estimation initiale obtenue à partir de la géométrie reconstruite. Nous utilisons un algorithme de vote dans l'espace de couleur qui identifie de façon robuste les valeurs de réflectance des nuages de points 3D reconstruits. Cette estimation de réflectance nous permet de déterminer avec précision la visibilité du soleil à ces points épars et de propager cette information aux pixels restants. Dans une dernière étape nous adaptons les algorithmes de propagation guidée par l'image, en remplaçant les indications manuelles de l'utilisateur par les données d'ombre et de réflectance estimées par notre algorithme aux points 3D reconstruits. Notre méthode permet l'extraction automatique des images intrinsèques à partir de multiples points de vue, permettant ainsi de nombreux types de manipulations d'images.

1. Introduction

La décomposition d'une image en composantes intrinsèques est un objectif de longue date en traitement d'image, vision par ordinateur et infographie [BT78, Wei01, FDL04, TFA05, BPD09]. Les images intrinsèques sont une décomposition d'une photographie en composantes de *réflectance* (aussi connue comme l'albédo) et d'illumination (également connue sous le nom d'ombrage ou d'éclairage). Ces décompositions peuvent être utilisées pour manipuler l'éclairage et les ombres dans les photographies, et permettent des manipulations d'image avancées. Un domaine étroitement lié en infographie est celui du ré-éclairage à partir de géométrie [YDMH99, LDR00, DTG*04], qui combine l'informa-

tion géométrique et des algorithmes d'illumination globale pour estimer la réflectance de la scène et effectuer le ré-éclairage.

La première classe d'approches est exclusivement basée sur les images. En conséquence, ces méthodes ont des difficultés à résoudre les ambiguïtés, comme celles dues à la texture et l'ombrage. De nombreuses interventions de l'utilisateur sont généralement appliquées pour résoudre ces ambiguïtés. En revanche, la deuxième classe d'algorithmes requiert des efforts manuels pour reconstruire une géométrie suffisamment exacte, et nécessite souvent une instrumentation complexe pour acquérir les valeurs de réflectance.

Nous proposons une approche intermédiaire qui couple

une géométrie acquise automatiquement (mais incomplète et inexacte) avec des algorithmes de manipulation d'images récents afin de résoudre les ambiguïtés avec une intervention minimale de l'utilisateur. Nous utilisons un algorithme de stéréo multi-vues [FP09] appliqué à un ensemble de photographies (entre 10 et 40) d'une scène en extérieur pour extraire automatiquement la géométrie de la scène. Nous exploitons l'information géométrique afin d'estimer l'illumination dans les zones où elle est fiable, puis nous propageons celle-ci à tous les pixels pour décomposer une photographie en images intrinsèques. Cet algorithme nous permet d'exploiter pleinement la puissance des approches de propagation guidées par l'image [LLW08, BPD09], sans la nécessité de multiples interventions de l'utilisateur.

Notre seconde contribution est un algorithme de vote pour identifier la visibilité du soleil en certains points, en utilisant l'information géométrique approximative pour initialiser le processus.

Enfin, la plupart des approches précédentes extraient la réflectance seulement, ou la réflectance et l'illumination totale. Dans notre cas, en plus d'extraire la réflectance de l'image originale, nous proposons de séparer l'illumination due au soleil et celle due au ciel, offrant ainsi des possibilités supplémentaires de manipulations d'images.

2. Travaux Précédents

Les méthodes de rendu inverse [YM98, YDMH99, LDR00, BG01, DTG*04, Deb06] estiment la réflectance et l'illumination d'une scène en inversant l'équation de rendu. Ces méthodes nécessitent un modèle 3D précis de la scène qui est soit modelé manuellement, soit capturé avec des scanners laser coûteux. Notre méthode en comparaison est robuste aux géométries incomplètes et gère des nuages de points clairsemés reconstruits à partir de quelques photographies de la scène.

Plusieurs méthodes ont été proposées pour estimer la réflectance et l'illumination à partir d'une seule image. Une telle décomposition est sévèrement mal posée et ne peut être résolue qu'avec des connaissances supplémentaires ou des hypothèses sur le contenu de la scène. L'algorithme Retinex [LM71, Hor86] suppose un éclairage doux et une réflectance constante par morceaux. Tappen *et al.* [TFA05] entraînent un algorithme de classification afin de discriminer les variations d'intensité dues à la réflectance et à l'illumination. Shen *et al.* [STL08] introduisent des contraintes de texture pour s'assurer que les pixels partageant une même texture ont la même réflectance. Malgré ces différentes approches, la décomposition automatique d'images complexes en extérieur reste un défi que nous proposons de résoudre en utilisant plusieurs images.

Weiss [Wei01] montre comment de multiples images d'une scène capturées sous différents éclairages peuvent être factorisées en une image de réflectance et un ensemble

d'images d'illumination. Sunkavalli *et al.* [SMPR07] décomposent des séquences d'images de scènes extérieures en un masque d'ombre et les images éclairées seulement par le soleil ou le ciel. Ces méthodes supposent un point de vue fixe et une illumination variable (i.e., des séquences de *chronophotographie*), tandis que notre méthode repose sur des images capturées depuis différents points de vue mais d'éclairage similaire. Le principal avantage de notre approche est de réduire le temps d'acquisition à quelques minutes tandis que la chronophotographie nécessite des temps de capture d'au moins une journée pour couvrir le plus grand nombre possible de directions d'éclairage.

Bousseau *et al.* [BPD09] s'appuient sur des indications utilisateur pour lever l'ambiguïté entre réflectance et illumination alors qu'Okabe *et al.* [OZM*06] propagent dans l'image un ensemble clairsemé de normales indiquées par l'utilisateur. Notre méthode partage des similitudes avec ces approches assistées par l'utilisateur, mais nous propageons des valeurs d'illumination calculée automatiquement à partir d'un nuage de points épars, à la place des indications de l'utilisateur.

L'intuition clé de notre approche est l'utilisation d'une interpolation guidée par l'image pour compléter l'information clairsemée générée par les algorithmes de stéréo multi-vue. Les méthodes d'interpolation guidée par l'image ont été introduites par Levin *et al.* et Lischinski *et al.* pour propager des couleurs [LLW04] et des ajustements de tons [LFUS06]. Notre méthode repose sur l'interpolation du *matting Laplacian* dérivé par Levin *et al.* [LLW08]. Bien que le *matting Laplacian* ait été initialement développé pour décomposer une image en une couche de premier plan et une d'arrière-plan, il a également été utilisé pour effectuer la balance des blancs sur des photographies prises sous un éclairage mixte [HMP*08] et pour propager des indications utilisateur pour la décomposition d'images intrinsèques [BPD09].

