ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium “Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future”, 25–30 June 2023, Florence, Italy

ROOM POINT CLOUDS SEGMENTATION: A NEW APPROACH BASED ON OCCUPANCY AND DENSITY IMAGES

C. Gourguechon [1], H. Macher[1], T. Landes[1]

[1] Université de Strasbourg, CNRS, INSA Strasbourg, ICube Laboratory UMR 7357,
Photogrammetry and Geomatics Group, 67000, France
(camille.gourguechon, helene.macher, tania.landes)@insa-strasbourg.fr

KEY WORDS: scan-to-BIM, as-built BIM, room segmentation, 3D point cloud, indoor, image processing, MLS.

ABSTRACT: The majority of buildings are existing and have not been constructed in a BIM process. That is why, the modelling of existing buildings becomes a major issue. Beyond the questions of maintenance or renovation with the promise of reducing the environmental impact, their modelling is of interest for documentation and valorisation. But today, while the acquisition techniques are significantly progressing, with the use of efficient laser scanners, the modelling remains manual and very time consuming. The literature is not empty of proposals to automate the process. Nevertheless, many studies are based on strict architectural hypotheses or restricted to unoccupied buildings free of furniture. This strongly limits their application field. In response to these limitations, this paper presents an innovative method retaining only verticality of walls as assumptions. It is based on occupancy and density image analysis. Tested on a wide variety of buildings, this method is very promising with very few classifications errors. Furthermore, the process is successful with dynamic laser scanning data, in cluttered environments, and applied on buildings with a non-Manhattan-World[1] scheme.

1. INTRODUCTION AND RELATED WORK

While BIM modelling is being generalised for new buildings, the modelling of existing buildings remains a major challenge. The potential applications are wide, from building management and maintenance, to security, accessibility, robotics, navigation, or heritage conservation. With the development, over the last ten years, of static laser scanners (SLS) and more recently of dynamic systems (MLS), significant progress has been made in terms of acquisition. However, modelling phase remains a laborious task. This is still mainly performed manually and is therefore very time consuming and error prone. To automate the modelling, the point clouds are generally segmented into storeys and then into rooms. This paper focuses on the latter: the segmentation of point clouds acquired in storey into rooms. Further such point clouds will be referred as storey point clouds. The final aim is the detection of structural elements based on the resulting segmentation.

Existing studies can be subdivided into two main groups: those that only consider the geometry of point clouds looking for holes formed by walls and those that rely on the trajectory of MLS data, as pointed out by Gourguechon et al. (2021).
Among the first group of authors, Armeni et al. (2016) deduce wall location by searching for "peak-gap-peak" patterns in density histograms computed along the main horizontal axes of buildings. The method is only applicable to scenes following the Manhattan-World scheme. This very restrictive assumption is overcome by many authors seeking rather to group the free spaces into rooms on 2D images or 3D voxel grids. For 2D methods, the transition from the image to the point cloud is performed by considering projected points forming each 2D room area. The challenge becomes to separate grid areas at doorways. In 2D, Macher et al. (2017) avoid this problem by considering only a slice of point clouds above the doors. However, this method cannot be applied to mansard rooms. Other works in the literature mainly propose the use of morphological operators or the evaluation of the potential field in free spaces. The first method separates free spaces by eroding them. The potential field methods are based on distance measurement to the nearest full pixel (or voxel). The room clustering is conducted from the central empty pixels (or voxels) with the longest distances. Once the areas are separated, the rest of the grid is then principally labelled with wavefronts. Bormann et al. (2016) apply them in 2D and fix the erosion width or the distance to select central pixels, from the doors' width and surface criteria of the rooms. Frias et al. (2020) perform something similar in 3D. The erosion is stopped based on a minimum area threshold and the final labelling is conducted with wavefronts. Either in 2D and 3D, these methods tend to gather corridors with adjacent rooms. Jung et al. (2017) also erode images of free areas, with a fixed width but avoid this phenomenon defining their room limits with the skeletonisation of empty pixels surrounding these areas. Bobkov et al. (2017), who work with potential fields, iteratively adjust the distance threshold to select the central pixels. Moreover, these methods are sensitive to room clutter. Thus, they work in 3D to compute their potential fields. Only the voxels lying in the half-space spanning positive z values are considered. Finally, Bormann et al. (2016) propose an alternative method based on geometry. This method relies on a Voronoi diagram, giving the image's skeleton. However, it tends to over-segment the corridors into many rooms.
The second group of authors consider the trajectory of MLS data. In this case, the aim is to cut the trajectory at each door crossing. The points' timestamp allows the segmentation of the point cloud. Diaz-Vilarino et al. (2017) and Zheng et al. (2018) apply this principle. The formers identify the doorways with the point cloud profile along the trajectory. The doorway corresponds to the points with a low average height. Zheng et al. (2018) use scanlines to detect holes in planes. As it is, these methods over-segment some rooms which are visited several times. Therefore, Diaz-Vilarino et al. (2017) propose the

[1] Manhattan-World: when the building follows a Cartesian system with walls, floors, and ceilings perpendicular to each other.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-M-1-2023-93-2023 | © Author(s) 2023. CC BY 4.0 License.

application of an energy function minimisation to merge room segments that were captured several times.
Existing methods have many limitations, especially concerning too strict architectural assumptions, or the presence of clutter. We propose to begin with the general definition of a room. Rooms are defined as enclosed spaces connected to the outside by one or more apertures. Rooms are bounded, horizontally, by solid or glazed walls, and in elevation, by a floor and a ceiling. Access to a room is at a point with a reduced width or height (transom). This is also the way followed by Bobkov et al. (2017). Like all the existing methods, we assume the verticality of walls, whatever on their orientation, or their shape (straight, curved...) or their thickness. No assumptions are made about the ceilings, which can be either horizontal or mansard.
The segmentation of storey point clouds into rooms is particularly challenged by the presence of furniture that looks like walls. Some authors avoid this consideration and only consider unoccupied scenes. But this greatly compromises the application of their solution since interior scenes are only rarely unoccupied. Moreover, the proposed method is designed to deal with MLS data (point cloud and trajectory), increasingly used for indoor scene surveying. Even if the trajectory is not mandatory, it improves the results.

