



# Indoor localisation and location tracking in indoor facilities based on LiDAR point clouds and images of the ceilings

Ioannis Dardavesis<sup>1</sup>, Edward Verbree<sup>1</sup>, and Azarakhsh Rafiee<sup>1</sup>

<sup>1</sup> GIS-technology, Section Digital Technologies, Department Architectural Engineering + Technology, Faculty of Architecture and the Built Environment, Delft University of Technology

Correspondence: Ioannis Dardavesis ([jdardave@hotmail.com](mailto:jdardave@hotmail.com)), this paper is based on a MSc-thesis (Dardavesis, 2022)

**Abstract.** Localisation and navigation technologies have vastly evolved during the last years, facilitating users' guidance in various environments. Unlike outdoor environments where GNSS comprises a universal solution, in indoor environments various localisation techniques have been used, each one with its drawbacks. Thus, this research investigates the reliability of the ceilings towards indoor localisation, by using components that are included in a simple mobile device. The choice of ceilings lies in their advantages, which include the incorporation of various characteristic components, as well as the absence of obstacles between them and the sensor. Indoor localisation is achieved based on LiDAR point clouds and images from RGB sensors of mobile devices. Additionally, this research involves location tracking of different users, to discover different movement patterns in an indoor facility. The proposed methodology revealed the robustness of the Coloured ICP algorithm for in-door localisation based on point clouds, both in terms of time efficiency and quality, while the combination of the SURF feature detector and SIFT descriptor provides the optimal indoor localisation results with image data. The proposed pipeline revealed encouraging results for use in emergencies, based on static data acquisition of a user, while it is also suitable for dynamic applications, in case a sensor is mounted on an automated device for indoor mapping operations. The captured point clouds of the ceilings can also be used as a reference to CAD and BIM models, to help the modelling of the existing utilities and their components in an indoor facility.

**Keywords.** indoor localisation; location tracking; LiDAR; iPad 12 pro; point cloud registration; feature matching

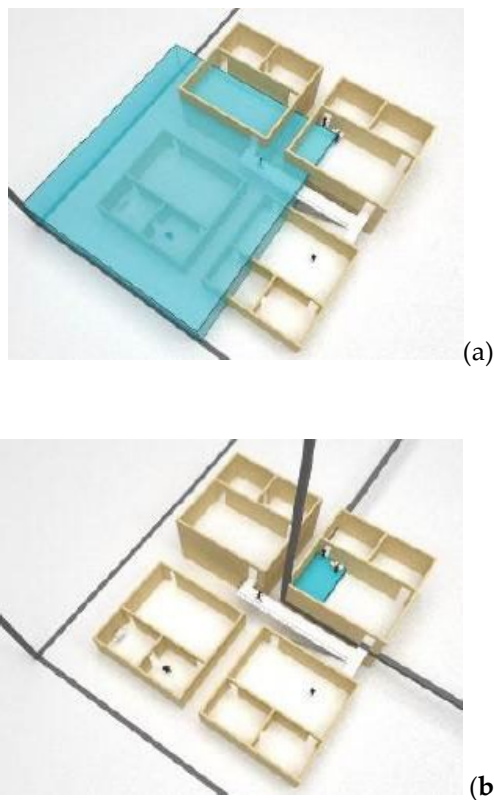
## 1. Introduction

### 1.1 Problem statement

Nowadays, the evolution of localisation and navigation technologies is vast, aiding towards facilitating users' guidance in various environments. Outdoor positioning can be easily achieved, with the widely used GNSS (Global Navigation Satellite Systems), which are included in every person's mobile device. However, the presence of high-rise structures in dense urban environments where there is no line-of-sight between the satellites and the receiver leads to signal attenuation. Additionally, there is poor reception in indoor environments, significantly degrading the performance of GNSS. Therefore, alternative ways of positioning and localisation respectively, need to be explored.

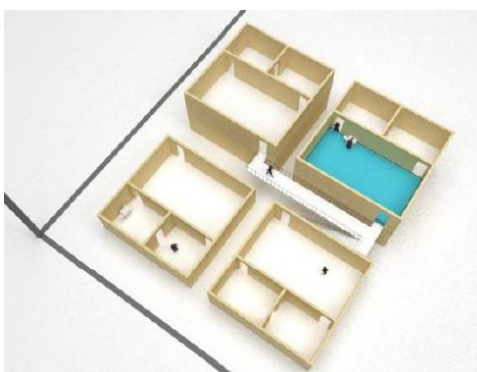
Position refers to the exact coordinates of a person or an object in a reference coordinate system. The position of individuals or objects in an indoor environment can be specified as a pin-point placement according to a global reference system of Cartesian coordinates that are specified for a building. Position can be also considered relative when it is relative to a local reference frame (Fig. 1).

Concerning indoor environments, GNSS signals are typically between 15 and 40 dB weaker compared to outdoors. A combination of different factors, such as the material of the building and the multipath interference can lead to signal blockage (Groves, 2013), creating the need for alternative solutions in indoor environments. The term location comes into life to bridge the gap between outdoor and indoor environments. In contrast to position, location does not refer to exact coordinates related to a global or



**Figure 1.** Absolute and relative position according to (Sithole & Zlatanova, 2016) (a) Absolute position according to a global reference frame (b) Relative position according to a local reference frame

local reference system but defines a general placement relative to the smallest defined physical space in an indoor facility, which could be a room, stairs or a corridor (Fig. 2). The uncertainty in the position of an individual is determined by the extent of the room (Sithole & Zlatanova, 2016). In that manner, localisation operations provide contextual information about a person's location in space, meaning the room or section of an indoor facility. In this research, an indoor map of the corresponding facility is required to acquire a location.



**Figure 2.** Location according to the smallest physically defined space in a building (Sithole & Zlatanova, 2016)

In indoor environments, there is lower landmark density and an absence of out-standing elements that can frequently result in easier loss of orientation compared to outdoors (Michon & Denis, 2001). As (Wadden & Scheff, 1983) mentioned, people spend around 80 % of their time indoors, thus localisation comprises an important problem, in public buildings, such as airports or train stations, that usually consist of chaotic spaces. Therefore, the necessity of an interactive indoor localisation system is apparent, so that a person can navigate in an indoor facility, especially if that person is exploring it for the first time. In that scenario, a location provider tool with adequate precision could be a significant aid. This tool could be applied in various indoor spaces, such as museums and art galleries (Gupta, et al., 2016).

Indoor localisation could also be applied during emergencies in complex indoor spaces. Persons in need could access the name of their current location, based on an indoor localisation application and transmit this information to the first-aid responders. The latter need guidance, related to the location of the person in need, as well as a way to reach that location (Yang & Worboys, 2011). Additionally, other applications of indoor localisation include the use of mobile autonomous units, to establish an indoor intelligent environment. Mobile service robots can be exploited for assisted living, setting up a smart living environment for elderly people and performing transportation and human interaction tasks.

Indoor localization has become increasingly important in recent years, with various technologies used for precise positioning. However, unlike outdoor environments where GNSS is a universal standard for positioning, indoor environments lack a universal solution (Lymberopoulos, et al., 2015). The most widely used technology for indoor positioning, Wi-Fi fingerprinting, is based on comparing Received Signal Strength (RSS) values with a reference radio map that translates signal values (Pérez-Navarro, et al., 2019). However, creating and maintaining an up-to-date radio map of signals in an indoor facility is a heavy and time-consuming task. Additionally, changes in the Wi-Fi infrastructure require the map to be created again, making it less practical. Poor WLAN planning in the facility can also result in irregular Wi-Fi signal availability. Similarly, alternatives such as Bluetooth-based positioning require costly installation of Bluetooth hotspots and suffer from similar issues (Pérez-Navarro, et al., 2019)

This has led to the development of new techniques that utilize camera and LiDAR sensors, which are increasingly available in mobile devices (Willems, 2017). While camera sensors exist in every mobile device, the use of LiDAR sensors has increased exponentially since their inclusion in the latest iPhone and iPad devices, making them a major part of mobile devices in the future.

Therefore, the challenge for achieving indoor localization in different environments is to find a technique that does not depend on costly and hard-to-access indoor sensor networks but uses features accessible to everyone on their mobile device.

Recent innovations include the use of Augmented Reality (AR) combined with the Simultaneous Localization and Mapping (SLAM) algorithm, which scans an indoor environment to find its position. These applications use various sensors such as RGB and depth cameras, with some requiring additional devices for a deeper understanding of the indoor environment (Oostwegel, 2020).