Notre travail est aussi lié aux méthodes d'extraction d'ombres [FHD02, FDL04, MTC07, WTBS07, SL08] qui visent à identifier et éliminer les ombres dans une image, soit automatiquement, soit avec l'assistance de l'utilisateur. En comparaison, bien que notre méthode soit capable d'identifier les ombres, notre objectif principal est d'extraire une image de réflectance, ainsi que des variations d'éclairage dépendant non seulement de la visibilité du soleil, mais également de l'orientation des surfaces et de la géométrie de la scène. Notre méthode permet également de séparer la contribution du soleil et celle du ciel dans l'illumination.

3. Vue d'Ensemble

Notre objectif est de calculer une *décomposition en images intrinsèques* [BT78] de photographies. Les composants intrinsèques que nous cherchons à récupérer sont la réflectance, l'illumination due au soleil (c'est-à-dire, une image de la scène éclairée seulement par le soleil, en supposant une réflectance blanche partout) et l'illumination due

au ciel (une image éclairée seulement par le ciel avec la réflectance blanche).

Notre méthode prend en entrée un petit ensemble de photographies de la scène ainsi que deux photographies d'une boule chromée (section 4). Nous utilisons d'abord un algorithme de stéréo multi-vue pour estimer des informations géométriques fiables sur un ensemble clairsemé de points 3D. Nous utilisons la géométrie reconstruite pour calculer des valeurs d'illumination à chaque point 3D (section 6). Nous proposons ensuite un nouvel algorithme de vote pour estimer la visibilité du soleil à chaque point 3D (section 7). Enfin, nous propageons l'information estimée sur toute l'image. Nous exprimons cette propagation sous la forme d'un problème d'incrustation que l'on résout en utilisant le *matting Laplacian* [LLW08] (section 8).

Notre méthode produit une décomposition de l'image qui peut être utilisée pour des manipulations avancées (section 9.4).

4. Capture et Reconstruction

Notre objectif est de proposer une approche de capture relativement légère, qui nécessite seulement un appareil photo numérique (de préférence sur un trépied), une carte grise de photographe, et une boule réfléchive (chromée ou en acier) pour capturer une carte d'environnement incluant le ciel. Aucun équipement supplémentaire de capture ou de mesure n'est nécessaire.

4.1. Photographie

Nous capturons d'abord un ensemble de photographies numériques classiques de la scène (*images LDR*, pour *Low Dynamic Range*), qui sont utilisées pour créer une reconstruction approximative de la scène. Afin de faciliter la reconstruction géométrique, cet ensemble doit avoir une bonne couverture des scènes grâce à des points de vue différents, et un bon recouvrement entre les points de vue voisins. Le nombre de photographies nécessaires pour obtenir une reconstruction acceptable dépend de la complexité de la scène et de la présence de zones texturées ; nous avons utilisé entre 10 et 40 photographies par scène. En outre, la scène doit être statique ; les objets fortement affectés par le vent (comme les arbustes) ne pourront pas être reconstruits.

Nous capturons ensuite deux images de la boule chromée, avec des points de vue orthogonaux. Nous utilisons des images à grande gamme dynamique (*images HDR*, pour *High Dynamic Range*), et appliquons la technique standard d'assemblage de Debevec *et al.* [DM97] pour obtenir une carte d'environnement. Enfin, nous capturons des images HDR, linéaires, à partir des points de vue pour lesquels nous souhaitons obtenir une décomposition en images intrinsèques.

L'ensemble du processus de capture prend entre 45 et 90

minutes par scène, en fonction du nombre d'images HDR que l'on souhaite décomposer, et du nombre d'images LDR nécessaires pour la reconstruction.

4.2. Reconstruction de la Géométrie et de l'Illuminant

Nous appliquons Bundler [SSS06] et l'algorithme de stéréo multi-vue PMVS [FP09], en utilisant les implémentations des auteurs respectifs (disponibles publiquement), avec les photos LDR et HDR comme entrée. Le résultat de ce processus est un nuage de points 3D orientés (où chaque point est associé à une position 3D et une normale) et un ensemble de caméras calibrées (paramètres intrinsèques et extrinsèques). En outre, le processus indique pour chaque point 3D, la liste des points de vue pour lesquels il est visible. Par la suite, lorsque nous cherchons à traiter une image, nous utilisons uniquement les points 3D visibles à partir du point de vue correspondant.

Nous avons trouvé cependant les normales estimées par PMVS bien trop bruitées pour être utilisées dans un calcul d'éclairage. Par conséquent, nous recalculons une normale à chaque point 3D en approximant son voisinage dans le nuage de points par un plan grâce à la méthode de Hoppe *et al.* [HDD*92], basée sur une Analyse en Composantes Principales (ACP). Nous supprimons les points 3D pour lesquels le nuage de point local est trop clairsemé ou trop déformé. La distorsion est considérée comme trop élevée si la première valeur singulière obtenue par ACP est deux fois plus grande que la seconde valeur singulière. Nous utilisons [JDZ04] pour lisser les normales restantes.

Nous construisons ensuite un *proxy géométrique*, approximation de la géométrie de la scène, à partir du nuage de points orientés en utilisant la reconstruction de Poisson de Kazhdan *et al.* [KBH06] implémentée dans *MeshLab* [CCR08]. Le résultat de ce processus n'est pas très précis, étant donné le faible nombre de points de vue que nous utilisons. En particulier, de fins détails géométriques sont souvent absents ou déformés, et de nombreuses régions du proxy présentent des bosses et creux absents dans la scène réelle. Cette reconstruction est trop imprécise pour estimer directement l'ombre du soleil (figure 2), mais nous permet toutefois de calculer une approximation de l'illumination due au ciel, aux points PMVS (voir la section 6).

Une rapide intervention manuelle (5-10 minutes) est nécessaire pour calibrer le soleil et le ciel, et aligner la carte d'environnement (voir la section 9.1). Ce processus nous permet d'obtenir la direction du soleil ainsi que son intensité et sa couleur C_{sun} (le soleil étant représenté comme une source de lumière directionnelle).

Le résultat de ce processus de capture et calibration est le suivant : a) un nuage moyennement dense de points orientés, que nous appellerons *points PMVS*, b) un *proxy géométrique* très approximatif pouvant contenir de larges erreurs géométriques, c) la direction et la couleur du soleil, et une

carte d'environnement correctement alignée représentant la radiance venant de toutes les directions du ciel.

