



Amélioration de la reconstruction 3D par Gaussian Splatting et stéréovision

August 21, 2025

Dorian Geraldès Pereira

Encadrant : M. Romain Dupont

CEA, LIST, LABORATOIRE DE VISION POUR LA MODÉLISATION ET LA LOCALISATION

Outline

Introduction

Données utilisées

Méthodologie

Outils

Implémentation et Résultats

Conclusion



Outline

Introduction

Données utilisées

Méthodologie

Outils

Implémentation et Résultats

Conclusion



Le laboratoire LVML - CEA List

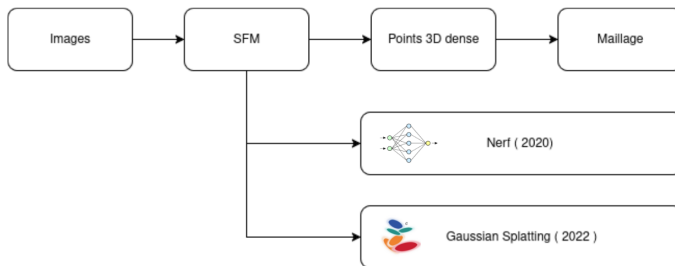
Le LVML (Laboratoire Vision pour la Modélisation et la Localisation) est un laboratoire de recherche appliquée du **CEA List**, spécialisé en vision par ordinateur et apprentissage automatique.

Domaines de recherche principaux :

- Localisation dans l'environnement
- Suivi d'objets en mouvement
- Reconstruction 3D de scènes
- Vision pour la robotique et la réalité augmentée

Contexte : reconstruction 3D à partir d'images

- Objectif général : reconstruire une scène 3D à partir d'un ensemble d'images prises sous différents points de vue.
- Deux objectifs principaux en contexte industriel :
 - **Rendu photoréaliste** : réalité virtuel, réalité augmentée.
 - **Maillage 3D** : modélisation, métrologie, analyse géométrique.



Contexte: Reconstruction 3D partir d'images



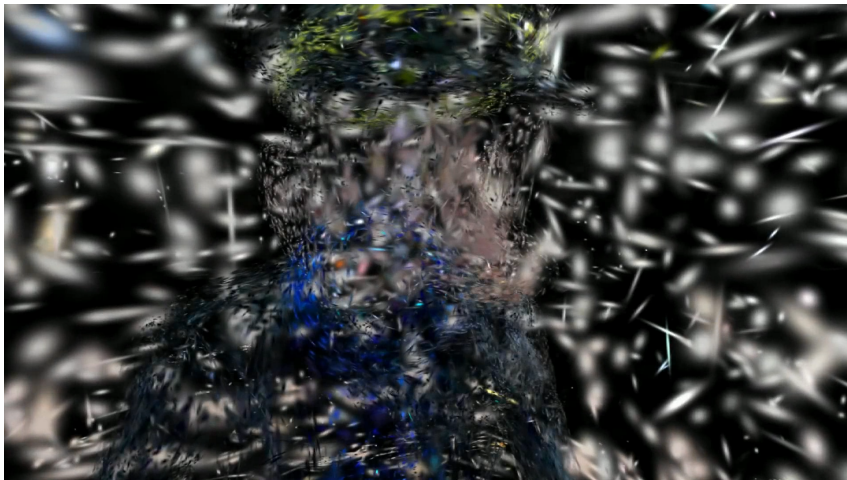
- Pourquoi viser un maillage ?
 - **Cohérence géométrique.**
 - **Mesures physiques fiables** (distances, volumes...).
 - **Topologie fine** (trous, cavités, connexité).
 - **Support pour la simulation** ou la planification robotique.

Le Gaussian splatting



Gaussienne d'un gaussian splatting

Le Gaussian splatting



Optimisation des gaussienne

Le Gaussian splatting



Rendu du Gaussian splatting

Objectif du stage

- Extraire une **géométrie maillée exploitable** à partir de *Gaussian Splatting*.
- Travailler sur des données **stéréo** acquises en **conditions réelles**, avec :
 - effets de spécularité et réflexions perturbantes,
 - calibrations souvent inconnues,
 - nécessité de rectification stéréo,
 - qualité variable (flou, bruit, éclairage).
- Exploiter la stéréo pour guider la reconstruction 3D, améliorer la cohérence géométrique et la robustesse face aux données bruitées.
- Construire un pipeline complet, de l'image au maillage, adapté à ce contexte réaliste et complexe.



Paire de caméras capturant la scène.

