

# COMPLÉTION D'IMAGE EXPLOITANT DES DONNÉES MULTISPECTRALES

Frédéric Bousefsaf, Mohamed Tamaazousti, Souheil Hadj Said, Rémi Michel

CEA LIST, Laboratoire de Vision et Ingénierie des Contenus (LVIC)  
CEA Saclay Nano-INNOV, Bat. 861-PC142, F-91191 Gif-sur-Yvette, France  
{frederic.bousefsaf, mohamed.tamaazousti, souheil.hadjsaid, remi.michel}@cea.fr

## Résumé

Nous proposons d'explorer le potentiel de l'imagerie multispectrale appliquée à la complétion d'image, une thématique classique de vision par ordinateur qui consiste à automatiquement remplacer une zone de l'image de manière plausible. Pour ce faire, nous avons utilisé une caméra multispectrale de type *snapshot mosaic* permettant d'acquérir en temps réel des images définies sous 16 canaux dans le spectre visible (475 à 650 nm). Nos premiers résultats indiquent que l'extension des algorithmes de complétion par l'exploitation directe de la donnée multispectrale ne traduit pas d'amélioration significative. Pour répondre à cette problématique, nous proposons une méthode dédiée qui repose sur une segmentation préalable de la scène. La chaîne algorithmique consiste à limiter la recherche des pixels de remplacement dans une zone prédéfinie par la segmentation spectrale de la scène. La méthode complète repose sur des techniques standard développées par la communauté de vision par ordinateur et d'imagerie hyperspectrale. Nous avons validé la méthode proposée sur différentes scènes à l'aide d'une évaluation visuelle subjective basée sur 20 observateurs.

**Mots clés :** Imagerie multispectrale, Complétion d'image, Segmentation spectrale

## Abstract

*We propose to explore the potential of multispectral imaging applied to image completion. In our experiments, multispectral images are acquired using an ultra-compact snapshot camera-recorder that senses 16 different spectral channels in the visible spectrum (475–650 nm). We show that direct exploitation of completion algorithms by extension of the spectral channels exhibits only minimum enhancement. A dedicated method that consists in a prior segmentation of the scene has been developed to address this issue. The segmentation derives from an analysis of the spectral data and is employed to constrain research area of exemplar-based completion algorithms. The full processing chain takes benefit of standard methods that were developed by both hyperspectral imaging and computer vision communities. We validate our method on a variety of scenes using a perceptual evaluation based on 20 volunteers.*

**Keywords :** Multispectral imaging, Image completion, Spectral segmentation

## 1. Introduction

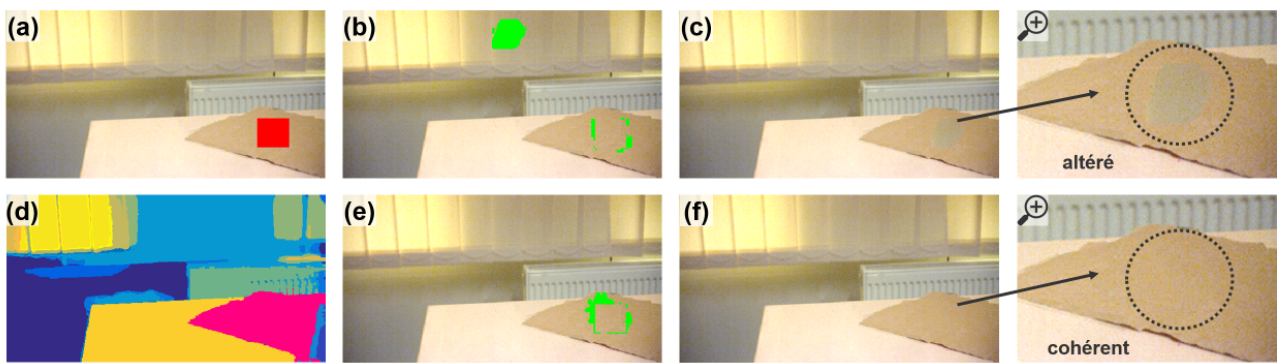
La complétion consiste à remplir ou restaurer une région manquante ou endommagée d'une image de manière cohérente. Cette manipulation est utilisée dans de nombreuses applications concrètes, telles que la suppression d'objets indésirables dans des photographies numériques, la restauration d'image (Guillemot et Le Meur, 2014) et la réalité diminuée (Kawai et al., 2016). Les recherches dans ce domaine ont atteint un niveau de maturité certain, les algorithmes les plus populaires ayant été intégrés dans des logiciels de retouche et de dessin assisté par ordinateur (par exemple l'algorithme PatchMatch (Barnes et al., 2009) dans Photoshop CS5<sup>1</sup>).

Des méthodes récentes ont été proposées par la communauté de vision afin de guider le remplissage de la région à compléter par des informations a priori sur les structures (He et Sun, 2014; Herling et Broll, 2014) et

perspectives (Huang et al., 2014). Le rendu synthétique délivré par la procédure de complétion étant visualisé sur des écrans et analysé par des personnes, la complétion est la plupart du temps effectuée sur des images acquises par des caméras RVB (rouge, vert, bleu). Parallèlement à ce type de capteur, les caméras multispectrales (Hagen et Kudenov, 2013) délivrent des informations plus détaillées sur les objets de la scène en fournissant un spectre par pixel de l'image. Ces capteurs peuvent donc présenter un intérêt particulier en vision par ordinateur (Chakrabarti et Zickler, 2011), notamment depuis l'introduction des caméras multispectrales permettant d'acquérir en temps réel des instantanés (contrairement aux systèmes à balayage) (Geelen et al., 2014).

Dans le domaine des sciences de la Terre et en planétologie, les images délivrées par les caméras multi ou hyperspectrales sont traitées et analysées en vue de fournir des informations pertinentes sur la composition chimique des scènes observées. Les phénomènes physiques, tels que l'absorption ou la réflectance spectrale,

1. <http://www.adobe.com/technology/projects/patchmatch.html>



**FIGURE 1:** (a) Image d'une scène acquise avec le capteur multispectral et convertie au format RVB. Le rectangle rouge sur le morceau de carton représente la zone à compléter. Cette dernière est sélectionnée manuellement par l'utilisateur. (b) Les pixels retenus par l'algorithme de complétion (dans notre cas PatchMatch (Barnes et al., 2009)) et qui seront par la suite copiés dans la zone à compléter sont indiqués en vert. L'algorithme utilise (par erreur) des pixels du rideau pour remplir le morceau de carton à compléter, leurs valeurs R, V et B étant très proches. (c) L'image complétée semble visuellement altérée. (d) Nous proposons une pré-segmentation spectrale de la scène par le biais d'une analyse en composantes principales sur l'image multispectrale. Notez que la segmentation spectrale produit des classes qui semblent corrélées à la géométrie et aux matériaux de la scène. (e) La recherche des pixels de copie est limitée aux classes présentes dans le voisinage de la zone à compléter (région magenta sur la figure d). (f) La complétion contrainte par la segmentation spectrale produit un résultat cohérent et compatible avec les standards en vision par ordinateur et en infographie.

peuvent être estimés pour chaque pixel de l'image. Les technologies multi et hyperspectrales sont utilisées pour détecter des modifications de la composition chimique des objets qui ne sont pas identifiables avec des capteurs monochromatiques ou RVB (Li et al., 2013). Les données spectrales ont typiquement été utilisées pour caractériser la couleur des océans (Gordon et Morel, 2012), classifier la surface des glaciers (Pope et Rees, 2014) ou encore pour détecter le gypse sur Mars (Gendrin et al., 2005). L'imagerie spectrale a aussi été utilisée dans le milieu biomédical pour, par exemple, détecter et identifier des anomalies sur les tissus physiologiques (Li et al., 2013).

A notre connaissance et à ce jour, aucune étude permettant d'analyser l'intérêt de l'imagerie multispectrale dans le contexte de la complétion d'image n'a été proposée. Analyser des images multispectrales plutôt que RVB revient à traiter un spectre par pixel de l'image plutôt qu'un groupe de valeurs R, V et B. Cette étude fournit des éléments permettant de conclure sur le potentiel de l'imagerie multispectrale appliquée à la complétion d'image, une application initialement dédiée aux images RVB.

La section 2 de ce document fournit une description synthétique des méthodes de complétion d'image. Les notions et rudiments de l'imagerie multispectrale (spécifications du capteur et pré-traitements) sont présentés en section 3. Les images complétées devant être visualisées par des personnes, cette section inclut également des détails sur le passage des canaux multispectraux à l'espace de couleur RVB. Nous décrivons, en section 4, le comportement d'un algorithme de complétion de référence sur les données multispectrales en étendant son vecteur d'entrée (passage d'une image RVB contenant 3 dimensions à une image multispectrale contenant 16 dimensions). La section 5 présente une méthode dédiée

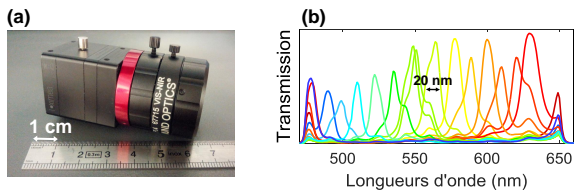
et reposant sur une pré-segmentation de la scène en classes d'intérêt afin de contraindre l'algorithme de complétion à rechercher des pixels de remplacement dans des zones spectralement cohérentes (voir figure 1 pour un exemple représentatif). Enfin, la section 6 est dédiée à l'analyse des résultats de complétion au travers d'une évaluation visuelle subjective. La méthode que nous proposons dans cette étude délivre des images complétées qui sont en meilleur adéquation avec les standards en vision par ordinateur et en infographie.

