



HAL
open science

Vers un accès plus inclusif à la Tapisserie de Bayeux par la génération semi-automatique d'objets 3D

Marjorie Redon, Matthieu Pizenberg, Yvain Quéau, Abderrahim Elmoataz

► To cite this version:

Marjorie Redon, Matthieu Pizenberg, Yvain Quéau, Abderrahim Elmoataz. Vers un accès plus inclusif à la Tapisserie de Bayeux par la génération semi-automatique d'objets 3D. ORASIS, 19ièmes journées francophones des jeunes chercheurs en vision par ordinateur, May 2023, Carqueiranne, France. hal-04075676

HAL Id: hal-04075676

<https://hal.science/hal-04075676>

Submitted on 20 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers un accès plus inclusif à la Tapisserie de Bayeux par la génération semi-automatique d’objets 3D

Towards a more inclusive access to the Bayeux Tapestry by semi-automated generation of 3D objects

Marjorie REDON

Matthieu PIZENBERG

Yvain QUÉAU

Abderrahim ELMOATAZ

Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, Caen, France

{Marjorie.Redon, Matthieu.Pizenberg, Yvain.Queau, Abderrahim.Elmoataz-billah}@unicaen.fr

Résumé

Les personnes aveugles ou partiellement aveugles n’ont souvent pas accès aux œuvres muséales qui ne peuvent être touchées et qui ne sont pas accessibles en audio-description. D’un autre côté, la création manuelle de représentations tactiles de ces œuvres s’avère généralement fastidieuse et onéreuse. Dans ce travail, nous proposons une méthodologie de création semi-automatique d’objets 3D afin de faciliter l’accessibilité de la Tapisserie de Bayeux, dans une démarche inclusive. Pour cela, nous utilisons une méthode de détection et segmentation automatique, couplée à une méthode de reconstruction 3D du micro-relief de la Tapisserie – qui est en réalité une broderie. Le dispositif proposé permet la création peu coûteuse et rapide de bas-reliefs imprimés en 3D, qui peuvent être explorés tactilement par les personnes aveugles ou partiellement aveugles.

Mots Clef

Accessibilité, Musée, Personnes présentant une incapacité visuelle, Apprentissage profond, Impression 3D

Abstract

Blind or partially blind people often do not have access to museum artworks which cannot be touched and which are not available in audio description. On the other hand, the manual creation of tactile representations of these artworks is usually tedious and expensive. In this work, we propose a methodology for semi-automatic creation of 3D objects in order to facilitate the inclusive accessibility of the Bayeux Tapestry. For this purpose, we use an automatic segmentation method, coupled with a 3D-reconstruction method of the micro-relief of the Tapestry – which is actually an embroidery. The proposed device allows the inexpensive and rapid creation of 3D-printed bas-reliefs, which can be explored tactilely by blind or partially blind people.

Keywords

Accessibility, Museum, Visually impaired people, Deep learning, 3D printing

1 Introduction

En 2020, M. J. Burton et al. [2] estimaient à 1.1 milliard le nombre de personnes présentant une incapacité visuelle (PPIV). Pour ces individus, la perception des œuvres visuelles (tableaux, sculptures ne pouvant être touchées, etc.) reste difficile. Afin de rendre l’accès aux œuvres visuelles plus inclusif, certains musées les font découvrir par le biais d’un autre sens – l’ouïe ou le toucher. Ainsi, la petite galerie du Louvre à Paris présente des audio-descriptions des œuvres. L’exposition “Prière de toucher, l’art et la matière”, mobile en Métropole Française jusqu’en 2024, invite à toucher des reproductions de sculptures. L’exposition “Touching the Prado” à Madrid propose de découvrir des tableaux à travers des plaques en trois dimensions accompagnées de descriptions en braille. Notre travail s’inscrit dans le cadre de ce troisième type d’exposition inclusive, fondé sur l’exploration tactile d’œuvres exclusivement visuelles imprimées en trois dimensions.

Parmi les différents dispositifs en relief mis à disposition des PPIVs, les représentations en bas-relief permettent le mieux de retranscrire tactilement les informations des images [4]. La fabrication d’une telle représentation tactile présente toutefois certaines limitations : son coût et temps de réalisation bien entendu, mais également la retranscription, forcément subjective, de l’œuvre originale par l’artiste. En outre, elle n’est possible que pour des œuvres de “petite” taille et dont la complexité en terme d’éléments est faible. Dans notre cas, nous nous intéressons au contraire à une œuvre de grande dimension : la Tapisserie de Bayeux. Il s’agit en réalité d’une broderie, mesurant 70 mètres de long sur 50 centimètres de haut. Elle conte la conquête de l’Angleterre par Guillaume le Conquérant à travers 58 scènes se succédant telle une “bande dessinée linéaire”. Au vu de la taille exceptionnelle de ce chef d’œuvre datant du 11ème siècle, il serait particulièrement intéressant d’avoir à disposition un système de génération semi-automatisé de représentations tactiles.

