A Novel Approach for Dynamic Vertical Indoor Mapping through Crowd-sourced Smartphone Sensor Data

Georgios Pipelidis¹, Omid Reza Mosleh Rad², Dorota Iwaszczuk², Christian Prehofer¹, and Urs Hugentobler³

¹Technical University of Munich Software- and Systems Engineering Research Group Munich, Germany
²Technical University of Munich Lehrstuhl für Astronomische und Physikalische Geodäsie Munich, Germany
³Technical University of Munich Fachgebiet Satellitengeodäsie Munich, Germany

In this paper we present our developed and evaluated method for the dynamic mapping of the vertical characteristics inside a building. For achieving that, we extract data from smart-phone sensors and use those data for altitude estimation via the barometric formula. We introduce a novel approach for the extraction of reference pressure during the outdoor-to-indoor-transition of the user inside a building, which is identified through sensor fusion. A combination of machine learning techniques is used for the identification of the number of floors and the unsupervised classification of the altitude of each floor. As far as we know, this is the first system able of mapping vertical characteristics inside a building autonomously. Finally, enhancements on the CityGML model are made for mapping those characteristic by following its given standards.


I. INTRODUCTION

INDOOR maps gain popularity in our days, mainly due to their increasing applications. Indoor maps are widely used today in robotics, augmented reality, location based services and mobile ad-hoc networks. Even though there are proposed approaches for the dynamic generation of indoor maps from crowd-sourced data [1], [2], there are no approaches focusing on the dynamic vertical mapping from crowd-sourced data. In our research we aim to automate the indoor vertical mapping process with crowd-sourcing, while enriching existing maps with indoor information. In this way, we enable those models to carry information regarding the number of floors in a building and the corresponding altitude of every floor. This information is identified using state of the art methods for unsupervised classification. Another contribution of this paper is also the novel way we identify reference locations which can be used for precise altitude estimation, through the identification of the human transition from outdoors to indoors. Our approach - allthough is infrastructure independent - performs equaly or even outperforms existing approaches such as [3] and [4] that are infrastructure dependent, with absolute average error at 0.3 m vertical disposition.

More precisely, in this paper we present our method and its evaluation for the dynamic generation of vertical maps, using crowd-sourced data extracted from users via a smartphone application. We first use this data to identify the user transition from outdoors to indoors. This transition serves as a landmark for the extraction of a reference pressure. This extracted reference pressure is then used to estimate the altitude differences of every users step with the barometric formula. From the estimated altitude, the vertical transitions (e.g. stairs or elevators) are then filtered out, since they do not belong to floors. Finally, the remaining altitude values from multiple users is aggregated for the identification of the number of floors in a building and the height of each floor.

Various studies attempt to vertically localize humans or objects via pressure sensors [5], [4], [6]. However they all assume reference sensor stations permanently installed on the building. Hence, these are highly infrastructure dependent approaches. Additionally, several studies attempt - and not restricted - to vertically localize objects or humans, mostly via triangulating them using the WiFi Received Signal Strength [7], Cellular Network Antennas [8] or even specialized hardware such as Bluetooth low energy beacons [9]. Unfortunately, every triangulation method highly depends on the assumption of the existence of particular infrastructure as well as the line of sight. This implies that sensed values are influenced when the observer is standing in front of them or behind them. Finally, approaches for the dynamic generation of maps have also been proposed [10], [11], [12]. However, those approaches suggest to use outdoor characteristics for mapping indoors, which most of the cases is not feasible due to the uniformed shape of various buildings which does not allow any subspace discretization, while most of the times the building contains underground structures which cannot be recognized through any outdoor model (e.g. a subway station).

Thus, the contribution of this paper is listed as follows:

- We introduce an infrastructure independent method for the dynamic vertical mapping.
- We introduce a novel approach for the reference pressure estimation through recognition of the outdoor to indoor transition of the user. In this way, the need of calibration between sensors becomes obsolete.
- We propose a new enhanced CityGML model LoD2+ that provides indoor geometry of buildings at lower levels of detail.

Paper Structure

In this paper, the background of this study is provided in the second chapter. The approach is described in the third chapter.
The evaluation is presented in the fourth chapter. The related work is listed in the fifth chapter and the paper concludes in the seventh chapter, while future work is presented in the eighth chapter.

II. BACKGROUND ON THE BAROMETRIC FORMULA

The atmospheric pressure is the weight exerted by the overhead atmosphere on a unit area of surface. The barometric formula describes how this atmospheric pressure is reduced when the altitude is increased and vice versa. The unit of pressure is 1hPa = 1mbar = 100Pa.