Overall, there is a need for reliable and accessible indoor localization techniques that do not depend on costly infrastructure (Willems, 2017). Thus, this research investigates the reliability of ceilings with characteristic details for indoor localisation purposes, by providing an accessible solution that makes use of components that are available in a mobile device. The focus includes indoor localisation, as well as near real-time location tracking of different users to discover different movement patterns in an indoor facility.

## 1.2 Research questions

Defining the main and secondary research questions is a crucial part of the re-search, aiming to address indoor localisation and ensure the concreteness of this project. Therefore, the primary research question is formed as follows:

*To what extent can ceilings with characteristic details be used for indoor localisation purposes?*

To obtain a better understanding of the concept and be able to answer the main research question robustly, some complementary research questions are formed.

- Which parameters (measuring angle, height, part of the room) should the user consider while acquiring point clouds and images of ceilings?
- Which is the optimal point cloud registration algorithm to achieve indoor localisation from ceiling data?
- Which is the optimal image-matching algorithm to achieve indoor localisation from ceiling data?
- Are LiDAR point clouds acquired by an iPhone device an accurate and accessible solution towards indoor localisation?
- Can the proposed pipeline aid towards facilitating localisation in emergencies?"
- How accurate is location tracking and does it respect user privacy?"

## 1.3 Contribution

This research demonstrates the versatility of using point clouds for indoor localisation and tracking, providing a dynamic aspect to indoor localization, especially with the use of ceilings, which are usually not altered over time. Ceiling data was used as an alternative way of performing indoor localization and tracking of users, with the implemented pipeline including two different localization techniques based on LiDAR and camera sensors available on recently released Apple mobile devices. The methodology can substitute the various localization methods that mostly involve Wi-Fi fingerprinting and Bluetooth sensors. The importance of this research lies in the fact that it offers a real-time indoor localization pipeline, available to a variety of users without the need for additional equipment, only requiring the existence of point clouds or images of ceilings as a reference for every room of the indoor facility. The point of cloud-based localization could be applied in buildings with large rooms, such as airports and train stations, where people can easily lose their orientation. The dynamic acquisition of point clouds allows users to perform data acquisition while moving between different rooms, setting the basis for navigation. Additionally, point clouds of ceilings can also be used as reference to CAD and BIM models, aiding in the modelling of existing utilities and their components in an indoor facility. In emergencies, the point cloud-based localisation can be used to transmit a user's location to first-aid responders. The LiDAR device could also be mounted on an automated device to map an indoor facility based on point cloud acquisition. The implementation of location tracking can provide daily to monthly statistics on the most used paths to a facility manager, helping to optimize the distribution of people inside the indoor facility.

## 2. Methods

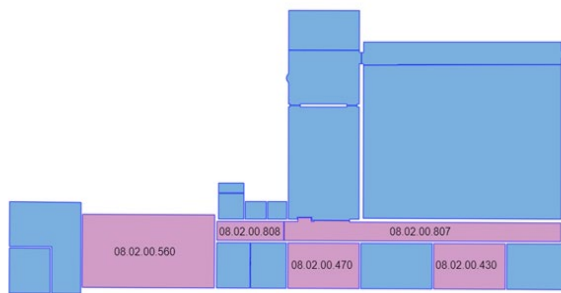
### 2.1 Overview

The overview of the methodology (Fig. 4) and the design of the experiments that were implemented to validate it will now be discussed.

The pipeline involves capturing ceilings, which include various protruding characteristics, to serve as reference points. Images of the rooms are acquired using camera sensors and point clouds are obtained from the LiDAR sensor of an iPad. Indoor localisation is achieved by comparing user data to reference data uploaded to a database. The LiDAR sensor gives the coordinates of the points in space, providing a 3D perspective of the ceiling, while the images from the camera sensor add colour.

Non-commercial software that combines the LiDAR and camera sensors is used for point cloud processing. Regarding image acquisition, their features were matched

using various matching techniques. Pre-processing and co-registration of point clouds were performed to achieve localisation. Results were stored in a database and visualised in a web application. An indoor model and network graph of the facility were combined with localisation results to provide information on users' current and previous locations. Movement patterns and paths are revealed through visualisation in the form of a heat map, which includes user paths at different times of the day. A dashboard with statistics on path usage is also created. This pipeline was applied to some rooms of the Faculty of Architecture and the Built Environment of TU Delft, the Netherlands, which are shown in Fig. 3.



**Figure 3.** Selected rooms from the Faculty of Architecture and the Built Environment

## 2.2 Data acquisition

### 2.2.1 Reference data

Reference data were created both in the case of point clouds and images, for each of the rooms that are shown in Fig. 3. Regarding point clouds, the LiDAR sensor of an iPad 12 pro and Pix4D Catch were used to acquire point clouds of ceilings that act-ed as reference. Each ceiling was captured with high detail, while walking inside a room at a steady pace, avoiding sudden changes in the measuring angle and height of the sensor. In rooms where the distance between the sensor and the ceiling is higher than 5 meters, an extensible accessory can be used, such as a monopod or a tripod. The reference point clouds were first pre-processed, as explained in the next section so that outliers and some wall parts are omitted.

The same rules regarding measuring height and angle apply to image reference data. However, there are some rooms such as long corridors where the whole ceiling cannot be captured by a single image. In this case, the camera sensor was placed al-most perpendicularly to the ceiling, to capture the largest possible area of the ceiling.

The reference data were then attached to the respective rooms in the created in-door model. Specifically, each room was represented by a polygon in the indoor map and the reference point clouds and images were its attributes.

### 2.2.2 Point clouds

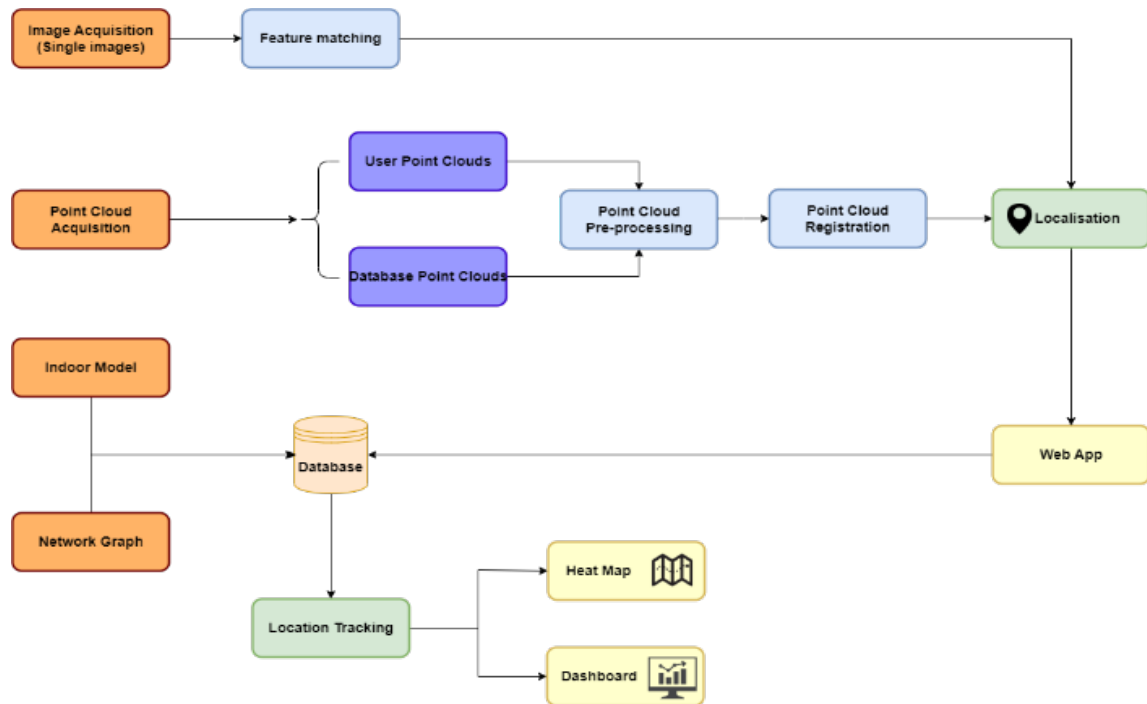
Regarding point cloud acquisition, two types of point clouds were acquired, from LiDAR sensors. There are point clouds that act as a reference and were stored in a database, as well as point clouds that are acquired by a user. The latter will be compared to these reference point clouds, so that indoor localisation is achieved, based on the best match. User point clouds were acquired with two different downscaling factors of 10 and 30 cm distance between each point. The point cloud acquisition was implemented in two ways: while a person is walking into a room, giving a dynamic perspective to the acquisition and while staying still, so that it is investigated if the final product of the research can be used during emergencies, in cases where an individual might be unable to move.