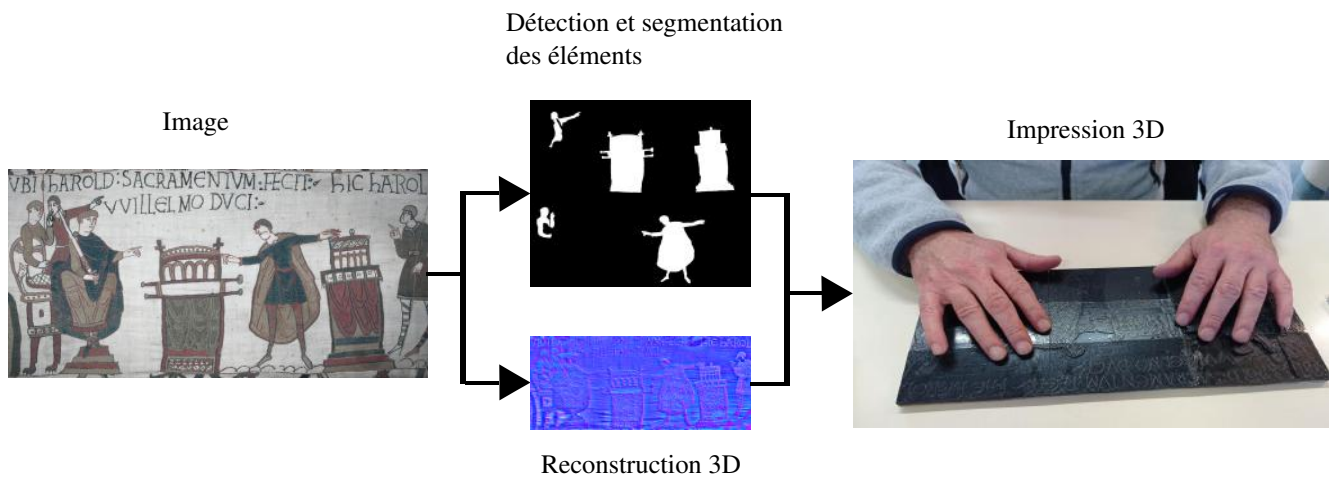


FIGURE 1 – Architecture générale du dispositif proposé, permettant de transformer une image de la Tapisserie de Bayeux (à gauche), en une impression 3D qui peut être explorée tactilement par les PPIVs (à droite). D’une part, un module de détection et segmentation permet d’isoler chaque élément constituant la scène. D’autre part, un module de reconstruction 3D transforme l’image en une carte de normales. Ces deux informations sont ensuite combinées en un modèle 3D où chaque élément est imprimé à une hauteur distincte et “texturé” grâce au micro-relief estimé de la broderie.

Dans cet article, nous présentons une méthodologie de création semi-automatisée de représentation en trois dimensions de scènes de la Tapisserie de Bayeux, afin de permettre aux PPIVs de “toucher” la broderie et de pouvoir en reconnaître les éléments lors d’expériences d’exploration tactile. Comme illustré sur la figure 1, le système que nous proposons prend en entrée une image de l’œuvre, et produit un modèle 3D imprimable sous forme de bas-relief. Dans cette représentation en trois dimensions, chaque élément constituant la scène (personnages, objets, etc.) est représenté à une hauteur particulière, afin de faciliter l’identification tactile de ces éléments. En outre, l’intérieur de chaque élément reproduit le micro-relief réel de la broderie, dans le but de le faire percevoir aux PPIVs.

La construction semi-automatique d’une telle représentation nécessite de résoudre deux problèmes bien connus en vision par ordinateur : la détection et segmentation des éléments de la scène, et la reconstruction 3D de la surface de la broderie à partir d’une seule photographie. Dans ce travail, nous recourons à des outils d’intelligence artificielle pour résoudre ces deux problèmes, à savoir des algorithmes de segmentation supervisée de type R-CNN, et de translation d’image par réseaux génératifs antagonistes (GANs).

Après avoir brièvement décrit quelques études liées à notre travail dans le paragraphe 2, nous détaillons dans le paragraphe 3 les différentes étapes de notre pipeline de création semi-automatisée de représentations en trois dimensions. Puis, nous présentons dans le paragraphe 4 les résultats des premières évaluations que nous avons menées auprès de PPIVs, avant de présenter nos conclusions et perspectives dans le paragraphe 5.

2 État de l’art

La création de représentations tactiles d’œuvres muséales passe majoritairement par la création manuelle d’objets par des artistes [4]. De telles représentations ont déjà été réalisées pour la Tapisserie de Bayeux, par exemple celle qui est présentée sur la figure 2. Rappelons toutefois que ce processus de création artistique est très fastidieux : les artistes bénéficieraient fortement de la mise à disposition d’outils semi-automatiques pour les accompagner dans leur travail. En outre, le relief qui est créé n’est pas le relief “réel” de la broderie, mais plutôt une représentation subjective par l’artiste. La nature intrinsèque de l’œuvre (la finesse de la laine, les points de broderie, etc.) reste donc inaccessible aux PPIVs. Notre travail vise à proposer des éléments de réponse à ces deux problématiques.



FIGURE 2 – Bas-relief réalisé manuellement par M. Closet : “Harold prête serment sur les reliques”. Il s’agit de la même scène de la Tapisserie de Bayeux que sur la figure 1. L’automatisation partielle de la génération de telles représentations tactiles constitue l’objectif de ce travail. Source de l’image : Association Valentin Haüy / Musée de Bayeux, 2019.

Plutôt que de choisir une représentation en relief arbitraire de l'œuvre, il est possible de réaliser un modèle en trois dimensions dans laquelle chaque élément est texturé au moyen d'un code en braille, comme ce fut fait dans le cadre de l'exposition de la bataille de Pavie [3]. Cependant, de telles représentations deviennent rapidement difficiles à appréhender dès lors que le nombre d'éléments représentés est important. À nouveau, cela souligne l'importance d'outils de segmentation semi-supervisés, et d'une représentation en relief qui soit à la fois pertinente en contenu et facile à percevoir au toucher.

J. Rodrigues et al. [17] proposent une méthode semi-automatique de construction d'objets tactiles en trois dimensions. Chaque élément de la scène est d'abord segmenté par l'algorithme semi-supervisé GrabCut [19]. Puis, un relief en trois dimensions est généré grâce à l'algorithme Lawson Flip [11]. Le relief est par la suite affiné par un algorithme de Ray Casting. Une texture fondée sur la cartographie des déplacements (Displacement mapping) [20] est ensuite générée. Après quoi, une interface graphique permet de choisir manuellement la taille des éléments et leur hauteur. Enfin, le tout peut être sauvegardé dans un modèle 3D. Ce dispositif présente plusieurs inconvénients pour des œuvres complexes tels que la Tapisserie de Bayeux. En effet, chaque élément doit être traité séparément, ce qui implique un temps considérable pour la segmentation de l'ensemble de l'œuvre. De plus, la génération des reliefs appliqués se fonde uniquement sur la couleur de l'image, et non sur la géométrie de la scène imagée. Or, dans notre cas, nous souhaitons nous rapprocher au plus près de l'œuvre originale, ce qui nous pousse à vouloir représenter ses micro-reliefs avec toutes leurs finesses. L'ajout des micro-reliefs de la Tapisserie de Bayeux sur les impressions nécessite de recourir à une méthode de reconstruction 3D plus élaborée.

