

---

# Colorisation de nuages de points 3D par recalage dense d'images numériques

**Nathan Crombez, Guillaume Caron, El Mustapha Mouaddib**

*Université de Picardie Jules Verne*

*Laboratoire MIS*

*33 rue Saint Leu*

*80039 Amiens CEDEX 1*

*{nathan.crombez,guillaume.caron,mouaddib}@u-picardie.fr*

---

*RÉSUMÉ. Le patrimoine architectural est composé de biens historiques et artistiques qui doivent être protégés, préservés, restaurés et exposés au plus grand nombre. Des appareils modernes tels que les scanners à laser 3D sont de plus en plus utilisés en documentation culturelle. Ces outils permettent de générer avec précision et rapidité des nuages de points de monuments historiques. Avec les données collectées, il est possible de créer un maillage afin de visualiser virtuellement les formes et/ou la surface de l'édifice. La plupart du temps, le scanner tridimensionnel est équipé d'un appareil photo numérique qui est utilisé pour coloriser les points relevés. Cependant, la qualité photométrique du nuage de points n'est pas toujours suffisante principalement à cause de problèmes de couleurs et de résolution. Des méthodes d'uniformisation d'intensités existent pour améliorer la colorimétrie mais ne permettent pas d'obtenir des rendus photo-réalistes et d'améliorer la résolution. C'est pourquoi, nous proposons une nouvelle méthode pour coloriser les nuages de points à l'aide d'images numériques de haute résolution acquises avec un appareil photo. Pour cela, nous avons développé une méthode permettant d'obtenir un recalage précis entre des images numériques acquises et un nuage de points, ce qui est une étape cruciale pour une bonne colorisation par projection de couleurs. Des résultats sur des jeux de données issus de la numérisation de la cathédrale d'Amiens à l'intérieur et à l'extérieur démontrent la validité de notre approche en obtenant des nuages de points de qualité et de résolution photométriques nettement meilleures.*

*ABSTRACT. Architectural heritage is a historic and artistic property which has to be protected, preserved, restored and must be shown to the public. Modern tools like 3D laser scanners are more and more used in heritage documentation. These systems allow fast generation of accurate point clouds of historical monuments. Using these data, we are able to create a polygon mesh that defines the shape of the edifice for virtual visualization. Most of the time, the 3D laser scanner is completed by a digital camera which is used to colorize the point cloud. However, the photometric quality of point clouds is generally rather low, mainly due to color and resolution problems. Intensity uniformization methods exist to improve the colorimetry but they do not permit to obtain a photo-realist rendering nor to globally improve the resolution. That is why this paper proposes a solution to colorize point clouds using high resolution digital images acquired with a camera from any viewpoint. For this, we have developed a new accurate method for*

*registering the photographs on the point cloud which is a crucial step for a good colorization by color projection. Results, on datasets of the cathedral of Amiens in France, highlight the success of our approach, leading to point clouds with better photometric quality and resolution.*

*MOTS-CLÉS : nuage de points, colorisation, optimisation visuelle virtuelle.*

*KEYWORDS: point clouds, colorization, visual and virtual optimization.*

---

DOI:10.3166/TS.31.81-106 © 2014 Lavoisier

### **Extended abstract**

Architectural heritage contains important remnants of our past, it is a historic and artistic property which has to be protected, preserved, restored and must be shown to the public. Cultural heritage includes monuments, works of art or archaeological landscapes. These sites are frequently faced with environmental conditions, human deterioration and can be victim of their age. It is in this spirit that E-Cathedrale, a research and development program, aims to construct a complete digital model of the Amiens Cathedral in France.

Modern tools like 3D laser scanners are more and more used in heritage documentation. These systems allow to scan quickly and accurately the spatial structure of these kind of sites. The scan result is a dense and panoramic 3D point cloud of the scanned scene. These measures may be useful for future restoration, analysis, preservation, reconstructions or virtual displaying. Using these data, we are able to create a polygon mesh that defines the shape of the scene for its virtual visualization. Most of the time, the 3D laser scanner is completed by a digital camera which is used to enrich the geometric information with the scanned objects color. However, the photometric quality of point clouds is generally rather low because of two main problems. The first problem is due to the digital camera resolution used for the points colorization. When the points sampling rate is upper than the digital camera resolution, a 3D point and its nearest neighbors are colorized with the same RGB value. This phenomenon creates a blur effect on the colored point clouds which are not representative of the real object aspect. The second problem comes from the fact that to entirely cover a scene, the laser scanner should be placed at different geographical positions. We obtain a point cloud for each scanner position. Despite fast acquisitions, these point clouds have been acquired at different time of the day, so at different sun exposures. Because of that, two points, side by side from two separate clouds, may have a completely different color.

Intensity uniformization methods exist to improve the colorimetry but they do not permit to obtain a photo-realist rendering nor to globally improve the resolution. That is why this paper proposes a solution to colorize point clouds using high resolution digital images acquired with a camera from any viewpoint. For this, we have developed a new accurate method for registering the photographs on the point cloud which is a crucial step for a good colorization by color projection.

Firstly, the point cloud colors from the different scans are homogenized. Thanks to that, we are able to detect and to match point features between a virtual image of the 3D scene and a digital image. A camera pose is computed using these sparse sets of corresponding points to register images and the 3D point cloud. Then, this initial camera pose is corrected. The pose optimization problem is formulated as the minimization of the difference of a digital image and a virtual one rendered from the initial pose. Finally, we have the real camera pose which took the digital image. We can now project the visible points of the cloud in the digital image frame and set new point colors.

Results, on datasets of the cathedral of Amiens, highlight the success of our approach, leading to point clouds with better photometric quality and resolution.

## 1. Introduction

Quoi de plus vrai que la célèbre citation de Confucius « Etudiez le passé si vous souhaitez définir le futur »? Le patrimoine architectural est composé de vestiges considérables de notre passé. Nous devons assurer leur protection. C'est pourquoi la sauvegarde du patrimoine est un axe de recherches important. Le patrimoine culturel comprend des monuments, des bâtiments et des lieux historiques. Ces sites peuvent régulièrement être soumis à des conditions environnementales difficiles, à des détériorations causées par l'homme et par le temps.

C'est dans cet esprit que le projet de recherche e-Cathédrale<sup>1</sup> a vu le jour. Il a pour objectif de réaliser et d'exploiter une maquette virtuelle fidèle et complète de la cathédrale Notre-Dame d'Amiens. C'est la cathédrale complète la plus haute de France avec une hauteur sous voûte de 42,30 mètres pour un volume intérieur de l'ordre de 200 000 mètres cube.

De récentes avancées technologiques ont permis le développement de méthodes et d'outils, comme les scanners à laser 3D, capables de mesurer avec précision la structure spatiale de ces monuments. La capacité de générer rapidement et précisément des nuages de points a fait de ces scanners à laser 3D des outils de plus en plus utilisés en documentation historique. La cathédrale d'Amiens a été partiellement scannée, les trois premières campagnes de numérisation ont permis de relever des milliards de points avec une résolution d'environ 2 millimètres sur les zones présentant des détails et 5 millimètres sur les parties moins détaillées (figure 1).

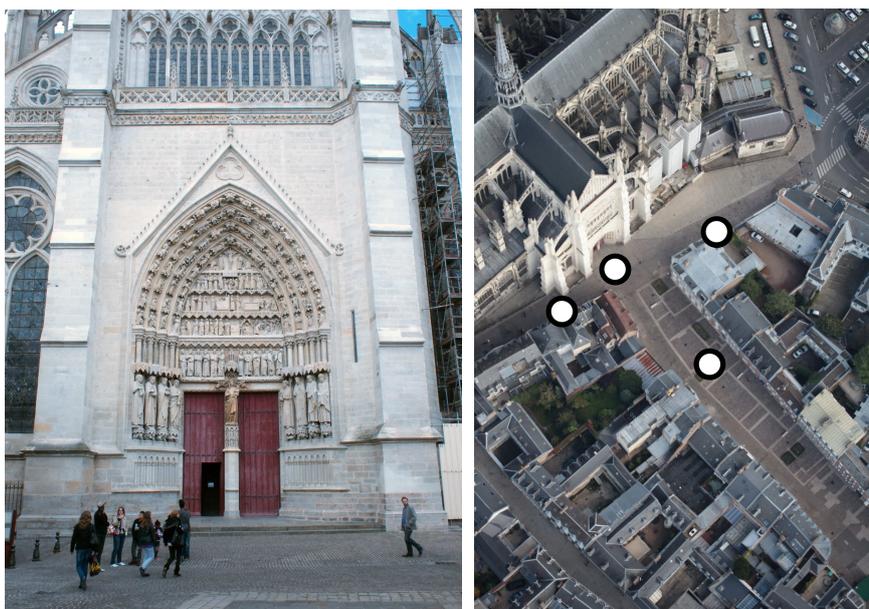
Pour cette étude, nous nous concentrerons principalement sur le portail sud de l'édifice appelé le portail de la Mère-Dieu ou de la Vierge (figure 2a). Ce portail mesure 20 mètres de haut pour 14 mètres de large. Pour couvrir le portail sud dans son intégralité, en évitant les occultations, le scanner a été placé à quatre positions comme le montre la figure 2b.