### 2.2.3 Images

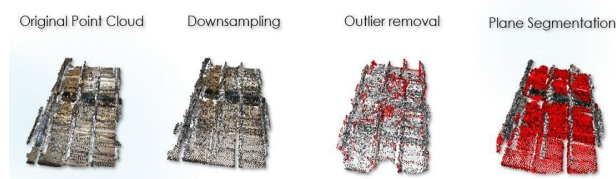
Single images of the tested rooms were acquired from the camera sensors. As in the case of point clouds, some images were used as reference to represent the room's ceiling in two dimensions. Images of a ceiling acquired by a user were then compared to the reference images of the rooms to reveal the user's location based on the optimal match.

## 2.2 Point cloud pre-processing

Pre-processing of the point clouds included voxel down sampling to reduce the processing time by manipulating a point cloud of smaller size (Miknis, Ware, Davies, & Plassmann, 2016). It must be mentioned that the reference point clouds were not downscaled, to preserve high details. However, this operation must be implemented carefully and until a certain threshold, because further down sampling might result in an important loss of information. Furthermore, when acquiring ceiling data, the point cloud may include adjacent wall parts that need to be excluded from the upcoming operations. To achieve their removal, a smaller part of the acquired point cloud was used, to discard parts that might exist in the corners of the point clouds. Additionally, some points were located and removed based on the number of their neighbours, to further improve the point cloud's quality and reduce processing time. These parts can be considered outliers (Han, et al., 2017). Finally, plane segmentation based on the RANSAC (Random Sample Consensus) algorithm was performed, to differentiate the flat surface of the ceiling with its protruding objects, such as lamps and other installations, which comprise the characteristic details of each room's ceiling.



These steps are visualised in Fig. 5, while the pseudo algorithm that was used, in Fig. 6.



**Figure 5.** Pre-processing steps

---

Input: Point cloud  $pc$  and voxel size  
 Output: Original point cloud, Cleaned point cloud and FPFH features

- 1  $pc \leftarrow$  bounding box of  $x,y,z$  coordinates
- 2 bounding box  $x_{min} = \max(points_x) - 0.1 * (\max(x)) - \min(points_x)$
- 3 bounding box  $x_{max} = \min(points_x) + 0.1 * (\max(x)) - \min(points_x)$
- 4 bounding box  $y_{min} = \max(points_y) - 0.2 * (\max(y)) - \min(points_y)$
- 5 bounding box  $y_{max} = \min(points_y) + 0.2 * (\max(y)) - \min(points_y)$
- 6 bounding box  $z_{min} = \max(points_z) - 0.1 * (\max(z)) - \min(points_z)$
- 7 bounding box  $z_{max} = \min(points_z) + 0.1 * (\max(z)) - \min(points_z)$
- 8  $pc \rightarrow$  downsample based on voxel size
- 9  $pc \rightarrow$  remove outliers(nearest neighbours, max standard deviation)
- 10  $pc \rightarrow$  segment plane (distance threshold, ransac points, iterations)
- 11 for points in outlier point cloud do
- 12    Calculate FPFH features(radius, nearest neighbours)
- 13 return  $pc$ , outlier cloud and FPFH

---

**Figure 6.** Algorithm for point cloud pre-processing

## 2.4 Point cloud registration

After acquiring and pre-processing reference and user point clouds, the next step was to create an algorithm that would aid towards comparing them. The main idea behind this is, that each point cloud taken by a user, would be compared with all the point clouds in the database and the best match will reveal the room where the user is located. This procedure works as follows for both types of point clouds. The comparison first included a global

registration, so that the user and the reference point clouds obtain an initial alignment and afterwards a local registration algorithm to refine the point cloud registration.

### 2.4.1 Global registration

First, the normal vectors of all the points were computed. Furthermore, points with a unique and descriptive neighbourhood were detected. The detection and description of these unique points for each point cloud were implemented based on FPFH (Fast Point Feature Histogram) feature calculation. The technique includes the afore-mentioned steps and then RANSAC, to select some random points from the reference point cloud and then find the corresponding points in the user point cloud, using a nearest neighbour query in the 33-dimensional FPFH feature space (Li, Hu, & Ai, 2021). Aside from the distance of the corresponding points in the compared point clouds, the similarity between two edges between the compared point clouds and the vertex normal affinity of the correspondences are also checked. In case the points satisfy the selected thresholds, the transformation of the user point clouds towards the reference point clouds is implemented.

### 2.4.2 Local refinement

Based on the results of the global registration, an attempt of improving the quality and time efficiency of the algorithm includes different variations of the ICP (Iterative Closest Point) algorithm. The further minimisation of the point cloud differences was performed by keeping one point cloud fixed, while the other is transformed towards it. Specifically, each point of

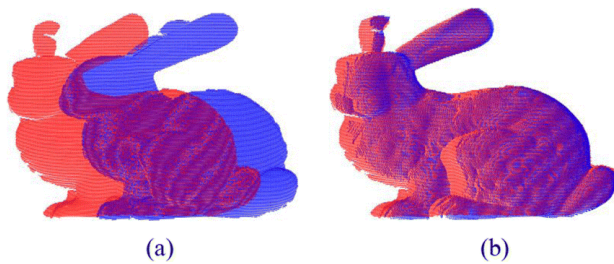


the user point cloud was matched to the closest point of each reference point cloud. Then, rotation and translation were estimated, and this process is iterated until the results converge (Li, Hu, & Ai, 2021). The user point cloud was compared to all the reference point clouds, based on the fitness value (1) and the RMSE (Root Mean Squared Error) value of the inlier correspondences which will result in the indoor localisation (2). Different variations of ICP were implemented and compared and more specifically Generalised, Point-to-Point, Point-to-Plane and Coloured ICP.

$$fitness = \frac{\# inliers}{\# points in target point cloud} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N \|R_f p_k + T_f - q_k\|_2^2} \quad (2)$$

Where  $p_k$  and  $q_k$  are the points of the reference and user point clouds respectively, while  $R_f$  and  $T_f$  are the rotation matrix and translation vector in the transformation matrix.



**Figure 7.** ICP registration. (a) Initial pose of two datasets (b) Result after ICP registration (Wan, et al., 2019)

The complete point cloud registration process, leading to indoor localisation is presented in Fig. 9.

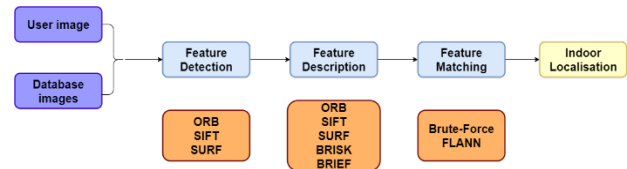
## 2.5 Feature matching

In this section feature matching based on the comparison of single images to examine the suitability of various feature detectors, descriptors and matching techniques.

### 2.5.1 Feature matching between single images

For each of the selected rooms, one image of a ceiling was acquired and acted as reference. For testing purposes, different user images were additionally acquired from different viewpoints and were compared with the reference images. This comparison included the use of different feature descriptors and detectors, such as ORB (Oriented FAST and Rotated BRIEF) (Rublee, Rabaud, Konolige, & Bradski, 2011), SIFT (Scale Invariant Feature Transform) (Lowe, 2004), and also two different feature matching techniques, brute-force and FLANN (Fast Library for Approximate Nearest Neighbours)

(Muja & Lowe, 2009). The number of matches between the user and the reference images was used to reveal the location of the user. The indoor localisation process that was based on the feature matching of images is presented in Fig. 8.