## 2. Complétion d'image

Les méthodes de complétion automatique d'image peuvent être classées en deux catégories (Guillemot et Le Meur, 2014) : (1) méthodes basées diffusion (Telea, 2004) et (2) méthodes basées sur l'analyse de patches (Criminisi et al., 2004; Barnes et al., 2009; Huang et al., 2014; He et Sun, 2014). Les méthodes de complétion d'une simple image ont récemment été étendues à des séquences vidéo (Newson et al., 2014) ou sur des flux temps réel pour des applications de réalité diminuée (Kawai et al., 2016). Plus récemment, Baek et al. (2016) proposent de compléter conjointement les canaux de couleur et de profondeur à partir d'images multivues.

### 2.1. Méthodes basées diffusion

Les techniques calquées sur les processus de diffusion ont été initialement développées pour remplir de petites régions manquantes en propageant les pixels adjacents vers la zone à compléter (?). Ces techniques rapides permettent de diffuser les structures locales mais ont, en contrepartie, tendance à délivrer des résultats flous et à mal se comporter sur des régions plus larges (propagation de textures par exemple).



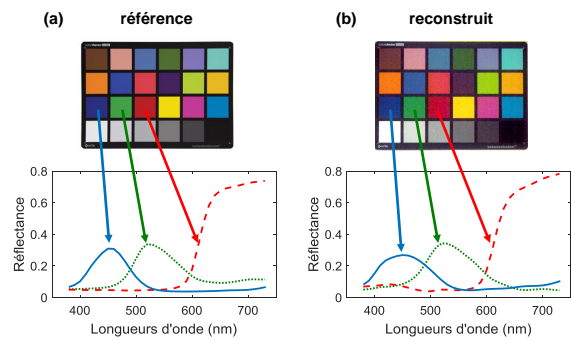
**FIGURE 2:** (a) Caméra multispectrale développée par IMEC (Geelen et al., 2014) de type *snapshot mosaic*. (b) Sensibilité spectrale des 16 canaux intégrés au capteur. Ces derniers englobent la plupart du spectre visible (de 475 à 650 nm). La largeur des bandes est d'environ 20 nm par canal.

Aussi, des techniques reposant sur la décomposition tensorielle (Chen et al., 2014) ont été développées afin de propager les structures en estimant les valeurs manquantes dans les tenseurs formés à partir de l'image. Ces méthodes reposent sur la complétion matricielle étendue au cas des tenseurs, la complétion tensorielle correspondant à une extension haut niveau de la complétion matricielle. La formulation repose sur une problématique d'optimisation convexe globale (Chen et al., 2014).

## 2.2. Méthodes basées sur l'analyse de patches

Les méthodes reposant sur l'analyse de patches s'inspirent des travaux relatifs à la synthèse de texture (Efros et Leung, 1999). Ces approches consistent à analyser les textures de l'image par le biais de patches, l'hypothèse sous-jacente étant que la région à compléter contient des textures similaires à celles du reste de l'image (Criminisi et al., 2004; Barnes et al., 2009; Huang et al., 2014; He et Sun, 2014). Les textures sont ainsi copiées de manière progressive à partir des bordures de la région à compléter. Ces méthodes permettent de répliquer des textures complexes à l'intérieur de la région à compléter. La propagation des structures reste néanmoins dépendante de l'ordre de copie. Pour résoudre cette problématique et en vue de délivrer des résultats plus cohérents, l'ordre de copie des patches peut être déterminé via des critères spécifiques. Par exemple, Criminisi et al. (2004) ont proposé de considérer les gradients afin d'encourager la propagation des textures le long des structures linéaires (arêtes franches de l'image).

En l'état, ces approches effectuent une copie progressive des patches qui ne permet pas d'assurer une cohérence globale au rendu. Ainsi, Wexler et al. (2007) ont proposé de contraindre le processus de complétion par patches en proposant de résoudre un problème d'optimisation globale. Sur la base de ces travaux et en vue de réduire les temps de calcul, Barnes et al. (2009) ont développé PatchMatch, un algorithme efficace qui permet de propager les textures de manière itérative en minimisant une fonction de coût. Cette dernière est basée sur une évaluation de la similarité entre les patches situés aux alentours de la région à compléter et les patches du reste de l'image. Des contraintes supplémentaires peuvent être apposées sur ces fonctions de coût. Par



**FIGURE 3:** (a) *ColorChecker Classic* (X-Rite). La mire contient 24 patches de couleur (McCamy et al., 1976). Les spectres de référence, définis entre 380 et 730 nm, sont fournis par le constructeur. (b) Image et spectres mesurés par la caméra multispectrale. Les spectres de référence (a) et reconstruits (b) sont très proches, même si de légères différences peuvent être observées. Ces dernières sont liées à des incertitudes sur la procédure d'étalonnage.

exemple, Huang et al. (2014) ont ajouté des contraintes géométriques permettant de segmenter les plans afin de considérer les perspectives de la scène. Aussi, He et Sun (2014) ont proposé de calculer différentes statistiques sur la position des patches afin d'améliorer la qualité des rendus.

Notons que l'apprentissage profond correspond à une piste intéressante qui a d'ores et déjà été considérée (Pathak et al., 2016). Les réseaux de neurones convolutifs sont entraînés afin de pouvoir générer le contenu de la région à compléter en capturant le contexte relatif de l'image. Ces approches, bien que prometteuses, sont actuellement moins performantes que les méthodes récentes de complétion basées sur l'analyse de patches.

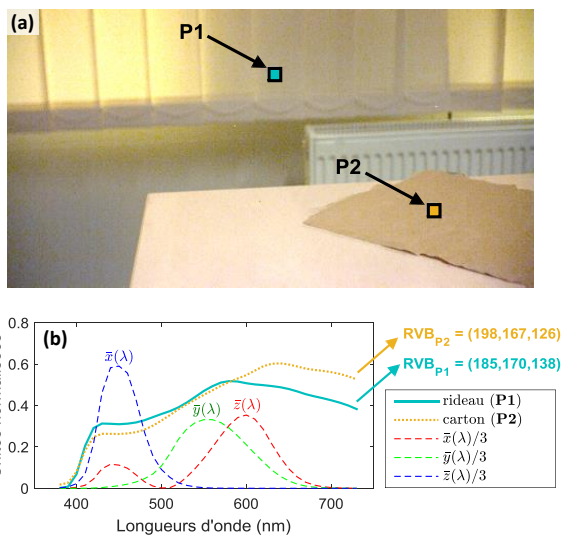
Dans cette étude, nous avons utilisé PatchMatch (Barnes et al., 2009) comme algorithme de complétion de référence. Une étape d'optimisation globale permet à l'algorithme d'assurer une cohérence sur le rendu final, le tout plus rapidement que les méthodes de complétion automatique d'image concurrentes.

## 3. Données multispectrales

Cette section fournit des détails sur le capteur multispectral utilisé dans cette étude ainsi que des informations sur les étalonnages et traitements effectués pour analyser les données multispectrales.

### 3.1. Spécifications

La technologie utilisée dans cette étude (figure 2a) a été développée par la société IMEC (Geelen et al., 2014). Le capteur est compact et permet d'acquérir en temps réel des instantanés (contrairement aux systèmes à balayage). La caméra mesure la luminance  $I(x, y, \lambda)$  d'une scène à travers une image multispectrale 3 dimensions (Hagen et Kudenov, 2013). Cette dernière est composée de plusieurs images 2 dimensions acquises à dif-



**FIGURE 4:** Pertinence de la finesse spectrale. La réflectance spectrale et les valeurs R, V et B des patches **P1** et **P2** (voir figure a) sont illustrées sur la figure b. L'échantillonnage multispectral permet de plus précisément observer les différences de chromatisme. La réduction de ces spectres au couple de valeurs R, V et B (par les fonctions colorimétriques  $\bar{x}$ ,  $\bar{y}$  et  $\bar{z}$ ) entraînent une atténuation de ces variations.

férentes bandes de longueur d'onde. Le capteur peut nominale- ment délivrer 170 images multispectrales par seconde en temps réel.

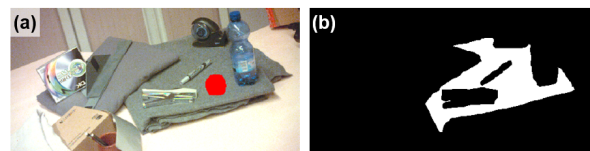
La caméra mesure 16 bandes spectrales comprises entre 475 et 650 nm avec une largeur de bande approxi- mativement égale à 20 nm (figure 2b). La résolution complète du capteur CMOS est de  $2048 \times 1024$  pixels, mais réduite à  $512 \times 256$  pixels pour chaque canal spectral (chaque cellule étant composée d'une mosaïque de  $4 \times 4$  pixels (Geelen et al., 2014)). La profondeur est signée sur 10 bits.

## 3.2. Pré-traitements

### 3.2.1. Étalonnage spectral

L'étalonnage spectral des capteurs multi ou hyper- spectraux consiste à transformer les mesures brutes en réflectance spectrale apparente (Mansouri et al., 2008a,b). Dans cette étude, l'étalonnage spectral a été effectué au moyen d'une mire comprenant 24 patches colorés. Ces derniers sont supposés donner une bonne représentation des spectres du milieu naturel (figure 3a).