---

1. [www.mis.u-picardie.fr/E-Cathedrale/](http://www.mis.u-picardie.fr/E-Cathedrale/)



Figure 1. Les nuages de points de la partie sud de la cathédrale d'Amiens assemblés



(a)

(b)

Figure 2. (a) Une image numérique acquise de haute qualité du portail sud de la cathédrale. (b) Les positions géographiques (points blancs) du scanner à laser pour couvrir le portail sud

Un appareil photo numérique (APN) est monté sur le scanner à laser afin d'affecter une couleur RVB (rouge, vert, bleu) à chaque point du nuage relevé. Cependant, à

cause de la faible définition et la faible dynamique d'intensités des APN utilisés, les nuages de points obtenus sont ternes par leur couleur et leur résolution en termes de texture. Malgré la grande précision de mesure des points, nous devons faire face à deux problèmes de colorisation des nuages de points. Le premier problème est dû à la faible définition du capteur de l'appareil photo utilisé pour la colorisation. En effet, la résolution du nuage de points est tellement élevée, par rapport à la définition de la caméra, qu'un seul unique triplet RVB colorise avec la même valeur tout un groupe de points 3D. Ce phénomène crée un effet de flou sur le nuage de points, qui est alors loin d'être représentatif de l'aspect de l'objet réel. Le second problème vient du fait que nous devons placer le scanner à différentes positions pour pouvoir couvrir entièrement une scène étendue comme la façade sud de la cathédrale d'Amiens (figure 2a). Nous obtenons un nuage de points pour chaque position du scanner. En dépit de la grande fréquence d'acquisition (quelques minutes par station), les nuages de points sont obtenus à différents moments de la journée et donc à différentes heures d'ensoleillement. A cause de cela, deux points côte à côte provenant de deux nuages distincts d'une même zone du monument, peuvent avoir une couleur complètement différente (figure 3).



*Figure 3. Quatre nuages de points bruts obtenus à partir des différentes positions du scanner à laser (à gauche) et le nuage de points issu de leurs recalage (à droite)*

L'image de droite de la figure 4 est générée avec le même nuage de points que celui de la figure 3 mais d'un autre point de vue et en niveaux de gris. Les problèmes de résolution et de couleur du nuage sont clairement visibles lorsque l'on compare l'image de synthèse et l'image réelle (figure 4) prise à une position similaire.

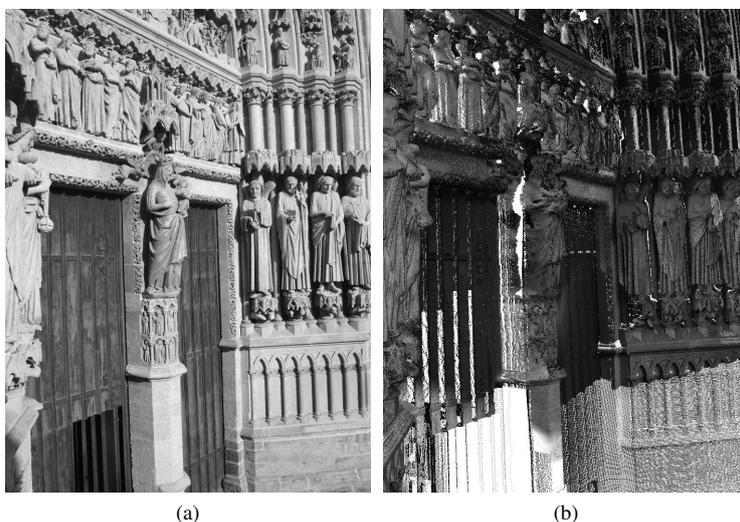


Figure 4. (a) une image numérique  $I$  acquise à partir d'une pose que l'on aimerait atteindre virtuellement. (b) Une image de synthèse  $I_v$  générée avec le nuage de points

## 2. Travaux connexes

Ces dernières années, les technologies de prélèvement 3D de surfaces sont devenues les principales méthodes pour mesurer avec précision des objets architecturaux ou historiques. Les objets à mesurer peuvent être petits comme la tombe de Sidamara (Altuntas *et al.*, 2007) ou de grande taille comme les palais du désert de Jordanie (Al-kheder *et al.*, 2009) ou encore le Todai-ji au Japon, plus grand édifice en bois du monde (Ikeuchi *et al.*, 2007). Dans tous les cas, l'objectif est d'obtenir des documentations précises avec un grand niveau de réalisme pour la protection, la conservation, de futures restaurations ou des études archéologiques et artistiques. Plusieurs méthodes d'acquisitions de mesures existent comme les méthodes de photogrammétrie terrestre ou aérienne, par photogrammétrie rapprochée (Moussa *et al.*, 2012), les scanners à laser terrestres (Haddad, 2011) ou en combinant ces méthodes (Al-kheder *et al.*, 2009) (Henry *et al.*, 2012).

Il paraît intéressant que le scanner à laser 3D et l'appareil photo numérique soient utilisés séparément. En effet, les mesures du scanner peuvent être combinées aux méthodes de photogrammétrie rapprochée ou à des images numériques pour améliorer la qualité et la résolution de la texture des nuages de points (par texture de nuage, on entend aspect lié aux couleurs de ses points) avec un photo-réalisme accru. Ces méthodes nécessitent de recalibrer les données collectées par les deux appareils (scanner à laser et appareil photo numérique) puisqu'ils n'ont pas de repère commun a priori. Ce recalibrage est une étape cruciale pour une bonne colorisation, c'est pourquoi il a lon-

guement été étudié et de nombreuses propositions ont été faites. Ces approches ont été catégorisées par Corsini (2012). La première catégorie contient les méthodes les plus simples, lorsque la pose des deux appareils est connue relativement l'une de l'autre. Cette méthode est utilisée dans (Abmayr *et al.*, 2004) où un nuage de points et des prises de vue à 360 degrés de haute qualité sont acquis à partir d'une même position géographique. Par conséquent, le passage entre les deux systèmes de coordonnées est facilement réalisé. Cependant, ces méthodes impliquent une préparation importante lors de l'acquisition des données, par exemple pour placer les appareils avec précision, et il est impossible d'utiliser des images numériques seules prises avec un APN à partir de positions quelconques.

Ce recalage peut être effectué en utilisant un appariement de primitives visuelles entre les deux jeux de données. Ces primitives visuelles peuvent être des points, des droites, des régions d'intérêt ou d'autres caractéristiques. Elles peuvent être sélectionnées manuellement ou détectées automatiquement. Dans (Adan *et al.*, 2012), des points sont choisis manuellement. Les nuages de points sont utilisés pour générer des ortho-images. Les ortho-images sont générées par projection parallèle ou orthographique des points du nuage, contrairement aux images « classiques » qui sont le fruit d'une projection perspective. Il en résulte une image à échelle constante et qui conserve les caractéristiques géométriques de la scène. Par conséquent, la sélection de points identiques dans l'image réelle et dans l'image virtuelle est facilitée. La matrice de transformation est calculée en utilisant ces points mis en correspondance afin de recaler l'image numérique acquise et le nuage de points 3D. Les points d'intérêt peuvent être détectés et appariés automatiquement en utilisant SIFT (Lowe, 2004), ASIFT (Morel, Yu, 2009) ou d'autres extracteurs et descripteurs de points d'intérêts. Dans l'article (Moussa *et al.*, 2012), les primitives sont détectées avec SIFT dans une image numérique acquise et extraites d'un modèle de l'environnement basé points. Ce modèle est constitué d'un nuage de points 3D dense de la scène et de l'intensité de ces points. Dans certains cas (Alshwabkeh, Haala, 2004), sélectionnées manuellement ou automatiquement, les droites peuvent être plus appropriées étant donné la géométrie rectiligne de la scène que l'on souhaite recaler.

Les scanners à laser attribuent à chaque point prélevé une valeur de réflectance. Cette réflectance est parfois utilisée pour rendre les données entre les deux appareils plus similaires (Mastin *et al.*, 2009) et ainsi faciliter la détection de caractéristiques visuelles. Le recalage peut être obtenu en comparant les formes du modèle 3D avec les contours de l'objet détectés dans l'image numérique acquise (Belkhouche *et al.*, 2012).