**Figure 8.** Indoor localisation based on feature matching

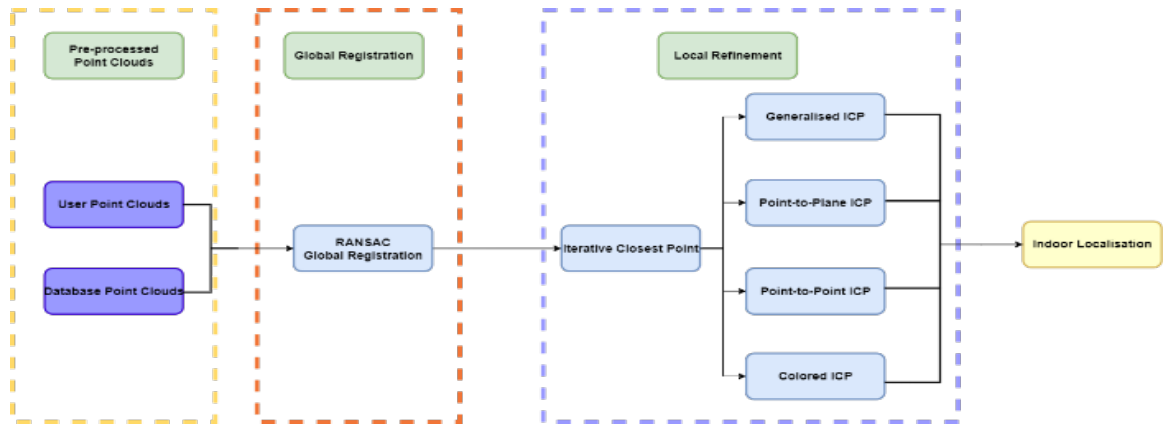
## 2.6 Storage

The setup of the whole system was organised in an online database, part of the ArcGIS Online Server. This database includes the indoor model of the case study and a network graph that connects all the rooms of the tested area. Except for the geometry of the rooms in the indoor model, each of them includes one pre-processed point cloud and an image that acts as a reference for the point cloud registration and feature matching operations respectively. Moreover, this indoor model serves as an embedded map in a web application that was created, allowing the users to have a visual insight into their location.

## 2.7 Location tracking

Each time the web application is used, the users' current and previous locations are stored in the ArcGIS Online Server, under an encrypted id. When users move between different rooms, it means that they used a certain path to achieve that. Based on the network graph of the indoor space that reveals all the connections between adjacent rooms, the current and previous locations of the users were translated to a line in the network graph, representing a specific route. The availability of this information is near-real time as the results appear in the online server after a few seconds. Based on the unique id of each user, a heat map that is based on the network graph was used to visualise the used routes.

Additionally, this information was used to reveal different movement patterns, during different times of the day. The visualisation is accomplished in the form of a heat map, where based on the usage of each path, different colours and widths were applied to the corresponding line of the network graph. Consequently, this information can reveal how much a path is used during a daily, weekly or even monthly period. Acquiring this knowledge is valuable, especially during the COVID-19 era, because it can be exploited by a building manager, who can achieve the



**Figure 9.** Indoor localisation based on point clouds

optimal distribution of people in an indoor facility (Spinoza Andreo, et al., 2021).

---

**Input:** Reference point cloud  $ref_{pc}$ , User point cloud  $user_{pc}$  and table with paths  
**Output:** List containing visited rooms

```

1 if length(localisation list) == 0 then
2   for  $ref_{pc}$  in  $room_{list}$  do
3     preprocess  $user_{pc}$  and  $ref_{pc}$ 
4     global_registration( $user_{pc}, ref_{pc}$ )
5     local_registration( $user_{pc}, ref_{pc}, result_{global}$ ) → calculate fitness value
6   return room(max fitness)
7 else if length(localisation list) ≥ 0 then
8   for  $i$  in range(length(result)) do
9     request path table
10    paths → query(previous room, current room)
11    query result → update path usage
12    preprocess  $user_{pc}$  and  $ref_{pc}$ 
13    global_registration( $user_{pc}, ref_{pc}$ )
14    local_registration( $user_{pc}, ref_{pc}, result_{global}$ ) → calculate fitness value
15  return room(max fitness)
16
  
```

---

**Figure 10.** Location tracking algorithm

## 2.8 Data and Software Availability

Input data, point clouds and images were acquired by the author and are available as open data in [https://github.com/jdardave/Indoor\\_Localisation\\_Dardavesis.git](https://github.com/jdardave/Indoor_Localisation_Dardavesis.git). Under the same link the code that was implemented for this research can be found. Aside from that, additional data, such as the CAD files of the floorplans of the Faculty of Architecture and the Built Environment were specifically requested from TU Delft Real Estate, therefore are not open to the public.

Different programming languages (Python, HTML, JavaScript) and open-source tools were used. Most of the software that was used were open source (CloudCompare, PhotoMatch), but the creation of the indoor model and network graph of the case study were implemented in ArcGIS pro.

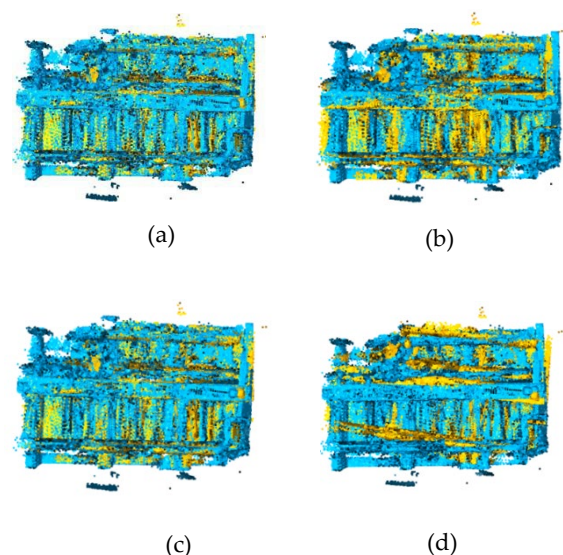
Concerning results, some of them are available in the corresponding section, however due to their number and size, they are not available online.

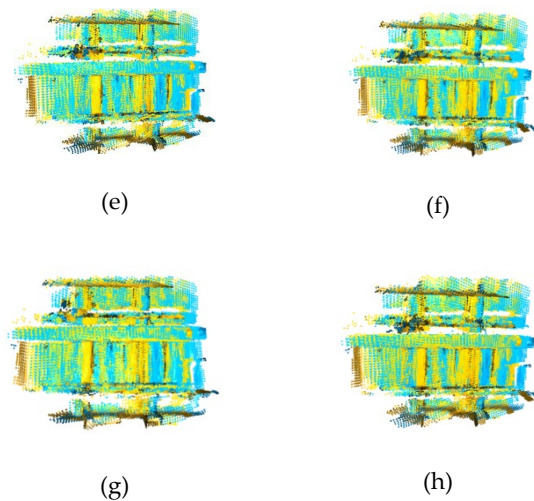
## 3. Results

### 3.1 Indoor localisation for point clouds

In this section, the indoor localisation results that were produced based on point cloud acquisition with the aforementioned downscaling factors will be presented and compared.

The first results emerge from point clouds with a 10 and 30 cm distance between each point. These results, based on different combinations of global and local registration algorithms are presented in Fig. 11. The point clouds with the blue colour represent the reference point clouds, while the ones with yellow, are the point clouds that were acquired by a user.





**Figure 11.** Global and Local registration from point clouds acquired with the 10 and 30 cm distance between each point for room 08.02.00.560 (a): RANSAC & Coloured ICP (fitness=0.974, RMSE=0.01) (b) RANSAC & Point-to-Plane ICP (fitness=0.955, RMSE=0.03) (c) RANSAC & Generalised ICP (fitness=0.953, RMSE=0.03) (d) RANSAC & Point-to-Point ICP (fitness=0.943, RMSE=0.06) (e) RANSAC & Coloured ICP (fitness=0.963, RMSE=0.01) (f) RANSAC & Point-to-Plane ICP (fitness=0.958, RMSE=0.02) (g) RANSAC & Point-to-Point ICP (fitness=0.954, RMSE=0.15) (h) RANSAC & Generalised ICP (fitness=0.954, RMSE=0.15)

The results for room 08.02.00.560 are promising, as in most cases all the point cloud registration methods match the tested room to its reference equivalent. The most accurate results are achieved when Coloured ICP was involved, producing accurate results when it was combined with global registration algorithms, as Fig. 11a and Fig. 11d indicate. It has to be noted, that the number of fitness is not important by itself, but it has to be higher compared to the reference point clouds of the remaining rooms.