Outline

Introduction

Données utilisées

Méthodologie

Outils

Implémentation et Résultats

Conclusion



Jeu de données et conditions d'acquisition

■ Séquence *Atelier électrique*

- 286 paires d'images stéréo (résolution : 1920×1200)
- Scène vaste avec surfaces peu texturées et réfléchissantes
- Paramètres caméra inconnus, estimés dans le pipeline
- Conditions réalistes, distances variables

■ Conditions de prise de vue

- Caméras non calibrées
- Mouvements complexes
- Éclairage et occlusions variables



Scène Atelier électrique

Types de données manipulées dans le pipeline



- **Images stéréo** : paires gauche/droite, base de la reconstruction.
- **Cartes de disparité** : décalage pixel à pixel entre les vues.
- **Cartes de profondeur** : distances métriques dérivées des disparités.
- **Nuages de points denses** : projection 3D des cartes de profondeur.
- **Poses de caméra** : position et orientation des images dans la scène.
- **Reconstruction 2DGS** : représentation par splats gaussiens optimisés.
- **Maillage 3D final** : surface reconstruite par post-traitement volumique.

Point de départ



Atelier Électrique



Atelier Électrique

Outline

Introduction

Données utilisées

Méthodologie

Outils

Implémentation et Résultats

Conclusion

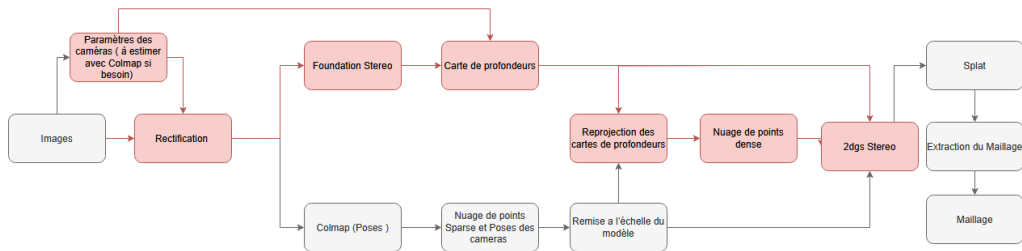


Pipeline Gaussian Splatting



Pipeline d'un Gaussian Splatting classique.

Méthodologie



Pipeline complet proposé : de la stéréo brute jusqu'au maillage optimisé.

Outline

Introduction

Données utilisées

Méthodologie

Outils

Implémentation et Résultats

Conclusion

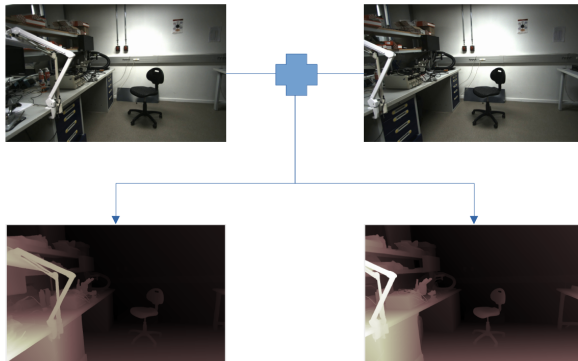




Trajectoire et nuage de
point sparse

Johannes Lutz Schönberger and Jan-Michael Frahm. [Structure-from-motion revisited.](#)
In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016

Foundation Stereo



Exemple de résultat de Foundation Stereo

Bowen Wen, Matthew Trepte, Joseph Aribido, Jan Kautz, Orazio Gallo, and Stan Birchfield. [Foundationstereo: Zero-shot stereo matching, 2025](#)

Outline

Introduction

Données utilisées

Méthodologie

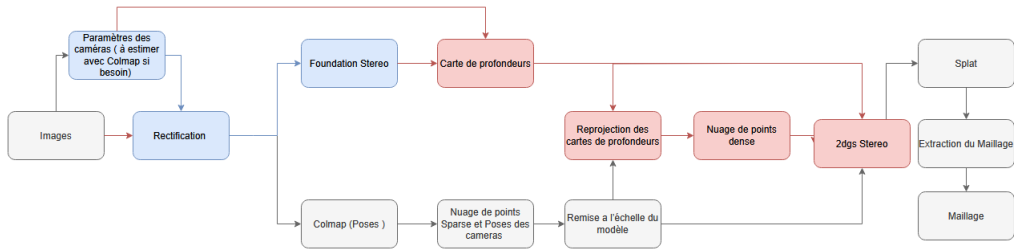
Outils

Implémentation et Résultats

Conclusion



Rectification et Cross-check



Ajout de la rectification et du cross-check



Rectification et Cross-check

Objectif : Garantir des cartes de profondeur précises et compatibles avec le pipeline 2D Gaussian Splatting.