Une dernière famille de méthode exploite des primitives visuelles denses. Elles visent, par exemple, à maximiser l'information mutuelle entre une image numérique acquise et des caractéristiques géométriques du modèle 3D. Différentes caractéristiques comme les normales, la réflexion ou les occultations ambiantes sont comparées dans (Corsini *et al.*, 2009). D'autres travaux utilisent plusieurs images numériques acquises afin de générer un nuage de points 3D épars, par exemple en Structure from Motion (SfM). Ensuite, ce nuage de points est recalé avec le modèle 3D ciblé. Une

fois ce recalage 3D/3D effectué, les images numériques acquises utilisées pour la SfM peuvent être alignées avec le modèle 3D et utilisées pour la colorisation. Cependant, cette méthode nécessite d'acquérir un grand nombre d'images numériques.

Les dimensions considérables de la cathédrale d'Amiens nous orientent vers les méthodes automatiques d'extraction de primitives. Nos jeux de données sont difficiles à cause des problèmes de couleurs des points (figures 3 et 4), par conséquent ces méthodes ne nous permettent pas d'obtenir un recalage aussi précis que nous le souhaitons. C'est pourquoi nous proposons une approche qui utilise dans un premier temps une image virtuelle de la scène et une image numérique acquise par un APN dans lesquelles nous détectons des points épars (points ASFIT). Ces points sont ensuite mis en correspondance, nous avons donc des appariements 2D/2D. Connaissant les paramètres intrinsèques de la caméra virtuelle nous pouvons remonter aux coordonnées 3D des points ASIFT de l'image virtuelle. Nous obtenons ainsi des appariements 2D/3D. Ces appariements sont utilisés pour estimer une première pose de la caméra virtuelle qui sera ensuite corrigée par une optimisation de pose basée sur un critère photométrique dense. On utilise l'ensemble des pixels de l'image numérique et des images virtuelles obtenues à chaque itération de l'optimisation de la caméra. La primitive photométrique a été utilisée pour du suivi visuel de plans (Benhimane, Malis, 2004). Dans ces travaux, le suivi de plans est effectué selon un modèle de mouvement 2D : l'homographie. La variation des intensités de l'image est reliée à celle des paramètres de l'homographie pour l'estimation dans un processus d'optimisation. Dans notre cas, nous proposons une méthode qui utilise directement les intensités d'une image pour estimer les six degrés de liberté cartésiens (translations et rotations) de la pose de la caméra dans un espace virtuel 3D. Pour cela, en s'inspirant de (Benhimane, Malis, 2004), nous avons besoin de mettre en relation la variation d'intensités d'une image à la variation de pose de la caméra. Cette problématique fut dans un premier temps abordée pour contrôler les mouvements d'un système robotique en utilisant une caméra. C'est ce que l'on appelle l'asservissement visuel photométrique (Collewet, Marchand, 2011). Depuis plus de vingt ans, l'asservissement visuel décline le contrôle de robot à l'aide de primitives visuelles variées (points, droites, moments, etc) (Chaumette, Hutchinson, 2006). Il a été également étendu à l'asservissement visuel virtuel (AVV) (Marchand, Chaumette, 2002) dans le but de calculer la pose d'une caméra à partir de mesures réalisées dans les images, par une optimisation non linéaire, et en exploitant les connaissances de l'asservissement visuel. Notre approche étend l'asservissement visuel photométrique (Collewet, Marchand, 2011) à l'AVV comme (Caron *et al.*, 2014) dans un autre cadre d'application. Cet article traite de la localisation de caméra basée modèle. L'environnement est matérialisé virtuellement par des formes géométriques simples (parallélogrammes) sur lesquelles sont plaquées des textures (images numériques). L'estimation de la pose de la caméra est exprimée comme un problème d'optimisation dans lequel on cherche à maximiser l'information mutuelle partagée par une image numérique acquise par la caméra et l'environnement virtuel.

Ce papier est une version étendue de l'article publié et présenté lors de la conférence ORASIS'13 (Crombez *et al.*, 2013). L'estimation automatique d'une pose pour initialiser l'optimisation photométrique, la comparaison de fonctions de coût obtenues

en utilisant d'autres critères de corrélation et de nouveaux résultats ont été ajoutés. La suite présente l'approche que l'on propose, en commençant par exposer la méthode d'estimation de la pose initiale suivie par le procédé d'optimisation photométrique de pose de caméra virtuelle. Notre approche est ensuite employée afin de coloriser des nuages de points. Finalement, des résultats expérimentaux sont présentés avant de conclure.

### 3. Approche proposée

Soit  $P_{i,j}(\mathbf{X}, \mathbf{C})$  un point 3D coloré  $j$  d'un nuage de points  $i$ , composé de ses coordonnées  $\mathbf{X} = (X, Y, Z)$  et de sa couleur  $\mathbf{C} = (R, V, B)$ . Un nuage de points acquis par la station  $i$  est exprimé par  $PC_i = \{P_{i,j}(\mathbf{X}, \mathbf{C})/\forall j \in \llbracket 0, N_i - 1 \rrbracket\}$ ,  $N_i$  étant le nombre de points acquis par la station  $i$ . Nous notons  $PC_m = \{PC_i/\forall i \in \llbracket 0, N_j - 1 \rrbracket\}$  l'assemblage de  $N_j$  nuages de points obtenus par des stations distinctes (figure 2b) à différents moments de la journée. Une image virtuelle d'une scène contenant le nuage de points est exprimée par  $I_v(pr_\gamma(PC_m), {}^cM_o)$ , où  ${}^cM_o$  représente la pose de la caméra virtuelle dans la scène et  $pr_\gamma$  la projection perspective et le passage dans le repère pixelique de l'image à partir des paramètres intrinsèques  $\gamma = \{\alpha_u, \alpha_v, u_0, v_0\}$  de l'APN utilisé (dans la suite du papier, nous simplifions l'écriture en utilisant  $I_v$  ou  $I_v(PC_m, {}^cM_o)$ ). Un Z-buffer détermine quels points projetés  $P_{i,j}(\mathbf{X}, \mathbf{C})$  sont visibles et les pixels de  $I_v$  sont colorés avec la couleur  $\mathbf{C}$  provenant de ces  $P_{i,j}$ . Connaissant les paramètres intrinsèques, nous pouvons aisément remonter au coordonnées 3D des points composants une image virtuelle.

Les nuages de points  $PC_i$  sont relevés à l'aide du scanner à laser Leica Geosystems TPS400. Ce scanner possède une portée d'acquisition de 200 m et une précision de mesure de plus ou moins 1,5 mm. L'assemblage  $PC_m$  de ces nuages est altéré par un certain nombre de problèmes de couleur (figure 3) : différences de luminosité, ombres, etc. S'ajoute à ces différences de couleur entre les zones de chevauchement des  $PC_i$ , un effet de flou sur les textures des nuages de points. Ce problème de résolution est lié à la distance entre le bâtiment et l'appareil photo combiné au scanner à laser qui a été utilisé pour coloriser les points.

Pour remédier à ce problème, nous utilisons des images numériques  $I$  de haute qualité acquises avec l'APN que nous cherchons à recalculer sur le nuage de points  $PC_m$ . Ces images numériques sont obtenues avec un Olympus PEN E-P3 équipé d'une lentille numérique Olympus M.Zuiko, offrant des images de  $4032 \times 3024$  pixels avec une focale de 17 mm pour un capteur de  $14 \times 42$  millimètres. L'appareil est calibré, nous connaissons donc ses paramètres intrinsèques  $\gamma$ .

Le cœur de l'approche réside dans l'estimation précise de la pose de l'appareil photo à l'origine des images numériques  $I$  dans le repère de  $PC_m$ . Le but est donc de calculer la matrice de transformation  ${}^cM_o$ , à partir de  $I_v(PC_m, {}^cM_o)$ , en considérant l'image numérique  $I$  acquise par l'APN comme référence. La matrice  ${}^cM_o$  exprime le

changement de repère partant du modèle  $PCm$  (objet ( $o$ )) vers le repère de la caméra ( $c$ ) :

$${}^c\mathbf{M}_o = \begin{pmatrix} {}^c\mathbf{R}_{o(3 \times 3)} & {}^c\mathbf{t}_{o(1 \times 3)} \\ \mathbf{0}_{(3 \times 1)} & 1 \end{pmatrix}_{(4 \times 4)} \quad (1)$$

Cette matrice  ${}^c\mathbf{M}_{o(4 \times 4)}$  est composée d'une matrice de rotation  ${}^c\mathbf{R}_{o(3 \times 3)}$  et d'un vecteur de translation  ${}^c\mathbf{t}_{o(1 \times 3)}$ . Notre approche pour calculer cette pose se décline en deux étapes : le calcul d'une pose approximative  ${}^c_0\mathbf{M}_o$  suivi de son optimisation. La première repose sur un pré-traitement des couleurs des  $PC_i$ , nécessaire pour l'utilisation des primitives géométriques menant à la pose approximative, et la seconde réalise l'optimisation de la pose sur un critère rassemblant toutes les intensités des images  $I_v$  et  $I$ . Ces deux étapes cruciales sont décrites ci-dessous.