Dans la suite de cet article, nous chercherons à décomposer en images intrinsèques chaque vue de la scène pour laquelle une photographie HDR a été capturée.

5. Modèle de Formation d'Image

D'après l'équation de rendu, la valeur de radiance en direction de la caméra \mathbf{I} (écrite sous forme d'un triplet RGB), à chaque point de la scène, peut être exprimée comme :

$$\mathbf{I} = \mathbf{R} * (\mathbf{L}_{sun} v_{sun} \cos \theta_{sun} + \int_{\Omega} (\mathbf{L}_{sky} v_{sky}(\omega) \cos \theta_{\omega} + \mathbf{L}_{ind}(\omega) \cos \theta_{\omega}) d\omega)$$

où \mathbf{R} est la réflectance RGB de l'objet ; * représente la multiplication composante par composante ; \mathbf{L}_{sun} , \mathbf{L}_{sky} et \mathbf{L}_{ind} correspondent respectivement à la radiance du soleil, du ciel, et de l'éclairage indirect ; v_{sun} , v_{sky} représentent la visibilité en direction du soleil et d'une portion de ciel ; θ_{sun} est l'angle entre la direction du soleil et la normale du point 3D de la scène ; et θ_{ω} est l'angle entre cette normale et la direction ω vers une portion de ciel, correspondant à l'intégration sur l'hémisphère Ω .

Nous ne modélisons par directement l'éclairage indirect, car il a généralement moins d'impact que le soleil et le ciel sur les scènes extérieures. D'autre part, puisque nous capturons l'éclairage dû au ciel à l'aide d'une carte d'environnement, certains objets à proximité de la scène (comme les bâtiments) contribuent implicitement à l'éclairage indirect distant. Cependant, l'éclairage indirect proche n'est pas explicitement pris en compte ; cette approche peut être trop restrictive dans le cas de points sur des surfaces verticales qui sont dans l'ombre du soleil, et nous prenons dans ce cas des mesures spécifiques (section 7).

Nous faisons l'hypothèse que les surfaces sont Lambertiennes, et exprimons la valeur de la radiance à chaque point de la scène comme le produit entre l'éclairage incident et la réflectance de l'objet. Nous séparons également la couleur de l'intensité de la lumière, en exprimant $\mathbf{L}_{sky} = \mathbf{C}_{sky} \mathcal{I}_{sky}$ (et de manière équivalente $\mathbf{L}_{sun} = \mathbf{C}_{sun} \mathcal{I}_{sun}$), où la norme de \mathbf{C}_{sky} est l'unité. Cette séparation est commode lorsque nous étendons la technique de balance des blancs de Hsu *et al.* [HMP*08] dans la section 8. Enfin, nous supposons \mathbf{C}_{sky} constant sur la voûte céleste. Cette approximation implique que la capture doit avoir lieu avec un ciel relativement clair, dans un délai de quelques heures autour de midi. Il en résulte un ciel majoritairement bleu dans la carte d'environnement.

Nous pouvons définir des relations à chaque point de la scène :

$$s_{sky} = \int_{\Omega} v_{sky} \mathcal{I}_{sky} \cos \theta_{\omega} d\omega \quad s_{sun} = \cos \theta_{sun} \mathcal{I}_{sun}. \quad (1)$$

ce qui nous permet d'utiliser un modèle simplifié de forma-

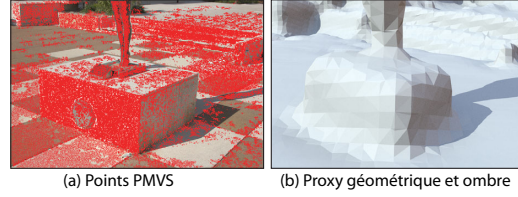


Figure 2: La stéréo multi-vue produit un nuage épars de points orientés (a, pixels rouges). Nous utilisons le proxy géométrique imprécis reconstruit à partir du nuage de points pour calculer une estimation initiale de l'ombre du soleil (b).

tion de l'image, que nous utilisons à présent :

$$\mathbf{I} = \mathbf{R} * (v_{sun} s_{sun} \mathbf{C}_{sun} + s_{sky} \mathbf{C}_{sky}) \quad (2)$$

où v_{sun} entre 0 et 1 représente la visibilité du soleil (0 s'il est complètement caché, 1 s'il est totalement visible). s_{sun} et s_{sky} sont des valeurs positives représentant l'intensité de l'éclairage incident (ou illumination) du soleil et du ciel respectivement, \mathbf{C}_{sun} et \mathbf{C}_{sky} sont la couleur RGB de l'éclairage dû au soleil et dû au ciel.

Nous cherchons à extraire chaque composante de ce modèle de formation d'image, comme illustré dans la figure 6. Nous allons montrer comment résoudre ce problème en mettant à contribution les informations géométriques éparses générées par les algorithmes de stéréo multi-vue.

6. Calcul Basé sur la Géométrie

La stéréo multi-vue produit un nuage épars de points orientés représentant la scène (section 4.2). Pour chaque point PMVS et en ignorant les ombres, nous pouvons calculer l'éclairage incident du soleil $s_{sun} \mathbf{C}_{sun}$ comme le produit scalaire entre sa normale et la direction vers le soleil.

Le calcul de l'éclairage incident du ciel $s_{sky} \mathbf{C}_{sky}$ est complexe, puisqu'il faut intégrer l'éclairage incident sur tout l'hémisphère centré sur la normale (équation 1). Nous calculons cette intégration en appliquant un échantillonnage préférentiel (*importance sampling*) sur la carte d'environnement, et en utilisant le proxy géométrique pour évaluer la visibilité de chaque échantillon. Ce calcul est robuste à la géométrie grossière du proxy, puisque l'intégrale prend en compte la visibilité du ciel dans toutes les directions de l'hémisphère. Nous avons modifié le moteur de rendu par lancer de rayons PBRT [PH10] afin d'effectuer ce calcul.

Enfin, nous calculons une estimation initiale de la visibilité du soleil v_{sun} à chaque point PMVS en traçant un rayon dans la direction du soleil et en testant s'il intersecte le proxy géométrique. Il faut cependant noter que ce test de visibilité est très sensible aux erreurs du proxy géométrique, comme illustré dans la figure 2b. Nous montrons dans la section suivante comment affiner cette estimation initiale.

7. Visibilité du soleil aux points PMVS

La contribution principale de notre approche est un nouvel algorithme pour identifier la visibilité v_{sun} du soleil pour chaque pixel.