Représenter le relief "réel" d'œuvres de grande dimension est possible grâce à des techniques de reconstruction par photogrammétrie, ce qui a par exemple été réalisé pour la "Table fontaine de Meissen" présentée au Victoria and Albert Museum à Londres [15]. Dans ce travail, le scannage 3D est suivi d'un algorithme d'aplatissement permettant *in fine* d'obtenir un bas-relief. Dans notre cas, nous souhaitons représenter les fines variations de relief induites par les fils de laine sur une toile de lin : les techniques usuelles de photogrammétrie ne permettent pas de reconstruire avec une résolution suffisante le relief [12]. Le recours à des techniques photométriques semble préférable, comme cela a déjà été proposé notamment dans [8, 9]. Cependant, ces travaux s'intéressent à donner une interprétation en trois dimensions de peintures, qui sont par construction soumises à l'ambiguïté du trompe l'œil, tandis que nous nous intéressons à un objet présentant de réelles micro-variations de relief à sa surface.

3 Présentation du dispositif

Le dispositif que nous proposons est fondé sur l'utilisation d'outils d'apprentissage profond. D'une part, il permet de segmenter les éléments présents dans la photographie, et d'autre part de reconstruire le relief de la surface de la scène imagée. Et ainsi, de rendre possible la perception des micro-reliefs de l'œuvre sur l'impression 3D. L'architecture générale du dispositif est résumée sur la figure 1, et nous en détaillons dans la suite les principales étapes.

3.1 Détection et segmentation

La première partie du processus concerne la détection et segmentation des différents éléments constituant l'image. Les éléments présents sur la Tapisserie de Bayeux (personnes, chevaux, bateaux, armes, etc.) ont déjà fait l'objet d'un recensement. Nous savons qu'ils sont au nombre de 1515 selon Simone Bertrand [1]. De plus, grâce à sa technique de conception, l'œuvre étudiée présente plusieurs avantages. D'une part, la stylisation des éléments tels que les personnages ou les animaux est la même sur l'ensemble de la broderie. Cette uniformité du style de broderie devrait faciliter la capacité de généralisation d'un réseau de neurones entraîné à la détection et la segmentation. D'autre part, le nombre de catégories d'éléments représentés peut se limiter à une dizaine.

Nous avons choisi de nous concentrer sur quatre de ces catégories : les lettres, les animaux, les hommes et les bateaux. L'étape de détection et de segmentation a alors pour objectif de localiser (via une boîte englobante) chacun de ces éléments dans l'image, de les segmenter précisément (afin de pouvoir par la suite séparer ces éléments du fond), et de les classifier. Nous utilisons pour cette tâche l'implémentation de l'algorithme Mask-RCNN [16] fournie dans Detectron2 [21]. Cet algorithme est illustré sur la figure 3.

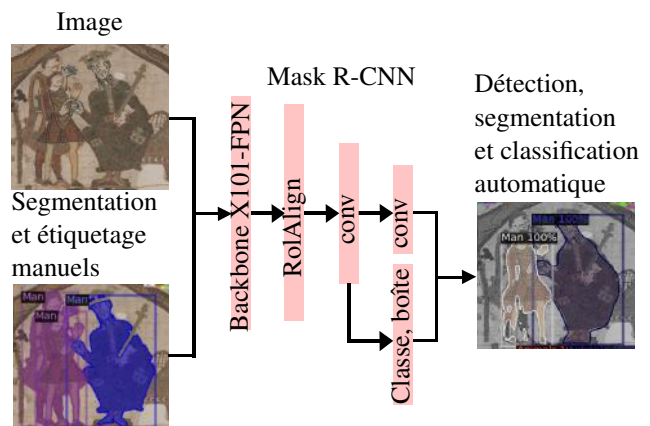


FIGURE 3 – Architecture du réseau de neurones Mask-RCNN. Des paires (images, annotations) sont mises en entrée du réseau de neurones, qui est entraîné à localiser, segmenter et classifier les éléments présents dans l'image. Dans cet exemple de la base d'apprentissage, la segmentation des trois personnages est difficile, car un des personnages sur la gauche est partiellement occulté par un autre.