The paper will be structured as follows. A first part presents theoretical elements on which the proposed method is based. Then the method is detailed and, finally, evaluated, with several datasets chosen to reflect a wide variety of environments.

2. THEORETICAL ELEMENTS

This section presents the different theoretical aspects on which the room segmentation method is based. Under the assumption of vertical walls, it will be shown that it is possible to use images to carry out this segmentation. The following paragraphs detail the image types and how walls can be identified with them.

2.1 Projection of the point cloud on a horizontal plane

Under the assumption of vertical walls, the point clouds of solid elements from floor to ceiling (e.g. walls or pillars) form voids, when projected to a horizontal plane. Moreover, points measured on a vertical wall and projected onto a horizontal plane, lead to denser areas relative to other types of geometry.

The method described in this paper was developed considering these two observations. Two important layers are produced: the occupancy and density images of the storey point cloud projected on a horizontal plane.
Occupancy images reflect the presence of points, in the vertical column above each pixel. It is formed by assigning the value 1 to pixels in which points of the cloud are projected, and 0 for the others (Figure 1, black = 0, white = 1).
Density images compare the number of points, in the vertical columns above each pixel. So, the density image is build counting the number of points of the cloud that are projected onto each pixel. To bring the pixel values into the [0,1] range, they are normalised by dividing them by the maximum number of points projected into a pixel.
To avoid holes and the closure of walls, the pixel size is chosen to be less than half of the minimum thickness of a wall and greater than the sampling distance of the point cloud.

2.2 Impact of density heterogeneity across point cloud

Note that input point clouds are irregular and non-uniform. They have some noise and, more importantly, a heterogeneous density of points due to the acquisition. Consequently, empty pixels (pixel value at 0) can be explained, not only by the solid elements from floor to ceiling, but also by a low density of points. It is especially the case for point clouds acquired with MLS. However, these pixels are rather isolated or make very fragmented areas (Figure 1d). This leads us to prefer working with the occupancy image, previously averaged, and its long discontinuities. On the contrary, thin walls, whose thickness is close to twice the pixel size, only create few or no empty pixels. Once again, this is particularly true for noisy point clouds.
In addition, variations in density or wall coverage across the scene make it difficult to directly exploit density values to infer the presence of a wall. As well as for the occupancy image, this leads us to work rather with the discontinuities of the density image. We will refer further to *o-discontinuities* and *d-discontinuities*, respectively, for discontinuities in occupancy images and discontinuities in density images.

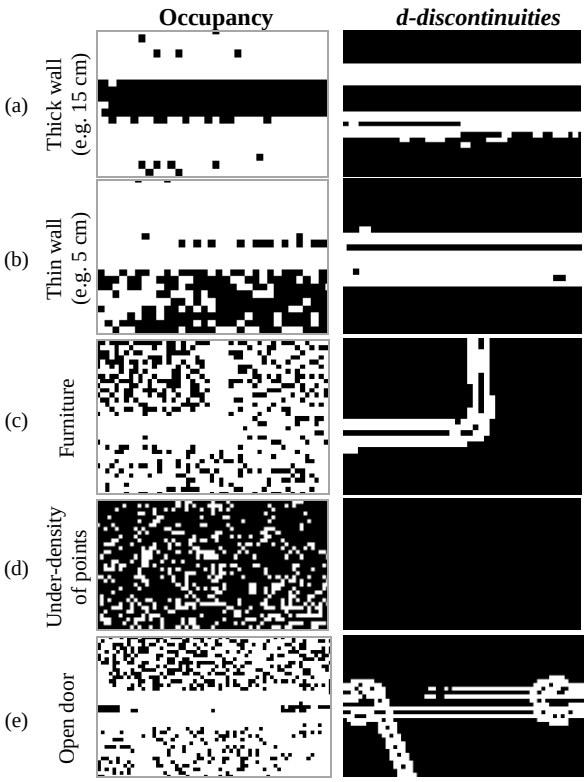

Figure 1. Samples of occupancy and *d-discontinuity* images for various scene elements (1 px. = 2 cm)

The discontinuities images are generated based on the method of Sobel et al., 1973. This method consists of calculating the intensity gradient at each pixel and selecting those with the highest gradient values. In practice, a convolution operation is performed on an image with two 3x3 matrices, which allows an approximation in each pixel of the horizontal and vertical derivative. The gradient norm is then deduced by combining these horizontal and vertical gradients.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-M-1-2023-93-2023 | © Author(s) 2023. CC BY 4.0 License.

2.3 Distinction between walls and furniture in images

The scenes studied are not empty of furniture and equipment. In addition to the walls, the furniture (e.g.: cabinets, privacy screens...) can also give rise to discontinuities in density.
Figure 1 shows the appearance of various scene elements in the occupancy and *d-discontinuity* images. A thick wall, about 15 cm in reality, (see Figure 1a) appears with a *"d-discontinuity"* / *"wide void"* / *"d-discontinuity"* pattern. A thin wall, about 5 cm, (see Figure 1b) or very slightly sloping wall appears with a *"d-discontinuity"* / *"d-discontinuity"* pattern. And finally, a piece of furniture adjacent to a wall appears as *"filled pixels"* / *"d-discontinuity"* / *"filled pixels"*. Lastly, note that at a door location, points on the ground are measured and lead to filled pixels (pixel value at 1) in the occupancy images. When doors are open, *d-discontinuities* may still appear, when the transom is solid. On the contrary, without a transom, nor a closed door, or in the presence of a glazed transom, no *d-discontinuity* appears because of a lack of points. Finally, it appears that walls are distinguishable from furniture and other vertical items by the contiguous presence of *d-discontinuities* with long empty pixel areas or consecutive *d-discontinuities*. These properties will be the basis of our approach.