### 3.1. Estimation de pose initiale

Une première pose de caméra virtuelle  ${}^{c_{init}}\mathbf{M}_o$  est nécessaire pour initialiser l'optimisation photométrique de pose. Différents détecteurs de points d'intérêts (ASIFT (Morel, Yu, 2009), KAZE (Alcantarilla *et al.*, 2012), Brisk (Leutenegger *et al.*, 2011)) ont été testés dans le but d'effectuer une mise en correspondance éparsée entre les images numériques acquises  $I$  et les images de synthèse  $I_v$  de la scène afin d'élargir la zone de convergence de l'optimisation de pose présentée par la suite. Cependant, aucun des points d'intérêts détectés ne sont correctement mis en correspondance comme, par exemple, pour la paire d'images de la figure 4. Les différences de couleurs entre les  $PC_i$  créent des perturbations sur les images de synthèses  $I_v(PCm, {}^c\mathbf{M}_o)$  générées. Ces dégradations jouent un rôle important sur les descripteurs des points d'intérêts détectés et rendent la mise en correspondance impossible.

Pour pallier ce problème, nous nous sommes inspirés de (Tian *et al.*, 2002) afin d'homogénéiser la couleur des nuages de points  $PC_i$  entre eux. Dans ces travaux, la correction de couleurs est effectuée sur des images numériques acquises afin de créer des panoramas par mosaïquage. Les images numériques utilisées peuvent avoir été prises sous différents points de vue et donc d'éclairage. Les différences d'illumination entre les images numériques acquises par un APN sont considérées comme une transformation linéaire  $\mathbf{T}_{(3 \times 3)}$  des couleurs. La comparaison de différentes formes de  $\mathbf{T}_{(3 \times 3)}$  par (Tian *et al.*, 2002) a permis de mettre en évidence les meilleures performances du modèle linéaire compensé par une transformation affine. Cette matrice de transformation est calculée par :

$$\mathbf{T}_{(3 \times 3)} = [\mathbf{J}_1^T \mathbf{J}_1]^{-1} \mathbf{J}_1^T \mathbf{J}_2 \quad (2)$$

$\mathbf{J}_1(3 \times N)$  et  $\mathbf{J}_2(3 \times N)$  contiennent les  $N$  valeurs RVB des pixels d'une partie de la scène que les deux images ont en commun, une zone de chevauchement (Tian *et al.*, 2002). Ce chevauchement permet de considérer que les pixels de  $\mathbf{J}_1$  et  $\mathbf{J}_2$  sont en corrélation, autrement dit les pixels aux coordonnées  $(u, v) \in \mathbf{J}_1 \wedge \mathbf{J}_2$  des deux images sont censés avoir la même valeur RVB.  $\mathbf{J}_1$  et  $\mathbf{J}_2$  sont donc de taille  $(3 \times N)$ ,

$N$  étant le nombre de pixels communs d'une zone des images. La matrice  $\mathbf{T}$  ainsi obtenue est de taille  $(3 \times 3)$ , chaque colonne permettant de corriger les 3 canaux RVB de l'image  $\mathbf{J}_1$ . S'ajoute à cela la compensation affine calculée par :

$$\begin{pmatrix} R_{affine} \\ V_{affine} \\ B_{affine} \end{pmatrix} = \begin{pmatrix} moy(R2) \\ moy(V2) \\ moy(B2) \end{pmatrix} - \mathbf{T}_{(3 \times 3)} \times \begin{pmatrix} moy(R1) \\ moy(V1) \\ moy(B1) \end{pmatrix} \quad (3)$$

où  $moy(R1)$  et  $moy(R2)$  correspondent respectivement à la moyenne des valeurs du canal rouge de l'image 1 et de l'image 2. Il en est de même pour les canaux vert et bleu.

Nous avons étendu cette approche pour corriger les couleurs  $\mathbf{C}_{(i,j)}$  des points 3D composant nos  $PC_i$ .  $\mathbf{J}_1$  et  $\mathbf{J}_2$  ne contiennent plus les valeurs RVB de portions d'images mais les couleurs  $\mathbf{C}_{(i,j)}$  de points qui se chevauchent appartenant à deux nuages  $PC_i$ . Les équations présentées précédemment restent inchangées. Nous choisissons visuellement parmi les  $PC_i$  un nuage de points de référence que l'on appelle  $PC_{ref}$  et les couleurs des autres nuages sont, l'un après l'autre, automatiquement uniformisées. Comme pour les images numériques de (Tian *et al.*, 2002), la matrice de transformation linéaire  $\mathbf{T}$  est calculée à partir des couleurs des points qui se chevauchent entre  $PC_{ref}$  et le nuage ciblé  $PC_t$ . Pour remplir  $\mathbf{J}_1$  et  $\mathbf{J}_2$ , nous cherchons pour chaque point de  $PC_{ref}$  son point voisin le plus proche appartenant à  $PC_t$  dans une sphère de rayon  $\lambda$ . S'il existe un point dans cette sphère, alors les deux points sont considérés comme confondu, le point du nuage de référence est ajouté à  $\mathbf{J}_1$  et le point du nuage cible est ajouté à  $\mathbf{J}_2$ .

A l'exception du nuage  $PC_i$  choisi pour être le nuage de référence  $PC_{ref}$ , les couleurs des  $N_i - 1$  nuages de points composant  $PC_m$  sont uniformisées avec la méthode décrite ci-dessus. Les couleurs  $\mathbf{C}_u$  des points des nuages uniformisées  $PC_u$  sont donc calculées par :

$$\mathbf{C}_{u,i,j} = \mathbf{C}_{i,j} \times T + RVB_{affine} \quad (4)$$

Les  $N_i - 1$  nuages de points aux couleurs uniformisées  $PC_u$  et le nuage  $PC_{ref}$  sont fusionnés pour former le nuage  $PC_{um}$ . Les images virtuelles  $I_v(PC_{um},^c M_o)$  sont donc moins perturbées par les couleurs des nuages provenant de différentes stations et permettent ainsi la mise en correspondance de leurs points d'intérêts ASIFT avec ceux détectés dans les images numériques (figure 5).

Cette mise en correspondance de points d'intérêt donne lieu à des paires de points 2D provenant de l'image numérique acquise et des points 3D provenant de l'image virtuelle  $I_v$ . Il nous est donc possible d'effectuer un calcul de pose basé RANSAC utilisant les points comme primitives géométriques, c'est-à-dire de déterminer les 6 degrés de liberté permettant de passer de la pose de la caméra virtuelle à la pose de l'appareil à l'origine de l'image numérique. Ce calcul de pose est effectué à l'aide des algorithmes d'asservissement visuel virtuel (AVV) de la bibliothèque ViSP (Marchand *et al.*, 2005). L'AVV est un algorithme d'optimisation destiné à minimiser la distance



Figure 5. Mise en correspondance de points d'intérêt ASFIT entre une image numérique acquise et une image de synthèse de nuages de points aux couleurs uniformisées  $I_v(PC_{um},^c M_o)$ . Sans cette correction des couleurs, aucun point n'est mis en correspondance

entre la projection des points 3D détectés dans  $I_v$  par rapport à la position de leurs appariements 2D de l'image numérique acquise. Nous obtenons ainsi une nouvelle pose de caméra virtuelle pouvant servir d'initialisation au procédé d'optimisation dense de pose.

### 3.2. étude de fonctions de coût photométriques

Cette section vise à déterminer la fonction de coût la mieux adaptée à notre cas d'étude pour un recalage visuel dense. Le nuage aux couleurs uniformisées  $PC_{um}$  n'est utilisé que pour permettre la mise en correspondance de points d'intérêts (section 3.1). Nous considérons le nuage de points fusionné aux couleurs d'origines  $PC_m$ , c'est à dire les données de position  $\mathbf{X}$  et de couleur  $\mathbf{C}$  des points retournées par le scanner comme vérité terrain. Nous ne souhaitons pas que les informations RVB uniformisées de  $PC_{um}$  influent négativement sur l'optimisation de la pose de notre caméra virtuelle. Nous cherchons une mesure de similarité robuste aux problèmes de couleur des points du nuage fusionné  $PC_m$  (figure 3). C'est pourquoi, nous étudions différentes mesures de similarité : la somme des différences au carré (ZNSSD), la corrélation croisée (ZNCC) sur les images  $I_v$  et  $I$  en niveaux de gris centrées et normalisées. Nous comparons également ces résultats avec l'information mutuelle (MI) entre les images numériques acquises et des cartes de normales et de réflexions. Ces cartes de normales et de réflexions ont été utilisées dans (Corsini *et al.*, 2009). Nous utilisons les mêmes méthodes pour générer les rendus et pour calculer l'information

mutuelle. Nous avons calculé ces différentes mesures en partant d'une pose de caméra virtuelle proche de la pose d'une image numérique et en faisant varier deux degrés de liberté de la caméra virtuelle. La caméra se déplace le long de ses axes  $\vec{X}$  et  $\vec{Y}$ . Pour chaque déplacement nous calculons les différentes mesures de similarité. Nous utilisons plusieurs images numériques de haute qualité et une caméra virtuelle possédant les mêmes paramètres intrinsèques que l'APN à l'origine des images numériques. Cette étude nous permet de comparer les formes des fonctions de coût ainsi obtenues selon ces quatre critères de corrélation (figure 6). Afin de faciliter les comparaisons, nous traçons l'opposé de la ZNCC et de l'information mutuelle.

