



Reconstruction et correspondance de formes par apprentissage

Thibault Groueix ¹

Thibault Groueix a soutenu sa thèse² le 22 octobre 2020 à l'université Paris-Est sous la direction de Mathieu Aubry et Renaud Marlet. Vous trouverez ci-dessous un résumé de cette thèse.

La création artistique en 3D numérique est affaire d'experts car elle est coûteuse en temps et nécessite une grande expérience technique. Afin de rendre cet art de niche accessible à tous, le travail de thèse a visé à développer des algorithmes permettant à un amateur d'effectuer des tâches complexes automatiquement. Par exemple, prendre un objet en photo avec un téléphone et éditer ensuite une estimation de son modèle 3D, ou encore modifier l'apparence d'un objet 3D en transférant la texture d'un autre modèle à l'aide d'un simple clic. Pour ce faire, nous proposons de nouvelles méthodes d'apprentissage profond pour modéliser et analyser les formes 3D en se concentrant sur deux tâches clefs : reconstruire un modèle 3D à partir d'une seule image et mettre des modèles 3D en correspondance.

Une méthode de reconstruction 3D à partir d'une seule image (SVR) est un algorithme qui prend en entrée une image et prédit un modèle virtuel en 3D du monde physique qui a engendré cette image. Ce problème remonte aux premiers jours de la vision par ordinateur, et il est très difficile. En effet, plusieurs configurations de formes, de textures et d'éclairages peuvent expliquer la même image et il y a donc une infinité de solutions au problème. La clef pour attaquer ce problème mal posé est donc de formuler des hypothèses sur ce qu'est un environnement 3D réaliste

1. thibault.groueix.2012@polytechnique.org.

2. <http://www.theses.fr/2020PESC1024>.

et d'utiliser ces hypothèses pour sélectionner une solution à notre problème parmi cette infinité. Dans ce travail, nous apprenons ces hypothèses directement à partir de grandes bases de données, au lieu de les concevoir manuellement. L'apprentissage sur les données permet de formuler des hypothèses très puissantes, qui permettent même de reconstruire les parties invisibles des objets. L'intérêt de développer des méthodes de SVR ne s'arrête pas à la modélisation 3D. En robotique par exemple, il est critique que les agents autonomes disposent d'une représentation 3D de leur environnement à partir de leurs capteurs visuels.

La mise en correspondance de formes vise à établir des correspondances entre des objets 3D. Cette tâche se pose traditionnellement sous forme d'un problème d'optimisation non-convexe, dont la convergence est sujette aux minimum locaux. Au lieu de cela, nous proposons de résoudre un problème d'apprentissage sur de grands jeux de données. La mise en correspondance de formes a de nombreuses applications en modélisation 3D telles que le transfert d'attribut, le gréement automatique pour l'animation ou l'édition de maillage. C'est l'une des briques élémentaires de l'analyse de forme.

Nous présentons à présent les contributions principales de ce travail pour avancer dans la compréhension et la résolution de ces deux tâches.

La première contribution technique est une nouvelle représentation paramétrique des surfaces 3D, que nous modélisons avec des réseaux neuronaux. Le choix de la représentation des données est un aspect critique de tout algorithme de reconstruction 3D. Jusqu'à récemment, la plupart des approches profondes prédisaient des grilles volumétriques de voxel ou des nuages de points, qui sont des représentations discrètes. Au lieu de cela, nous présentons une approche qui prédit une déformation paramétrique de surface. Notre approche, baptisée AtlasNet (cf. figure 1), est la première approche profonde de SVR capable de reconstruire des maillages à partir d'une seule image sans s'appuyer sur un post-traitement, et peut le faire à une résolution arbitraire sans problème de mémoire. Une analyse plus détaillée d'AtlasNet révèle qu'en sus, par rapport aux autres approches par apprentissage, il généralise mieux sa proposition aux catégories pour lesquelles il n'a pas été entraîné.

Notre seconde contribution principale est une nouvelle approche de correspondance de formes, appelée 3D-CODED. Entièrement basée sur des reconstructions par déformation de surface, la clef de voûte de notre raisonnement est de lier correspondances et reconstructions : nous montrons que la qualité des correspondances prédites dépend de la qualité des reconstructions 3D. Nous mettons ainsi en relation deux grands problèmes de 3D. Sur le plan technique, nous introduisons également une optimisation au moment de l'inférence pour affiner les déformations apprises, combinant ainsi apprentissage et optimisation. 3D-CODED (cf. figure 2) se décline en deux versions. Tout d'abord, pour les humains et d'autres catégories qui diffèrent d'une quasi-isométrie, notre approche peut tirer parti d'un modèle de forme. Pour les autres catégories, celles présentant des variations non isométriques, telles

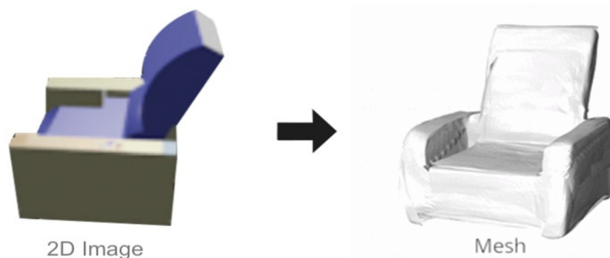


FIGURE 1. Reconstruction 3D à partir d'une seule image : notre approche, AtlasNet, reconstruit un maillage paramétrique de l'objet. Notez que même les parties non visibles dans l'image sont reconstruites.



FIGURE 2. Correspondances de formes : notre approche, 3D-CODED, établit automatiquement des correspondances entre deux formes, suggérées par la couleur.

des chaises, nous apprenons à déformer n'importe quelle forme en n'importe quelle autre et introduisons des contraintes de cohérence du cycle pour apprendre des correspondances respectant la sémantique. 3D-CODED fonctionne directement sur les nuages de points ; elle est robuste à de nombreux types de perturbations et a surpassé l'état de l'art de 15 % sur des scans d'humains réels.

Les algorithmes développées dans ce travail ont été présentés en conférences internationales, et ont suscité l'intérêt dans plusieurs groupes de recherches. Le code est libre d'accès sur GitHub³ et activement utilisé. À ma connaissance, aucun transfert technologique vers un produit commercialisé n'a encore abouti, car de nombreuses questions de recherche restent ouvertes. Toutefois, plusieurs brevets industriels connexes à ces travaux ont été déposés, notamment par l'entreprise Adobe.

3. <https://github.com/ThibaultGROUEIX>.