Université Claude Bernard Lyon 1

DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **27 mars 2024**

Nom de famille et prénom de l'auteur : **Monsieur PRADELLE Olivier**

Titre de la thèse : *« Analyse et enrichissement des scènes 3D numérisées »*

Résumé

Au fil des dernières décennies, les technologies de numérisation 3D LiDAR ont gagné en popularité, permettant de capturer de nombreuses scènes du monde réel, que ce soit en intérieur ou en extérieur. Ces dispositifs de capture génèrent une quantité considérable de données, comprenant des nuages de points comportant des centaines de millions d'éléments, ainsi que des dizaines de milliers d'images. En raison de la croissance constante du volume de données résultant de ces acquisitions et de la complexité croissante des scènes numérisées, les méthodes algorithmiques traditionnelles voient leur efficacité décroître.
Depuis quelques années, les méthodes d'apprentissage automatique sont devenues de plus en plus populaires et ont permis au domaine de l'analyse de scènes d'accomplir des tâches de plus en plus complexes, telles que la reconnaissance et la segmentation d'objets.

En raison des conditions d'acquisition, certaines images ne peuvent pas être utilisées pour extraire des informations sur la scène. En effet, ces dernières peuvent être difficilement exploitables en cas de variations de luminosité, qu'elles soient sous-exposées ou surexposées. De la même façon, le nuage de points obtenu peut présenter des erreurs liées à des réflexions sur les surfaces, des objets mobiles ou des zones peu accessibles. Toutefois, nuages de points et images peuvent être mis en correspondance et utilisés pour compléter l'information manquante ou altérée par les conditions d'acquisition. Cette thèse propose plusieurs contributions sur l'analyse jointe de nuages de points et d'image.

Partant d'images et de nuages de points dont la pose est connue, nous proposons d'incorporer les informations géométriques présentes dans le nuage de points en tant que données complémentaires pour la tâche de segmentation d'images. Ces informations permettent de compléter les données présentes dans l'image, ce qui améliore les performances des réseaux neuronaux de segmentation.

Contrairement à d'autres méthodes, notre approche requiert peu de mémoire supplémentaire pour intégrer les informations 3D, tout en évitant la nécessité d'étiqueter chaque point 3D individuellement pour l'apprentissage.

Dans le cadre de la segmentation sémantique de scène, l'obtention d'annotation sur les nuages de points et les images est un travail coûteux en temps. Pour réduire le temps nécessaire à l'annotation d'image, nous proposons de mettre à profit les annotations présentes dans le nuage de points en les projetant sur le plan image. Cependant, la projection d'un nuage de points sur une image engendre une image peu dense, ce qui diminue les performances d'un réseau de segmentation d'image en raison du manque d'informations pour guider le réseau. Pour pallier cette perte d'information, nous proposons de définir un terme de régularisation lors de l'apprentissage, exploitant la cohérence locale de l'image.

Cette thèse s'intéresse également à l'établissement d'une correspondance entre les nuages de points et les images. Pour ce faire, les paramètres de la caméra sont essentiels. Cependant, il est important de noter que ces paramètres peuvent ne pas être accessibles ou précis dans certains cas. Nous explorons également dans cette thèse l'utilisation de réseaux de type Transformer pour retrouver cette correspondance, étant donné que ces réseaux sont capables d'identifier la redondance d'informations présente à la fois dans les nuages de points et les images.