D'après notre modèle de formation d'image de l'équation 2, nous pouvons exprimer la réflectance à chaque point PMVS en fonction de la visibilité du soleil :

$$\mathbf{R}(v_{sun}) = \frac{\mathbf{I}}{(v_{sun}s_{sun}\mathbf{C}_{sun} + s_{sky}\mathbf{C}_{sky})}. \quad (3)$$

Pour un point PMVS, faire varier v_{sun} entre 0 et 1 définit une courbe dans l'espace RGB, correspondant aux candidats de réflectance pour ce point. L'intuition de notre approche est que si plusieurs points PMVS partagent la même réflectance, alors leurs courbes candidates se croisent dans l'espace couleur. En trouvant ces intersections, nous pouvons trouver de manière robuste la valeur correspondante v_{sun} pour chaque point PMVS. Nous faisons l'hypothèse que chaque scène contient peu de réflectances différentes ; grâce à cette hypothèse, nous pouvons trouver les intersections de courbes. Cette hypothèse a également été utilisée dans des travaux précédents, comme par exemple [HMP*08].

Nous utilisons un algorithme de vote pour identifier les intersections entre courbes candidates de réflectance, dans un espace de couleur $L*a*b*$ discrétisé sous forme de grille 3D. Pour chaque point PMVS nous construisons un échantillonnage dense de sa courbe candidate en espace RGB ; les coordonnées des échantillons sont ensuite converties vers l'espace $L*a*b*$; nous pouvons alors identifier les cellules de la grille $L*a*b*$ que traverse la courbe. Pour chaque cellule de la grille traversée par la courbe, nous obtenons la valeur de visibilité correspondante v_{bin} en projetant la réflectance de la cellule (c'est-à-dire le centre de la cellule dans la grille) sur la courbe.

Chaque courbe correspondant à un point PMVS vote une fois pour chaque cellule qu'elle traverse. Plus précisément, le vote de la courbe pour chaque cellule qu'elle couvre est affecté d'un poids gaussien centré sur la valeur de visibilité v_{curr} :

$$w_{bin} = \exp(-(v_{curr} - v_{bin})^2 / (2\sigma_v)) \quad (4)$$

Le vote pour chaque cellule est incrémenté de w_{bin} , normalisé par la somme des poids de toutes les cellules couvertes par la courbe. Nous utilisons $\sigma_v = 0.1$ dans tous les exemples présentés ici, nous n'avons pas remarqué d'influence significative sur les résultats lors de la modification de cette valeur. Nous initialisons v_{curr} avec l'ombre du proxy géométrique, comme décrit dans la section précédente et illustré dans la figure 2b.

Grâce aux normales connues aux points PMVS, nous pouvons identifier les points PMVS orientés dans la direction opposée au soleil (ombres attachées). Pour ces points, l'influence relative de l'éclairage indirect est assez élevée. Comme notre modèle de l'équation 2 ne prend pas en

compte l'éclairage indirect, cette contribution est absorbée dans l'éclairage du ciel et du soleil. A ces points, $v_{sun} = 0$, ce qui entraînerait des valeurs de réflectance incorrectes (trop élevées). Nous ignorons donc les points PMVS dans les ombres attachées lors du calcul des votes.

Une fois que tous les points PMVS ont été pris en compte, et les contributions des courbes correspondantes traitées, nous sélectionnons la cellule de la grille avec le nombre maximal de votes. Pour chaque courbe passant par cette cellule, nous calculons la valeur v_{sun} correspondant au centre de la cellule et l'assignons au point PMVS associé. Les votes de ces courbes sont ensuite retirés de toutes les cellules de la grille. Nous répétons cette étape tant qu'au moins une cellule de la grille possède un nombre élevé de votes. A l'issue de ce processus itératif, les points PMVS correspondants aux courbes restantes (pour laquelle la visibilité n'a pas pu être estimée) sont éliminés.

Dans tous nos exemples, nous avons discrétisé l'espace $L*a*b*$ uniformément en $30 \times 20 \times 20$ cellules, avec $L*$ compris entre 5 et 95, et $a*$ et $b*$ entre -25 et 25. Entre 100000 et 200000 points PMVS sont visibles depuis chaque point de vue, chacun de ces points étant associé à une courbe candidate. Le seuil contrôlant la fin du processus itératif (i.e., le nombre minimal de votes passant par la cellule contenant le plus grand nombre de votes) est fixé à 80 votes.

A l'issue de cette étape, l'illumination due au ciel, l'illumination due au soleil, et la réflectance, ont été estimées à chaque point PMVS. Ces informations seront utilisées dans la section suivante afin de permettre la décomposition en images intrinsèques.

8. Décomposition en Images Intrinsèques

Contrairement aux précédentes méthodes de décomposition en images intrinsèques qui calculent une valeur unique d'illumination pour chaque pixel, nous cherchons à propager deux valeurs d'illumination : une pour le soleil, et une pour le ciel. Nous calculons cette propagation en deux étapes. Dans la première étape nous procédons à une *séparation des sources lumineuses* de l'image d'entrée pour obtenir deux images — une image représente la scène éclairée seulement par le soleil tandis que l'autre image représente la scène éclairée seulement par le ciel. Dans une seconde étape, nous traitons les deux images de manière indépendante afin d'estimer l'illumination due à chaque source de lumière.

8.1. Séparation des Sources Lumineuses

Nous cherchons à décomposer l'image \mathbf{I} en deux images \mathbf{I}_{sun} et \mathbf{I}_{sky} , suivant l'équation 2 :

$$\mathbf{I}_{sun} = \mathbf{R} * v_{sun}s_{sun}\mathbf{C}_{sun} \quad (5)$$

$$\mathbf{I}_{sky} = \mathbf{R} * s_{sky}\mathbf{C}_{sky}. \quad (6)$$

Inspirés par les travaux antérieurs sur la balance des blancs sur des photographies prises sous un éclairage

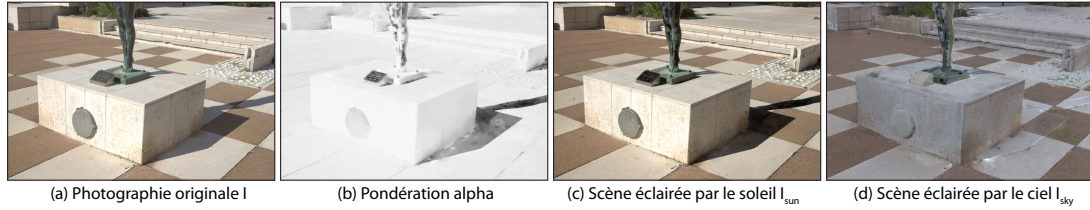


Figure 3: Nous exprimons l'image originale (a) comme une combinaison de deux images : la première représente la scène éclairée seulement par le soleil (c) ; la seconde représente la scène éclairée seulement par le ciel (d).