3. DEVELOPED METHOD

The developed segmentation method of storey point clouds is based on the creation of 2D planimetric masks for each room. Room point clouds are then deduced by selecting the points of the input point cloud projected vertically on each mask. The challenge is to define the planimetric boundaries of the rooms. These limits must be defined in the wall position separating the rooms. The aim is to isolate the image of the walls. This is not an easy task in cluttered environments where furniture can look like walls.

The image of the walls is first initialised based on the analysis of the pattern in occupancy and density images and the images of their discontinuities. This approximate image of the walls is then used to initialise the room masks using morphological operators. Finally, as the space is then over-segmented, regions are gathered when they correspond to the same rooms and their boundaries are refined. The developed method is summarised in Figure 2.

3.1 Initialisation of a first approximate image of the walls

First, an approximate image of the walls is initialised. For this purpose, the occupancy and density images and the images of their discontinuities are used. Previous section describes their creation from the storey point cloud projected to a horizontal plane.

3.1.1 Approximative image of the walls from empty pixels of the occupancy image: We will first use the following property: the point clouds of solid vertical walls, when projected onto a horizontal plane, form voids (pixel value at 0) in the occupancy image.

First, the broadest region of empty pixels in the occupancy image is selected. This corresponds to the outdoors. Not all empty pixels are considered because of the point cloud heterogeneity. Thus, to select the other empty pixel regions of interest, several criteria are introduced. To be included in the first image of the walls, small regions must meet the following criteria:

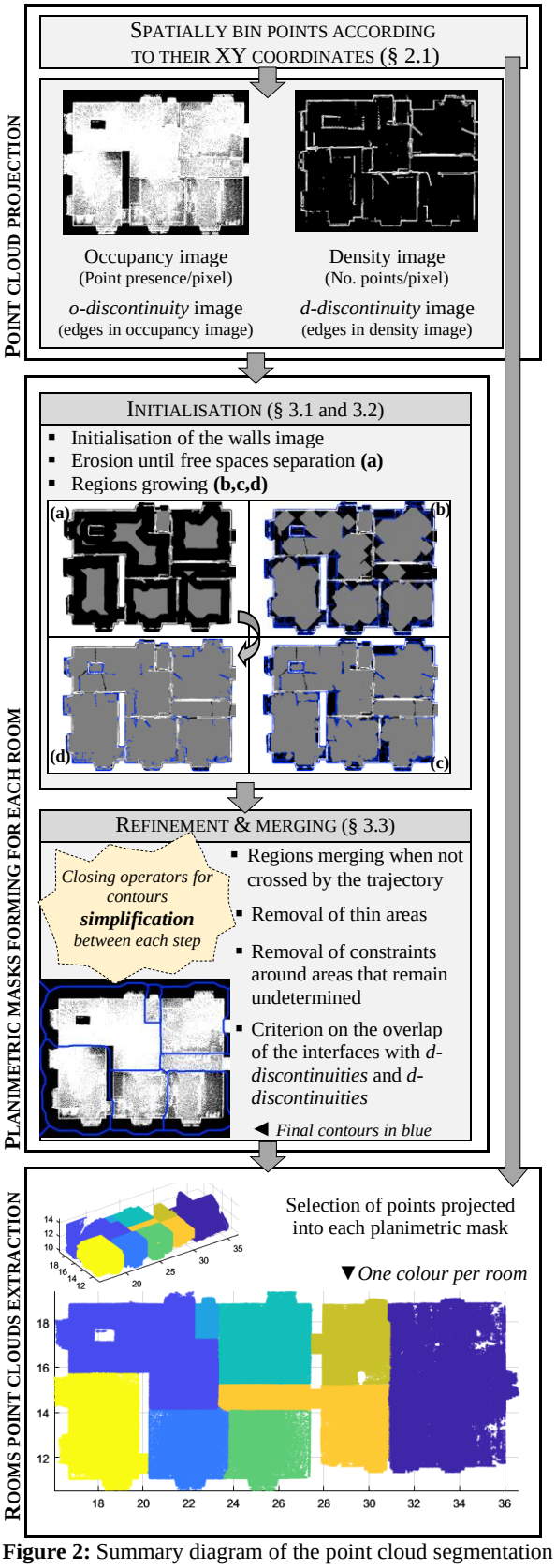

Figure 2: Summary diagram of the point cloud segmentation approach from storey to rooms

- The main axis of the region should be longer than 0.5 m. This threshold has been fixed to keep columns and wall axes and allow eliminating isolated pixels or small groups of pixels.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-M-1-2023-93-2023 | © Author(s) 2023. CC BY 4.0 License.

- The region should be almost hole-free, i.e. the image of the area should cover at least 95 % of the image of the hole-free area. Moreover, the region must be "solid", i.e. the image of the area must cover at least 98 % of the area image that has undergone a morphological closure of 0.20 m and the removal of its holes. These two criteria, dealing with the presence of holes in the images, are intended to eliminate broader but fragmented areas.

These selected empty pixel regions form the first approximate image of the walls (white pixels, in Figure 3). Later, this image will be referred to as $\mathbf{I}_{w\text{-}1}$. The eliminated empty pixel regions appear in red in Figure 3.