**Table 1.** Number of correct matches per point cloud registration algorithm

10 cm	RANSAC global registration			
	Point-to-Point	Point-to-Plane	Coloured ICP	Generalised ICP
Dynamic acquisition	9/10	9/10	10/10	9/10
Static acquisition	7/10	8/10	10/10	8/10
30 cm				
Dynamic acquisition	9/10	7/10	9/10	7/10
Static acquisition	8/10	7/10	9/10	7/10

Tab. 1 shows the number of correct matches for each combination of global and local registration algorithms that were applied. The testing includes twenty point clouds per method and specifically ten for the ceilings that a user acquired while walking, and ten more while the user remained static. RANSAC is a non-deterministic algorithm, however, the high number of iterations that was selected, increases the probability that the result is reasonable.

The results are better when users are walking inside a room during data acquisition, in contrast to when they remain static. This is a reasonable outcome, as while a user is walking, the entire ceiling of a room can be captured. On the contrary, while users remain static, they can only capture a specific part of a room's ceiling, in case the room is considerably large since the range of the LiDAR sensor is approximately five meters. Therefore, in cases where users are unable to move, there are higher chances that the localisation is correct when they capture a part of a ceiling that has characteristic details. The quality of the point clouds that were acquired with a 30 cm distance between each point is slightly worse compared to the previous results. The result is reasonable, due to the lower density of point clouds that was chosen for the acquisition. However, regarding the Coloured ICP, its results are at a similar level as before, showing the importance of adding colour information that the other algorithms do not include. The worst results are presented for Point-to-Plane and Generalised ICP when they are combined with global registration algorithms, with 7/10 correct indoor localisation results.

The wrong point cloud matches for some registration techniques appear between rooms 08.02.00.430 and 08.02.00.470. This confusion arises from the fact that these rooms have an almost identical size in squared meters and similar characteristic details in their ceilings, as they are both lecture rooms. Additionally, the second wrong set is mostly between rooms 08.02.00.808 and 08.02.00.807. This happens, because they are both corridors and room 08.02.00.808 is significantly smaller than room 08.02.00.807. Thus, this room may be wrongly matched as a part of 08.02.00.807. Some rooms, such as 08.02.00.807, which is a long corridor, have a significantly different shape than the common rectangular rooms, hence the possibility that the localisation is wrong is significantly reduced.

Concerning wall parts that were acquired along with ceilings, small areas did not affect the results, as some minor wall parts remained in the tested point clouds even after the pre-processing operations. However, in cases where a significant part of a wall is captured, the plane segmentation could be implemented in the wrong way, as

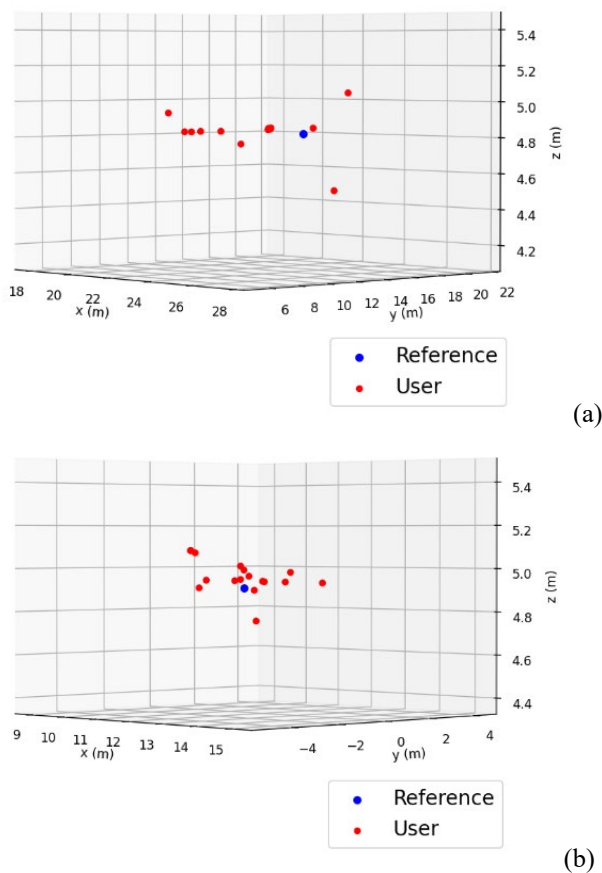


the main plane that is computed, might be the wall instead of the ceiling's upper flat part.

### 3.1.1 Performance parameters

This section presents some performance parameters that were calculated to test the robustness of the results.

Fig. 12a and Fig. 12b show the centres of the respective reference point cloud with blue colour, as well as the centres of different user point clouds after the implementation of the point cloud registration algorithms and specifically RANSAC based global registration and Coloured ICP local refinement.



**Figure 12.** Scatter plot with centres of reference and user point clouds after point cloud matching. (a) Room 08.02.00.808 (b) Room 08.02.00.807

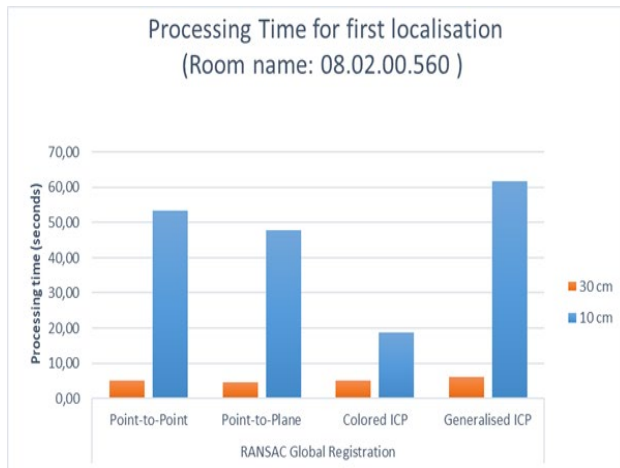
The results concerning room 08.02.00.808 reveal good accuracy, as most of the centres of the user point clouds are a few centimetres away from the centre of the reference point cloud, while at the same time the precision is adequate, as most of the centres of the user point clouds are close to each other. On the contrary, the same results for room 08.02.00.807 are worse concerning the accuracy and precision, since the centres of the user point clouds are further away from the centre of the reference point cloud and at the same time far from each other. This has to do with the size and length of room 08.02.00.807, which is a corridor with similar and lengthy protruding

installations on the ceilings, therefore the user point clouds may be matched to the reference point cloud on a different part of those installations further away from the centre of the point cloud. However, in both cases, there is good accuracy and precision regarding the height dimension, which shows that the flat part of the ceilings of the user and reference point clouds is in most cases correctly matched. The developed point cloud-based localisation method is possible, in buildings that include a database of reference point clouds of ceilings for each room of the indoor facility. Satisfactory solutions for ceilings with characteristic details are shown, such as the ones in the Faculty of Architecture and the Built Environment. However, the quality of the solution, might not be the same when applied to primarily flat ceilings, with fewer characteristic details, or ceilings that include glass, whose reflective abilities might affect the indoor localisation result.

The creation of a database with reference point clouds and images for every ceiling of an indoor facility requires some devices. Regarding point cloud acquisition, an Apple device such as an iPhone 12 pro or an iPad 12 pro is required, as well as some non-commercial software. The cost for these devices is approximately 1000 euros. Alternatively, a laser scanner could be rented to perform the acquisition. Concerning image acquisition, camera sensors are included in each mobile device, so no further devices are required.

### 3.1.2 Time processing

Fig. 13 shows the processing time of different combinations of global and local registration algorithms, based on point clouds acquired with 10 and 30 cm point distance. It is noticeable that when point clouds with 30 cm of minimum point distance are used, indoor localisation is computed within 3-6 seconds depending on the algorithm. However, concerning the point clouds with smaller point distance, the time complexity augments exponentially and reveals the time efficiency of Coloured ICP, which produces indoor localisation in approximately 18 seconds. In this case, there is a significant difference in processing time, as other algorithms such as Generalised ICP take approximately 60 seconds to result in indoor localisation. This difference will be even greater when a database with a higher number of point clouds is used to perform indoor localisation.



**Figure 13.** Processing time of different registration algorithms

### 3.2 Indoor localisation from images

This subsection includes techniques that were implemented based on image acquisition. The indoor localisation result is based on the number of matches between the user and the reference images. Additionally, different combinations of feature detection, description and matching techniques are analysed.