mixte [HMP*08], nous montrons comment exprimer notre séparation de sources de lumière comme un problème d'incrustation (*matting*). Tout d'abord, nous définissons $\mathbf{I}' = \mathbf{R}(v_{sun}s_{sun} + s_{sky})$ comme une image de la scène telle qu'elle apparaîtrait si éclairée par un soleil et un ciel blancs, c'est-à-dire $\mathbf{C}_{sky} = \mathbf{C}_{sun} = (1, 1, 1)/\sqrt{3}$. Cette image intermédiaire nous permet de réécrire les équations 5 et 6 comme :

$$\mathbf{I}_{sun} = \frac{v_{sun}s_{sun}}{v_{sun}s_{sun} + s_{sky}} \mathbf{I}' * \mathbf{C}_{sun} \quad (7)$$

$$\mathbf{I}_{sky} = \left(1 - \frac{v_{sun}s_{sun}}{v_{sun}s_{sun} + s_{sky}}\right) \mathbf{I}' * \mathbf{C}_{sky}. \quad (8)$$

En notant $\alpha = \frac{v_{sun}s_{sun}}{v_{sun}s_{sun} + s_{sky}}$ nous pouvons exprimer les valeurs de l'image originale à chaque pixel comme une combinaison de deux images, pondérée par α :

$$\mathbf{I} = \mathbf{I}_{sun} + \mathbf{I}_{sky} \quad (9)$$

$$= \alpha \mathbf{I}' * \mathbf{C}_{sun} + (1 - \alpha) \mathbf{I}' * \mathbf{C}_{sky} \quad (10)$$

Nous pouvons alors obtenir les images \mathbf{I}_{sun} et \mathbf{I}_{sky} en résolvant un problème classique de *matting*, que nous allons maintenant décrire.

Nous calculons α à chaque pixel correspondant à un point PMVS, à partir de l'éclairage du ciel, de l'éclairage du soleil, et de la visibilité du soleil, estimés précédemment. Nous propageons ensuite α sur toute l'image en utilisant le *matting Laplacian* développé par Levin *et al.* [LLW08].

A partir de l'équation 10, l'image intermédiaire \mathbf{I}' est calculée comme :

$$\mathbf{I}' = \frac{\mathbf{I}}{\alpha \mathbf{C}_{sun} + (1 - \alpha) \mathbf{C}_{sky}}. \quad (11)$$

puis \mathbf{I}_{sun} et \mathbf{I}_{sky} sont calculés à partir de α , \mathbf{I}' , \mathbf{C}_{sun} et \mathbf{C}_{sky} , en utilisant les équations 7 et 8. Ces images sont illustrées dans la figure 3. Notez que les pixels de l'image éclairée par le soleil uniquement (figure 3c) sont trop lumineux sur le côté du socle de la statue. Cela est dû au fait que nous ne modélisons pas explicitement l'éclairage indirect. Ce problème est discuté dans la section 9.3.

8.2. Décomposition en Réflectance et Illuminations

Séparer les contributions de la lumière du soleil et du ciel facilite grandement l'estimation des images intrinsèques d'illumination. Pour chaque image \mathbf{I}_{sun} et \mathbf{I}_{sky} , nous appliquons l'algorithme de décomposition en images intrinsèques de Bousseau *et al.* [BPD09], qui décompose l'image en une composante d'illumination et une de réflectance. Dans leur article, Bousseau *et al.* utilisent des indications utilisateur définies manuellement pour guider leur algorithme. Les utilisateurs peuvent fournir directement les valeurs de l'illumination à quelques pixels dans l'image, ou indiquer des ensembles de pixels qui partagent les mêmes valeurs de réflectance et d'illumination. Par contraste, notre méthode automatique s'appuie sur les informations géométriques fournies par le nuage de points reconstruit et sur notre nouvel algorithme de vote, afin d'éliminer le besoin d'intervention de l'utilisateur. Dans la pratique, nous utilisons les valeurs de l'illumination estimées aux points PMVS pour imposer des contraintes douces aux pixels correspondants. L'algorithme propage ensuite ces contraintes sur l'image en résolvant un système d'équations similaire au *matting Laplacian*. Nous renvoyons le lecteur à l'article original de Bousseau *et al.* pour les détails techniques.

A partir des deux images d'illumination $v_{sun}s_{sun}\mathbf{C}_{sun}$ et $s_{sky}\mathbf{C}_{sky}$, nous pouvons ensuite calculer l'image de réflectance en utilisant l'équation 3.

Ce processus de décomposition donne des résultats satisfaisants dans les scènes que nous avons testées. Cependant, une petite bordure résiduelle peut persister aux limites des ombres dures (voir figure 4a), artefact commun avec certaines méthodes d'enlèvement d'ombre [FDL04, WTBS07]. Pour détecter ces limites d'ombre nous propageons tout d'abord les normales sur l'image en utilisant la méthode de Okabe *et al.* [OZM*06]. Nous appliquons ensuite un détecteur de contours sur l'image α , et sélectionnons uniquement les pixels qui sont des contours dans α mais ne présentent pas de discontinuité de normales. Ces pixels et leurs voisins sont supprimés dans les images de réflectance et d'illumination due au ciel, puis les trous sont comblés automatiquement par complétion d'image (*inpainting*) [BSCB00]. Le résultat final est illustré sur la figure 4b.

L'ensemble du processus de décomposition en images in-

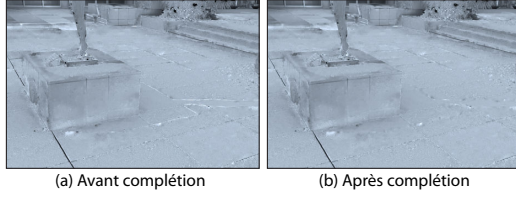


Figure 4: La réflectance et l'illumination du ciel estimées par notre algorithme peuvent contenir des résidus le long des limites d'ombre dures (a). Nous détectons automatiquement ces artefacts et utilisons la complétion d'image pour les atténuer (b).

trinsèques dure environ 2 minutes avec notre implémentation non-optimisée, sur des images de 3.2 megapixels. L'étape d'*inpainting* prend environ 60 secondes supplémentaires.

9. Implémentation, Résultats et Discussion

Nous décrivons tout d'abord les détails de calibration qui nous ont permis de déterminer la direction et couleur du soleil, et de calibrer la couleur du ciel. Nous présentons ensuite nos résultats et discutons des limites de notre approche.