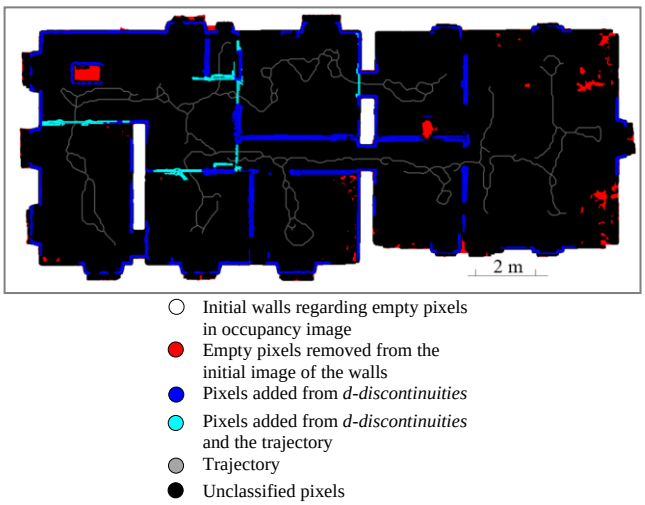

Figure 3. Initialisation of the first approximate image of the walls

3.1.2 Complement from density discontinuities: Besides empty pixels in the occupancy image, walls also cause *d-discontinuities*. The approximate image of the walls will be supplemented using this property. However, as shown in paragraph 2.3, it is not only walls that form *d-discontinuities* but also furniture. In contrast to furniture (Figure 1c), when they are thick enough, the walls also produce empty pixels in the occupancy image, bordered by two *d-discontinuities* (Figure 1a). At the very least, if the walls are thin, two consecutive discontinuities are formed with no empty pixels in between (Figure 1b).

According to these observations, the approximate image of the walls is supplemented with the *d-discontinuities* meeting the following criteria:

- The *d-discontinuity* covers more than 0.5 m long.
- The *d-discontinuity* is in the immediate neighbourhood (< 0.05 m) of consecutive *d-discontinuities* or long *o-discontinuities* (≥ 0.5 m).

As above, the length thresholds are set to avoid density or occupancy discontinuities due to a lack of points. The proximity threshold, set at 0.05 m, is justified by the immediate relationship, in the case of a wall, between consecutive *d-discontinuities* or long *o-discontinuities* and *d-discontinuities*.
These additions to the approximate walls' image are shown in blue in Figure 3.

3.1.3 Additional processing based on the trajectory: For datasets with a known trajectory, the door passing allows estimating the door's localisation. It helps therefore to complement the initial image of the walls. The door localisation is deduced from the intersection of the trajectory with the *d-discontinuity* image, filtered to keep only long segments (≥ 0.2 m). It should be noticed that doors with glazed transoms are not detected, because they do not produce significant d-discontinuities.

These door points are extended with the contiguous *d-discontinuities* within a radius of 0.2 m. This allows an estimation of the orientation of walls containing the doors.
Only the door points located within 3 m of two walls (according to the approximate image of the walls created so far), following the orientation derived previously, are retained. This filter makes it possible to eliminate a large part of the points wrongly detected as doorway points. This can happen, for example, when the trajectory passes over a piece of furniture. The approximate image of the walls is supplemented by the addition of contiguous *d-discontinuities* to the filtered door points, within a radius of 3 m and following the orientation defined previously (in cyan in Figure 3).

3.2 Rooms image initialisation

At this stage, an approximate image of the walls is available. In this binary image, the pixels with a value of 0 (in black on Figure 3) correspond to the free spaces of the scene, namely the rooms and the walls' openings. The aim is now to segment these free spaces to lead to the segmentation of individual rooms. Two main steps are followed: (1) kernels for each room are defined by successive erosions of the free spaces, (2) then a progressive region growing algorithm is applied on these kernels to produce a first approximate image of the rooms. Both steps are described in the next paragraphs and illustrated in Figure 2. Figure 4 summarises the algorithms in pseudo-code.

3.2.1 Rooms kernel definition: The first step consists of initialising the regions for each room. The goal is to isolate a zone of contiguous pixels for each room, which we refer to as "region kernels".
The starting point is the image of free spaces, i.e. the negative of the initial image of the walls (Figure 4, line 2). This is simplified by applying a closure operator to the image of the walls, to avoid the creation of very small areas. The closure distance, set at 0.5 m, was chosen to be long enough to eliminate all empty pixels inside the approximated walls, but small enough to allow the initialisation of regions in narrow rooms.
The progressive erosion of the simplified free space image allows the separation of kernels for each room (Figure 2a). This erosion is carried out until all spaces are separated or until a maximum erosion size is reached (Figure 4, lines 3 to 5). More precisely, a space is separated from the others (and forms a kernel), when its image reaches a minimum area threshold (4 m) or a minimum width threshold (0.8 m).

3.2.2 Region growing algorithm: The second phase extends the region kernels to label the whole image. The unknown areas are labelled from the region kernels. A region growing algorithm is used (Figure 4, lines 7 to 21). It is limited by a set of pixels that cannot be crossed, which are called "constraint pixels". The labelling is applied iteratively. The constraint image evolves with each iteration. The constraints are initially large and then progressively reduced to allow the labelling of all unknown areas. By beginning with important constraints and then reducing them, it is possible to avoid that a region of a small room, which requires less erosion to isolate it than a large one, " pours into another ".

The constraint pixels are based, in the first step, on the approximate image of the walls, combined with the long *o-discontinuities* (≥ 0.20 m) and the *d-discontinuities* occurring in the unknown areas after the free spaces partitioning. The application of smaller and smaller morphological closure operators (1.00 m, 0.50 m, 0.20 m, 0.10 m, 0.05 m) gives the successive images of increasingly smaller constraints. From the 0.20 m width closure, the empty pixel areas in the remaining unknown areas are added to the constraints. The occupancy and density discontinuities used as a first approximation of the constraints are gradually eliminated. The last constraint image is deduced only by the approximated image of the walls and the empty pixel areas with a 0.05 m width closure. We obtain a first approximation of the rooms segmentation in 2D.

```
% Rooms kernel definition
kernels = negative of (closure 1rst wall image)
While some kernels are too wide (>0.8 m) & extended (>4 m² )
    kernels = small kernels + eroded oversized kernels;
End

% Region growing algorithm
UnknownArea = negative of kernels − the 1rst wall image;
Im_discInUArea = (o-discontinuities ≥ 0.20 m +
                  d-discontinuities) overlapping UnknownArea;
WidthList = [1.00, 0.50, 0.20, 0.10, 0.05];
For i=1: length(WidthList)
    % Constraints evolution
    Constraints = closure of Im_discInUArea image with
                  WidthList(i) width − kernels;
    % Room kernels growing
    While some kernels evolve between two iterations
        kernels = dilation of kernels − areas encountering
                  constraints;
    end
end
```