**Table 2.** Number of correct matches per feature detection, description and matching techniques

5 MP Camera		
	ORB-ORB	SIFT-SIFT
Brute-force	9/10	9/10
FLANN	7/10	8/10
8 MP Camera		
Brute-force	9/10	7/10
FLANN	8/10	7/10

The results are based on images that were taken from two different cameras with 5 and 8 MP resolutions respectively. Both cameras perform similarly resulting in 18/20 correct room matches. Additionally, the two feature-matching techniques have similar efficiency when they are combined with the two different detectors and descriptors, while brute-force performs slightly faster than FLANN. However, the latter can be more efficient than brute-force, when large datasets are involved. FLANN results in a higher number of matches between the user image and the reference image of the correct room in most cases.

The same can be mentioned about SIFT, which results in more matches between the images compared to ORB however the indoor localisation is calculated with worse time efficiency. In terms of quality, the suitability of SIFT lies in the fact that it is scale and rotation invariant,

whereas ORB is only rotation invariant and robust to noise. As a result, in case SIFT is used, the height and angle of the device do not affect the result. The time efficiency of SIFT could be improved, by implementing the SURF (Speed-ed-Up Robust Features) detector and descriptor. The ratio test that was applied in each experiment was strict, to avoid false correspondences, due to the common installations between the different rooms. The clearest results were noticed concerning a test image of room 08.02.00.470, where approximately 400 matches were observed between the user and reference image, a number which is significantly higher compared to the other reference images. This is an outcome of the similarity of the user and reference images, as they were acquired from a similar angle and cover approximately the same part of the ceiling. In other cases where the viewpoints of the user and reference images were different, the indoor localisation results were correct, as the user image had the most matches with its corresponding reference image, however, the number of matches was significantly lower, between 50 and 100.

The wrong localisation results were related to room 08.02.00.807, which cannot be entirely captured from a single image, due to its length. Therefore, in terms of size, it appears to be similar to the different rooms of the case study. However, this result can be partially solved, in case the data acquisition is performed, by holding the sensor almost perpendicular to the ceiling, so that a bigger part of the ceiling is captured.

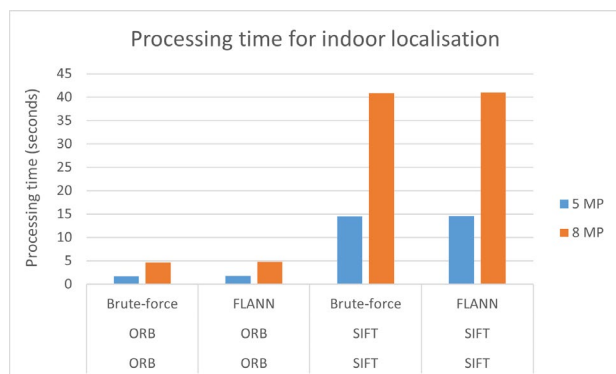
In this testing, there are no differences between the two different cameras regarding the quality of the results. However, certain illumination changes that create blurry areas, may significantly affect the intensity of each pixel of the tested images, due to the ceiling lights that are on, during most of the day in the Faculty of Architecture and the Built environment. In this situation, a high-resolution camera could better capture reality and avoid these blurry parts in the images. However, a drawback of using cameras with high resolution, is that they tend to produce bigger image files that are not suitable for real-time applications, due to the necessity of a time-efficient solution. The intensity values of these areas might appear similar to the windows, resulting in wrong matches between the windows and the lights, when two images are compared. Hence, during the acquisition, windows should be avoided as much as possible, due to their reflective ability.

Overall, indoor localisation based on the comparison of the features of an image seems promising, however additional testing regarding lighting conditions and viewpoints, has to be implemented to produce safe conclusions about this method.

### 3.2.1 Time processing

Fig. 14 shows the processing time while using different combinations of feature detection, description and matching algorithms with two different cameras of 5 and 8 MP resolution respectively. It is clear that the resolution of the camera affects the time efficiency of the calculation. Moreover, the ORB detector and descriptor is faster than SIFT, while Brute force performs faster than FLANN, as the dataset is small.

Additionally, some images were chosen, so that additional combinations of feature detectors, descriptors and feature matching techniques could be tested. The testing that was performed in an open-source software called Photomatch (González-Aguilera, et al., 2020), showed that the combination of SURF (Bay, Tuytelaars, & Gool, 2006) as a detector and descriptor detects the maximum number of key points (5000) with both brute force and FLANN matching techniques. The opposite is observed for the combination of SIFT and BRIEF (Binary Robust Independent Elementary Features) (Calonder, et al., 2011) are combined with almost 3500 thousand key points. The latter happens due to the simplicity of the BRIEF descriptor which targets in the fast description from simple intensity difference tests. Regarding the percentage of key points that are used for matching, the SURF detector with SIFT descriptor and FLANN matching take into advantage approximately 13 % of the detected points, while the combination of SIFT detector, SURF descriptor and FLANN uses less than 1 % of the detected key-points for feature matching. This is a result of the size of the vectors of SIFT and SURF descriptors, which have a size of 128 and 64 elements, showing that SIFT entails more details concerning the description of the sub-region of the tested key points. In most cases, FLANN uses a higher percentage of key points for matching, compared to brute force except when the SIFT detector and SURF descriptor are combined, however, the difference is minor. This information is presented in detail in Tab. 3.

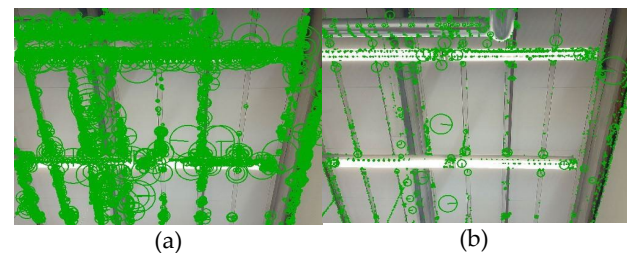


**Figure 14.** Processing time of different feature matching combinations

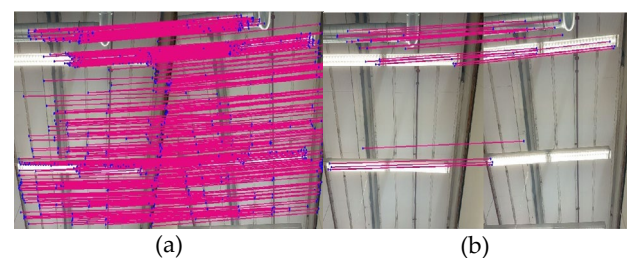
**Table 3.** Number of key points and their percentage used for matching per combination.

Detector	Descriptor	Matcher	Average Number of Key-points	Percentage of key points used for matching (%)
SURF	SURF	Brute-force	5000	10.60
		FLANN	5000	9.82
SURF	SIFT	Brute-force	5000	12.06
		FLANN	5000	12.72
SIFT	BRIEF	Brute-force	3495	4.46
		FLANN	3495	6.67
SIFT	SURF	Brute-force	2896	1.04
		FLANN	3671	0.68

Fig. 15 and Fig. 16 show the feature extraction and matching for the best and worst combinations respectively, where the diameter of the circles, indicates the meaningful key-point neighbourhood.



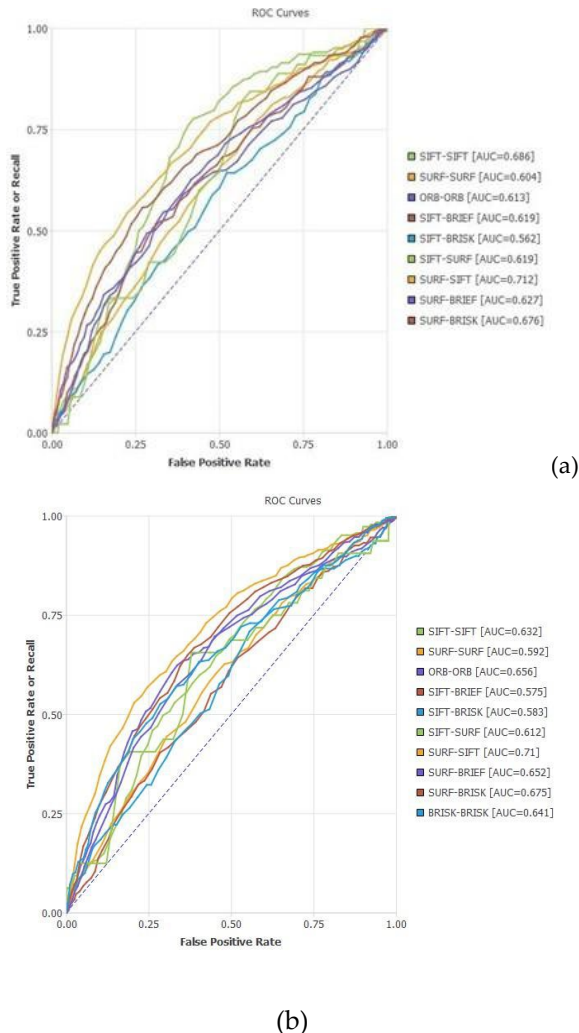
**Figure 15.** Feature extraction from images. (a) SURF detector and SIFT descriptor (b) SIFT detector and SURF descriptor