The barometric formula reads:

\[ P = P_b \times \left[ \frac{T_b}{T_b + L_b \times (h - h_b)} \right]^{\frac{g_0 \times M}{R \times T_0}} \]  

(1)

where \( h_b \) is the reference altitude, \( T_b \) is the reference point temperature, \( L_b \) is the standard temperature lapse rate of 6.49K/km, \( P \) is the current pressure, \( P_b \) is the pressure in the reference point, \( R \) is the universal gas constants 8.3144621J/K/mol, \( g_0 \) is the earth’s acceleration 9.80665m/s² and \( M \) is the molar mass of Earth’s air 0.0289644kg/mol.

(1) can be altered for estimating altitude to the following:

\[ h = h_p + \frac{T_b}{L_b} \times \left[ \left( \frac{P}{P_b} \right)^{-\frac{g_0 \times L_b}{R \times T_0 \times M}} - 1 \right] \]  

(2)

According to the above formula 1mbar difference in pressure, with 15° ambient temperature, leads to 8.33m altitude change, while 1m change of altitude leads to 0.1201mbar change in pressure.

III. RELATED WORK

Enhancing CityGML models with indoor geometry is discussed in [10]. In this study, the level of detail two plus (LoD2+) is defined. This method is for automatic generation of indoor geometries based on CityGML LoD2. The method is robust and implemented successfully using Nef Polyhedra. However, they used some prior knowledge, such as building facades and available data modeled following the LoD2 format. As a result, this method is not applicable to general cases because not all the buildings contain sufficient information that can be used for mapping indoor areas.

In [3], they point out the need of vertical transition discovery due to the limitation of the ZUPT algorithm to identify vertical displacements. To solve this problem they introduce a moving platform detection module. It works by combining accurate sensors, and not those available on a smart-phone, such as accelerometer, barometer and magnetometer. They use ZUPT for localization and a SLAM algorithm for reducing the remaining drift. They estimate altitude using the barometric sensor, while they are also using it to identify instance phases. In addition, they attempt to identify the boundaries of vertical movements. The intuition for the use of acceleration for the detection of MPD is that the acceleration caused by external factors is weaker than the one caused by the pedestrian. However, their approach focuses on correcting real time localization and assumes the existence of indoor maps.

In [4], they suggest to use barometers for 2.5D (floor level) localization. They examine how barometric formula performs on height determination. They researched the robustness of altitude estimation on different devices where they record differences from 2.1hPa to 2.5hPa, which is translated to multiple floors offset. They noted that the variation of pressure in two hours could reach an equivalent of 10m height change. They also examine latency robustness as well as stability in short term, where they have noticed changes of 0.1hPa every 10min. Based on their experiments, they argue that it is impossible to accurately determine height, using barometer in indoor environment in an absolute manner. They strongly point out, the necessity of a reference station. In their study, they used a reference station from 5km away. However, a reference station is most of the times not available, while using other devices for reference requires calibration among those devices, which is not realistic to do in a real world scenario.

In [5], the authors propose the use of multiple barometers as reference points for the floor positioning of smart-phones with build-in barometric sensors. This method does not require knowledge of the accurate heights of buildings and stores. It is robust against temperature and humidity and it considers the difference in the barometric-pressure change trends and different floors. The intuition is that atmospheric pressure decreases as the altitude increases. Using a reference and the barometric formula, it is possible to calculate the altitude changes corresponding to the pressure. As they argue, humidity does not significantly affect the accuracy of the system in indoor altitude estimation; so, they use the gas constant for dry air and air molar mass of dry air instead of humid air. Based on the barometric formula and using built-in barometric sensors of smart phones as well as information from a local weather station, they are capable of achieving a good discretization between different floor levels. For the current temperature they consult local weather station online service. However, this approach is heavily depends on dense existing infrastructure, while it focuses only on localization and assumes the existence of maps which describe the location of each sensor.