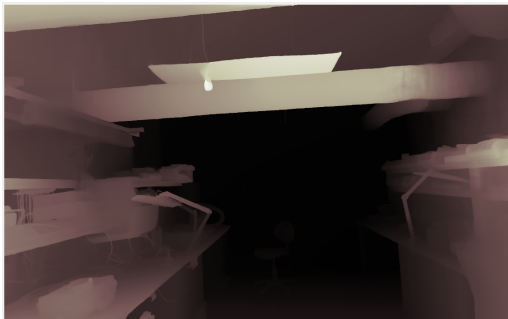
Modifications apportées :

- **Rectification stéréo modifiée :** Recentre systématiquement le point principal via un patch du code OpenCV.
- **Cross-check :** Invalide les pixels incohérents entre disparité gauche et droite.

Impact mesuré sur la séquence de l'atelier électrique :

Sans rectification	Avec rectification
78.28% de pixels invalides	38.66%

Cartes de profondeur sans rectification



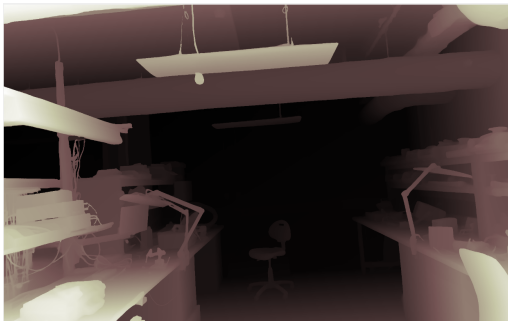
Sans rectification
sans cross-check



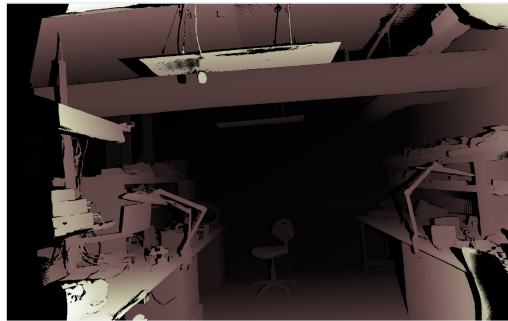
Sans rectification
avec cross-check

Les cartes sont bruitées et présentent un fort taux de pixels invalides, en particulier avec le cross-check.

Cartes de profondeur avec rectification



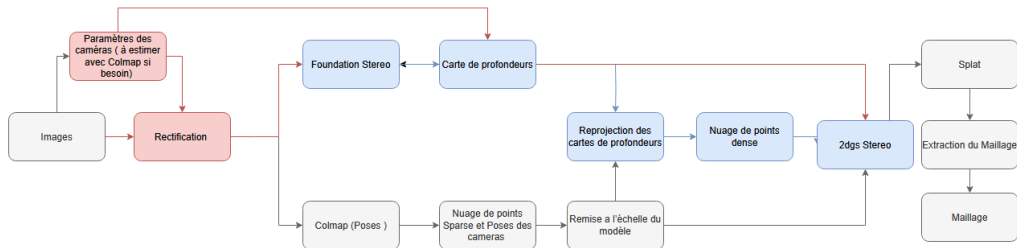
Avec rectification
sans cross-check



Avec rectification
avec cross-check

La rectification améliore la cohérence des cartes et réduit drastiquement les pixels invalides.

Amélioration 1 : Initialisation dense du Gaussian Splatting



Ajout de l'initialisation dense

Résultats en rendus 3D



Sans initialisation dense



Initialisation avec nuage dense

Résultat en maillage



Sans initialisation dense

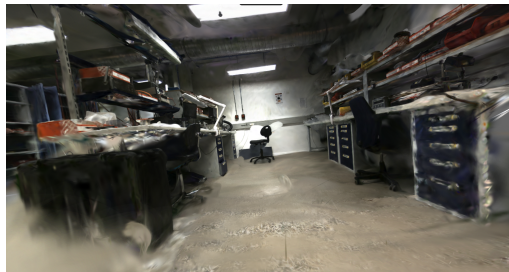


Initialisation avec nuage dense

Résultats en rendus 3D



Sans initialisation dense



Initialisation avec nuage dense

Résultat en maillage



Sans initialisation dense



Initialisation avec nuage dense

Conclusion : Initialisation dense des splats

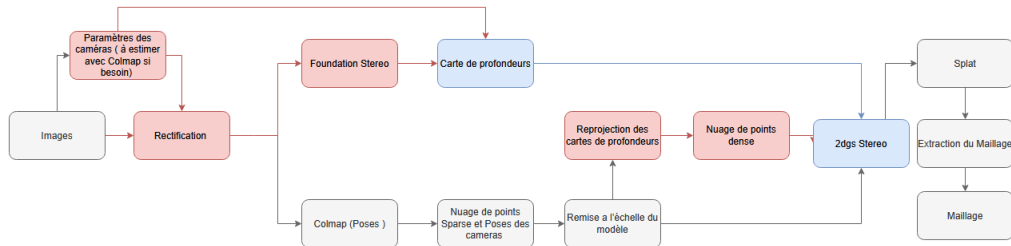
Avantages :