**Figure 16.** Feature matching between images. SURF detector and SIFT descriptor (b) SIFT detector and SURF descriptor.

Figure 17a and Fig. 17b show the ROC (Receiver Operating Characteristics) curves for an image set when brute-force matching and FLANN are used respectively. Overall, both feature-matching techniques perform similarly, with a recall between 50 - 70%, revealing the percentage of matches that were true and not mistakenly matched by the algorithms. If the information of these graphs is combined with Table 5.5, the finer results can be noticed when the SURF detector and SIFT descriptors

are combined, with approximately 71 % of true to false positive ratio. On the other hand, the worst performance is observed when SIFT detector is combined with BRIEF and BRISK (Binary Robust Invariant Scalable Key points) (Leutenegger, Chli, & Siegwart, 2011) descriptors.

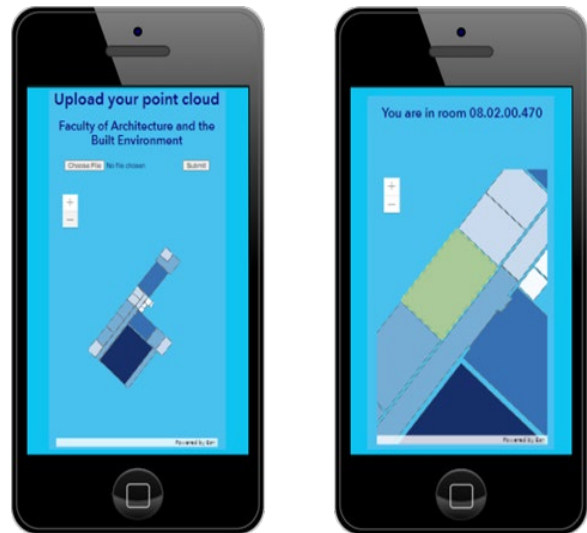


**Figure 17.** ROC curves between 2 images with (a) Brute-force matching (b) FLANN

### 3.3 Web-app

The indoor localisation results were visualised in a web application, as shown in Fig. 18. The app works by requesting the reference point clouds from the database, so that they can be compared based on the discussed algorithms to the user data in near real-time. Users can post their data in the application and after a few seconds, the room they are located in is revealed. Additionally, the app includes the indoor model of the case study, so aside from the name of the room, the app also highlights the polygon that represents the

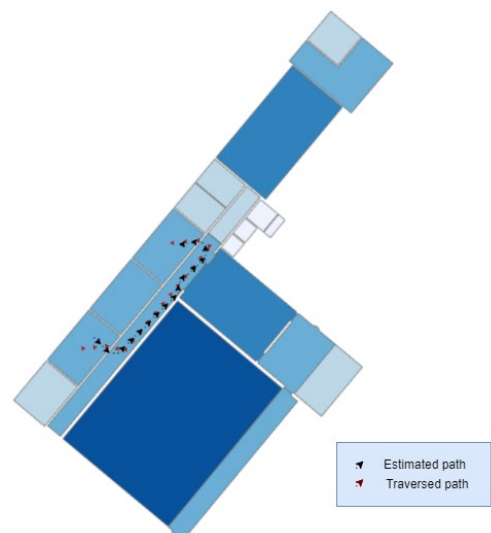
room in the indoor model of the indoor facility and zooms in on it.



**Figure 18.** Web-app interface

### 3.3 Location tracking

The location tracking results are based on the different indoor locations of different users at different times of the day. Therefore, the quality of the followed paths is a direct outcome of the indoor localisation quality. The results are available in the ArcGIS online Server and can be seen in near real-time on a map, that is updated every 30 seconds. To test the accuracy of the location tracking algorithm, ground truth was set, based on the path that the user originally followed and was compared to the path, as it is visualised in the final product. This is shown in Fig. 19.

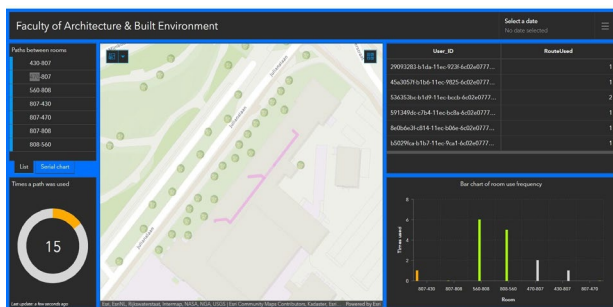


**Figure 19.** Estimated vs Traversed path between rooms 08.02.00.430, 08.02.00.807 and 08.02.00.470



Fig. 19 shows the path of a user that moved between rooms 08.02.00.430, 08.02.00.807 and 08.02.00.470. The indoor localisation was performed correctly for these three rooms; therefore, the ground truth is similar to the path as it is visualised in ArcGIS Pro. Some differences exist due to the indoor network that is used to visualise the paths, as the centre of each room is the representative node and the fact that the rooms are connected with lines, therefore small deviations when the user is not moving completely straight cannot be detected.

Finally, a dashboard was created (Fig. 20) to visualise daily, weekly and monthly statistics about the use of each path, to discover different movement patterns during different times of a day.



**Figure 20.** ArcGIS Dashboard

## 4. Conclusion

This research aimed to investigate the reliability of ceilings as indoor landmarks, examining an alternative way of achieving indoor localisation and in extent location tracking of users. It focuses on LiDAR and camera sensors, which are incorporated in up-to-date mobile devices, to substitute the varied used localisation methods that mostly involve Wi-Fi fingerprinting and Bluetooth sensors. In that manner, indoor localisation becomes possible for a variety of users, without the need for additional equipment. The only requirement of this pipeline is the existence of point clouds of ceilings that will act as reference for every room of the indoor facility.

The indoor localisation pipeline showed promising results, both in terms of quality as well as time efficiency, as the scope of the research was to be able to perform real-time localisation of large indoor environments, focusing on ceilings with characteristic details. Based on the results, a point cloud acquisition of a few seconds is enough to indicate the room that users are in, especially when the whole ceiling can be captured. In case a ceiling is partly acquired, the indoor localisation result depends on the uniqueness of the captured part.

Regarding data acquisition, the user should not perform any sudden movements and changes in the measuring angle and height. Additionally, small wall parts do not

affect the localisation results, while larger parts should be avoided. Due to the range of the current LiDAR sensors which is approximately 5 meters, some rooms 'ceilings cannot be captured; thus, the mobile device should be mounted on an extensible monopod or tripod. However, this unavailability in the acquisition can be also translated into information that a person is in a room with a high ceiling. During image acquisition, the sensor should be placed almost perpendicularly to the ceiling to capture a larger part of it.

The Coloured variation of the ICP, proved to be the optimal solution. In contrast to the other implemented local refinement algorithms, Coloured ICP adds colour information to the geometry as its name indicates, hence this additional information is the reason behind the suitability of the algorithm. The multi-registration scheme of Coloured ICP significantly improves the time efficiency of the algorithm, making it a concrete choice for real-time applications that use point clouds of ceilings for indoor localisation.

Concerning the image-based indoor localisation techniques, the combination of the SURF feature detector with the SIFT descriptor provides the most optimal results, when combined with brute-force and FLANN. The scale and rotation invariant character of SIFT makes it adaptable and robust to different types of distortion, illumination and noise. However, the combination of the ORB feature detector and descriptor proved to be the most time efficient, making it suitable for real-time applications. Brute force performs slightly faster than FLANN, due to the small size of the dataset that was used. FLANN resulted in a higher number of matches between the user and reference images, compared to brute force, with a higher match difference between the correct and wrong reference images, ensuring the quality of indoor localisation.

Promising results were also shown regarding emergencies, however, some improvements especially concerning time efficiency must be implemented. Concerning location tracking, the resulting quality is based on the succession of the indoor localisation results. The accuracy of location tracking is at room level as the centre of each room was chosen as a representing point.

