



Traitement d'Image  
Jérôme DESSEAUX &  
Pierre-Alain EMO &  
Clément GOMME

Detection, segmentation and classification of 3D  
urban objects using mathematical morphology  
and supervised learning

**Rapport**



### **Résumé**

Harum trium sententiarum nulli prorsus assentior. Nec enim illa prima vera est, ut, quem ad modum in se quisque sit, sic in amicum sit animatus. Quam multa enim, quae nostra causa numquam faceremus, facimus causa amicorum ! precari ab indigno, supplicare, tum acerbius in aliquem invehi insectarique vehementius, quae in nostris rebus non satis honeste, in amicorum fiunt honestissime ; multaeque res sunt in quibus de suis commodis viri boni multa detrahunt detrahique patiuntur, ut iis amici potius quam ipsi fruantur.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Objectifs . . . . .	2
1.2	Organisation . . . . .	2
<b>2</b>	<b>Travaux relatifs</b>	<b>3</b>
2.1	Méthodes en 2D . . . . .	3
2.2	Méthodes en 3D . . . . .	3
<b>3</b>	<b>Méthodologie</b>	<b>4</b>
3.1	Images d'élévation . . . . .	4
3.2	Détection . . . . .	6
3.3	Segmentation . . . . .	6
3.3.1	Sol . . . . .	6
3.3.2	Objets . . . . .	6
3.4	Classification . . . . .	6
<b>4</b>	<b>Résultats</b>	<b>7</b>
<b>5</b>	<b>Conclusion</b>	<b>8</b>

# Chapitre 1

## Introduction

### 1.1 Objectifs

Dans ce rapport, nous étudions une approche automatique et robuste pour la détection, segmentation et classification des objets urbains sur des nuages de points 3D via un scanner laser ainsi qu'une imagerie digitale par morphologie mathématique et apprentissage supervisé.

Le traitement est effectué via des images d'élévation, appelées « modèles numériques de terrain » et permet une projection du résultat final sous forme de nuage de points 3D.

### 1.2 Organisation

Nous suivrons la structure de l'article, en reprenant dans un premier temps, l'état de l'existant, puis l'étude des méthodes permettant le traitement d'images 3D, l'étude comparative par rapport à l'existant, et enfin, une conclusion.

# Chapitre 2

## Travaux relatifs

Autant la détection 3D est aboutie, autant la détection des éléments reste un sujet de recherche. Il existe donc deux approches différentes sur ce sujet :

### 2.1 Méthodes en 2D

Projection de l'information en 2D se rapprochant de la méthode ARP, étudiée en ASI [Architecture des Systèmes d'Information] en quatrième année dans le cadre du cours de Data-Mining, pour réduire le temps de traitement. Cette méthode transforme une matrice de données de taille  $\mathbb{R}^N$  en une matrice de données de taille  $\mathbb{R}^d$  où  $d < N$ .

### 2.2 Méthodes en 3D

Le nuage de points 3D qui donne une vision panoramique. Proposé par Zhu et al. (2010) utilisant les données MLS [Multiple Listings Services]<sup>1</sup> pour générer une « zone » panoramique dans laquelle chaque ligne représente le temps d'acquisition de chaque ligne laser, les colonnes représente l'ordre séquentiel de mesure et la valeur du pixel code la distance du capteur au point. Ils proposent une segmentation-classification à l'aide de graphes, SVM [Séparateurs à Vaste Marge]<sup>2</sup> et arbres de décision. Hernandez et Marcotegui (2009b) proposent une méthode projectant les données MLS en « image d'élévation », c'est à dire, une vue de scène. Le sol, et les objets sont segmentés utilisant la transformation morphologique et classifiés en 4 catégories via la méthode SVM :

- Voitures
- Lampes
- Piétons
- Autres

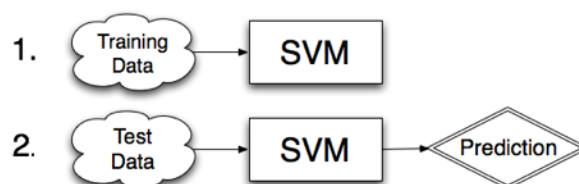


FIGURE 2.1 – Schéma montrant l'utilisation de la méthode SVM dans le cadre de la prédiction via apprentissage supervisé

1. Partage de fichiers utilisés initialement pour le secteur immobilier et la communication entre vendeurs.

2. Les séparateurs à vaste marge sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination et de régression

# Chapitre 3

## Méthodologie

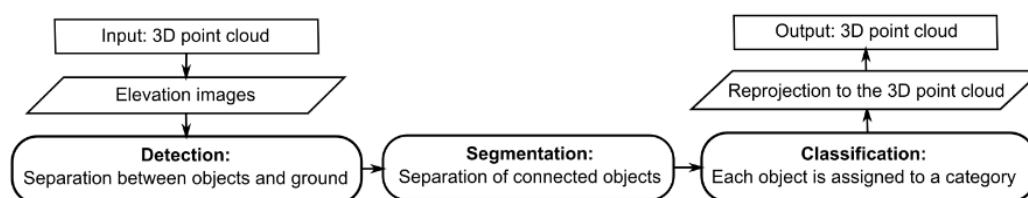


FIGURE 3.1 – Schéma montrant les différentes étapes du processus de modélisation du nuage 3D

**Elévation :** Le nuage de point 3D est projeté sur des images d'élévation.

**Détection :** Un DTM [Digital Terrain Model] ou MNE [Modèle Numérique d'Elévation] en français, est automatiquement créé et les objets hypothétique sont générés comme des discontinuités sur le sol.