Barometers are also used on medical applications [4]. In such applications, precise altitude estimation of the patient’s body is needed. Because all triangulation techniques require direct line of sight which is not easily predicted. The only alternative technique is the use of barometers. Exiting challenge in this case is the disturbances due to macroscopic flow, such as ventilation influence, opening and closing of doors or weather. Calibration between sensors is also needed, in order to compensate for the offset between different sensors. In their research, they created a small sensor network, with sensors are attached to the patient body as well as a reference stationary sensor. They measure a maximum error of 21cm, but they suggest that a second sensor might reduce the maximum error into 10cm. However, in our application scenario, we are not focusing on so accurate vertical localization, while we are looking for an infrastructure independent approach.
is applicable only outdoors, while even there its error can be 2.5 times the error of horizontal location. As a result they suggested barometers for vertical localization. Their main limitation is the lack of reference points, since the only available reference stations are meteorological stations, which are often sparsely located, while they broadcast periodically, usually at one-hour interval. As a result, they introduce the concept of ad-hoc reference points. They integrate information from multiple points, while they also use forecast models to estimate air pressure on demand.Besides reference meteorologic stations, they additionally use other smart-phones when the elevation indication is accurate enough. In order to retrieve better accuracy from other phones, first, they take into account all the reference points which are within a specified distance and time period and then they give higher weights to reference stations which are closer in distance as well as in time. They also assign different credibility to different reference stations. As a result a reference station will be more reliable if its location is known and can report better pressure. They score errors less that 3m in outdoor walking, 6m in mountain climbing, 0.9m in indoor floor localization. However, ad-hoc reference point will constantly have the need of calculating, which is not clear on how it can be achieved, especially without maps that describe those reference locations.

### IV. Approach

In this chapter, the main components of our approach are presented. As visualized in figure 1, the approach is composed of the **Sensor Data Collection** module, where data are collected from smart-phone users via an application which is developed for the purpose of this research and can be found here [15]. After smart-phone pressure sensor data are collected, in the **Signal Filtering** module, noise from collected data is filtered out. The **Reference Pressure Area** module has two roles, (1) to filter out data that belong outdoors and (2) to identify locations where pressure readings can be extracted. In the **Stair Removal** module the rejection of features that belong to intermediate heights (i.e. stairs or elevators) is made. Those pressure readings are later used in the barometric formula for **Altitude Estimation**. In the **Data Aggregation** module data from multiple users are combined, while the **Floor Estimation** module has two roles, (1) to identify the number of floors in a set, and (2) to estimate the altitude of each floor. Finally, in the **CityGML Generator** module, a **CityGML Model** is dynamically generated.

#### A. Sensor Data Collection

The sensor data collection module is collecting sensor data from pressure, light and GPS sensors. Data collected during different temperatures, days, times and humidity situations, labeled with timestamp and a unique user identifier are streamed on a server developed for this purpose, through a client server approach via HTTP protocol, in JSON format. The data, the precise collected date, time, the humidity, the temperature indicated by two open weather applications can be seen in the table I, while they are also openly available here [16].

#### B. Signal Filtering

For filtering out outliers the Savitzky-Golay filter [17] is used. Savitzky-Golay is a moving average filter which applies local regression to a subset of our entire data set. More specific, it smooths data by replacing each data point with the average of the neighboring data points, within a defined span. This approach is equivalent to:

\[
y_{s}(i) = \frac{1}{2N+1} \cdot \left( y(i+N) + y(i+N-1) + \ldots + y(i-N) \right)
\]

where \(y_{s}(i)\) is the smoothed value for the \(i^{th}\) data point, \(N\) is the number of neighboring data points on either side of \(y_{s}(i)\), and \(2N + 1\) is the span. A sample of this process can be seen in figure 2.