Toutes les propriétés et paramètres optiques ne pouvant être estimés au préalable, une méthode indirecte a été employée pour calibrer le capteur multispectral. Nous invitons le lecteur à se référer au mémoire de Ziemer (2013) pour disposer de tous les détails sur l'implé- mentation technique de la procédure. La reconstruction de la réflectance spectrale des patches rouge, vert et bleu de la mire d'étalonnage est présentée en figure 3b. Les spectres de référence et reconstruits sont très proches, même si de légères différences peuvent être observées. Ces dernières sont liées à des incertitudes comprises dans la procédure d'étalonnage :



**FIGURE 5:** (a) Illustration d'une scène de la base de données, initialement acquise avec la caméra multispectral mais convertie au format RVB afin de pouvoir être affichée. La région en rouge représente la zone à compléter par l'algorithme. (b) La zone correspondant au matériau le plus proche de la région à compléter ( $\Omega$ ) est représentée par un masque binaire.

- La camera multispectral effectuant des mesures dans un sous-ensemble du spectre visible (475 à 650 nm au lieu de 380 à 730 nm), des données extrapolées sont introduites lors de la procédure de reconstruction des réflectances spectrales.

- Les courbes de sensibilité spectrale du capteur (figure 2b) indiquent que certaines bandes sont partiellement corrélées entre elles, notamment dans les courtes longueurs d'onde (entre 475 et 500 nm) et aux alentours de 550 nm. Ces corrélations partielles peuvent entraîner une chute de la qualité de la reconstruction spectrale.

- La matrice de passage permettant de convertir les mesures brutes en réflectances spectrales est estimée à l'aide d'un ensemble d'observations tiré d'une seule image spectrale de la mire étalon. L'introduction d'un plus grand nombre d'observations (en faisant varier temps de pose, éclairage et point de vue par exemple) permettrait d'améliorer la qualité des réflectances et des images RVB reconstruites.

### 3.2.2. Conversion de l'espace multispectral vers RVB

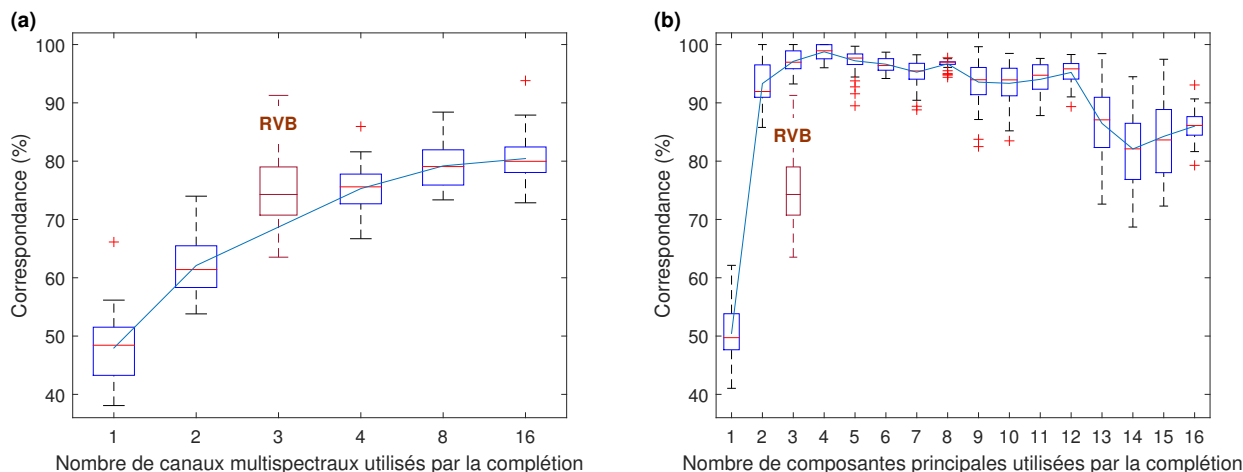
Les algorithmes de complétion produisent des images qui sont affichées sur un écran avant d'être observées par des personnes. Une conversion des images multispectrales vers le standard RVB est donc nécessaire.

En pratique, cette conversion est effectuée sur chaque pixel via la réflectance reconstruite lors de l'étape d'étalonnage spectral (voir section précédente) et au moyen des fonctions colorimétriques de l'observateur standard (CIE XYZ 1931) (Poynton, 2012). Ces dernières sont données entre 380 et 780 nm par pas fixe de 5 nm (voir figure 4b). Elles ont été développées par la Commission Internationale de l'Éclairage (CIE) et permettent de convertir les réflectances apparentes en un couple de trois valeurs  $X$ ,  $Y$  et  $Z$  :

$$\begin{cases} X = \frac{1}{N} \sum \bar{x}(\lambda) S(\lambda) I(\lambda) \\ Y = \frac{1}{N} \sum \bar{y}(\lambda) S(\lambda) I(\lambda) \\ Z = \frac{1}{N} \sum \bar{z}(\lambda) S(\lambda) I(\lambda) \end{cases} \quad (1)$$

$N$  correspond au nombre de valeurs intégrées dans la somme.  $\bar{x}$ ,  $\bar{y}$ ,  $\bar{z}$  correspondent aux fonctions colorimétriques de la CIE (figure 4b) et  $S(\lambda)$  à la réflectance apparente reconstruite par étalonnage spectral (voir section précédente). La réponse spectrale de l'illuminant est





**FIGURE 6:** Comportement de l’algorithme de complétion en termes de respect des matériaux qui entourent la région à compléter. Le pourcentage de correspondance a été calculé entre le masque binaire définissant le matériau le plus proche de la zone à compléter (figure 5b) et la position des pixels sélectionnés par l’algorithme de complétion. Chaque boîte inclut 500 tirages (10 scènes  $\times$  50 tirages par scène). (a) Correspondance calculée en utilisant les canaux multispectraux bruts. La finesse spectrale a été réduite en moyennant progressivement et deux par deux les canaux. (b) Pourcentage de correspondance calculé à partir des composantes principales.

caractérisée par  $I(\lambda)$ . Les valeurs  $X$ ,  $Y$  et  $Z$  sont in fine converties au format *sRGB* (Poynton, 2012). Un exemple de conversion est présenté en figure 3b.

## 4. Analyses préliminaires

### 4.1. Pertinence de l’échantillonnage spectral

La complétion d’image repose sur l’analyse de couleur et de luminosité de patches présents dans l’image. La figure 4a présente un exemple typique, où **P1** et **P2** correspondent à des patches présentant des couleurs RVB similaires. Augmenter la finesse spectrale peut se révéler particulièrement intéressant, notamment afin de relever des informations additionnelles qui ne sont pas forcément observables en RVB. La figure 4b présente la réflectance spectrale ainsi que le couple de valeurs R, V et B des patches **P1** et **P2**. Ici, d’importantes différences chromatiques sont observables entre 590 et 730 nm. Ces disparités très caractéristiques sont atténuées par l’intégration des fonctions colorimétriques de l’observateur standard ( $\bar{x}(\lambda)$ ,  $\bar{y}(\lambda)$  et  $\bar{z}(\lambda)$  sur la figure 4b) utilisées pour calculer les valeurs R, V et B. Ainsi, il semble pertinent d’augmenter la finesse spectrale pour permettre aux algorithmes de complétion de mieux considérer les variations de chromatisme entre les patches.

### 4.2. Procédure expérimentale

Un ensemble de 10 images multispectrales a été utilisé pour évaluer la pertinence de la donnée multispectrale appliquée à la complétion d’image. Les images ont été acquises avec le capteur multispectral présenté en section 3.1. Les scènes ont été sélectionnées pour mettre en avant les limites de l’algorithme de complétion d’image via des objets et matériaux de couleur proche (figure 5a). Les zones à compléter ont été sélectionnées

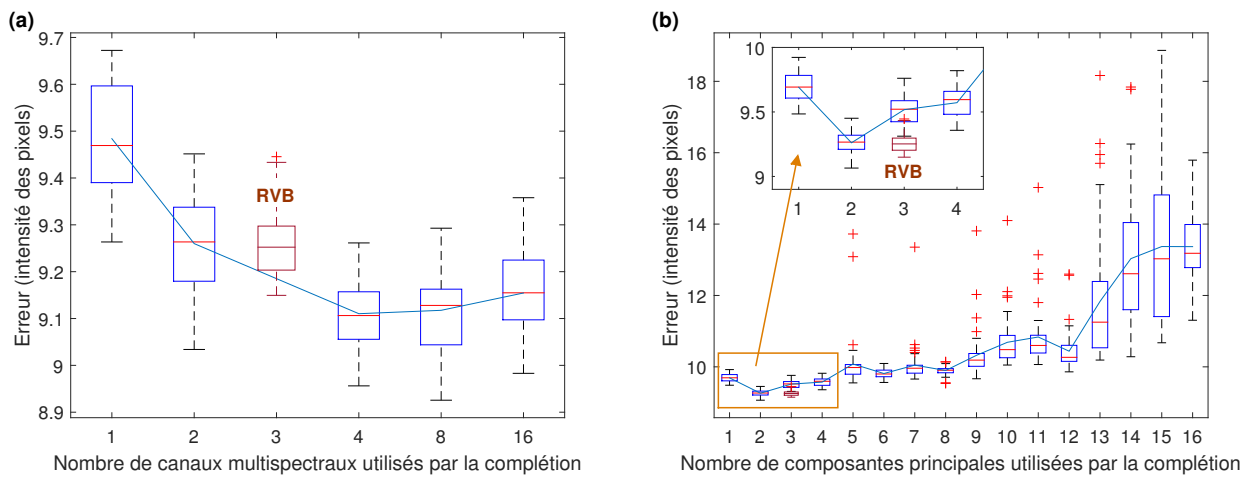
manuellement. Afin d’évaluer le comportement de l’algorithme de complétion, nous avons détourné la zone correspondant au matériau le plus proche de la région à compléter sur chacune des images de la base. Le résultat a été placé dans un masque binaire (figure 5b).