9.1. Calibration

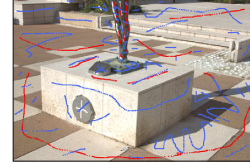
Les pixels correspondant au soleil sont saturés dans la carte d'environnement capturée dans la section 4.2. Le barycentre des pixels saturés correspond à la position du soleil dans la carte d'environnement. Ces pixels sont éliminés et remplacés par la valeur de leurs voisins (par complétion d'image), pour obtenir une carte d'environnement ne contenant pas le soleil.

Une rapide étape de réglage manuel est nécessaire afin d'aligner la carte d'environnement avec le proxy géométrique, et d'obtenir la direction du soleil : l'utilisateur marque une arête verticale dans la scène reconstruite, puis il applique une rotation à la carte d'environnement jusqu'à ce que les ombres projetées par un soleil virtuel (placé au barycentre des pixels saturés dans la carte d'environnement) sur le proxy géométrique soient alignées avec les ombres se trouvant dans l'image.

Enfin, nous pouvons calibrer grossièrement la puissance du soleil et du ciel en utilisant une carte grise de photographie. Nous plaçons une carte de couleur neutre et dont la réflectance est connue dans une image où elle est dans l'ombre, et dans une image où elle est éclairée par le soleil. La carte est placée dans des zones où nous nous attendons à ce que la géométrie soit bien reconstruite. Grâce à la première image, nous pouvons déduire de notre modèle de formation d'image :

$$\mathbf{I} = \mathbf{R} * (s_{sky} \mathbf{C}_{sky} * \kappa_{sky}) \quad (12)$$

où κ_{sky} est un vecteur RGB inconnu, que nous pouvons



(a) Photographie originale et indications manuelles [BPD09]



(b) Réflectance assistée par l'utilisateur [BPD09]



(c) Illumination assistée par l'utilisateur [BPD09]



(d) Notre réflectance sans intervention manuelle



(e) Notre illumination (sun+sky) sans intervention manuelle

Figure 5: La méthode de Bousseau et al. [BPD09], assistée par l'utilisateur, requiert une importante intervention manuelle pour produire des résultats de qualité comparable à notre approche automatique. Dans cet exemple, l'utilisateur avait dessiné plus de 80 indications sur l'image (a).

trouver en résolvant cette équation. A partir de la seconde image :

$$\mathbf{I} = \mathbf{R} * (v_{sun} s_{sun} \mathbf{C}_{sun} + s_{sky} \mathbf{C}_{sky} * \kappa_{sky}) \quad (13)$$

et puisque $v_{sun} = 1$ et κ_{sky} a été estimé dans l'équation précédente, nous pouvons déterminer \mathbf{C}_{sun} .

9.2. Résultats

Dans la figure 6 nous présentons nos résultats sur deux scènes très différentes : une place urbaine, et une paroi rocheuse. Dans le premier cas, le principal générateur d'ombre est une statue sur son socle projetant une ombre dure, tandis que dans la seconde nous avons placé un parapluie devant la paroi pour générer une ombre douce. Nous présentons également un cas d'échec de notre méthode, lorsqu'un nouveau personnage projette une ombre sur la paroi mais n'apparaît pas dans les images utilisées pour la reconstruction. Dans ce dernier cas, le proxy géométrique donne une estimation initiale de la visibilité du soleil trop erronée pour que notre algorithme ne parvienne à la corriger.

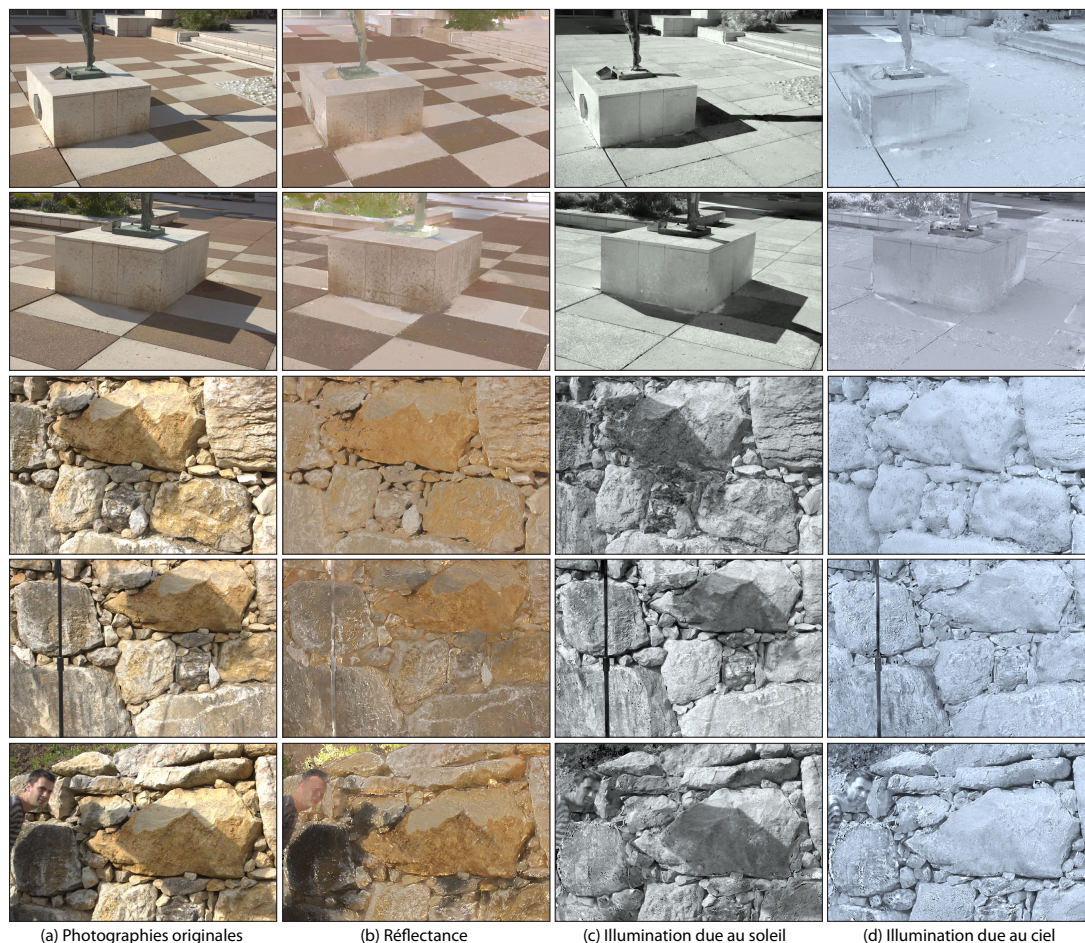


Figure 6: Résultats de notre décomposition appliquée sur deux scènes différentes. Dans la plupart des cas, notre méthode est capable de décomposer de manière plausible la réflectance et l'éclairage. La dernière ligne illustre un cas où un personnage n'apparaît que dans une seule photographie. Le proxy reconstruit ne contient pas de géométrie correspondant à cette personne, et par conséquent l'estimation initiale de la visibilité du soleil est trop imprécise pour permettre à notre algorithme de retrouver une bonne réflectance.