- Meilleure couverture des zones planes (sols, murs)
- Réduction des trous et meilleure densité géométrique initiale
- Moins de dérive au cours de l'optimisation

Limites persistantes :

- Zones peu texturées ou brillantes → dérives résiduelles
- Supervision photométrique seule : non fiable en cas de reflets ou ombres

Amélioration 2 :Supervision par carte de profondeur pondérée



Ajout de la supervision par carte de profondeur

Amélioration 2 :Supervision par carte de profondeur pondérée

Méthode :

- Ajout d'une fonction de coût basée sur la différence entre la profondeur rendue D_r et la profondeur stéréo D_{gt} .
- Pondération par l'inverse du carré de la profondeur stéréo pour compenser la perte naturelle de précision avec la distance :

$$\mathcal{L}_{\text{depth}} = \sum_{i \in \mathcal{P}} \frac{1}{D_{gt}(i)^2} |D_r(i) - D_{gt}(i)|$$

Résultats en rendus 3D



Initialisation avec nuage dense



Supervision par carte de profondeur

Résultat en maillage

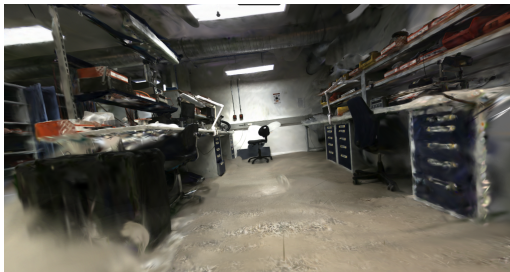


Initialisation avec nuage dense

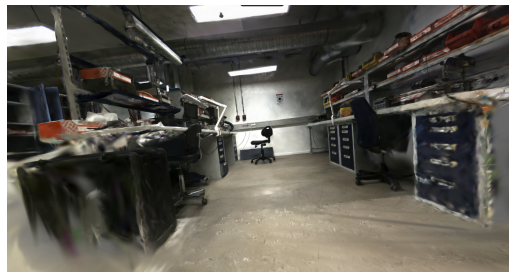


Supervision par carte de profondeur

Résultats en rendus 3D



Initialisation avec nuage dense



Supervision par carte de profondeur

Résultat en maillage



Initialisation avec nuage dense



Supervision par carte de profondeur

Conclusion : Supervision par carte de profondeur pondérée

Avantages :

- Renforce la supervision dans les zones proches, où la géométrie est mieux contrainte.
- Compense la variabilité naturelle de précision des cartes de profondeur stéréo.
- Améliore la stabilité de l'optimisation et la fidélité géométrique des reconstructions.

Outline

Introduction

Données utilisées

Méthodologie

Outils

Implémentation et Résultats

Conclusion



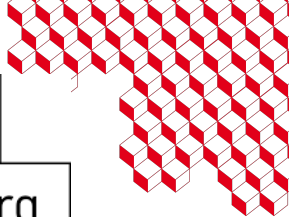
Conclusion : Apports du pipeline proposé



- **Pipeline hybride** combinant apprentissage profond (2D Gaussian Splatting) et supervision géométrique explicite par **cartes de profondeur stéréo**.
- **Rectification stéréo modifiée** pour recentrer le point principal, essentielle pour garantir une projection cohérente dans le modèle 2DGS.
- **Supervision en profondeur** stabilise l'optimisation et améliore la cohérence géométrique, en particulier dans les zones peu texturées ou complexes.
- **Qualité des reconstructions améliorée**, tant sur le plan de la fidélité géométrique que celui du rendu visuel.



- **Comparaison quantitative** à approfondir, dès que les méthodes concurrentes (état de l'art) seront disponibles publiquement.
- **Extension au monoculaire** : génération de paires pseudo-stéréo à partir de vues successives d'une vidéo. Encore limité par la rectification.
- **Rectification augmentée** : créer des paires d'images non adjacentes mais spatialement compatibles pour renforcer la supervision multi-vues.



Merci pour votre attention. Des questions ?

Références I



Johannes Lutz Schönberger and Jan-Michael Frahm.

Structure-from-motion revisited.

In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.



Bowen Wen, Matthew Trepte, Joseph Aribido, Jan Kautz, Orazio Gallo, and Stan Birchfield.

Foundationstereo: Zero-shot stereo matching, 2025.