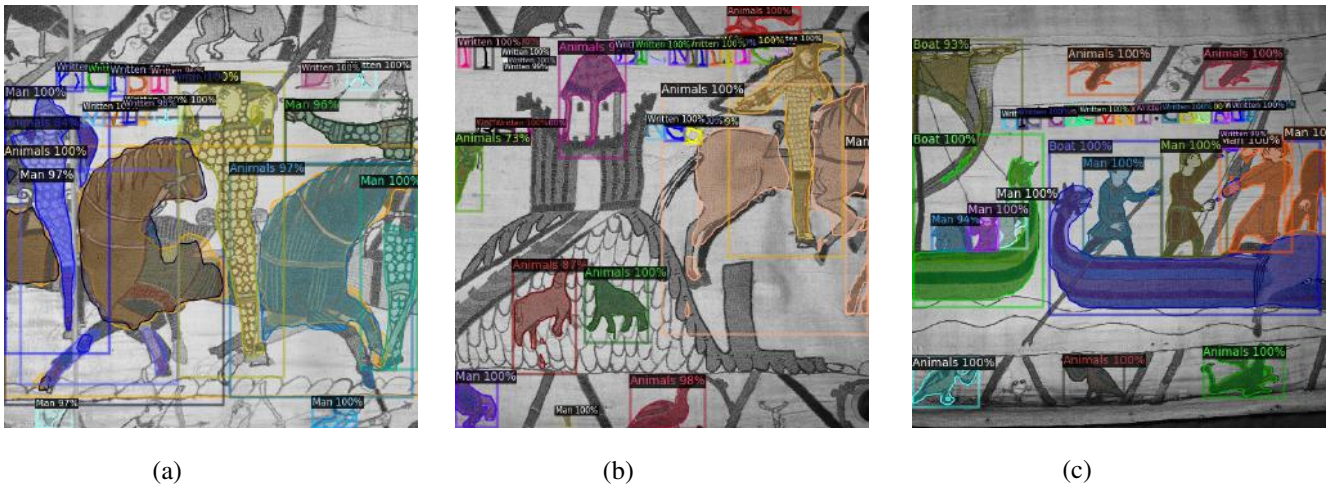


FIGURE 4 – Résultats de détection et de segmentation sur trois images de la base de test. Les quatre catégories d'éléments (lettres, animaux, hommes et bateaux) sont globalement bien retrouvées. Toutefois, on peut remarquer certaines limitations : en (a), la tête du personnage de droite n'est pas segmentée (sous-segmentation) ; en (b) le haut de la tour est faussement détecté comme étant un animal (erreur de classification) ; et en (c) le bateau de gauche est détecté en deux parties – une première partie avec la coque du navire et une seconde avec la voile (sur-segmentation).

Pour constituer la base de données, nous avons choisi 12 images représentatives de la broderie, toute de taille 8100×5600 pixels. Sur chacune, nous avons annoté les éléments appartenant aux quatre classes grâce à l'outil CVAT [7], qui permet de sélectionner les contours des objets et de les étiqueter via une interface graphique. Il autorise en outre la réalisation de groupements d'annotations. Ainsi, deux éléments se superposant auront des annotations qui ne recouperont pas, cela est illustré sur la figure 5.



FIGURE 5 – Exemples d'annotations simples (à gauche) et complexes (à droite).

Afin d'augmenter la base de données, nous simulons d'abord des translations de la caméra. Pour cela, nous coupons aléatoirement les images d'entrées en imagerie de taille 5500×5500 pixels. Elles sont ensuite redimensionnées par interpolation linéaire à une résolution de 1000×1000 pixels. La diminution de résolution des images nous autorise à entraîner le réseau de neurones sur des petits serveurs. Ceci nous permet d'obtenir 560 images annotées pour l'apprentissage et 50 pour le test. Une procédure d'augmentation de données (modifications aléatoires de la brillance, de la saturation, de la luminosité et du contraste) est enfin appliquée sur la base d'apprentissage. Ces étapes nous permettent de disposer finalement de 56000 couples (image, annotations) pour l'entraînement du réseau.

La plateforme Detectron2 [21] propose 8 modèles de détection et segmentation pré-entraînés sur la base de donnée COCO par le réseau de type Mask-R-CNN. Une étude comparative des performances sur nos données de chacun des modèles disponibles nous a menés à choisir celui nommé X101-FPN. Nous avons recouru à une approche de type apprentissage par transfert. En partant du modèle pré-entraîné, nous avons ré-affiné les poids du réseau de neurones à partir de nos données d'apprentissage. Pour l'optimisation, nous utilisons un algorithme de descente de gradient stochastique inertielle, que nous arrêtons après 50000 époques. La taille du lot est de 5 images par époque. Le processus de ré-affinement des poids requiert environ 30 heures de calcul sur un processeur Intel Xeon E5-26-40 v4 à 2.4 GHz. Le processeur est équipé d'un GPU MSI Ge Force GTX 1080 Ti 3584 Cores avec 11GB de RAM.

Les résultats que nous obtenons, illustrés sur la figure 4, sont encourageants tant en termes de détection que de segmentation des éléments. En effet, sur des éléments simples, comme les lettres et les éléments non superposés, le résultat est conforme à ce qui est attendu pour chacune des catégories. Cependant, dès que les éléments se superposent, les limites de l'approche deviennent visibles, et des problèmes de sous-segmentation peuvent apparaître (voir la figure 4-a). Une parade possible consisterait à affiner dans un second temps la segmentation automatique par une technique semi-supervisée comme GrabCut [19]. Il se peut aussi que des éléments soient faussement détectés ou non détectés, comme dans l'exemple de la figure 4-b. On peut alors imaginer une vérification manuelle a posteriori. Il arrive enfin qu'un élément soit sur-segmenté (figure 4-c). Un affinage a posteriori serait là aussi nécessaire.

3.2 Reconstruction du relief réel de l'œuvre

Nous souhaitons pouvoir intégrer le relief de l'œuvre originale à nos impressions en trois dimensions, de telle sorte qu'en plus de pouvoir localiser les éléments via leurs contours, les PPIVs puissent également percevoir les micro-reliefs de la broderie. Ceci nécessite de reconstruire le relief de la scène à partir d'une seule photographie, ce qui est un problème inverse mal posé puisqu'il est impossible de séparer la couleur du relief sans information a priori. Afin de lever les ambiguïtés inhérentes à ce problème de reconstruction 3D à partir d'une seule image, nous avons constitué une grande base de données de couples (image, relief), et entraîné un réseau à transformer une image en un relief. Dans la mesure où toutes les images sur lesquelles nous allons travailler représentent des fils de laine (de différentes couleurs) sur une toile de lin, nous espérons que le réseau parviendra à généraliser correctement et, parmi l'infinité de reliefs expliquant l'image, à en trouver un qui soit qualitativement satisfaisant (on rappelle que notre travail a une visée pédagogique et ludique, et non métrologique).

Afin de capturer les plus fines variations du relief de la Tapisserie de Bayeux, nous choisissons de le représenter sous forme d'une carte de normales RGB. Chaque composante de couleur de cette carte encode une des composantes du vecteur normal à la surface, dans un repère orthonormé ayant pour centre l'appareil photographique. Nous considérons le problème d'association entre image et carte de normales comme un problème de translation d'image, que nous résolvons grâce à l'algorithme Pix2Pix [10] fondé sur un réseau génératif antagoniste (GAN). L'architecture de ce réseau est illustrée sur la figure 6.