Nous proposons d’évaluer le comportement de l’algorithme de complétion (détaillé en section 4.3) envers les matériaux qui entourent la région à compléter (section 4.4) en faisant, notamment, varier de manière progressive la finesse spectrale. Nous proposons ensuite d’évaluer empiriquement la qualité du rendu en comparant la zone synthétisée par l’algorithme de complétion avec son contenu original (section 4.5).

### 4.3. Détails d’implémentation

PatchMatch, initialement proposé par Barnes et al. (2009), est utilisé en tant que méthode de complétion d’image de référence. Une étape d’optimisation globale permet à l’algorithme d’assurer une cohérence sur le rendu final, le tout plus rapidement que les méthodes de complétion automatique d’image concurrentes. Des étapes spécifiques ont été développées par les auteurs et nous invitons le lecteur à se référer à l’article original (Barnes et al., 2009) pour disposer de tous les détails sur l’implémentation de la procédure.

La méthode est multi-échelle (pyramidale) et repose sur trois grandes étapes : (1) initialisation : Barnes et al. (2009) proposent de tirer aléatoirement les premiers pixels de remplacement candidats sur la plus petite échelle (image la plus réduite en résolution). Dans notre implémentation de l’algorithme, cette phase n’est pas réalisée aléatoirement mais de manière exhaustive, tel que cela a été présenté par Criminisi et al. (2004); (2) propagation : la position des pixels retenus en étape (1) est propagée à l’échelle supérieure; (3) une recherche aléatoire est introduite afin de sélectionner un pixel voisin, potentiellement plus intéressant, par l’évaluation de



**FIGURE 7:** Évaluation empirique de la qualité du rendu en comparant la zone synthétisée par l’algorithme de complétion avec son contenu original. L’erreur est calculée sur les images RVB où seule la position des pixels de remplacement varie. Sur les figures, chaque boîte inclut 500 tirages (10 scènes  $\times$  50 tirages par scène). (a) Erreur calculée en utilisant les canaux multispectraux bruts. La finesse spectrale a été réduite en moyennant progressivement et deux par deux les canaux. (b) Erreur calculée à partir des composantes principales.

sa fonction de coût. Cette dernière consiste en une distance euclidienne calculée entre les trois composantes R, V et B.

Le nombre d’échelles a été fixé à 4 dans cette étude. La plus grande échelle consiste en l’image pleine résolution. L’image est ensuite divisée par 2, par 4, puis par 8 pour la plus petite échelle. La taille des patches a été fixée à  $5 \times 5$  pixels à l’initialisation (et donc pour la plus petite échelle) et à  $13 \times 13$  pour les autres échelles. Dans le cadre de la complétion d’images multispectrales, la fonction de coût (initialement basée sur une distance euclidienne entre les valeurs R, V et B des deux pixels comparés) a été étendue pour intégrer un nombre arbitraire de canaux spectraux. L’algorithme étant composé de phases de recherche aléatoire des pixels de remplacement, 50 tirages par image de la base ont été lancés (un seul tirage par image étant non représentatif).

Les résultats des sections 4.4 et 4.5 sont représentés par des diagrammes de Tukey en figures 6 et 7 respectivement. Sur chacune des boîtes, la barre centrale correspond à la médiane, les extrémités de la boîte aux premier et troisième quartiles, les moustaches s’étendent jusqu’aux points extrêmes (sans considérer les valeurs aberrantes, indiquées par des croix rouges). La ligne bleue représente l’évolution de la valeur moyenne par rapport au nombre de canaux ou de composantes principales.

#### 4.4. Respect des matériaux

Dans cette section, nous proposons d’évaluer le comportement de l’algorithme de complétion face aux matériaux qui entourent la région à compléter en faisant varier de manière progressive la finesse spectrale. Pour cela, nous avons calculé le pourcentage de bonne correspondance ( $\alpha$  dans l’équation 3) entre la région définie par le matériau le plus proche de la zone à compléter et la position des pixels sélectionnés par l’algorithme de

complétion :

$$\gamma(p) = \begin{cases} 1, & \text{si } f \in \Omega \\ 0, & \text{sinon} \end{cases} \quad (2)$$

$\Omega$  correspond au masque binaire définissant le matériau le plus proche de la zone à compléter (figure 5b).  $\gamma$  est défini pour chacun des pixels  $p$  de la zone à compléter.  $f$  correspond à la position du pixel sélectionné par l’algorithme de complétion.  $\gamma$  vaut 1 si ce pixel se trouve dans le masque binaire.

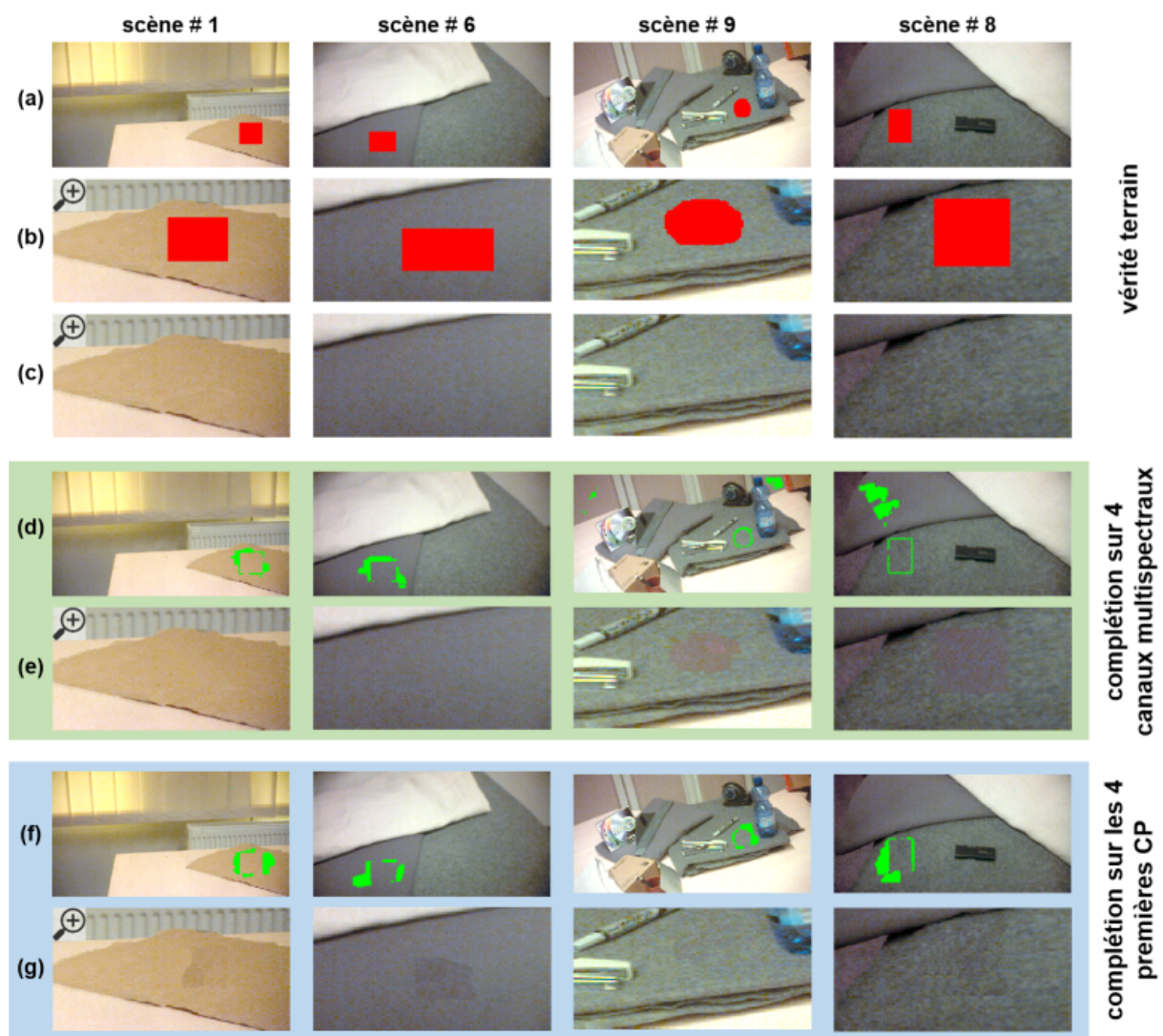
$$\alpha = \frac{100}{N} \sum_{i=1}^N \gamma(p_i) \quad (3)$$

$N$  correspond au nombre de pixels à compléter et  $\alpha$  au pourcentage de bonne correspondance.

Les résultats sont présentés en figure 6a et 6b. Chaque diagramme inclut 500 correspondances (10 images de la base  $\times$  50 tirages par image). À des fins de comparaison, les correspondances calculées à partir des images RVB sont présentées (boîtes rouges).

La figure 6a présente les correspondances calculées lorsque la complétion est effectuée sur les canaux multispectraux bruts. La finesse spectrale a été réduite en moyennant progressivement et deux par deux les canaux. La figure 6b présente le pourcentage de correspondance calculé à partir d’une complétion effectuée sur les composantes principales de l’image multispectrale. Ces dernières ont été calculées à partir de l’analyse en composantes principales avec ajustement d’un modèle de bruit sur les 16 canaux bruts. La transformée a initialement été développée pour l’analyse d’images hyperspectrales (Lee et al., 1990).