### 4.1 Recommendations and future research

Some additional experiments could be applied to enhance the current pipeline. Image acquisition during different times of the day could be performed to further investigate the influence of lighting conditions. Furthermore, point clouds and images should be combined with a Wi-Fi fingerprinting approach, to reduce the search area during the first localisation of a user.

Additional research could involve the use of machine learning algorithms, which could automatically detect the large wall planes that negatively affect the indoor localisation results based on ceilings. Additionally, feature matching based on monocular depth estimation could be tested, as an alternative way of image-based indoor localisation. The protruding installations of the ceilings could be used in combination with an AR platform to recognise the different utilities to develop a landmark-based localisation approach. Furthermore, important research could be implemented on navigation, after incorporating a robust solution for its basis, indoor localisation, focusing on closing the gap between the planned towards the estimated trajectory of a followed path. Additional research could include the establishment of navigational instructions for humans and also robots, as well as navigation for specific user groups, such as people with partial or severe blindness, by incorporating the braille language in a real-time application, or navigational applications that focus on people with movement disorders, who need to follow specific paths as they navigate to their desired destination.

## References

- Bay, H., Tuytelaars, T., & Gool, L. V. (2006). Surf: Speeded up robust features. *European conference on computer vision*, (pp. 404–417). Retrieved from [https://link.springer.com/chapter/10.1007/11744023\\_32](https://link.springer.com/chapter/10.1007/11744023_32)
- Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., & Fua, P. (2011). BRIEF: Computing a local binary descriptor very fast. *IEEE transactions on pattern analysis and machine intelligence*, 34, 1281–1298. Retrieved from <https://ieeexplore.ieee.org/document/6081878>
- Dardavesis, I. (2022). Indoor localisation and location tracking in semi-public buildings based on LiDAR point clouds and images of the ceilings, Master's thesis, Geomatics, Delft University of Technology, pp. 131, 2022. <http://www.gdmc.nl/publications/2022/MScThesisIoannisDardavesis.pdf>
- González-Aguilera, D., Oña, E., López Fernández, L., Farella, E., Stathopoulou, E., Toschi, I., . . . Nex, F. (2020, August). PHOTOMATCH: AN OPEN-SOURCE MULTI-VIEW AND MULTI-MODAL FEATURE MATCHING TOOL FOR PHOTOGRAMMETRIC APPLICATIONS. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B5-2020, 213-219. doi:10.5194/isprs-archives-XLIII-B5-2020-213-2020
- Groves, P. (2013, March). *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, Second Edition.
- Gupta, G., Kejriwal, N., Pallav, P., Hassan, E., Kumar, S., & Hebbalaguppe, R. (2016). Indoor Localisation and Navigation on Augmented Reality Devices. *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, (pp. 107-112). doi:10.1109/ISMAR-Adjunct.2016.0052
- Han, X.-F., Jin, J., Wang, M.-J., Jiang, W., Gao, L., & Xiao, L. (2017, May). A review of algorithms for filtering the 3D point cloud. *Signal Processing: Image Communication*, 57. doi:10.1016/j.image.2017.05.009
- Leutenegger, S., Chli, M., & Siegwart, R. (2011, November). BRISK: Binary Robust invariant scalable keypoints., (pp. 2548-2555). doi:10.1109/ICCV.2011.6126542
- Li, J., Hu, Q., & Ai, M. (2021, January). Point Cloud Registration Based on One-Point RANSAC and Scale-Annealing Biweight Estimation. *IEEE Transactions on Geoscience and Remote Sensing*, PP, 1-14. doi:10.1109/TGRS.2020.3045456
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60, 91–110. Retrieved from <https://link.springer.com/article/10.1023/B:VISI.0000029664.99615.94>
- Lymberopoulos, D., Liu, J., Yang, X., Choudhury, R. R., Handziski, V., & Sen, S. (2015). A Realistic Evaluation and Comparison of Indoor Location Technologies: Experiences and Lessons Learned. *Proceedings of the 14th International Conference on Information Processing in Sensor Networks* (pp. 178–189). New York, NY, USA: Association for Computing Machinery. doi:10.1145/2737095.2737726
- Michon, P.-E., & Denis, M. (2001). When and Why Are Visual Landmarks Used in Giving Directions? In D. R. Montello (Ed.), *Spatial Information Theory* (pp. 292–305). Berlin: Springer Berlin Heidelberg. Retrieved from [https://link.springer.com/chapter/10.1007/3-540-45424-1\\_20](https://link.springer.com/chapter/10.1007/3-540-45424-1_20)
- Miknis, M., Ware, J., Davies, R., & Plassmann, P. (2016, June 1). Efficient Point Cloud Pre-processing using The Point Cloud Library. *International Journal of Image Processing*, 10, 63–72. Retrieved from <https://www.cscjournals.org/manuscript/Journals/IJIP/Volume10/Issue2/IJIP-1063.pdf>
- Muja, M., & Lowe, D. G. (2009). Flann, fast library for approximate nearest neighbors. *International Conference on Computer Vision Theory and Applications (VISAPP'09)*, 3. Retrieved from

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.192.5378&rep=rep1&type=pdf>

- Oostwegel, L. (2020). Indoor Positioning using Augmented Reality. Retrieved from <https://repository.tudelft.nl/islandora/object/uuid%3Ab50e234d-3fe0-44a4-9ec0-f5664ca79844>
- Pérez-Navarro, A., Torres-Sospedra, J., Montoliu, R., Conesa, J., Berkvens, R., Caso, G., . . . Wilk, P. (2019). 1 - Challenges of Fingerprinting in Indoor Positioning and Navigation. In J. Conesa, A. Pérez-Navarro, J. Torres-Sospedra, & R. Montoliu (Eds.), *Geographical and Fingerprinting Data to Create Systems for Indoor Positioning and Indoor/Outdoor Navigation* (pp. 1-20). Academic Press. doi:<https://doi.org/10.1016/B978-0-12-813189-3.00001-0>
- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. 2011 International Conference on Computer Vision, (pp. 2564-2571). doi:[10.1109/ICCV.2011.6126544](https://doi.org/10.1109/ICCV.2011.6126544)
- Sithole, G., & Zlatanova, S. (2016). POSITION, LOCATION, PLACE AND AREA: AN INDOOR PERSPECTIVE. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-4, 89–96. doi:[10.5194/isprs-annals-III-4-89-2016](https://doi.org/10.5194/isprs-annals-III-4-89-2016)
- Spinoza Andreo, G., Dardavesis, I., de Jong, M., Kumar, P., Prihanggo, M., Triantafyllou, G., . . . Verbree, E. (2021). Building Rhythms: Reopening the Workspace with Indoor Localisation., (pp. 106–116). Retrieved from <https://repository.tudelft.nl/islandora/object/uuid%3A060d104f-bce9-4608-9aa9-a73132317254?collection=education>
- Wadden, R. A., & Scheff, P. A. (1983). Indoor air pollution. Retrieved from <https://www.bookdepository.com/Indoor-Air-Pollution-Peter-Scheff/9780471876731>
- Wan, T., Du, S., Xu, Y., Xu, G., Li, Z., Chen, B., & Gao, Y. (2019, December). RGB-D point cloud registration via infrared and color camera. *Multimedia Tools and Applications*, 78. doi:[10.1007/s11042-019-7159-6](https://doi.org/10.1007/s11042-019-7159-6)
- Willems, O. (2017, November). Exploring a pure landmark-based approach for indoor localisation. Master's thesis. doi:[10.13140/RG.2.2.29510.32329](https://doi.org/10.13140/RG.2.2.29510.32329)
- Yang, L., & Worboys, M. (2011, January). Similarities and differences between outdoor and indoor space from the perspective of navigation. Retrieved from [https://www.researchgate.net/publication/228843732\\_Similarities\\_and\\_differences\\_between\\_outdoor\\_and\\_indoor\\_space\\_from\\_the\\_perspective\\_of\\_navigation](https://www.researchgate.net/publication/228843732_Similarities_and_differences_between_outdoor_and_indoor_space_from_the_perspective_of_navigation)
- Zhou, Q.-Y., Park, J., & Koltun, V. (2016, October). Fast Global Registration., 9906. doi:[10.1007/978-3-319-46475-6\\_47](https://doi.org/10.1007/978-3-319-46475-6_47)