**Segmentation :** Les facades sont segmentées automatiquement comme les structures verticales les plus hautes sur les images d'élévation. Ensuite les petites régions isolées sont éliminées, et les objets connectés sont segmentés. Le résultat de l'étape de segmentation crée une image d'étiquettes, où chaque objet segmenté contient un identifiant unique.

**Classification :** Par la suite, de nombreuses caractéristiques géométriques sont calculées et un classement est opéré. A partir de l'étape de classement, une image de classes est créée, contenant une catégorie pour chaque objet segmenté. Le fait d'avoir les étiquettes et les classes dans deux images différentes est utile dans le cas des objets connectés appartenant à la même classe (par exemple un alignement de voiture garées). **Reprojection :** Finalement, les images de classes et d'étiquettes sont reprojettées sur le nuage de points 3D afin d'obtenir le résultat final. Cette étape transforme les images 2D en nuage de points 3D. Chaque pixel du nuage de point prend l'étiquette et la classe de ce pixel.

### 3.1 Images d'élévation

Les images d'élévation sont des structures en  $2.5D$  qui contiennent une information sur l'altitude de chaque pixel. Les nuages de points 3D sont projetés sur des images d'élévation car elles sont des structures pratiques pour visualiser et traiter les données. On peut utiliser tous les outils de traitements d'image existants, en particulier la morphologie mathématique (Mathero, 1975 ; Serra ; 1988 ; Soille, 2003). De plus,



les images peuvent être traitées rapidement, implicitement définissent les relations de voisinage, et requièrent moins de mémoire que les données 3D.

Les images d'élévation sont générées par une **projection orthographique** du nuage de points 3D en utilisant une caméra virtuelle. La caméra virtuelle est placée sur le plan horizontal et croise le point le plus bas dans le nuage de point  $(0, 0, z_{min})$ . Par conséquent, chaque pixel de l'image d'élévation contient l'élévation de la cellule de grille au dessus de  $z_{min}$ . Le seul paramètre libre dans cette projection est la taille du pixel spatial, qui doit être choisi avec attention.

- Si il est trop large, trop de points seront projetés sur le même pixel, faisant ainsi perdre les détails.
- Si il est trop petit, cela implique des problèmes de connectivité et des grandes tailles d'images, ce qui ne justifierait plus l'utilisation d'images d'élévation à la place des nuages de points 3D.

Pour éviter les problèmes de connectivité et la perte d'information, la taille du pixel spatial est choisi selon la résolution du nuage de points. En général, de nombreux points sont projetés sur le même pixel. Par conséquent, quatre images sont définies :

1. Image d'élévation maximum (ou simplement image d'élévation), stocke l'élévation maximale entre tous les points projetés sur le même pixel ;
2. Image d'élévation minimal, stocke l'élévation minimum entre tous les points projetés sur le même pixel ;
3. Image de différence de hauteurs, contient la différence entre les deux images précédentes ;
4. Image d'accumulation, qui stocke le nombre de points projetés sur chaque pixel.

En général, les étapes opératoires sont effectuées sur l'image d'élévation maximum. Les autres images sont utilisées pour supporter certaines décisions lors des analyses ou pour traiter les caractéristiques des objets. Après la création de l'image, une interpolation morphologique est effectuée afin de remplir les trous causés par les occlusions et les lignes de balayage manquantes. Une technique d'interpolation basée sur l'opération morphologique « fill holes » est préférée puisque cette transformation ne crée pas un nouveau maximum de région dans l'image. Un trou est une région sombre (= entourée par des pixels plus clairs), non connectée au bord de l'image. Cette méthode d'interpolation a été proposée par Hernandez and Marcotegui (2009a) et une explication détaillée peut être trouvée dans Serna and Marcotegui, 2013b.

Quand la détection, la segmentation et la classification ont été effectuées, les images sont projetées sur le nuage de points 3D.



## **3.2 Détection**

## **3.3 Segmentation**

### **3.3.1 Sol**

### **3.3.2 Objets**


## **3.4 Classification**

# Chapitre 4

## Résultats

# Chapitre 5

## Conclusion



INSA de Rouen  
Avenue de l'Université - BP 08  
76801 Saint-Etienne-du-Rouvray Cedex  
Tél : 02 32 95 97 79  
Fax : 02 32 95 97 08  
<http://asi.insa-rouen.fr>



*Architecture des Systèmes d'Information*

**INSA**

INSTITUT NATIONAL  
DES SCIENCES  
APPLIQUÉES  
**ROUEN**