Les résultats présentés en figure 6a montrent que le pourcentage de correspondance augmente en fonction du nombre de canaux spectraux considéré. Au contraire, la variance diminue (les boîtes sont de plus en plus compactes). Accroître la finesse spectrale en augmentant



**FIGURE 8 :** Résultats de différentes complétions d'images multispectrales. (a et b) Image d'une scène acquise avec le capteur multispectral. La région en rouge représente la zone à compléter. (c) Gros plan de l'image originale (non modifiée). (d et f) Pixels de remplacement (en vert) sélectionnés par l'algorithme de complétion. Ces derniers ont été calculés (d) à partir de quatre canaux multispectraux et (f) à partir des quatre premières composantes principales (CP). (e) et (g) illustrent en gros plan les images automatiquement complétées.

le nombre de canaux semble permettre à l'algorithme de complétion de mieux considérer les matériaux de la scène. La complétion basée sur les composantes principales (figure 6b) délivre les meilleures performances : le maximum de correspondance est atteint lorsque la complétion est effectuée sur les quatre premières composantes principales (99 % de correspondance au lieu de 80 % au maximum lorsque la complétion est basée sur les canaux spectraux bruts). Utiliser un nombre plus important de composantes principales tend à réduire les scores. Cette observation traduit la chute de qualité dans les composantes élevées, l'algorithme de complétion sélectionnant de plus en plus aléatoirement les pixels de remplacement.

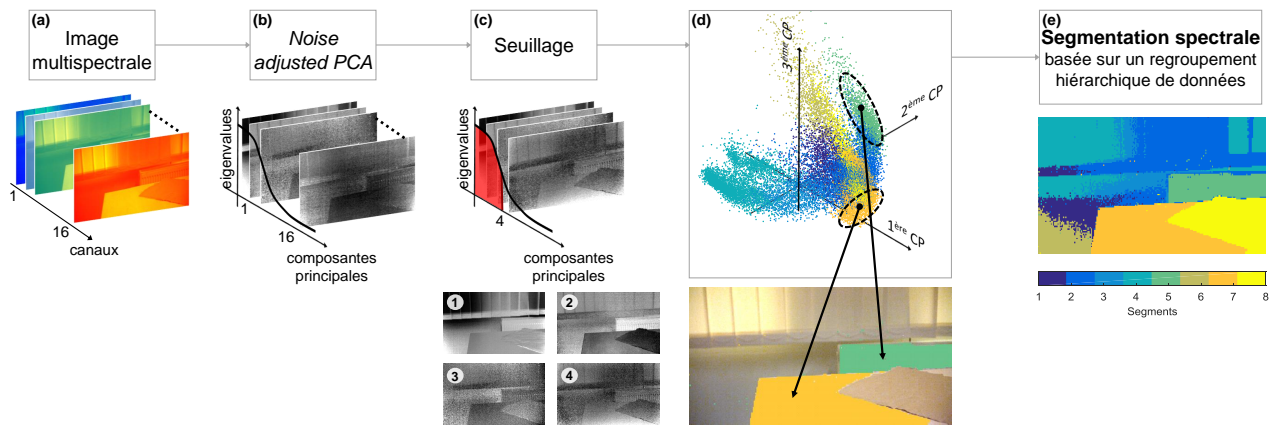
#### 4.5. Qualité des rendus

Dans cette section, nous proposons d'évaluer empiriquement la qualité du rendu en comparant la zone syn-

thétisée par l'algorithme de complétion avec son contenu original. Une fonction d'erreur correspondant à la distance euclidienne RVB entre chaque pixel de la zone à compléter a été utilisée. Notez que seules les positions des pixels de remplacement sélectionnés par l'algorithme de complétion varient. La figure 7a présente l'erreur de rendu de l'image synthétisée par la complétion basée sur les canaux multispectraux bruts.

Le minimum d'erreur est obtenu lorsque quatre canaux sont utilisés. La figure 7b présente le même type d'information, mais lorsque la complétion est effectuée sur les composantes principales de l'image multispectrale (calculées à partir d'une analyse en composantes principales sur les 16 canaux bruts). Comme cela a été présenté en section 4.4, l'augmentation du nombre de composantes principales tend à dégrader la qualité des images synthétisées.

En plus de ces tendances statistiques et afin d'illus-



**FIGURE 9:** Segmentation spectrale. (a) Données d'entrée : image multispectrale (16 canaux). (b) Analyse en composantes principales (Lee et al., 1990). (c) Seuillage afin de ne conserver que les quatre premières composantes principales (section 4.4). (d) Projection des pixels dans l'espace spectral défini par les quatre composantes principales (seul les trois premières sont représentées). Dans cet espace, les regroupements de pixels semblent s'associer avec la géométrie et les matériaux de la scène. (e) Segmentation résultante, basée sur une approche de type regroupement hiérarchique de données (*hierarchical data clustering*).

trer le comportement de la complétion, des résultats sur différentes scènes sont présentés en figure 8. La complétion basée sur quatre canaux spectraux produit des résultats plausibles lorsque les matériaux de la scène sont convenablement respectés. Cette observation est illustrée sur les scènes 1 et 6 de la figure 8. Sur ces mêmes scènes, les images délivrées par la complétion basée sur les quatre premières composantes principales sont de moins bonne qualité : le chromatisme (couleurs) du carton semble respecté mais l'intensité (luminosité) n'est pas bien distribuée. À l'inverse, la complétion basée sur quatre canaux spectraux délivre des images truffées d'inconsistances chromatiques lorsque les matériaux ne sont pas convenablement pris en compte. Cet effet est illustré sur les scènes 8 et 9 de la figure 8. Cette fois, la complétion basée sur les quatre premières composantes principales délivre des résultats plus plausibles, même si de légers artefacts de luminosité sont à noter.

## 5. Complétion contrainte par pré-segmentation spectrale

### 5.1. Motivation

Les analyses préliminaires (section 4) permettent de dégager deux aspects importants :

- La complétion basée sur les composantes principales assure une meilleure stabilité envers les matériaux de la scène (figure 6b).
- Utiliser les canaux multispectraux bruts permet à l'algorithme de produire des images complétées plus plausibles lorsque les matériaux sont convenablement pris en compte (figures 7a et 8).

Il est donc possible de conclure que les composantes principales doivent être considérées pour permettre à la complétion d'utiliser des pixels tirés d'un même matériau. Aussi, l'utilisation des canaux multispectraux permet à la complétion de produire des images synthétisées de meilleure qualité et plus proches des images originales.

La méthode que nous proposons dans cette section est basée sur ces deux observations : la complétion est effectuée sur quatre canaux multispectraux (voir résultats de la section 4.5) mais limitée à une zone prédéfinie et cohérente en termes de matériaux. Cette zone est calculée à l'issue d'une première phase de segmentation spectrale basée sur les quatre premières composantes principales (voir résultats de la section 4.4).

### 5.2. Segmentation spectrale

#### 5.2.1. Méthode

La méthode est décrite graphiquement en figure 9. Une analyse en composantes principales (Lee et al., 1990) est calculée sur l'image multispectrale d'entrée (figure 9b). La dimensionnalité est réduite par une opération de seuillage (figure 9c) où seules les quatre premières composantes sont conservées (voir section 4.4).

La segmentation spectrale est basée sur une approche de type regroupement hiérarchique de données (*agglomerative hierarchical clustering*) qui consiste à regrouper les pixels sous forme d'une arborescence binaire et à l'aide de dendrogrammes. La similarité entre chaque paire de pixels est tout d'abord évaluée au moyen d'une distance euclidienne standard. Notez que chaque pixel est défini par quatre coordonnées, une coordonnée par composante principale. La distance est alors utilisée pour lier chaque paire de pixels proche dans un arbre binaire. Chaque nœud de l'arborescence est constitué de deux pixels. Les nœuds nouvellement formés sont à leur tour liés les uns aux autres afin de constituer un étage supplémentaire de l'arborescence. L'aggrégation est effectuée via la méthode de Ward Jr (1963). L'étape est répétée jusqu'à ce que tous les pixels soient liés à l'arbre binaire, formant ainsi une arborescence sous forme de regroupement hiérarchique.

L'arbre peut contenir des groupes denses de pixels qui soient naturellement séparés du reste des données. Ces groupes peuvent contenir, dans notre cas, des pixels



**TABLE 1:** Évaluation de la segmentation spectrale proposée dans cette étude (section 5.2.1). La précision, le rappel et la distance de Jaccard ont été calculés en utilisant les masques binaires de vérité terrain. Ces derniers ont été manuellement segmentés (section 4.2).

Scène	1	2	3	4	5	6	7	8	9	10	toutes les scènes	
											moyenne ± écart type	
<b>Précision (%)</b>	92	88	90	87	91	89	96	88	58	72	85 ± 11	
<b>Rappel (%)</b>	100	100	96	97	36	91	98	63	89	95	86 ± 21	
<b>Distance de Jaccard (%)</b>	92	88	87	85	35	82	94	58	54	70	74 ± 20	

de matériaux similaires. Par exemple, reprenons la seconde composante principale présentée en figure 9c. L'arbre binaire formé à partir de cette image contiendra quatre groupes très denses de données : un groupe incluant les pixels du carton (pixels noirs), un autre pour les pixels de la table (pixels gris foncé), un troisième groupe pour les pixels du radiateur (pixels blancs) et un dernier groupe pour les pixels du fond (pixels gris clair).

In fine, la segmentation en classes est produite en élaguant l'arbre binaire (figure 9e). Le nombre de classes souhaité doit être choisi afin de ne pas sur ou sous-segmenter la scène. Dans notre cas, la sous-segmentation est à éviter : les pixels qui appartiennent à différents matériaux vont être rassemblés dans un même groupe, risquant ainsi d'altérer la complétion résultante. Les classes situées aux alentours de la zone à compléter étant fusionnées (figure 10), la sur-segmentation est tolérée. Une seule condition doit néanmoins être respectée : les classes fusionnées (figure 10b) doivent contenir une quantité raisonnable de pixels pour que la procédure de complétion puisse être lancée.