9.3. Limites

Comme le montrent les figures 1 et 6, l'illumination due au soleil n'est pas complètement nulle dans les régions d'ombres portées (sur le sol) et attachées (côté du socle de la statue). Dans le premier cas, au cours du processus de vote la valeur de v_{sun} aux points PMVS est déterminée en résolvant l'équation 3. Cependant, I_{sun} contient des résidus d'éclairage indirect, ce qui résulte en une valeur non nulle de v_{sun} . Pour les ombres attachées de cette image, la valeur de v_{sun} est légèrement plus élevée par rapport à celle des ombres portées. Au cours de la propagation, l'éclairage indirect est réparti dans s_{sun} et s_{sky} ; de plus, en raison de la position ver-

ticale de ces points, la moitié du ciel n'est pas visible, ce qui réduit la contribution relative de s_{sky} . En conséquence, pendant la propagation de l'image α (section 8), les valeurs de I_{sun} sont plus élevées qu'elles ne devraient. Il serait possible de corriger cela en modélisant explicitement l'éclairage indirect. Toutefois, cela nécessiterait une procédure de calibration plus complexe pour l'éclairage indirect, ainsi qu'une modification du modèle de formation de l'image et de la méthode de propagation.

Les décompositions de la figure 6 ont été générées sans aucune interaction de l'utilisateur. Toutefois, certaines petites ombres résiduelles persistent (par exemple, derrière la

statue). Cela est dû à l'absence de points PMVS dans ces régions de l'image. Dans d'autres cas, une réflectance incorrecte peut être estimée lorsque des courbes candidates s'intersectent plusieurs fois (par exemple, la plaque ronde à l'avant de la statue est estimée avec la même réflectance que le socle, alors qu'elle devrait être grise). Une simple indication utilisateur permet d'éliminer ces artefacts, à comparer aux 80 indications nécessaires pour la méthode de Bousseau *et al.* sur la figure 5a.

9.4. Applications

Nous montrons deux applications de notre décomposition en images intrinsèques. Dans la figure 7b, nous imitons un coucher de soleil, ou un ré-éclairage avec un ciel rougeâtre. Comme l'illumination du soleil se trouve dans une image séparée, nous sommes en mesure de brouiller les ombres manuellement dans un programme de retouche d'image, sans affecter la réflectance de la scène. Nous avons ensuite modifié séparément l'illumination du soleil et du ciel, pour donner une couleur rougeâtre au ciel et une couleur orange au soleil. Dans la figure 7c, nous avons manipulé l'image de réflectance en ajoutant un logo sur le sol, et ajouté un objet synthétique. La séparation des images de réflectance et d'illumination du soleil nous permettent d'obtenir des ombres plausibles dans l'image recomposée. L'objet synthétique est éclairé par la carte d'environnement, et l'ombre qu'il projette sur le proxy géométrique est intégrée à l'illumination du soleil, ce qui permet d'obtenir un résultat convaincant.

Des applications supplémentaires dans un logiciel de retouche d'image sont illustrées dans les vidéos accompagnant cet article.

10. Conclusion et Travail Futur

Nous avons présenté une méthode pour décomposer des photographies de scènes extérieures en images intrinsèques, c'est-à-dire en images de réflectance, d'illumination due au soleil et d'illumination due au ciel. Pour ce faire, nous capturons un ensemble de 10 à 40 images, et appliquons une étape de reconstruction géométrique entièrement automatique. Cette reconstruction fournit des informations géométriques éparpillées, et approximatives. Nous utilisons ces informations dans un algorithme de vote qui estime la visibilité du soleil à certains points de la scène. Cette estimation est ensuite combinée avec une propagation utilisant le *matting Laplacian*, pour obtenir deux images de la scène éclairée uniquement par le soleil ou uniquement par le ciel. L'algorithme d'images intrinsèques de Bousseau *et al.* [BPD09] est séparément appliqué sur chacune de ces images et contraint par les valeurs d'illumination estimées aux points PMVS. L'utilisation de l'information géométrique évite la nécessité d'indications manuelles comme dans l'algorithme original. Au final, nous obtenons une décomposition de la photographie originale en trois images intrinsèques : réflectance, illumination due au soleil, et illumination due au ciel.

Dans les travaux futurs, nous allons chercher à améliorer l'estimation de l'illumination due au soleil, en utilisant un modèle explicitant l'éclairage indirect. Une extension évidente et importante serait d'offrir la possibilité de déplacer les ombres d'une manière plausible, afin de simuler l'aspect de la scène à différents moments de la journée. Nous allons également chercher à éliminer ou simplifier la dépendance à la calibration pour la couleur du soleil et du ciel. Enfin, nous souhaiterions étendre cette approche à un algorithme de rendu basé image : une fois notre approche appliquée sur toutes les photographies de la scène, nous pourrions modifier l'éclairage et nous déplacer librement dans un environnement virtuel 3D.

Remerciements Nous remercions Peter Vangorp, Gaurav Chaurasia et Carles Bosch pour leur aide lors de la capture, Adobe (donation), Autodesk (Research donation program) et NVIDIA (Professor partnership program) pour leur soutien à notre recherche.