Figure 4. Rooms image initialisation as a pseudo-algorithm

3.3 Mask refinement

At this stage some rooms are over-segmented, especially rooms that are narrow and long, such as corridors, or those that are very cluttered. Some areas may also remain undefined, and regions are not contiguous with each other, forming large gaps between rooms. The challenge is then to group the regions corresponding to the same room and to refine their contours so that the limits between rooms correspond to wall locations. Three criteria are considered for that purpose: (1) the crossing of the trajectory with each region, for datasets with an available trajectory, (2) the width of the regions, (3) and for the remaining limits, their overlap with wide empty pixel areas, d-discontinuities and occupation discontinuities.

Between each refinement step, a region simplification is performed. It allows the elimination of the grooves in areas mainly appearing due to open doors or furniture. So, this is particularly relevant in the first stages. This simplification is carried out with a morphological closure operator on each region. The approximated image of the walls is then redefined between each step, in the neighbourhood of the simplified regions.

3.3.1 Grouping areas on a trajectory criterion: On the assumption that the MLS necessarily visits all the rooms during the acquisition, a criterion based on the trajectory is introduced to merge over-segmented areas.
Thus, when the trajectory is known, areas that are not crossed by the latter are eliminated. Then, they are labelled by means of a region growing. As for the rooms image initialisation, this region growing is carried out iteratively by gradually reducing the constraints.

3.3.2 Removing narrow rooms and undefined areas: Based on the principle that rooms cannot be narrower than a certain threshold, the narrow areas of the regions defined so far and simplified are eliminated. To do this, a morphological opening operator is used on each region. The removed narrow areas are relabelled by a region growing.
Similarly, the wide (≥ 0.2 m) and remaining undefined areas (black pixels with a 0 value) are labelled by region growing. However, as these have so far remained undefined, due to the presence of an approached wall, constraints are eliminated in their neighbourhood. It allows surrounding areas to grow.

3.3.3 Refining room limits: Before evaluating the limits between rooms and defining the final masks, the room contours are refined. Two simplifications of each region obtained so far are then carried out, followed by region growing constrained by the approximate image of the walls in the neighbourhood of these regions. A final non-constrained region growing completes the process. This allows obtaining thin limits between regions at wall location.

3.3.4 Removing the last non-relevant limits: Since the image is still over-segmented, a final test aims to evaluate the limits between regions. Those that are not significant, i.e. that don't really reflect a room change are eliminated. The elimination of the border, then leads to the merging of the areas on either side. To achieve this evaluation, three criteria are considered. A limit is removed if it does not meet any of these criteria.
The first criterion considers the overlap of the limits with the narrow areas (< 0.80 m) of the $\mathbf{I_{w-1}}$ image. These areas correspond to the thick walls and pillars of the storey. The test involves counting the number of pixels of the limits overlapping the narrow areas of the $\mathbf{I_{w-1}}$ image. Limits whose ratio between the number of overlapping pixels and the total number of pixels exceeds a threshold of 50 % are retained. This criterion is particularly relevant for limits on thick walls.
The second and third criteria consider the overlap of limits with *d-discontinuities*, *o-discontinuities*, and wide empty pixel areas. To measure this overlap, these two criteria consider the images of the limits reduced to the parts out of the $\mathbf{I_{w-1}}$ image. As before, a ratio is calculated, this time between the number of pixels overlapping and the total length of the reduced limits. For the second criterion, this ratio is compared to a threshold set at 75 %. The third criterion selects limits with a 50 % overlap ratio. and evaluates whether they are straight lines or not. For this purpose, a limit is considered as straight when the second principal axis of the enclosing ellipse is smaller than 0.20 m.

4. RESULTS AND DISCUSSIONS

To evaluate the developed method, it has been applied on several datasets. These datasets were chosen to reflect a wide variety of environments. Particular attention has been paid to demonstrate the strengths and limitations of the approach through several indicators and figures.

4.1 Datasets

Our method of point cloud segmentation was tested on five datasets acquired in three buildings with MLS scanners. The five datasets are point clouds of storey, accompanied by the corresponding acquisition trajectory. They were previously

spatially subsampled with a regular 1 cm grid to homogenise the point cloud (Figure 5).