En pratique, nous avons utilisé la fonction *clusterdata* incluse dans Matlab (The MathWorks Inc.). La distance euclidienne et la méthode de Ward Jr (1963) ont été utilisées pour créer l'arborescence hiérarchique. Une étape de post-traitement sur le masque binaire résultant des classes fusionnées (figure 10b) a été développée. Des opérations morphologiques sont appliquées afin de sup-

primer les petits groupes de pixels isolés (aire  $\leq 200$  pixels) et afin de remplir les trous (fermeture morphologique basée sur un élément structurant en forme de disque, rayon : 3 pixels). Un exemple de résultat est présenté en figure 10c.

### 5.2.2. Évaluation

La segmentation spectrale est évaluée via la distance de Jaccard et au moyen des indices de précision et rappel (Jardine et van Rijsbergen, 1971). Les métriques ont été calculées entre le masque binaire délivré à l'issue de la procédure (figure 10c) et le masque de vérité terrain correspondant (figure 10d). Les résultats sur l'ensemble des 10 scènes de notre base sont présentés en table 1. De manière générale, la valeur moyenne de la précision et du rappel est supérieure à 85 %.

### 5.3. Contrainte de la procédure de complétion

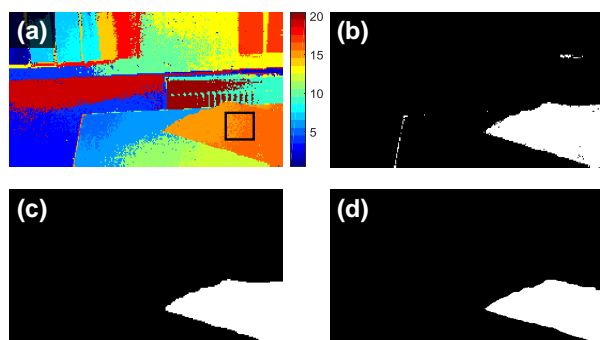
La complétion est contrainte en désactivant le processus de recherche sur les pixels situés en dehors du masque binaire délivré à l'issue de la segmentation spectrale (figure 10c). Pour rappel, cette dernière est basée sur les quatre premières composantes principales (section 4.4). Le processus de complétion consistant à sélectionner les pixels de remplacement utilise quatre canaux multispectraux (section 4.5). Des résultats sur certaines scènes de la base sont illustrés en figure 14.

## 6. Évaluation visuelle subjective

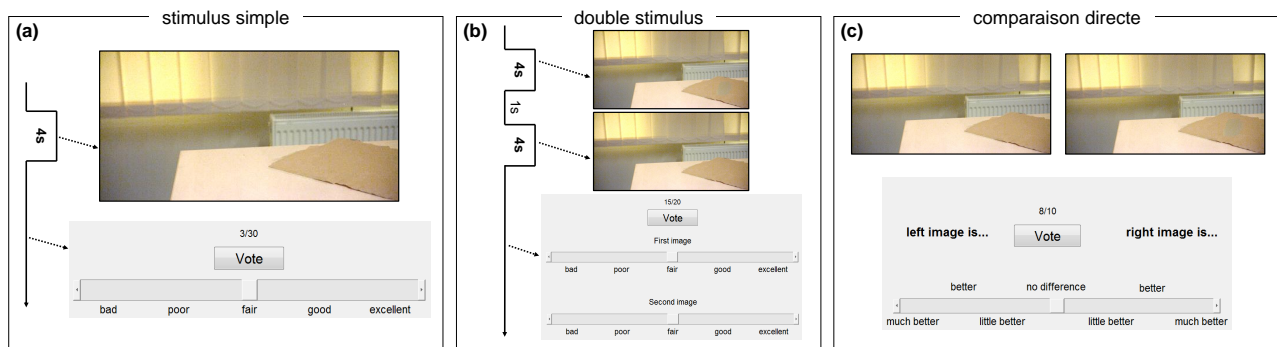
### 6.1. Introduction

Nous avons comparé les rendus produits par la méthode que nous proposons à ceux délivrés par la méthode de référence (complétion reposant sur les trames RVB) au travers d'une évaluation visuelle subjective basée sur 20 observateurs.

Nous proposons, dans cette section, d'étudier la qualité visuelle au travers de méthodes développées par la communauté d'infographie (Mantiuk et al., 2012) à travers un ensemble de questionnaires subjectifs. Les méthodes de synthèse d'image ont pour objectif de produire des résultats plausibles et acceptables visuellement par les personnes. L'erreur calculée entre les images complétées et originales (section 4.5) ne reflète pas nécessairement la qualité telle qu'elle serait perçue par des observateurs humains (Lavoué et Mantiuk, 2015). L'évaluation subjective de la qualité des images repose sur l'analyse de questionnaires permettant aux observateurs



**FIGURE 10:** (a) Résultat d'une segmentation spectrale (nombre de classes : 20). Le rectangle noir représente la zone à compléter. (b) Les classes situées aux alentours de la zone à compléter sont fusionnées. Le résultat est présenté sous la forme d'un masque binaire. (c) Des opérations morphologiques sont appliquées afin de supprimer les groupes de pixels isolés et remplir les trous. (d) Vérité terrain (voir section 4.2).



**FIGURE 11:** Interface permettant d'évaluer les images synthétisées par les algorithmes de complétion. (a) Stimulus simple. Les observateurs notent la qualité de l'image affichée. (b) Double stimulus. Les observateurs notent la qualité d'une première puis d'une deuxième image. (c) Comparaison directe. Les observateurs indiquent leur préférence en évaluant la différence de qualité entre les deux images.

d'exprimer leurs préférences par des sélections et classements (Mantiuk et al., 2012). Les temps de vote sont également enregistrés. Ils reflètent le degré de difficulté rencontré par les observateurs pour effectuer la tâche.

Les 10 scènes de la base (section 4.2) et acquises par la caméra multispectrale ont été évaluées par un ensemble de 20 observateurs (3 femmes et 17 hommes, leur âge étant compris entre 25 et 37 ans). Trois images ont été utilisées pour évaluer chaque scène : (1) image complétée par la méthode de complétion de référence (PatchMatch sur l'image RVB) ; (2) image complétée par la technique proposée dans cet article (section 5) ; (3) image originale (non modifiée).

## 6.2. Méthodes d'évaluation

Chaque observateur a reçu un ensemble d'indications sur le déroulement des expériences. Une première session d'apprentissage leur a été proposée afin qu'ils puissent se familiariser avec les contrôles de l'interface graphique. Les images sont présentées de manière aléatoire. L'expérience dure en moyenne une quinzaine de minutes. Trois méthodes ont été mises en œuvre pour évaluer la qualité des images :

**Simple stimulus.** Une interface (figure 11a) permet à l'observateur de noter la qualité d'une image via une échelle de Likert (continue sur 5 points). Chaque image est affichée pendant 4 secondes. L'utilisateur doit évaluer 30 images : 10 images originales (non modifiées) ; 10 images complétées par notre méthode (complétion spectrale) et 10 images complétées par l'algorithme de référence (complétion basée sur les images RVB).

**Double stimulus.** L'expérience est similaire à celle basée sur le simple stimulus, excepté que deux images (complétée et originale) sont successivement présentées à l'observateur de manière aléatoire (figure 11b). L'utilisateur doit noter 40 images ( $2 \times 10$  images liées à la complétion de référence et  $2 \times 10$  images liées à la complétion spectrale).

**Comparaison directe.** Les observateurs indiquent leur préférence (figure 11c) en quantifiant la différence de qualité entre les images synthétisées (de la même scène) par chacune des méthodes de complétion. Une

échelle continue sur 7 points a été utilisée. Les utilisateurs ont la possibilité de positionner le marqueur au centre de l'échelle si aucune différence n'est perçue.

## 6.3. Résultats et analyses

**Simple et double stimulus.** Il n'est pas conseillé d'exploiter directement les notations liées aux expériences simple et double stimulus (Mantiuk et al., 2012), mais plutôt de présenter les scores sous forme différentielle (*difference mean opinion scores*) :

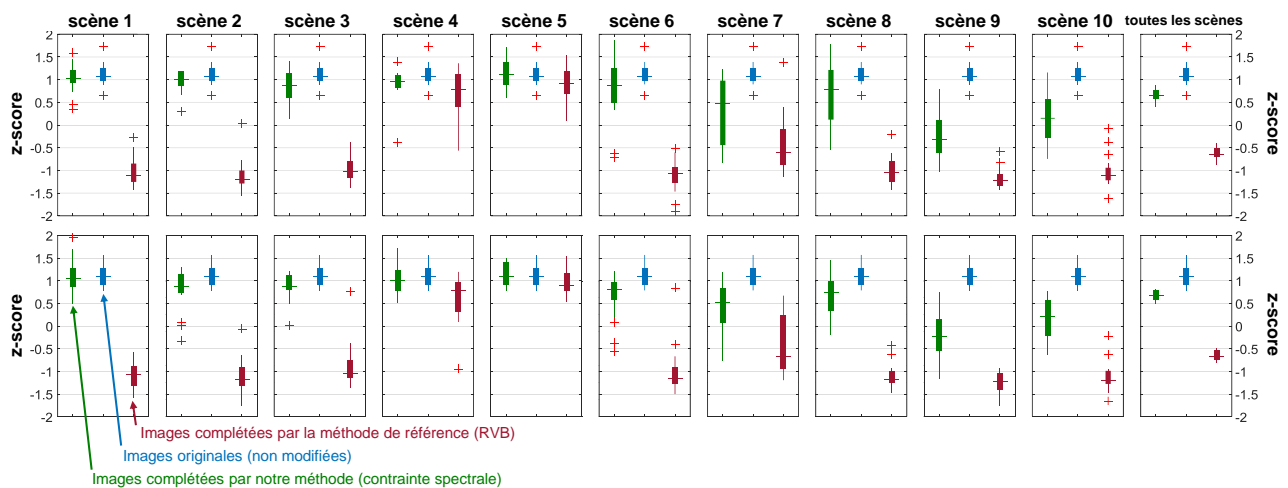
$$d_{i,j,k} = r_{i,\text{ref}(k),k} - r_{i,j,k} \quad (4)$$

$$z_{i,j,k} = \frac{d_{i,j,k} - \bar{d}_i}{\sigma_i} \quad (5)$$

$r_{i,j,k}$  correspond à la note pour une image donnée (originale ou complétée). Les indices correspondent à l'observateur  $i$ , la méthode de complétion  $j$  (référence ou spectrale) et à la scène  $k$ .  $\text{ref}(k)$  correspond à l'image originale pour la scène  $k$ . Les scores centrés (en annulant la moyenne  $\bar{d}_i$ ) et normalisés (par l'écart-type  $\sigma_i$  calculé entre les observateurs) sont présentés en figure 12. Les résultats indiquent que les observateurs ont globalement mieux apprécié les images complétées par la méthode que nous proposons dans cette étude. Les scores des scènes 4 et 5 indiquent que les observateurs ont fourni des opinions similaires et particulièrement positives dans les deux cas.