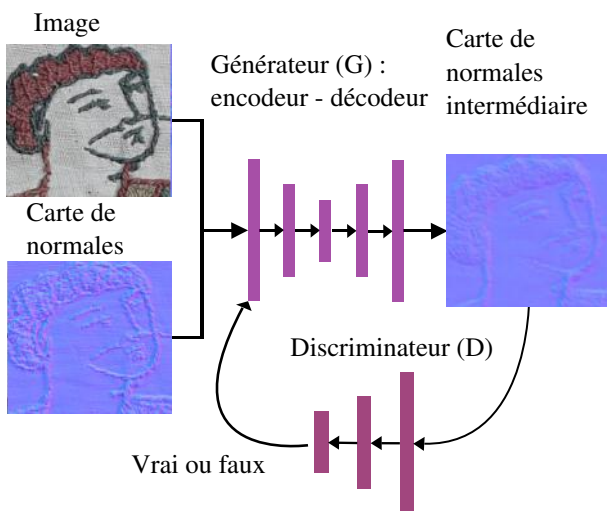


FIGURE 6 – Architecture du réseau génératif antagoniste. Des couples (image, carte de normales) sont envoyés dans un générateur (G) de type encodeur-décodeur. Celui-ci crée une nouvelle carte de normales qui est comparée avec la carte de normales réelle fournie en entrée, par le discriminateur (D), dans le but de minimiser la distance L1 entre les deux cartes de normales.

Pour créer la base d'apprentissage, nous avons besoin de disposer de nombreuses cartes de normales "vérité terrain". Pour cela, nous avons réalisé des acquisitions de plusieurs scènes de la broderie, depuis le même point de vue mais sous des angles d'éclairage différents. Ceci permet de décomposer une image en une carte de réflectance (la "couleur" réelle de la scène, indépendamment des effets d'ombrage) et une carte de normales (qui encode le relief), selon le principe de la reconstruction 3D par stéréophotométrie. La figure 7 illustre ce processus, et davantage de détails sur cette numérisation 3D peuvent être trouvés dans [12].



FIGURE 7 – Création des cartes de normales "vérité terrain" par stéréophotométrie [12]. À partir d'un ensemble de photographies prises depuis le même point de vue mais sous des éclairages différents (à gauche), l'image (à droite, bord gauche) peut être décomposée en une carte de réflectance (à droite, bord supérieur) et une carte de normales (à droite, bord inférieur).

Suite à cette numérisation, nous disposons de 16 paires (image, carte de normales) en haute définition, parmi lesquelles nous en choisissons 13 pour l'entraînement et 3 pour le test. Chaque couple d'images est ensuite coupé de façon aléatoire en 500 imageries de taille 256×256 pixels, ce qui nous donne finalement 6500 données d'entraînement et 1500 données de test. Le réseau est ensuite entraîné à minimiser l'écart en norme 1 entre les cartes de normales réelles et celles simulées, au moyen de l'algorithme RMSPProp, que nous laissons itérer durant 50 époques. Ceci nécessite environ 12 heures de calcul dans les mêmes conditions que dans le paragraphe 3

Les résultats obtenus sur trois images de la base test sont présentés sur la figure 8. Dans l'ensemble, les cartes générées par le réseau sont qualitativement très proches des cartes réelles, ce qui valide l'approche. Quelques imperfections peuvent toutefois être remarquées. Par exemple, sur la deuxième image de la figure 8, un relief absent sur la carte réelle a été généré à l'endroit où se situe une tache. Le réseau n'a pas été capable dans ce cas d'interpréter la variation de couleur autrement que par une variation de relief, probablement par ce que la base d'apprentissage ne contenait pas de telles tâches. Agrandir la base d'apprentissage permettrait donc sans doute d'éviter ces artefacts. Nous constatons par ailleurs que les cartes générées, bien que réalistes, sont plus floues que les cartes réelles (voir par exemple la troisième image de la figure 8), ce qui est probablement dû au fait que l'encodeur-décodeur se fonde sur des couches de convolution.

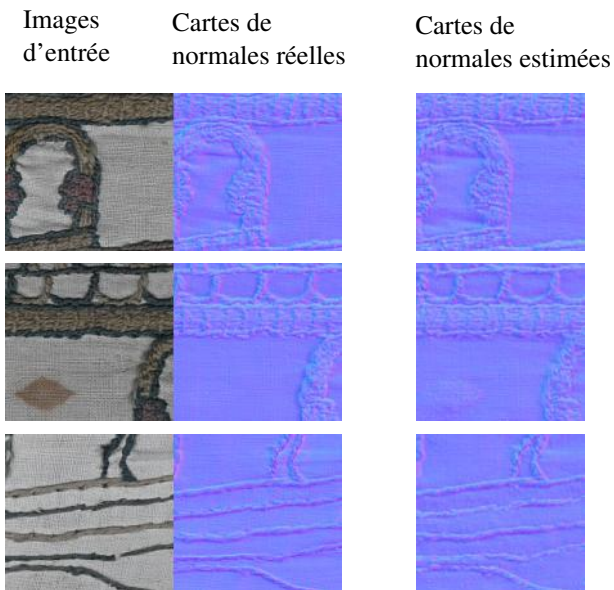


FIGURE 8 – Reconstruction de cartes de normales à partir de photographies issues de la base de test. Les cartes de normales estimées à partir d'une seule photographie sont très proches des cartes réelles obtenues à partir de plusieurs photographies par stéréophotométrie.