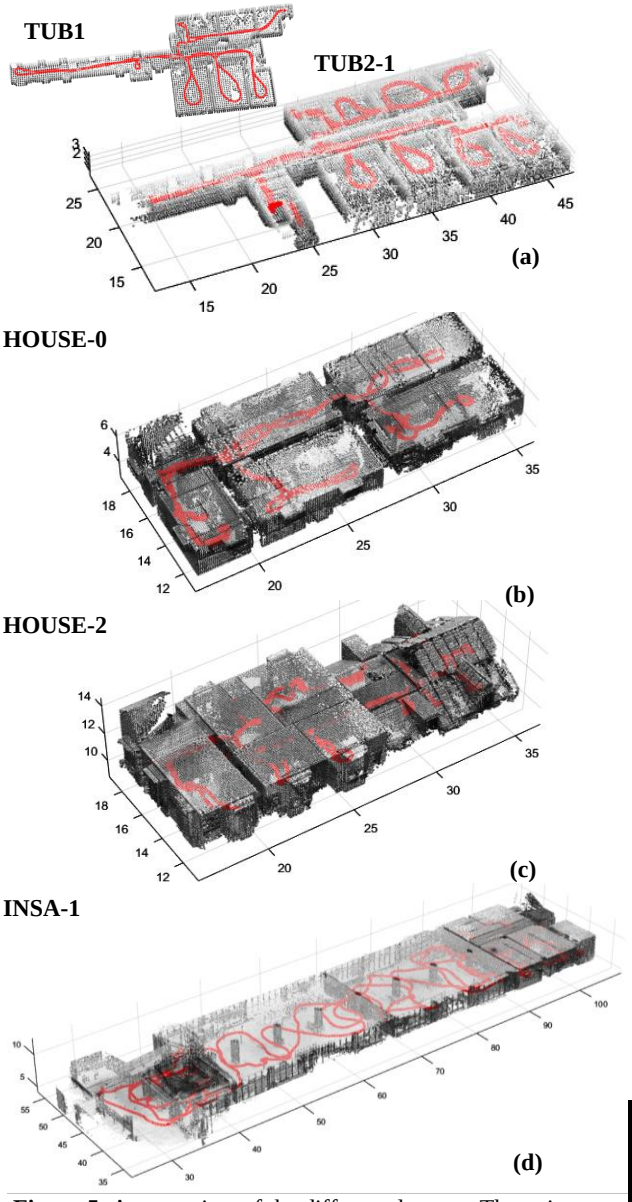

Figure 5. An overview of the different datasets. The point clouds (grey level) have been downsampled to highlight the trajectories (red).

TUB1 and TUB2-1. These datasets are from the ISPRS Benchmark (Khoshelham et al., 2017). They were acquired in a building of the Technische Universität Braunschweig, in Germany, respectively, with a Viametris iMS3D system, and Geoslam Zeb-Revo RT (Figure 5a). They correspond mainly to the same storey. TUB1 has 11 rooms and TUB2-1 has 16 rooms including the 11 rooms from TUB1. Rooms are separated by walls of different widths and by single or double doors, both open and closed during the acquisition. These datasets differ from the others by the very small amount of furniture.

HOUSE-0 and HOUSE-2. These two datasets correspond to two floors of a single-family house, built in the 1920s. They were acquired with a Geoslam Zeb-Revo RT. The house was occupied at the time of the acquisition, hence the presence of furniture in the point clouds. HOUSE-0 (Figure 5b) is the cellar of this house with 6 rooms, separated by very thick walls. It is a particularly cluttered set of data with shelves along the walls and a large tank. HOUSE-2 (Figure 5c) corresponds to the 1st floor of the house, with 10 rooms. This storey presents both very thin and very thick walls. It also has 3 attic rooms, including storage rooms that are very cluttered.

INSA-1. The last dataset was acquired in a building of the National Institute of Applied Sciences of Strasbourg, in France, with a Geoslam Zeb-Revo RT (Figure 5d). It is an administrative building from the 1950s and is composed of 13 rooms. This dataset was chosen to challenge the proposed method. Indeed, it has the particularity to present three large halls and several storage rooms. The buildings were occupied during acquisition, and therefore furniture was present. The rooms are separated both by solid and glazed walls. Some walls are not straight or directed along the main axes of the building. Doors between rooms can be opened or closed, single or double, crossed or not by the trajectory and surmounted by a solid or glass transom. Finally, the dataset presents two staircases, under which three rooms pass.

4.2 Results

The results are presented in Table 1. For each dataset, the number of rooms that were over- or under-segmented is calculated, as well as the number of points concerned. In addition to over- or under-segmentation, some points may also be locally misclassified (e.g. face of a wall assigned to the neighbouring room). This is referred to as a "slight shift" in room contours. These "slight shifts" have also been quantified by measuring the quantity of points involved. Note that the default parameters have been kept for all the datasets.

Table 1: Results of the segmentation into rooms applied on five datasets (*under-segmentation ratio deduced). The lower the values, the better it is.

	Over-segmentation		Under-segmentation		Slight boundary shift	TOTAL point misclassification
	Nb. Rooms	Amount misclass. points (%)	Nb. Rooms	Amount misclass. points (%)	Amount misclass. points (%)	Amount misclass. points (%)
TUB1	3/11	5.3*	1/11	6.4	0.2	11,9
TUB2-1	2/16	3.6	0/16	0,0	0.4	4.0
HOUSE-0	0/6	0,0	0/6	0,0	1,1	1,1
HOUSE-2	0/10	0,0	1/10	3.4	3.6	7.0
INSA-1	5/13	5.2	0/13	0,0	0.4	5.6
ALL (EXCEPT TUB1)	7/45	3.3	1/45	0.6	1.1	5.0

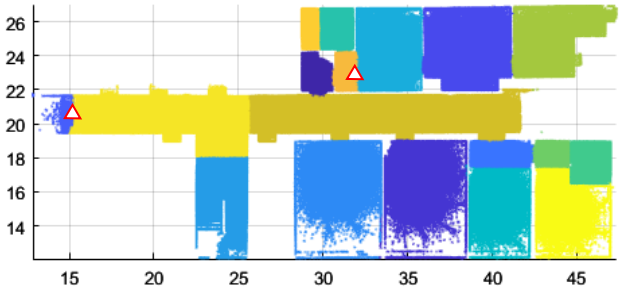

Figure 6. TUB2-1 point cloud segmented into rooms. Red triangles localise over-segmentations.

Let us first consider TUB1 and TUB2-1 corresponding to the same scene but acquired with different mobile systems. With our method applied on TUB2-1, only two rooms are over-

segmented, representing 3.6 % of the points (Table 1). The corridor, at the very left of the image, is over-segmented at the location of a piece of furniture (Figure 6). The entrance to the room, at the top centre, is over-segmented due to a ceiling beam. Similarly, over-segmentations appear on TUB1, also at the entrance to the room, at the top centre of the image, as well as in the room at the top right of the image, due to the presence of pipes (Figure 7). Moreover, the corridor is divided into two parts at the wrong place. The phenomenon is more frequent than for TUB2-1, with 11.7 % of points corresponding to over- or under-segmentation (Table 1). At first glance, these poor results are surprising, since the original scene is the same.