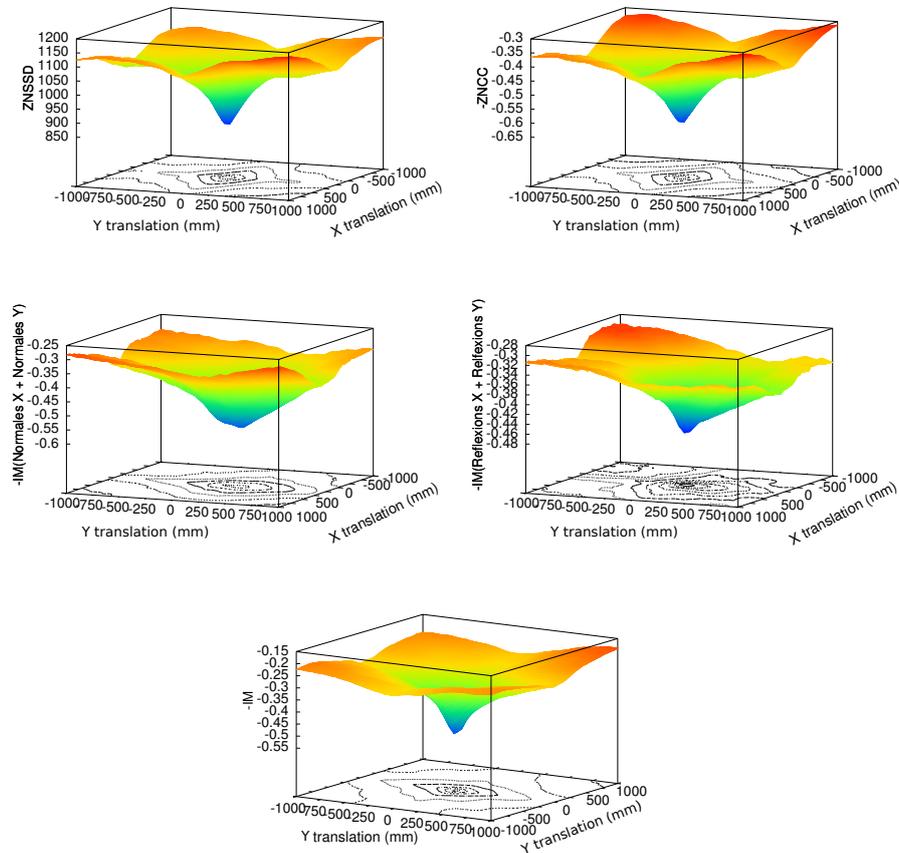


Figure 6. Exemples de fonctions de coût pour 5 critères de corrélation

Nous pouvons observer que la fonction de coût de la ZNSSD, la ZNCC et l'information mutuelle basée sur les intensités des images ont une forme plus convexe que les autres et un minimum plus prononcé. Les fonctions de coût de l'information mutuelle basé sur les normales et sur la réflexion sont plus bruitées que les deux autres.

Ces observations indiquent qu'il est préférable d'utiliser la ZNSSD ou la ZNCC ou l'information mutuelle basée sur les intensités afin de faire converger la pose d'une caméra virtuelle vers une pose réelle plus facilement et plus rapidement. Enfin, la complexité de calcul réduite de la ZNSSD par rapport à la ZNCC ou l'information mutuelle, que ce soit pour le critère à minimiser/maximiser ou pour la matrice interaction, confirme la pertinence du choix de la ZNSSD. Le fait de centrer et normaliser les intensités de l'image numérique  $I$  et de  $I_v(PCm, {}^cM_o)$  permet de faire face aux problèmes de couleur du nuage fusionné  $PCm$ . L'erreur à minimiser étant calculée sur ces intensités, la forme de la fonction de coût est très lisse, convexe et présente un minimum clairement identifiable.

### 3.3. Optimisation de pose sur critère photométrique

La mise en correspondance présentée en 3.1 nous permet d'obtenir une pose de caméra virtuelle approximative. Nous exprimons la correction de ces approximations comme un problème d'optimisation dans le cadre d'un asservissement visuel virtuel. Cependant, contrairement à l'asservissement visuel photométrique (Collewet, Marchand, 2011), notre méthode ne cherche pas à minimiser la différence entre deux images réelles provenant d'un même capteur de vision. Ici, on minimise la différence entre une image numérique  $I$  et l'image virtuelle  $I_v$  d'une scène 3D acquise par une caméra virtuelle ayant les mêmes paramètres intrinsèques que l'appareil à l'origine de  $I$ .  $I$  et  $I_v$  sont utilisées en niveaux de gris centrés et normalisés. Le critère que l'algorithme doit minimiser est :

$$e_r = \|\mathbf{I}_v(PCm, {}^cM_o) - \mathbf{I}^*\| \quad (5)$$

où  $\mathbf{r} = (t_x, t_y, t_z, \theta_{u_x}, \theta_{u_y}, \theta_{u_z})$  est la pose actuelle de la caméra virtuelle (3 éléments de translation  ${}^c\mathbf{t}_o$  et 3 de rotations extraits de  ${}^c\mathbf{R}_o$  à l'aide des formules de Rodrigues, avec les rotations représentées par un angle et un vecteur unitaire),  $\mathbf{I}_v(PCm, {}^cM_o)$  est le vecteur contenant les intensités de  $I_v(PCm, {}^cM_o)$  et  $\mathbf{I}^*$  est le vecteur des intensités désirées provenant de l'image numérique acquise avec l'APN.

La loi de commande du mouvement de la caméra virtuelle est obtenue grâce à l'algorithme de Levenberg-Marquardt qui estime l'incrément de pose suivant :

$$\dot{\mathbf{r}} = -\lambda(\mathbf{H} + \mu \text{diag}(\mathbf{H}))^{-1} \mathbf{L}_{\mathbf{I}_v}^T (\mathbf{I}_v - \mathbf{I}^*) \quad (6)$$

avec  $\mathbf{H} = \mathbf{L}_{\mathbf{I}_v}^T \mathbf{L}_{\mathbf{I}_v}$ , avec  $\mathbf{L}_{\mathbf{I}_v}$ , la matrice d'interaction liée à la luminance de l'image  $I_v$ . Cette matrice d'interaction, également appelée Jacobienne de l'image, relie les variations de l'image virtuelle aux variations de la pose de la caméra (voir (Collewet, Marchand, 2011) pour plus de détails). Ensuite, la pose  ${}^cM_o$  est itérativement mise à jour grâce à l'incrément de la pose  $\dot{\mathbf{r}}$  en utilisant l'application exponentielle  $e$  du groupe spécial euclidien  $se(3)$ .

$${}^{c[t+1]}M_o = (e^{\dot{\mathbf{r}}})^{-1} \times {}^{c[t]}M_o = ({}^{c[t]}M_{c[t+1]})^{-1} \times {}^{c[t]}M_o = {}^{c[t+1]}M_{c[t]} \times {}^{c[t]}M_o \quad (7)$$

La mise à jour de la pose s'arrête lorsque l'erreur résiduelle devient constante à un epsilon près. Nous considérons alors avoir atteint la pose de l'APN à l'origine de  $I$ , nous appelons cette pose  ${}^{c_{opt}}\mathbf{M}_o$ . L'optimisation est rendue encore plus robuste à l'aide d'un M-estimateur (Comport *et al.*, 2003) qui est basé sur des statistiques robustes (Huber, 1981).