Les données "vérité terrain" ont été acquises avec un éclairage direct (flash) qui diffère de l'éclairage normal (lumière diffuse) de la Tapisserie de Bayeux. Afin d'évaluer la robustesse de notre approche à des variations de conditions d'éclairage, nous avons réalisé des tests sur des images acquises en conditions normales d'éclairage. Les résultats présentés sur la figure 9 sont plus flous que dans le cas précédent. En outre, nous constatons que certaines structures ne sont pas reconstruites (les deux arcs de cercle situés en haut à gauche dans la première image, ou la ligne horizontale dans la seconde). L'approche proposée s'avère donc assez peu robuste aux changements d'éclairage. Une parade possible consisterait à transférer les couleurs des images à reconstruire vers celles qui ont servi à constituer la base d'apprentissage, par exemple grâce à des outils de transport optimal [14].

3.3 Création des reliefs

À l'issue de la phase de segmentation et de détection, nous disposons d'un masque par élément. Ces masques nous servent à constituer les différentes hauteurs de l'objet en trois dimensions qui sera imprimé. Chacune de ces régions est ensuite hiérarchisée manuellement, en fixant la hauteur du plan moyen de l'élément correspondant, comme cela est illustré sur la figure 10. À l'avenir, une hiérarchisation automatique, construite en fonction des besoins et attentes des PPIVs, pourrait être envisagée, par exemple en hiérarchisant les hauteurs selon les classes (les lettres à la hauteur 1, les bateaux à la hauteur 2, etc.). Enfin, un contour permettant de correctement délimiter les bords des éléments est ajouté.

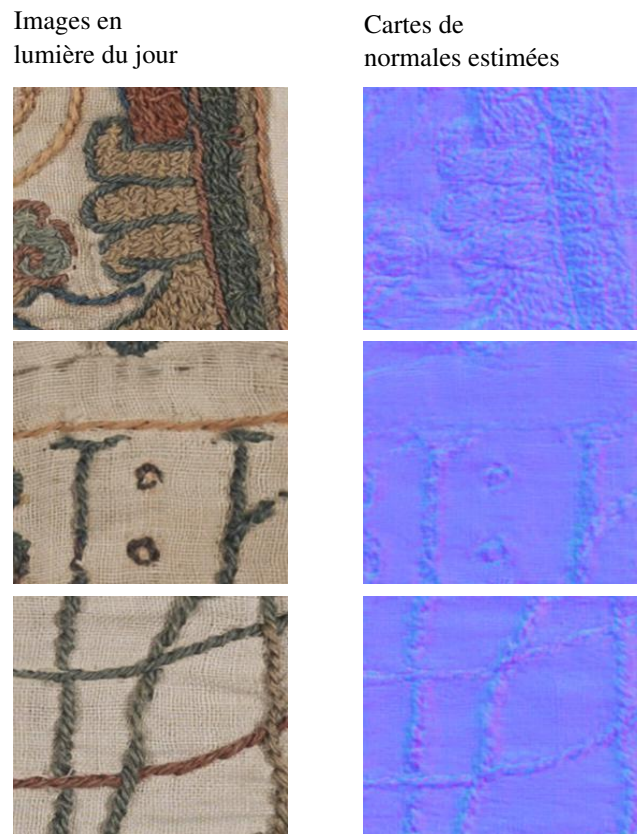


FIGURE 9 – Reconstruction des cartes de normales à partir de photographies acquises en condition d'éclairage non contrôlé. Bien que prometteurs, ces résultats restent inférieurs à ceux obtenus en lumière contrôlée (figure 8). En effet, la reconstruction est globalement plus floue, et certaines structures ne sont pas reconstruites.

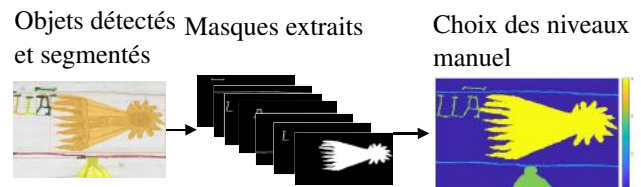


FIGURE 10 – Attribution manuelle d'une hauteur moyenne à chaque élément, à partir des masques issus de la phase de détection et de segmentation. Dans cet exemple, le fond est placé au niveau zéro, les lettres à 2 mm, et la comète à 3 mm.