**Comparaison directe.** Les observateurs ne votant pas nécessairement de la même manière, les scores ont été normalisés par l'écart-type calculé sur les votes de chaque observateur. Les résultats sont similaires aux scores calculés pour les expériences basées sur les simple et double stimulus, excepté que la moyenne  $\bar{d}_i$  n'est pas enlevée. Le signe du vote de comparaison nous indique quelle image a été préférée par les observateurs.

Les résultats sont présentés en figure 13a. Les votes de comparaison laissent transparaître une préférence envers les images complétées par la méthode que nous proposons. En particulier, 186 sur 200 votes vont en faveur de la complétion par contrainte spectrale (valeurs positives sur la figure 13a), 7 votes sur 200 indiquent



**FIGURE 12:** z-scores pour chaque scène (voir section 6.3) pour l'expérience basée sur les simples stimuli (première ligne) et sur les doubles stimuli (deuxième ligne). Sur chacune de ces figures, la boîte de gauche donne la répartition des scores pour les votes liés aux images complétées par la méthode proposée dans cette étude. La boîte centrale correspond aux votes liés à l'image originale. La boîte de droite traduit les votes de l'image complétée par la méthode de référence (basée RVB).

qu'aucune différence n'a été observée entre les deux images et 7 votes sur 200 vont en faveur de la méthode de référence (valeurs négatives sur la figure 13a).

Les résultats semblent moins tranchés pour les scènes 4 et 5. La qualité des résultats délivrés par chacune des deux méthodes de complétion ayant été perçus identiquement par les observateurs (figure 12), ces résultats indiquent que les observateurs n'ont pas réussi à percevoir de différences notables de qualité entre les deux rendus. Cette donnée est corrélée avec le temps passé par les observateurs à évaluer ces deux scènes (figure 13b) qui est, en moyenne, plus important que le temps passé sur les autres scènes.

## 7. Discussion

Préférer l'utilisation de caméras multispectrales aux capteurs hyperspectraux permet aujourd'hui d'envisager une exploitation temps réel des méthodes. Ce pré-requis est nécessaire pour bien des applications pratiques, en particulier dans le domaine de la vision par ordinateur.

L'intégration directe de l'information spectrale, signée sur 16 canaux dans notre cas, entraîne une augmentation drastique des temps de calcul. Cette limitation a été considérée dans notre étude via l'intégration de technique de réduction de dimensionnalité. De part les résultats présentés en section 4.4 et 4.5, seules les quatre premières composantes principales sont utilisées pour former les cartes de segmentation spectrale tandis que quatre canaux multispectraux sont utilisés pour compléter l'image par détermination de la position des pixels de remplacement.

### 7.1. Limitations

#### 7.1.1. Augmentation du nombre d'échantillons de la base

10 images multispectrales ont été utilisées pour générer les résultats présentés dans cette étude. Les scènes ont été sélectionnées pour mettre en avant les limites des algorithmes de complétion d'image via des objets et matériaux de couleur proche. En raison des phases de recherche aléatoire introduites dans la procédure de complétion, 50 tirages par image de la base ont été lancés. Malgré le faible nombre d'images introduites dans la base, nous pensons que les statistiques présentées en sections 4.4 et 4.5 sont suffisamment représentatives.

#### 7.1.2. Fusion des dimensions spatiales et spectrales

La segmentation spectrale proposée dans cette étude (section 5.2) est basée sur une recherche de groupes de pixels similaires dans l'espace spectral. Ces groupes ont tendance à respecter la géométrie et le matériaux des objets de la scène mais aucune information explicite, comme les textures ou les formes des objets, n'est actuellement incorporée à la méthode.

### 7.2. Travaux futurs

Au regard des limitations présentées en section 7.1, des travaux seront menés pour augmenter le volume de la base de données d'images multispectrales en intégrant, entre autres, des scènes tirées du milieu naturel.

Aussi, les techniques permettant d'utiliser conjointement les dimensions spatiales et spectrales doivent être intégrées pour améliorer la qualité des segmentations produites. Pour ce faire, les techniques reposant sur les *Schrödinger Eigenmaps* (Cahill et al., 2014), sur les machines à tenseurs de support (*Support Tensor Machines*) (Guo et al., 2016), ou encore sur l'apprentissage profond

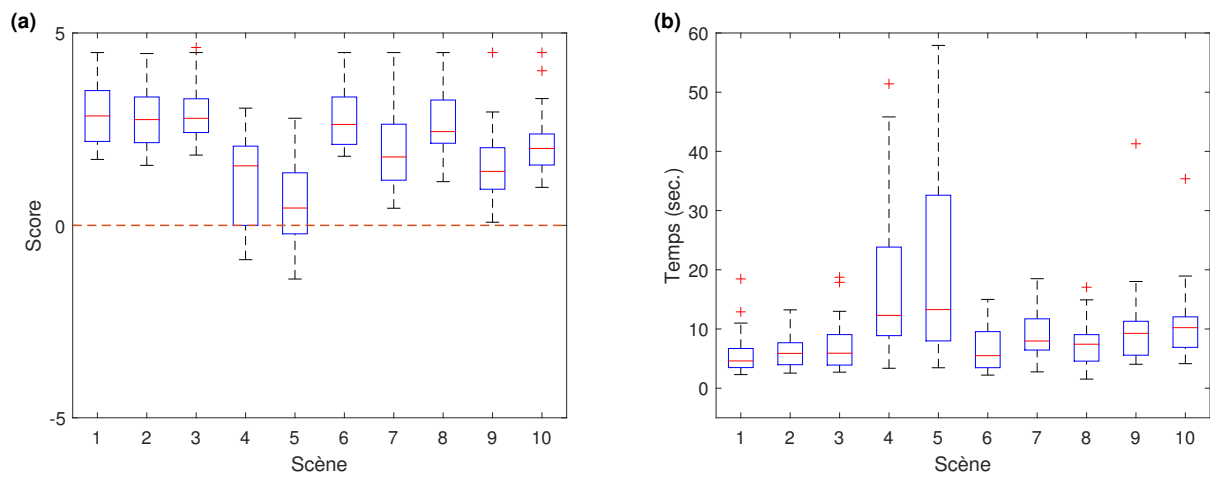
(Romero et al., 2016) correspondent à des pistes prometteuses qu'il sera important d'approfondir.