### 3.4. Résultats expérimentaux

#### 3.4.1. Expérimentation de l'approche

Nous avons expérimenté notre approche sur le portail sud de la cathédrale. Nous cherchons à recaler avec précision le nuage de points  $PC_m$  avec l'image numérique  $I$  de la figure 7a. Une image de synthèse  $I_v$  du portail aux couleurs uniformisées  $PC_{um}$  est générée de manière à obtenir un maximum d'informations visuelles, c'est-à-dire face au portail et à une distance permettant de le couvrir dans sa totalité (figure 7b). La mise en correspondance de descripteurs ASIFT entre ces deux images, puis l'asservissement visuel virtuel à partir de ces points d'intérêts permettent de calculer une nouvelle pose de la caméra virtuelle  ${}^{c_1}\mathbf{M}_o$ . La figure 7c montre l'image  $I_v({}^{c_1}\mathbf{M}_o)$  générée à partir de cette nouvelle pose. Expérimentalement, nous nous sommes aperçus qu'il peut être intéressant d'effectuer une seconde fois la mise en correspondance et le calcul de pose épars entre  $I$  (figure 7a)  $I_v({}^{c_1}\mathbf{M}_o)$  (figure 7c). En effet, la pose de la caméra  ${}^{c_1}\mathbf{M}_o$  étant plus proche de la pose de l'appareil photo, l'image virtuelle  $I_v({}^{c_1}\mathbf{M}_o)$  est plus similaire à image numérique  $I$ . Par conséquent, plus de points d'intérêts sont détectés et mis en correspondance et le calcul d'une nouvelle pose de caméra  ${}^{c_2}\mathbf{M}_o$  sera moins approximatif que le précédent, et donc encore plus proche de la pose que l'on souhaite atteindre (figure 7d).

La pose  ${}^{c_2}\mathbf{M}_o$  précédemment calculée est utilisée en tant que  ${}^{c_{init}}\mathbf{M}_o$  afin d'initialiser l'optimisation photométrique de pose pour corriger les décalages et obtenir la pose précise à laquelle l'image numérique  $I$  a été acquise. Par exemple, pour l'image numérique de la première ligne de la figure 7, l'optimisation photométrique a positionné la caméra virtuelle en calculant itérativement ses 6 degrés de liberté (figure 8) jusqu'à atteindre un état stable après 17 itérations. La figure 9 montre la différence entre l'image numérique à atteindre et l'image virtuelle à la première et à la dernière itération de l'optimisation de pose  $I_v(PC_m, {}^{c_{init}}\mathbf{M}_o)$  et  $I_v(PC_m, {}^{c_{opt}}\mathbf{M}_o)$ . La distance entre la pose de la caméra virtuelle initiale  ${}^{c_{init}}\mathbf{M}_o$  et la pose optimale  ${}^{c_{opt}}\mathbf{M}_o$  est de 674 millimètres.

#### 3.4.2. Application à la colorisation

Après l'estimation de la pose optimale  ${}^{c_{opt}}\mathbf{M}_o$  permettant de recaler le nuage de points 3D sur l'image 2D  $I$  de haute qualité, nous sommes en mesure de procéder à la colorisation du nuage de points  $PC_m$ .

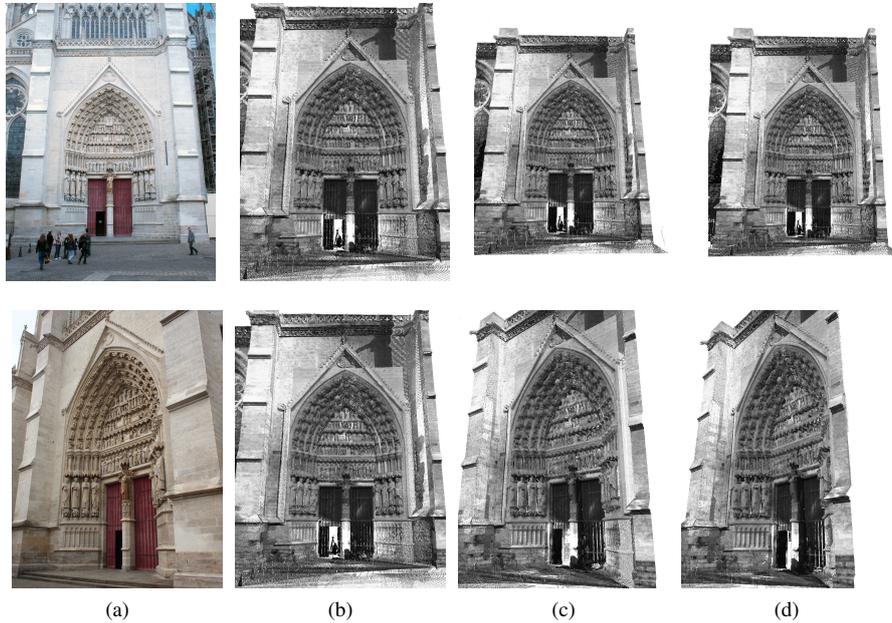


Figure 7. (a) Image numérique  $I$  de haute qualité du portail sud de la cathédrale. (b) Image de synthèse  $I_v(PCum,^{c_0} M_o)$  apportant un maximum d'informations visuelles du portail. (c) Image de synthèse  $I_v(PCum,^{c_1} M_o)$  générée à partir d'une première pose estimée par mise en correspondance éparsse et asservissement visuel virtuel entre (a) et (b). (d) Image de synthèse  $I_v(PCum,^{c_2} M_o)$  générée à partir d'une seconde pose estimée par mise en correspondance éparsse et asservissement visuel virtuel entre (a) et (c)

#### 3.4.2.1. Méthodologie

La pose  $^{c_{opt}} M_o$  estimée permet de transformer et projeter chaque point  $P_{i,j}(X, C)$  du nuage  $PCm$  dans le repère de l'image numérique  $I$ . Nous utilisons une interpolation bilinéaire pour définir la couleur  $C$  de chaque point projeté. Cette valeur est calculée à partir de la moyenne pondérée des valeurs RVB des quatre pixels voisins dans l'image numérique  $I$ .

Dans la plupart des cas, la colorisation ou la texturisation est réalisée lorsque le nuage de points est maillé, sur le modèle 3D représentant les surfaces du nuage. Notre approche consiste à coloriser directement les nuages de points originaux  $PCm$ . De ce fait, nous ne pouvons nous attendre à de bons résultats de colorisation sans prendre en considération la gestion des occultations du nuage. En effet, certains points provenant de certaines parties de la scène qui sont invisibles d'une position de caméra peuvent être considérés comme visibles et colorisés. Ce phénomène peut se produire dans deux situations : lorsque deux (ou plusieurs) points de différentes profondeurs sont sur la

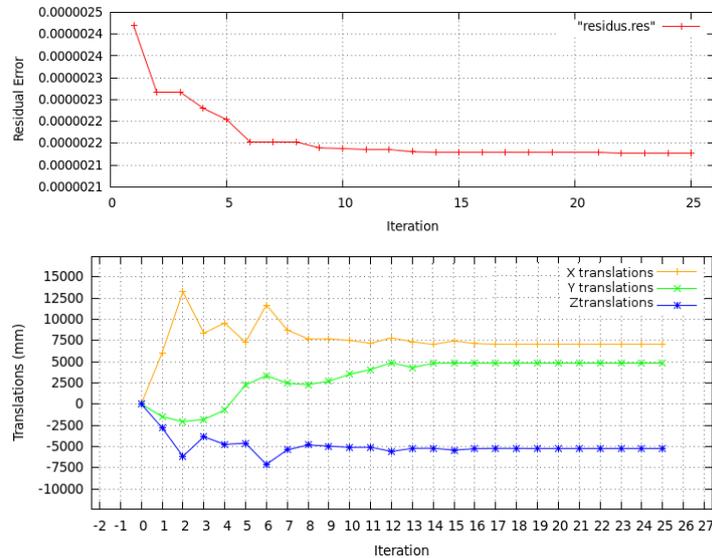


Figure 8. L'évolution de l'erreur résiduelle pondérée moyenne et des trois translations de la caméra virtuelle en fonction des itérations de l'optimisation de pose

même ligne de vue et sont donc projetés dans un même pixel ou lorsque tous les points, projetés dans un pixel, proviennent de parties invisibles de la scène. Dans les deux cas, des points sont colorisés avec une couleur RVB qui, en réalité, n'est pas censée leur être attribuée.

Pour résoudre ces problèmes, nous utilisons un maillage 3D du nuage de points dont les surfaces sont estimées avec l'algorithme de reconstruction de surface de Poisson (Kazhdan *et al.*, 2006). Cet algorithme possède la particularité de remplir les trous lorsque certaines mesures sont manquantes dans le nuage de points. Le modèle surfacique ainsi obtenu permet de générer une carte de profondeurs qui, à l'inverse du nuage de points, contient les différentes occultations de la scène. Ainsi, nous pouvons comparer la profondeur des points 3D du nuage à coloriser et celle du maillage pour ne coloriser que les points visibles du nuage à partir de la pose estimée.