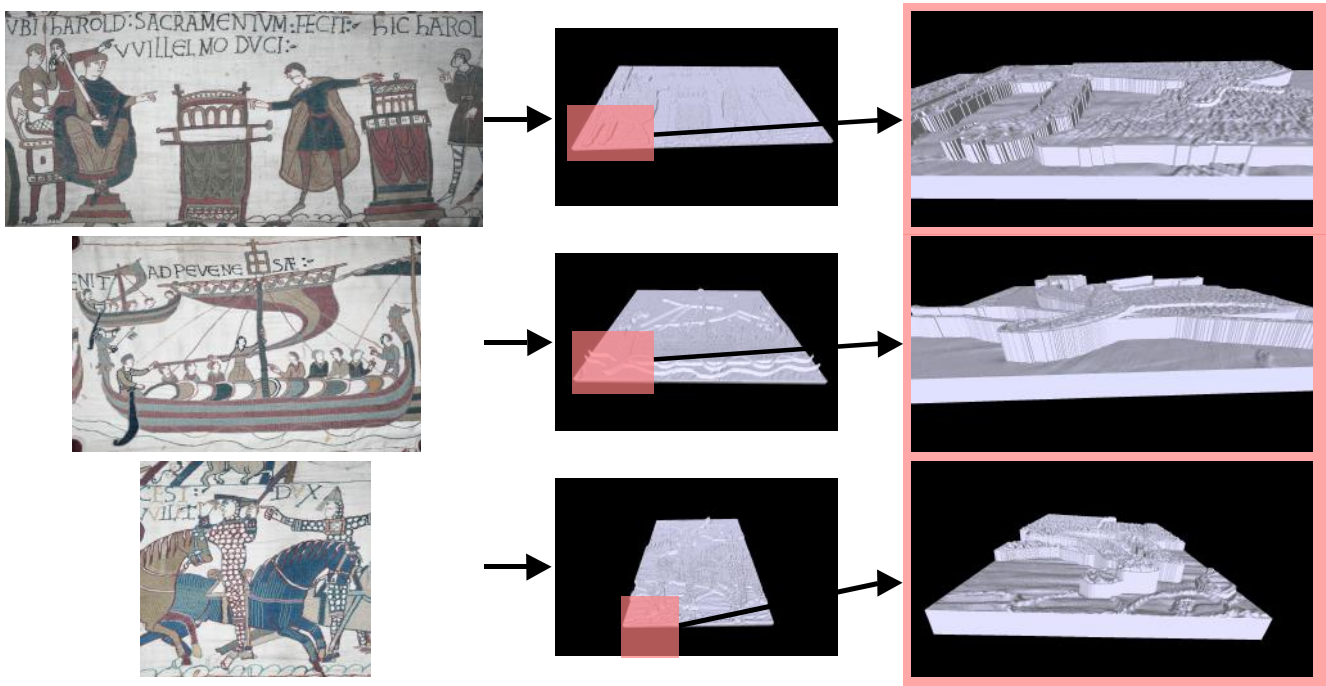


FIGURE 11 – Photographies de trois scènes de la Tapisserie de Bayeux, et modèles 3D générés en suivant l’approche proposée. Attribuer une hauteur différente à chacun des éléments segmentés permet de les mettre en évidence, tandis que le transfert de la géométrie de la broderie sur le modèle 3D permet de faire percevoir les micro-reliefs de l’œuvre.

Une fois la hauteur de chaque plan moyen fixé, il reste à leur transférer le relief haute fréquence précédemment estimé. La carte de normales dégauchie est intégrée en une carte de profondeur par transformation en cosinus [13]. Ceci fournit le relief à une constante additive près, que nous fixons grâce à la hauteur choisie par l'utilisateur. On obtient ainsi une représentation 2.5D où chaque élément est positionné à une hauteur différente, et où l'intérieur de chaque élément est "texturé" selon le relief réel de la broderie. La figure 11 montre quelques exemples de modèles 3D qui ont été réalisés ainsi.

4 Retours d'expérience

Afin de valider empiriquement l'intérêt du dispositif proposé, nous avons imprimé plusieurs bas-reliefs "idéaux". Par "idéal", nous entendons ici que la segmentation a été affinée manuellement pour éviter les phénomènes de sous- et sur-segmentation mentionnés plus haut, et que nous avons restitué le relief "vérité terrain" plutôt que l'estimation fournie par le réseau génératif antagoniste. Dans ces impressions, une hauteur distincte a été attribuée aux éléments ayant des bords communs afin de faciliter leur identification. Enfin, des contours lisses ont été ajoutés aux éléments. Certains des modèles 3D ainsi générés sont présentés sur la figure 11. Pour obtenir un retour d'expérience, comme illustré sur la figure 1, nous avons mis les bas-reliefs à disposition de 7 personnes : une "voyante", trois partiellement aveugles et trois autres aveugles.

4.1 Impressions générales

Les premiers retours d'expérience qui nous ont été communiqués sont très positifs. En effet, ces supports ont été perçus comme très complémentaires des plaques tactiles réalisées manuellement par Rémy Closset (figure 2). En effet, ces bas-reliefs manuels sont lissés, à l'instar de sculptures, tandis que nos impressions 3D retranscrivent le relief de l'œuvre. Toutefois, pour ces deux supports d'exploration tactile, l'ajout d'une audio-description est indispensable. Celle-ci permet de mieux guider la découverte de l'objet touché et d'en expliquer le contenu sémantique, comme cela est aussi analysé dans [18].

Il est également ressorti de cette expérience que les micro-reliefs doivent être très marqués. En effet, si leur hauteur n'est pas suffisante, on perçoit une "texture" assez lisse au toucher. On risque alors de confondre les reliefs de la broderie de laine avec ceux du fond (qui est une toile de lin). Dans nos futures impressions, nous amplifierons le relief de la broderie avant de l'intégrer, ce qui est d'autant plus facile à réaliser que nous disposons d'une représentation du relief sous forme de cartes de normales.

Une difficulté majeure reste la perception des objets qui sont partiellement occultés, comme la main du personnage s'accrochant au mât sur le deuxième exemple de la figure 11. Il nous faut dans ce cas décider entre respecter l'œuvre originelle ou au contraire la transformer, par exemple en prenant la liberté de ne pas représenter l'élément occultant. Cette seconde solution a été plus appréciée, car elle facilite la compréhension des éléments.

4.2 Extensions possibles du dispositif

Il apparaît également très important de faire attention à bien espacer les éléments représentés. Pour l’instant, nous avons conservé la disposition originelle de l’œuvre, afin de ne pas la trahir. Ceci n’est toutefois pas si pertinent, car il est parfois difficile de ressentir la séparation entre deux éléments si l’espacement est insuffisant. Dans nos futures impressions, nous ferons donc en sorte de mieux séparer les différents objets. Une autre piste possible consiste à permettre de détacher les éléments de leur support, à la manière d’un puzzle. Cette idée, qui a rendu les PPIVs participant à l’expérience très enthousiastes, permettrait à la fois de mieux discerner les contours de ces éléments et en même temps rendre l’expérience plus interactive.

Nous avons également proposé l’ajout de couleurs sur les impressions. Même si l’utilisation de couleurs pour les PPIVs ne semble pas pertinente au premier abord, elle l’est pour les personnes qui ne sont que partiellement aveugles, et l’utilisation de couleurs très contrastées leur permettrait de mieux distinguer les objets. En outre, l’intégralité de la Tapisserie de Bayeux a été réalisée en utilisant seulement 10 couleurs, ce qui rend la colorisation des impressions réalisable. Même si les couleurs de la broderie actuelle sont assez ternes, car dégradées par le temps. Des travaux en cours [6] ont pour objectifs d’identifier la composition des laines utilisées lors de la création de la Tapisserie de Bayeux, et donc les couleurs originelles. Il serait intéressant, à l’avenir, de réaliser des impressions avec ces couleurs originelles, afin de faire percevoir ce à quoi pouvait ressembler l’œuvre à la fin de sa création.