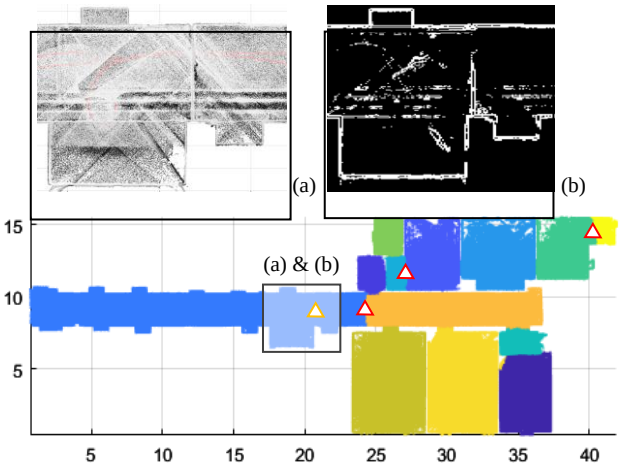

Figure 7. TUB1 point cloud segmented into rooms. Red and yellow triangles localise, respectively, over-segmentations and under-segmentation. (a) and (b) focus on an area with rough density variations. (a) is a zoom on point cloud. (b) is the *d-discontinuity* image of the area.

If we look closely, TUB1 and TUB2-1 highlight the impact of the quality of the point cloud and, in particular, a regular sampling. Indeed, it appears quite clearly that the TUB1 point cloud presents a very heterogeneous point density with abrupt variations, even after having applied a regular subsampling (Figure 7, a and b details). This appears to be due to the operator's walking speed, in relation to the Viametris iMS3D acquisition system used for TUB1. This dataset is the only one acquired with this system and presenting this type of problem. This very irregular density creates disturbing *d-discontinuities* that compromise the application of our method from the first step, and the walls’ image generation.

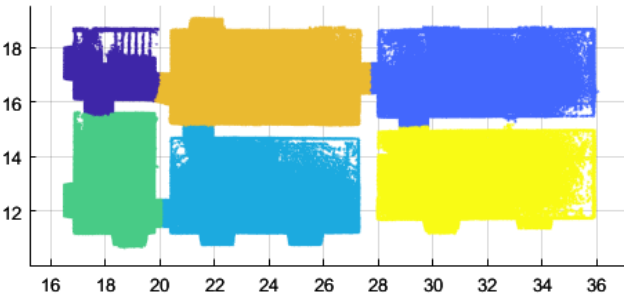

Figure 8. HOUSE-0 point cloud segmented into rooms.

Regarding the 1920s house, although very cluttered and with a more complex architecture (significant difference in wall size and attic presence), only one room was sub-segmented, out of the 16 rooms constituting the dataset (Figure 9a). The corridor and the adjacent room in HOUSE-2 were not separated. This is because there is very little shrinking at the end of the corridor and no door. As for the "slight shifts", these are minimal for HOUSE-0 with 1.1 % measured (Table 1). This mainly concerns the interface between the purple and orange rooms at the top left in Figure 8, where the transition from one room to the other is not clear. These "slight shifts" represent 3.6 % of the points for HOUSE-2 (Table 1). They can be particularly, observed where the walls are very thin (approx. 5 cm), as in Figure 9a. This can also be observed in the presence of inclined beams as in Figure 9b. This last case shows that our method does not consider alcoves or overhangs between rooms.

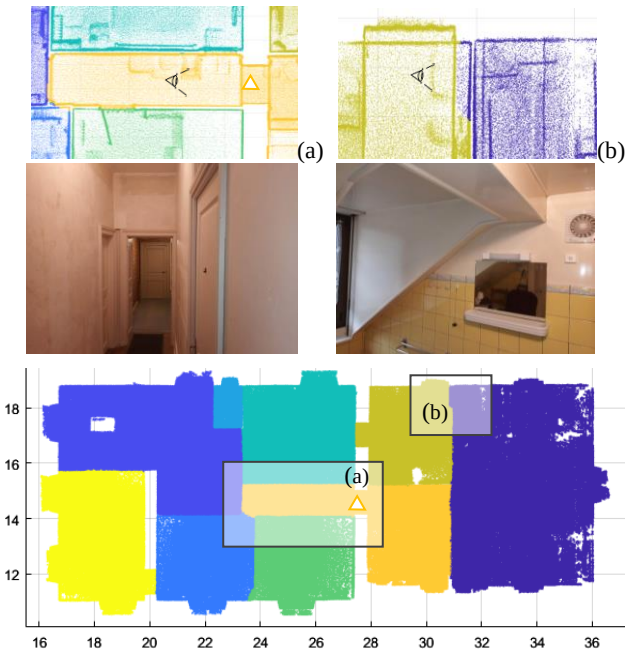

Figure 9. HOUSE-2 point cloud segmented into rooms. Yellow triangles localise under-segmentations. (a) focuses on the under-segmentation locus, (b) on "slight shifts".

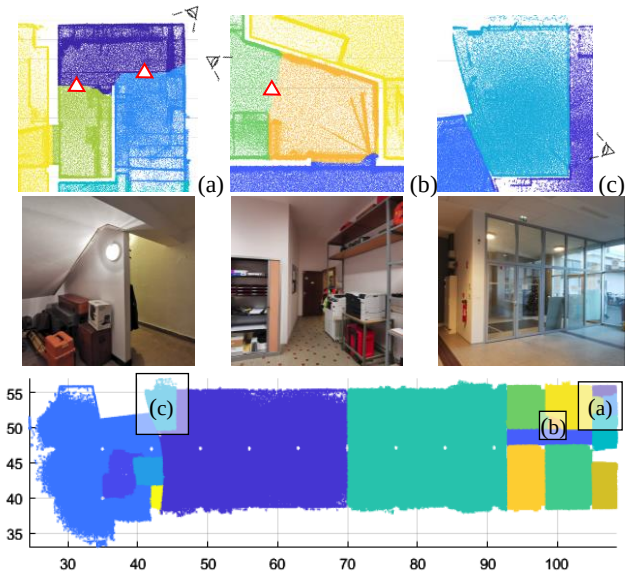

Figure 10. INSA-0 point cloud segmented into rooms. Red triangles localise over-segmentations. (a) and (b) focus on over-segmentation locus, and (c) on glazed walls.