Les résultats présentés ensuite ont été obtenus sur une machine possédant huit processeurs cadencés à une fréquence de 3.40 GHz, 16 Go de mémoire vive et une carte graphique GeForce 8400 GS. Pour donner un ordre d'idée des temps de calcul nécessaires de l'approche proposée, nous allons considérer un nuage de plus de 54 millions de points (taille du nuage du portail sud de la cathédrale) que l'on souhaite coloriser avec une image numérique de  $3042 \times 4032$  pixels. L'estimation de la pose initiale nécessite environ sept secondes et est répétée trois fois de suite. Chaque itération du Levenberg-Marquardt de l'optimisation photométrique de la pose nécessite environ 20 secondes. En moyenne, la pose de la caméra virtuelle converge vers la pose

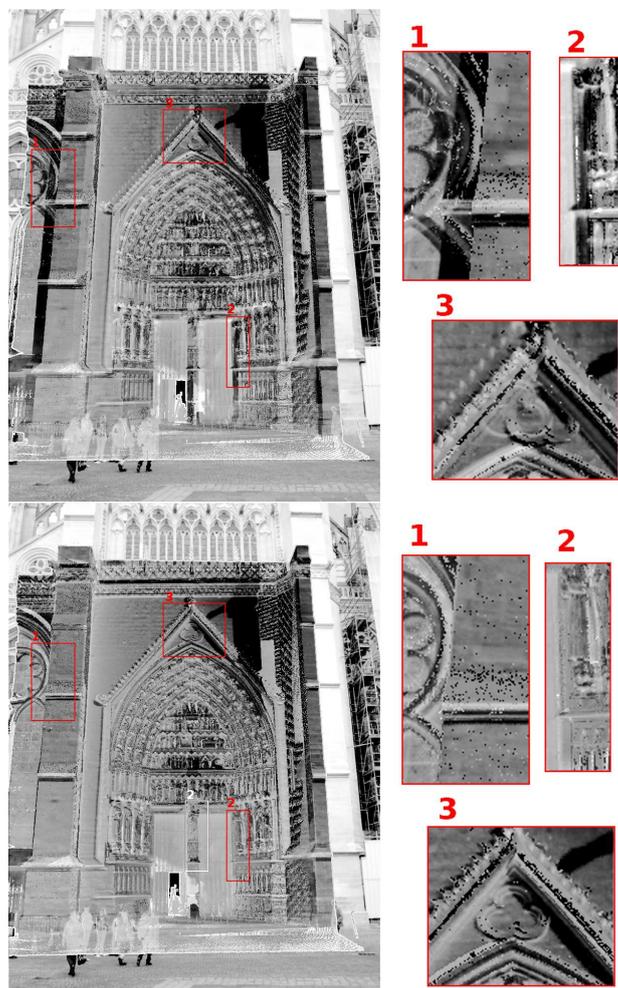


Figure 9. Les images de différence entre l'image numérique et les images virtuelles avant (en haut) et après (en bas) l'estimation de la pose

de la caméra réelle en 15 itérations du Levenberg-Marquardt. Enfin, pour l'étape de colorisation du nuage, la projection des 54 millions de points du nuage dans l'image numérique, la vérification de la visibilité des points à l'aide du modèle surfacique puis l'assignation des nouvelles couleurs aux points nécessite 40 secondes. Au total, la colorisation d'un nuage de 54 millions de points avec une image numérique de  $3042 \times 4032$  prend environ 6 minutes.

### 3.4.2.2. Résultats de colorisations

Le nuage de points du portail sud de la cathédrale d'Amiens (figure 3)  $PC_{cm}$  en sortie de notre procédé semble parfaitement recoloré (figure 10). Les couleurs sont

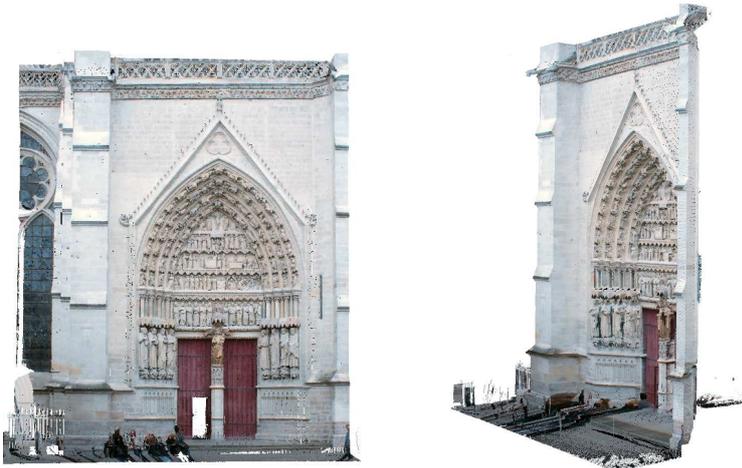


Figure 10. Le nuage de points  $PC_{cm}$  du portail sud après sa colorisation en utilisant une image numérique acquise avec un APN

beaucoup plus uniformes mais si on regarde de plus près (figure 11), les points colorés en rouge montrent que nous avons évidemment besoin de plusieurs images numériques pour couvrir entièrement le portail. Nous pouvons également remarquer le manque de précision de la texture de la statue de la Vierge Marie du fait d'un point de vue mal choisi pour prendre l'image numérique  $I$ . En effet, pour améliorer la visibilité des détails, nous devons prendre des images numériques à partir de positions plus proches de la statue.

Nous avons utilisé notre méthode complète pour coloriser entièrement une plus petite partie du portail sud : la statue de la Vierge Marie. Pour parvenir à un résultat optimal, l'idéal serait que chaque point 3D  $P_{i,j}$  visible à partir de la pose de la caméra soit projeté dans un pixel différent de l'image numérique acquise. Connaissant les caractéristiques intrinsèques de l'appareil photo et la résolution du nuage de points, nous avons calculé la distance nécessaire entre la statue et l'appareil photo pour s'assurer qu'un pixel de l'image numérique ne reçoive qu'un unique point 3D projeté. Quatre mètres sont nécessaires entre la statue et notre appareil photo en considérant une scène plane fronto-parallèle à l'APN.

Nous avons donc pris trois images numériques  $I_1, I_2, I_3$  autour de la statue à environ quatre mètres de distance pour une couverture maximale (figure 12). Les trois poses de l'appareil photo à l'origine de ces images numériques ont été calculées avec notre approche et nous avons utilisé le processus complet de colorisation décrit ci-dessus. Pour le moment, les points sont projetés dans une première image numérique  $I_1$ , les points couverts par cette image sont colorisés, les autres sont marqués comme

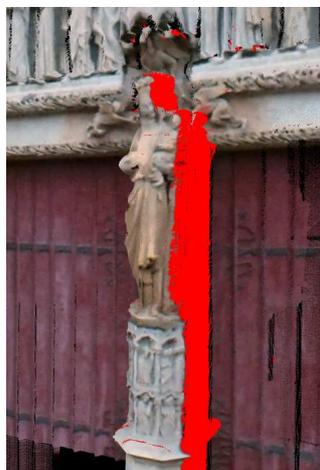


Figure 11. Zoom sur la statue de la Vierge du portail sud recoloré. Les points colorés en rouge sont invisibles à partir de la pose à l'origine de l'image numérique



Figure 12. Les images numériques de la Vierge Marie utilisées pour la colorisation de la statue

non-colorés. Les points sont ensuite projetés dans une seconde image numérique  $I_2$  et les points marqués comme étant non-colorés et couverts par cette image sont cette fois colorisés. L'opération est répétée sur l'ensemble des images numériques ou jusqu'à ce que la totalité des points du nuage soit colorée. Une comparaison entre le nuage de points de la Vierge Marie avant ( $PC_m$ ) et après ( $PC_{cm}$ ) notre colorisation est visible en figure 13.

Nous avons pris une autre image numérique  $I_4$  de la statue à partir d'un angle de vue différent pour une validation visuelle. Cette dernière image n'a pas été utilisée pour la colorisation du nuage de points. La figure 14 montre  $I_v(PC_m, {}^cM_o)$  et



Figure 13. Le nuage de points de la statue de la Vierge Marie avant (au-dessus) et après (en-dessous) la colorisation

$I_v(PCc_m, {}^c M_o)$  : le nuage de la Vierge Marie depuis la même pose de caméra que la nouvelle image numérique acquise  $I_4$ , avant et après le processus de colorisation. Cette comparaison met en évidence les améliorations indéniables apportées par la colorisation. Les deux images centrées sur le visage de la Vierge (figure 14) montrent que la résolution de la texture du nuage de points est très précise après sa colorisation. L'image numérique  $I_4$  et  $I_1, I_2, I_3$  utilisées pour la colorisation ont été prises avec cinq mois d'intervalle (différentes saisons). C'est ce qui explique les différences de colorimétrie entre la nouvelle  $I_4$  et le nuage de points recoloré avec  $I_1, I_2$  et  $I_3$ .