## 8. Conclusion

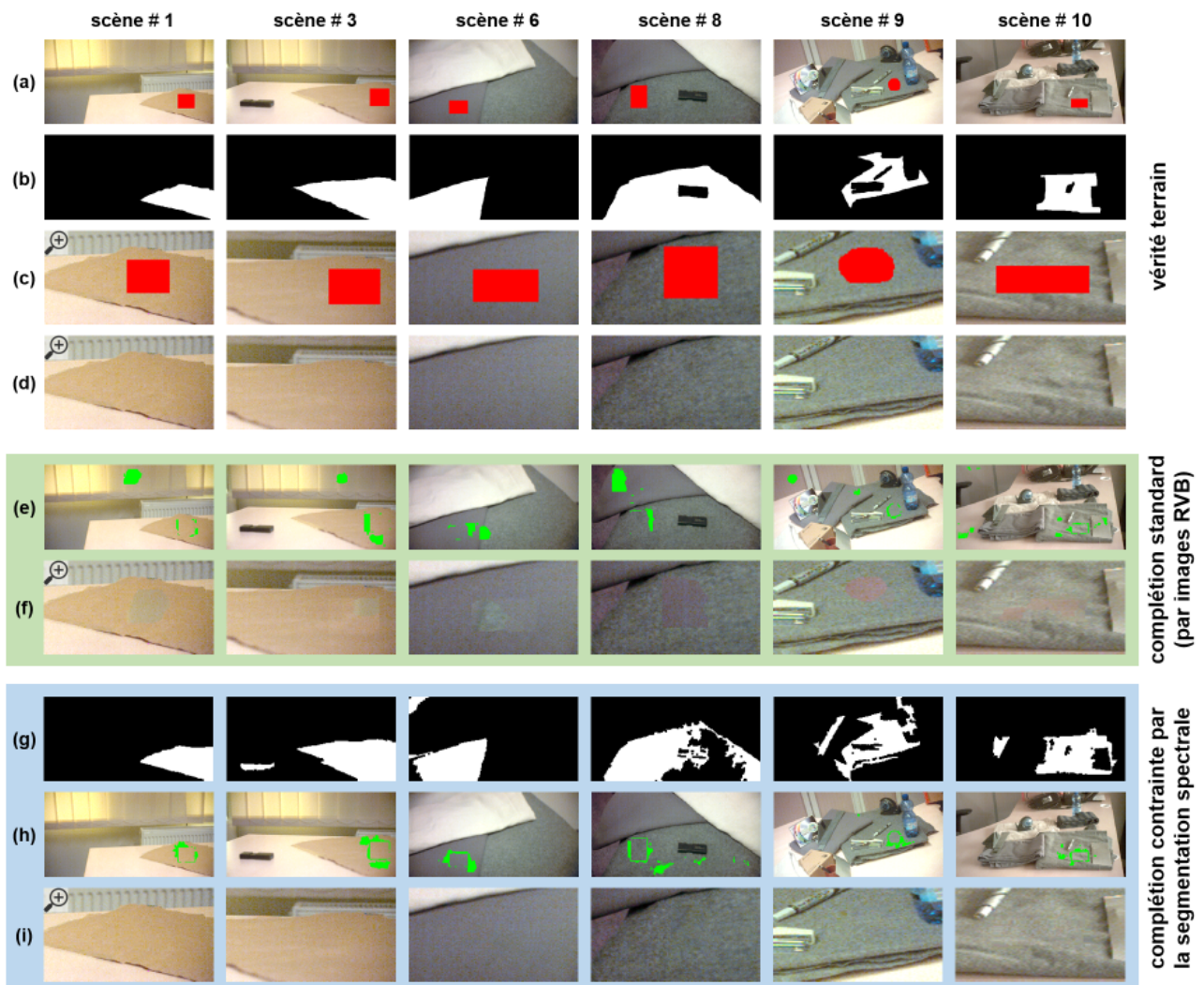
Nous avons proposé, dans cette étude, d'évaluer le potentiel de l'imagerie multispectrale appliquée à la complétion automatique d'image. Les analyses préliminaires ont indiqué que l'intégration directe des données multispectrales ne produit pas d'amélioration significative. Au regard de ces observations, nous avons proposé une méthode permettant d'exploiter la donnée multispectrale en pré-segmentant la scène en classes d'intérêt. Seules les classes situées aux alentours de la zone à compléter sont considérées lors de la recherche de pixels de remplacement. Les résultats indiquent que la méthode que nous proposons produit des images de qualité supérieure tout en respectant les matériaux de la scène.

Les résultats ont été validés à l'aide de critères numériques et par une évaluation visuelle subjective. La méthode que nous proposons produit des images cohérentes et compatibles avec les standards en vision par ordinateur et en infographie. Les caméras multispectrales temps réel correspondent à des systèmes très prometteurs qui permettent de mieux estimer les propriétés physiques des objets visionnés. La technologie est donc tout à fait adaptée aux problématiques de vision par ordinateur.





**FIGURE 13:** (a) Comparaison directe pour chacune des scènes. Les résultats positifs indiquent que l'image complétée par la méthode que nous proposons dans cette étude a été préférée à l'image complétée par la méthode de référence (basée RVB). (b) Temps de vote pour chaque scène.



**FIGURE 14:** Comparaison des rendus entre les différentes méthodes de complétion. (a) Image d'une scène acquise avec le capteur multispectral. (b) Masque binaire de vérité terrain illustrant la région contenant le matériau le plus proche de la zone à compléter (manuellement segmentée). (c) Gros plan de (a). Dans (a) et (c), la région en rouge représente la zone à compléter. (d) Gros plan de l'image originale (non modifiée). (e et h) Pixels de remplacement (en vert) sélectionnés par la procédure de complétion. Ces derniers ont été calculés (e) à partir de la méthode de référence (basée sur les images RVB) et (h) à partir de la méthode que nous proposons (section 5). (g) Masque binaire calculé à l'issue de la procédure de segmentation spectrale. La recherche des pixels de remplacement est géométriquement contrainte à l'aire définie sur ce masque (en blanc). (f et i) Résultats des complétions respectives (gros plan).

## Références

- Baek, S.-H., Choi, I., Kim, M. H., 2016. Multiview image completion with space structure propagation. Dans : Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D., 2009. Patchmatch : A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics* 28 (3), 24.
- Cahill, N. D., Czaja, W., Messinger, D. W., 2014. Schroedinger eigenmaps with nondiagonal potentials for spatial-spectral clustering of hyperspectral imagery. Dans : SPIE Defense+ Security. International Society for Optics and Photonics, p. 908804.
- Chakrabarti, A., Zickler, T., 2011. Statistics of real-world hyperspectral images. Dans : Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 193–200.
- Chen, Y.-L., Hsu, C.-T., Liao, H.-Y. M., 2014. Simultaneous tensor decomposition and completion using factor priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36 (3), 577–591.
- Criminisi, A., Perez, P., Toyama, K., 2004. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing* 13 (9), 1200–1212.
- Efros, A. A., Leung, T. K., 1999. Texture synthesis by non-parametric sampling. Dans : The Proceedings of the Seventh IEEE International Conference on Computer Vision. Vol. 2. pp. 1033–1038.
- Geelen, B., Tack, N., Lambrechts, A., 2014. A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic. Dans : SPIE MOEMS-MEMS. International Society for Optics and Photonics, p. 89740L.
- Gendrin, A., Mangold, N., Bibring, J.-P., Langevin, Y., Gondet, B., Poulet, F., Bonello, G., Quantin, C., Mustard, J., Arvidson, R., et al., 2005. Sulfates in martian layered terrains : the omega/mars express view. *Science* 307 (5715), 1587–1591.
- Gordon, H. R., Morel, A. Y., 2012. Remote assessment of ocean color for interpretation of satellite visible imagery : A review. Vol. 4. Springer Science & Business Media.
- Guillemot, C., Le Meur, O., 2014. Image inpainting : Overview and recent advances. *IEEE Signal Processing Magazine* 31 (1), 127–144.
- Guo, X., Huang, X., Zhang, L., Zhang, L., Plaza, A., Benediktsson, J. A., 2016. Support tensor machines for classification of hyperspectral remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing* 54 (6), 3248–3264.
- Hagen, N., Kudenov, M. W., 2013. Review of snapshot spectral imaging technologies. *Optical Engineering* 52 (9), 090901.
- He, K., Sun, J., 2014. Image completion approaches using the statistics of similar patches. *IEEE Transactions on Pattern Analysis and Machine Interlligence* 36 (12), 2423–2435.
- Herling, J., Broll, W., 2014. High-quality real-time video inpainting with pixmix. *IEEE Transactions on Visualization and Computer Graphics* 20 (6), 866–879.
- Huang, J.-B., Kang, S. B., Ahuja, N., Kopf, J., 2014. Image completion using planar structure guidance. *ACM Transactions on Graphics* 33 (4), 129.
- Jardine, N., van Rijsbergen, C. J., 1971. The use of hierarchic clustering in information retrieval. *Information storage and retrieval* 7 (5), 217–240.
- Kawai, N., Sato, T., Yokoya, N., 2016. Diminished reality based on image inpainting considering background geometry. *IEEE Transactions on Visualization and Computer Graphics*, 1236–1247.
- Lavoué, G., Mantiuk, R., 2015. Quality assessment in computer graphics. Dans : Visual Signal Quality Assessment. Springer, pp. 243–286.
- Lee, J. B., Woodyatt, A. S., Berman, M., 1990. Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform. *IEEE Transactions on Geoscience and Remote Sensing* 28 (3), 295–304.
- Li, Q., He, X., Wang, Y., Liu, H., Xu, D., Guo, F., 2013. Review of spectral imaging technology in biomedical engineering : achievements and challenges. *Journal of biomedical optics* 18 (10), 100901.
- Mansouri, A., Sliwa, T., Hardeberg, J. Y., Voisin, Y., 2008a. An adaptive-pca algorithm for reflectance estimation from color images. Dans : 19th International Conference on Pattern Recognition. IEEE, pp. 1–4.
- Mansouri, A., Sliwa, T., Hardeberg, J. Y., Voisin, Y., 2008b. Representation and estimation of spectral reflectances using projection on pca and wavelet bases. *Color Research & Application* 33 (6), 485–493.
- Mantiuk, R. K., Tomaszewska, A., Mantiuk, R., 2012. Comparison of four subjective methods for image quality assessment. Dans : Computer Graphics Forum. Vol. 31. Wiley Online Library, pp. 2478–2491.
- McCamy, C. S., Marcus, H., Davidson, J., 1976. A color-rendition chart. *J. App. Photog. Eng* 2 (3), 95–99.
- Newson, A., Almansa, A., Fradet, M., Gousseau, Y., Pérez, P., 2014. Video inpainting of complex scenes. *SIAM Journal on Imaging Sciences* 7 (4), 1993–2019.
- Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A. A., 2016. Context encoders : feature learning by inpainting. Dans : Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- Pope, A., Rees, W. G., 2014. Impact of spatial, spectral, and radiometric properties of multispectral imagers on glacier surface classification. *Remote Sensing of Environment* 141, 1–13.
- Poynton, C., 2012. The cie system of colorimetry. Dans : Digital Video and HD : Algorithms and Interfaces, 2ème Edition. Morgan Kaufmann Publishers Inc., Ch. 25, pp. 265–286.
- Romero, A., Gatta, C., Camps-Valls, G., 2016. Unsupervised deep feature extraction for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing* 54 (3), 1349–1362.
- Telea, A., 2004. An image inpainting technique based on the fast marching method. *Journal of graphics tools* 9 (1), 23–34.
- Ward Jr, J. H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* 58 (301), 236–244.
- Wexler, Y., Shechtman, E., Irani, M., 2007. Space-time completion of video. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (3), 463–476.
- Ziemer, P., 2013. Design and implementation of a multispectral imaging system. Rapport de Master.