5 Conclusion et perspectives

Nous avons proposé un dispositif de génération semi-automatique d’objets 3D. En combinant des algorithmes de reconstruction 3D et de détection et segmentation automatiques, ce dispositif permet de produire plus rapidement et à moindre coût des bas-reliefs qui permettent de percevoir tactilement les éléments de la Tapisserie de Bayeux. En intégrant le relief réel de la broderie aux impressions, l’exploration tactile est également plus enrichissante. Outre les pistes d’amélioration évoquées dans le paragraphe 4, nous envisageons d’intégrer le dispositif dans un système inclusif multimodal, à la manière de Q. Cavazos et al. [5]. Un couplage de la segmentation et de la prédiction de la profondeur est aussi envisagé. Il permettrait l’amélioration de la qualité des deux procédés.

Remerciements

Ce travail a été soutenu par la subvention ANR “Guide Muséal Inclusif” (ANR-20-CE38-0007). Les autrices et auteurs remercient C. Berthelot du Musée de la Tapisserie de Bayeux pour avoir supervisé la campagne d’acquisition d’images de la Tapisserie de Bayeux, ainsi que l’association Valentin Haüy pour les expériences réalisées avec les PPIVs.

Références

- [1] S. BERTRAND : *La Tapisserie de Bayeux et la manière de vivre au onzième siècle*. Zodiaque, 1966.
- [2] M. J. BURTON et AL : The Lancet Global Health Commission on Global Eye Health : vision beyond 2020. *The Lancet Global Health*, 9(4):489–551, 2021.
- [3] V. CANTONI, P. DONDI, L. LOMBARDI, N. NUGRAHANINGSIH, M. PORTA et A. SETTI : A multi-sensory approach to cultural heritage : the battle of Pavia exhibition. In *IOP Conference Series : Materials Science and Engineering*, volume 364, page 012039, 2018.
- [4] M. CARFAGNI, R. FURFERI, L. GOVERNI, Y. VOLPE et G. TENNIRELLI : Tactile representation of paintings : an early assessment of possible computer based strategies. In *Euro-Mediterranean Conference*, pages 261–270, 2012.
- [5] L. CAVAZOS QUERO, J. IRANZO BARTOLOMÉ et J/ CHO : Accessible visual artworks for blind and visually impaired people : comparing a multimodal approach with tactile graphics. *Electronics*, 10(3):297, 2021.
- [6] C. CHAVANNE, V. VERNEY, M. PAQUIER-BERTHELOT, C. and-Bostal, P. BULÉON et P. WALTER : Bayeux Tapestry : First use of early synthetic dyes for the restoration of a masterpiece. *Dyes and Pigments*, 208:110798, 2022.
- [7] CVAT.AI CORPORATION : Computer Vision Annotation Tool (CVAT), 9 2022.
- [8] R. FURFERI, L. GOVERNI, Y. VOLPE, L. PUGGELLI, N. VANNI et M. CARFAGNI : From 2D to 2.5D i.e. from painting to tactile model. *Graphical Models*, 76(6):706–723, 2014.
- [9] L. GOVERNI, R. FURFERI, Y. VOLPE, L. PUGGELLI et N. VANNI : Tactile exploration of paintings : an interactive procedure for the reconstruction of 2.5 D models. In *Mediterranean Conference on Control and Automation*, pages 14–19, 2014.
- [10] P. ISOLA, J. ZHU, T. ZHOU et A. A. EFROS : Image-to-Image Translation with Conditional Adversarial Nets. In *CVPR*, pages 5967–5976, 2017.
- [11] C. L. LAWSON : Transforming triangulations. *Discrete Mathematics*, 3(4):365–372, 1972.
- [12] M. PIZENBERG, A. ELMOATAZ et Y. QUÉAU : Image Processing for Cultural Heritage Accessibility : Digitizing the Bayeux Tapestry. <https://hal.science/hal-03684035>, 2022.
- [13] Y. QUÉAU, J.-D. DUROU et J.-F. AUJOL : Normal integration : a survey. *JMIV*, 60:576–593, 2018.
- [14] J. RABIN, S. FERRADANS et N. PAPADAKIS : Adaptive color transfer with relaxed optimal transport. In *ICIP*, pages 4852–4856, 2014.
- [15] A. REICHINGER, H. G. CARRIZOSA et C. TRAVNICEK : Designing an interactive tactile relief of the meissen table fountain. In *International Conference on Computers Helping People with Special Needs*, pages 209–216, 2018.
- [16] S. REN, K. HE, R. GIRSHICK et J. SUN : Faster R-CNN : Towards Real-Time Object Detection with Region Proposal Networks. *PAMI*, 39(6):1137–1149, 2015.
- [17] J. RODRIGUES, A. FERREIRA, I. MAIA, G. J. BRAZ et J. de ALMEIDA : Image Processing of Artworks for Construction of 3D Models Accessible to the Visually Impaired. *Advances in Manufacturing, Production Management and Process Control*, pages 243–253, 2019.
- [18] K. ROMEO, H. THOMPSON et M. CHOTTIN : Inclusive Multimodal Discovery of Cultural Heritage : Listen and Touch. In *Computers Helping People with Special Needs*, pages 278–285, 2022.
- [19] C. ROTHER, V. KOLMOGOROV et A. BLAKE : Grabcut : interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 23:309–314, 2004.
- [20] L. SZIRMAY-KALOS et T. UMENHOFFER : Displacement Mapping on the GPU - State of the Art. *Computer Graphics Forum*, 27(6):1567–1592, 2008.
- [21] Y. WU, A. KIRILLOV, F. MASSA, W. LO et R. GIRSHICK : Detectron2. <https://github.com/facebookresearch/detectron2>.