Figure 14. De gauche à droite : vue en contre-plongée de la statue de la Vierge acquise avec l'APN (a), en image virtuelle avant (b) et après (c) la colorisation

Le scanner à laser utilisé possède son propre visualiseur de nuage de points appelé Cyclone. Afin d'améliorer l'affichage des nuages de points, Cyclone utilise des traitements sur les couleurs pour homogénéiser l'aspect des nuages de points (figure 15). Cependant, même si les détails sont légèrement accentués, ces traitements ne permettent pas d'obtenir une meilleure résolution de texture des nuages de points.

Pour évaluer quantitativement la colorisation du nuage de points de la Vierge Marie, nous calculons la somme des différences au carré (SSD) ainsi que la corrélation croisée normalisée centrée (ZNCC) entre l'image numérique  $I_4$  et les deux images de synthèse  $I_v(PCm,^c M_o)$  et  $I_v(PCcm,^c M_o)$  (figure 14).

La SSD diminue d'environ 22 %. Les résultats des ZNCCs nous montrent une corrélation améliorée de 20 %. La ZNCC entre l'image numérique acquise et les images de synthèse avant et après la colorisation sont respectivement 0,53 et 0,74. Ces résultats démontrent explicitement que la colorisation a permis de rendre l'aspect du nuage de points de la Vierge Marie plus proche de la réalité (terrain) en termes de couleurs et de résolution. Nous avons également utilisé notre méthode sur un nuage de points de la chapelle Saint Sébastien. La figure 16 montre une comparaison entre le nuage de points aux couleurs prélevées par le scanner (partie basse) et après avoir été colorisé avec notre approche (partie haute).



Figure 15. Image du nuage de points de la Vierge Marie générée par Cyclone (Leica)

#### 4. Conclusion et perspectives

Nous avons proposé une méthode alternative pour le recalage automatique entre nuages de points et images numériques acquises avec un appareil photo numérique. Les méthodes manuelles existantes, utilisées pour sélectionner et pour mettre en correspondance des primitives géométriques de ces deux ensembles de données, peuvent amener à de bons résultats. Cependant, ces méthodes peuvent s'avérer très fastidieuses lorsque l'on doit gérer une quantité importante de données. Les méthodes automatiques de primitives géométriques quant à elles, peuvent s'avérer être délicates car elles imposent d'avoir des similitudes certaines entre les images numériques et les images virtuelles générées à partir des nuages 3D.

Dans notre approche, après avoir utilisé des primitives géométriques détectées dans les images, nous utilisons les intensités de l'image complète afin de rendre le recalage plus robuste et précis. Il serait intéressant, par la suite, d'utiliser les informations RVB de couleur de l'image plutôt que la luminance uniquement. Nous souhaitons également expérimenter notre approche en utilisant ou en combinant des caractéristiques telles que les normales, la réflexion ou la saillance en chaque point du nuage.

L'optimisation visuelle virtuelle d'une position de caméra 3D permet d'obtenir un recalage précis automatiquement. Les résultats présentés montrent que l'approche



*Figure 16. Nuage de points de la Chapelle Saint Sebastien. Les points de la partie basse ont les couleurs prélevées par le scanner, ceux de la partie haute ont été colorisés via notre approche*

proposée permet d'obtenir des nuages de points avec une texture de qualité réaliste en termes de couleur et de résolution.

### **Bibliographie**

- Abmayr T., Härtl F., Mettenleiter M., Heinz A., Neumann B., Fröhlich C. (2004). Realistic 3d reconstruction - combining laserscan data with rgb color information. *XXth ISPRS Congress: Proceedings of Commission V*.
- Adan A., Merchan P., Salamanca S. (2012). Creating Realistic 3D Models From Scanners by Decoupling Geometry and Texture. *International Conference on Pattern Recognition (ICPR)*, p. 457–460.

- Alcantarilla P. F., Bartoli A., Davison A. J. (2012). Kaze features. In *Proceedings of the 12th european conference on computer vision - volume part vi*, p. 214–227. Berlin, Heidelberg, Springer-Verlag.
- Al-kheder S., Al-shawabkeh Y., Haala N. (2009, février). Developing a documentation system for desert palaces in Jordan using 3D laser scanning and digital photogrammetry. *Journal of Archaeological Science*, vol. 36, n° 2, p. 537–546.
- Alshawabkeh Y., Haala N. (2004). Integration of digital photogrammetry and laser scanning for heritage documentation. *IAPRS*, p. 12–23.
- Altuntas C., Yildiz F., Karabork H., Yakar M., Karasaka L. (2007). Surveying and documentation of detailed historical heritage by laser scanning. *XXI International CIPA Symposium*, p. 01-06.
- Belkhouche Y., Buckles B., Duraisamy P., Namuduri K. (2012). Registration of 3d-lidar data with visual imagery using shape matching. *Int. Conf. on Image Processing, Computer Vision, and Pattern Recognition*, p. 749–754.
- Benhimane S., Malis E. (2004). Real-time image-based tracking of planes using efficient second-order minimization. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, p. 943–948.
- Caron G., Dame A., Marchand E. (2014). Direct model based visual tracking and pose estimation using mutual information. *Image and Vision Computing*, vol. 32, n° 1, p. 54–63.
- Chaumette F., Hutchinson S. (2006). Visual servo control part 1: Basic approaches. *IEEE Robotics and Automation Mag*, vol. 13, p. 82–90.
- Collewet C., Marchand E. (2011, août). Photometric Visual Servoing. *IEEE Transactions on Robotics*, vol. 27, n° 4, p. 828–834.
- Comport A., Marchand E., Chaumette F. (2003). A real-time tracker for markerless augmented reality. In *ACM/IEEE Int. Symp. on Mixed and Augmented Reality, ISMAR*, p. 36-45. Tokyo, Japon.
- Corsini M., Dellepiane M., Ponchio F., Scopigno R. (2009). Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties. *Computer Graphics Forum*, vol. 28, n° 7, p. 1755-1764.
- Crombez N., Caron G., Mouaddib E. (2013). Colorisation photo-réaliste de nuages de points 3D. *ORASIS*.
- Haddad N. A. (2011, juin). From ground surveying to 3D laser scanner: A review of techniques used for spatial documentation of historic sites. *Journal of King Saud University - Engineering Sciences*, vol. 23, n° 2, p. 109–118.
- Henry P., Krainin M., Herbst E., Ren X. (2012). Rgb-d mapping : Using kinect-style depth cameras for dense 3d modeling of indoor environments. *The International Journal of Robotics Research*, vol. 31, n° 5, p. 647–663.
- Huber P. J. (1981). *Robust statistics*. John Wiley and Sons.
- Ikeuchi K., Oishi T., Takamatsu. (2007, octobre). The great buddha project: Digitally archiving, restoring, and analyzing cultural heritage objects. *Int. J. Comput. Vision*, vol. 75, n° 1, p. 189–208.

- Kazhdan M., Bolitho M., Hoppe H. (2006). Poisson surface reconstruction. In *Eurographics symposium on geometry processing*, p. 61–70. Aire-la-Ville, Suisse.
- Leutenegger S., Chli M., Siegwart R. Y. (2011). Brisk: Binary robust invariant scalable keypoints. *Computer Vision, IEEE International Conference on*, vol. 0, p. 2548-2555.
- Lowe D. G. (2004, novembre). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, vol. 60, n° 2, p. 91–110.
- Marchand E., Chaumette F. (2002). Virtual visual servoing: A framework for real-time augmented reality. In *Eurographics Conference Proceeding*, vol. 21(3), p. 289-298.
- Marchand E., Spindler F., Chaumette F. (2005, December). Visp for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics and Automation Magazine*, vol. 12, n° 4, p. 40–52.
- Mastin A., Kepner J., Fisher J. (2009). Automatic registration of lidar and optical images of urban scenes. *IEEE Conference on Computer Vision and Pattern Recognition*, p. 2639-2646.
- Morel J.-M., Yu G. (2009, avril). Asift: A new framework for fully affine invariant image comparison. *SIAM J. Img. Sci.*, vol. 2, n° 2, p. 438–469.
- Moussa W., Abdel-Wahab M., Fritsch D. (2012). An automatic procedure for combining digital images and laser scanner data. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, p. 229–234.
- Tian G. Y., Gledhill D., Taylor D., Clarke D. (2002). Colour correction for panoramic imaging. *Proc. 6th Int. Conf. Inf. Vis.*, p. 483–488.

Article reçu le 27/09/2013

Accepté le 14/04/2014