Références

- [BG01] BOIVIN S., GAGALOWICZ A. : Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *SIGGRAPH '01* (2001), pp. 107–116.
- [BPD09] BOUSSEAU A., PARIS S., DURAND F. : User-assisted intrinsic images. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2009)*. Vol. 28, Num. 5 (2009).
- [BSCB00] BERTALMIO M., SAPIRO G., CASELLES V., BALLESTER C. : Image inpainting. In *SIGGRAPH '00* (2000), pp. 417–424.
- [BT78] BARROW H., TENENBAUM J. : Recovering intrinsic scene characteristics from images. *Computer Vision Systems* (1978).
- [CCR08] CIGNONI P., CORSINI M., RANZUGLIA G. : Meshlab : an open-source 3d mesh processing system. *ERCIM News*, Num. 73 (2008), 45–46.
- [Deb06] DEBEVEC P. : Virtual cinematography : Relighting through computation. *IEEE Computer*. Vol. 39, Num. 8 (2006), 57–65.
- [DM97] DEBEVEC P. E., MALIK J. : Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97* (1997), pp. 369–378.
- [DTG*04] DEBEVEC P., TCHOU C., GARDNER A., HAWKINS T., POUILLIS C., STUMPFEL J., JONES A., YUN N., EINARSSON P., LUNDGREN T., FAJARDO M., MARTINEZ P. : *Estimating surface reflectance properties of a complex scene under captured natural illumination*. Tech. rep., USC Institute for Creative Technologies, 2004.
- [FDL04] FINLAYSON G., DREW M., LU C. : Intrinsic images by entropy minimization. In *ECCV* (2004).
- [FHD02] FINLAYSON G., HORDLEY S., DREW M. : Removing shadows from images. In *ECCV* (2002).

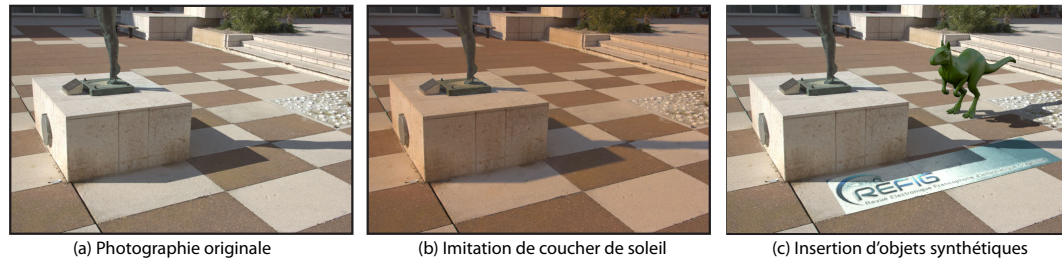


Figure 7: Nous utilisons notre méthode de décomposition pour modifier indépendamment l'illumination du soleil et du ciel, pour imiter un coucher de soleil (b). En (c) nous insérons un logo et un objet virtuel tout en préservant la cohérence des ombres.

- [FP09] FURUKAWA Y., PONCE J. : Accurate, dense, and robust multi-view stereopsis. *IEEE Trans. PAMI*. Vol. 32, Num. 8 (2009), 1362–1376.
- [HDD*92] HOPPE H., DE ROSE T., DUCHAMP T., MCDONALD J., STUETZLE W. : Surface reconstruction from unorganized points. *Computer Graphics (Proceedings of SIGGRAPH '92)*. Vol. 26 (1992), 71–78.
- [HMP*08] HSU E., MERTENS T., PARIS S., AVIDAN S., DURAND F. : Light mixture estimation for spatially varying white balance. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2008)*. Vol. 27, Num. 3 (2008), 70.
- [Hor86] HORN B. K. : *Robot Vision*. MIT Press, 1986.
- [JDZ04] JONES T. R., DURAND F., ZWICKER M. : Normal improvement for point rendering. *IEEE Computer Graphics Applications*. Vol. 24 (2004), 53–56.
- [KBH06] KAZHDAN M., BOLITHO M., HOPPE H. : Poisson surface reconstruction. In *Proc. of Eurographics Symposium on Geometry Processing* (2006), pp. 61–70.
- [LDR00] LOSCOS C., DRETTAKIS G., ROBERT L. : Interactive virtual relighting of real scenes. *IEEE Transactions on Visualization and Computer Graphics*. Vol. 6 (2000), 289–305.
- [LFUS06] LISCHINSKI D., FARBMAN Z., UYTENDAELE M., SZELISKI R. : Interactive local adjustment of tonal values. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*. Vol. 25 (2006), 646–653.
- [LLW04] LEVIN A., LISCHINSKI D., WEISS Y. : Colorization using optimization. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2004)*. Vol. 23 (2004), 689 – 694.
- [LLW08] LEVIN A., LISCHINSKI D., WEISS Y. : A closed-form solution to natural image matting. *IEEE Trans. PAMI* (2008).
- [LM71] LAND E. H., MCCANN J. J. : Lightness and retinex theory. *Journal of the optical society of America*. Vol. 61, Num. 1 (1971).
- [MTC07] MOHAN A., TUMBLIN J., CHOUDHURY P. : Editing soft shadows in a digital photograph. *IEEE Computer Graphics and Applications*. Vol. 27, Num. 2 (2007).
- [OZM*06] OKABE M., ZENG G., MATSUSHITA Y., IGARASHI T., QUAN L., YEUNG SHUM H. : Single-view relighting with normal map painting. In *Proc. Pacific Graphics* (2006), pp. 27–34.
- [PH10] PHARR M., HUMPHREYS G. : *Physically Based Rendering : From Theory to Implementation, second edition*. Morgan Kaufmann Publishers Inc., 2010.
- [SL08] SHOR Y., LISCHINSKI D. : The shadow meets the mask : Pyramid-based shadow removal. *Comput. Graph. Forum*. Vol. 27, Num. 2 (2008), 577–586.
- [SMPR07] SUNKAVALLI K., MATUSIK W., PFISTER H., RUSINKIEWICZ S. : Factored time-lapse video. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2007)*. Vol. 26, Num. 3 (2007), 101.
- [SSS06] SNAVELY N., SEITZ S. M., SZELISKI R. : Phototourism : exploring photo collections in 3D. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*. Vol. 25, Num. 3 (2006), 835–846.
- [STL08] SHEN L., TAN P., LIN S. : Intrinsic image decomposition with non-local texture cues. In *CVPR* (2008).
- [TFA05] TAPPEN M. F., FREEMAN W. T., ADELSON E. H. : Recovering intrinsic images from a single image. *IEEE Trans. PAMI*. Vol. 27, Num. 9 (2005).
- [Wei01] WEISS Y. : Deriving intrinsic images from image sequences. In *Proc. IEEE International Conference on Computer Vision (ICCV)* (2001), vol. 2, p. 68.
- [WTBS07] WU T.-P., TANG C.-K., BROWN M. S., SHUM H.-Y. : Natural shadow matting. *ACM Transactions on Graphics*. Vol. 26 (2007).
- [YDMH99] YU Y., DEBEVEC P., MALIK J., HAWKINS T. : Inverse global illumination : recovering reflectance models of real scenes from photographs. In *SIGGRAPH '99* (1999), pp. 215–224.
- [YM98] YU Y., MALIK J. : Recovering photometric properties of architectural scenes from photographs. In *SIGGRAPH '98* (1998), pp. 207–217.