Lastly, applied to the INSA-0 dataset, five over-segmentations are observed (Figure 10). Four of them are formed under stairs

(e.g. Figure 10a). The vertical walls of the stairwells thus favour over-segmentation. Although the place was occupied, only one over-segmentation is related to the presence of a piece of furniture (Figure 10b), this time again occurring in a narrow area. No sub-segmentation has occurred. In particular, the glazed entrance (Figure 10c) was well detected despite a lack of points on its boundaries compared to solid walls. Finally, there were very few "slight offsets" to report. The rounded shape of the walls did not affect the segmentation.

Finally, we consider all the datasets that meet the initial hypotheses. So, TUB1 is excluded since it presents brutal variations of data density. All the datasets together represent 45 theoretical rooms. Seven over-segmentations, and one under-segmentation have occurred. This represents 3.9 % of the total points (Table 1). Four over-segmentations are due to the stairs. Considering in addition the slight offsets, 95 % of the point clouds are correctly segmented.

5. CONCLUSION

In this paper, a new method for segmentation of storey point clouds into rooms is proposed. It is based on the use of two maps deduced from datasets: *d-discontinuity* and *o-discontinuity* images. This method has been developed to overcome many assumptions found in the literature. Only the assumption of vertical walls is retained.

Tests carried out on a set of buildings, both old and new, with varying architecture and degrees of clutter validate the overall segmentation approach. It is particularly well adapted to the presence of furniture, except sometimes in narrow areas. Long corridors are not over-segmented thanks to the overlap criteria. The tests have been performed on very different environments, using the same default parameters. The satisfying results led us to conclude that the threshold values are adapted for a large variety of scenes.
It seems promising for cultural heritage buildings. Although specific tests would be required on castles, temples, or military barracks to confirm its applicability.

However, the tests reveal some weaknesses, particularly in the presence of rooms under stairs. In these cases, over-segmentation occurs. The condition of a point cloud without too many rough variations is also unavoidable. We can already suggest some ways to improve the process and overcome these limitations. For example, the development of a filter on the *d-discontinuity* images might improve the problems of density changes. This filter would be based on the vertical distribution of the points.

Another advantage of the proposed methodology is that it works successfully on MLS data. It must be noticed that the trajectory is not mandatory for the segmentation process, it is only used as an add-on. In future experiments, the methodology will be assessed by comparing the produced results with a reference dataset from SLS data.

REFERENCES

Armeni, I., Sener, O., Zamir, A.R., Jiang, H., Brilakis, I., Fischer, M., Savarese, S., 2016. 3D semantic parsing of large-scale indoor spaces. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem, 1534–1543. https://doi.org/10.1109/CVPR.2016.170

Bobkov, D., Kiechle, M., Hilsenbeck, S., & Steinbach, E. (2017). Room segmentation in 3D point clouds using anisotropic potential fields. Proceedings - IEEE International Conference on Multimedia and Expo, July, 727–732. https://doi.org/10.1109/ICME.2017.8019484

Bormann, R., Jordan, F., Li, W., Hampp, J., Hägele, M., 2016. Room segmentation: Survey, implementation, and analysis. Proc. - IEEE Int. Conf. Robot. Autom. 2016-June, 1019–1026. https://doi.org/10.1109/ICRA.2016.7487234

Díaz-Vilariño, L., Verbree, E., Zlatanova, S., Diakité, A., 2017. Indoor modelling from SLAM-based laser scanner: Door detection to envelope reconstruction. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42, 345–352. https://doi.org/10.5194/isprs-archives-XLII-2-W7-345-2017

Frías, E., Balado, J., Díaz-Vilariño, L., Lorenzo, H., 2020. Point Cloud Room segmentation based on indoor spaces and 3D mathematical morphology. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 44, 49–55. https://doi.org/10.5194/isprs-archives-XLIV-4-W1-2020-49-2020

Gourguechon, C., Macher, H., & Landes, T. (2022). Automation of as-built BIM creation from point cloud: an overview of research works focused on indoor environment. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 43(B2-2022), 193–200. https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-193-2022

Jung, J., Stachniss, C., & Kim, C. (2017). Automatic room segmentation of 3D laser data using morphological processing. ISPRS International Journal of Geo-Information, 6(7). https://doi.org/10.3390/ijgi6070206

Khoshelham, K., Vilariño, L. D., Peter, M., Kang, Z., & Acharya, D. (2017). The ISPRS benchmark on indoor modelling. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 42(2W7), 367–372. https://doi.org/10.5194/isprs-archives-XLII-2-W7-367-2017

Macher, H., Landes, T., & Grussenmeyer, P. (2017). From point clouds to building information models: 3D semi-automatic reconstruction of indoors of existing buildings. Applied Sciences (Switzerland), 7(10), 1–30. https://doi.org/10.3390/app7101030

Sobel, I., & Feldman, G. (1973). A 3x3 isotropic gradient operator for image processing. In Hart, P. E. & Duda R. O. Pattern Classification and Scene Analysis, January 1973, 271–272.

Zheng, Y., Peter, M., Zhong, R., Elberink, S.O., Zhou, Q., 2018. Space subdivision in indoor mobile laser scanning point clouds based on scanline analysis. Sensors (Switzerland) 18, 1–20. https://doi.org/10.3390/s18061838