#### C. Reference Pressure Area

The reference pressure is essential for estimating the altitude differences with the barometric formula. The reference pressure is extracted from areas that cover the following preconditions: (1) are common for all user data of each building, (2) are located indoors and (3) the pressure fluctuations are low. Such area, is the one that follows the Outdoor-Indoor Transition (OITransition), since everyone inside a building was at some point in time outside, while it is located indoors where the pressure disturbances are low.
TABLE I: Collected Data used for Evaluation. The table shows the date of collecting the data, the time, the indicated temperature from AccuWeather (T A) and Google (T G) in °C, the humidity from the same two sources (H A) and (H G), the ambient pressure from AccuWeather (P A) and the building ID. While the buildings belong to the TUM main campus area and they are (1) Agness 27, (2) Adelheid 13A, (3) Agness 33 and (4) TUM Main Campus.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>T A</th>
<th>T G</th>
<th>H A</th>
<th>H B</th>
<th>P A ID</th>
<th>Date</th>
<th>Time</th>
<th>T A</th>
<th>T G</th>
<th>H A</th>
<th>H B</th>
<th>P A ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-May</td>
<td>10:20-10:30 AM</td>
<td>9</td>
<td>10</td>
<td>70%</td>
<td>74%</td>
<td>1011</td>
<td>9-May</td>
<td>10:10-10:20 AM</td>
<td>8</td>
<td>9</td>
<td>75%</td>
<td>60%</td>
<td>1017</td>
</tr>
<tr>
<td>10-May</td>
<td>18:20-18:30 PM</td>
<td>21</td>
<td>19</td>
<td>40%</td>
<td>52%</td>
<td>1004</td>
<td>10-May</td>
<td>16:40-16:50 PM</td>
<td>10</td>
<td>9</td>
<td>49%</td>
<td>59%</td>
<td>1016</td>
</tr>
<tr>
<td>9-May</td>
<td>17:00-17:10 PM</td>
<td>10</td>
<td>9</td>
<td>75%</td>
<td>60%</td>
<td>1017</td>
<td>9-May</td>
<td>17:50-18:00 PM</td>
<td>21</td>
<td>19</td>
<td>40%</td>
<td>43%</td>
<td>1004</td>
</tr>
<tr>
<td>9-May</td>
<td>10:40-10:50 AM</td>
<td>8</td>
<td>9</td>
<td>75%</td>
<td>72%</td>
<td>1017</td>
<td>11-Feb</td>
<td>14:30-15:00 PM</td>
<td>6</td>
<td>2</td>
<td>70%</td>
<td>72%</td>
<td>1019</td>
</tr>
<tr>
<td>9-May</td>
<td>17:30-17:40 PM</td>
<td>10</td>
<td>11</td>
<td>49%</td>
<td>55%</td>
<td>1016</td>
<td>12-May</td>
<td>19:00-20:00 PM</td>
<td>7</td>
<td>0</td>
<td>87%</td>
<td>80%</td>
<td>1028</td>
</tr>
<tr>
<td>10-May</td>
<td>22:00-22:10 PM</td>
<td>11</td>
<td>11</td>
<td>61%</td>
<td>65%</td>
<td>1006</td>
<td>12-Feb</td>
<td>21:30-00:00 PM</td>
<td>7</td>
<td>0</td>
<td>93%</td>
<td>83%</td>
<td>1017</td>
</tr>
<tr>
<td>12-May</td>
<td>18:30-18:40 PM</td>
<td>21</td>
<td>19</td>
<td>40%</td>
<td>45%</td>
<td>1004</td>
<td>21-Mar</td>
<td>13:30-14:00 PM</td>
<td>13</td>
<td>8</td>
<td>58%</td>
<td>64%</td>
<td>1010</td>
</tr>
</tbody>
</table>

As [18] describes, the OITransition can be identified by aggregating multiple smart-phone sensor data. For example the ambient light sensor can be used to identify such transition since there is a difference on the light intensity during the day as well as in the frequency due to the alternate current variation, during the night. For identifying the outdoor-indoor transition, in our research we used the following:

- The rapid increase of the GPS uncertainty with approximately 1 Hz sampling frequency. As can be seen in figure 3 at the moment of the transition (100th sample), the GPS uncertainty can be increased from less than 10 m to almost 60 m. Hysteresis thresholding [19] is applied for the maximization of the margin between low GPS accuracy (indoors) and high accuracy data (outdoors) for better classification. Finally, steep curves of the signal are estimating from the derivative between successive sensor readings as follows:

\[
x'[k] = \frac{x[k] - x[k - 1]}{\Delta t}
\]

\[
\Delta t = \text{time between two samples}
\]

Fig. 3: GPS uncertainty data from five OITransitions. As can be seen, the moment of the transition at the 100th sample, the uncertainty rapidly increases.

- Light and proximity sensor, with 7 Hz and 25 Hz respectively, are fused to identify the transition. As can be seen in the figure 4, light intensity drops when crossing indoors during the day and increases during the night, while the proximity sensor indicates whether to trust the light sensor, due to various phone poses (e.g. phone in pocket). Hysteresis thresholding is used for maximizing the margin of signal that belong outdoors and indoors. Finally binary classification is made based on high and low frequencies, while for the decision of whether the collected data belong to day or night is taken as follows:

\[
\cos \omega_0 = -\tan \phi \cdot \tan \delta
\]

where \(\omega_0\) is the hour angle at either sunrise (when negative value is taken) or sunset (when positive value is taken), \(\phi\) is the latitude of the observer on the Earth and \(\delta\) is the sun declination.

Fig. 4: Light data from five OITransitions. As can be seen it the 70th sample, which corresponds to the outdoor-indoor transition, the light intensity is rapidly dropping.

Finally, the results are fused following the majority approach. Hence the decision is taken based on the result only when both of our sensors agreed.

D. Stair Removal

In the stair removal phase, sets of features with high disturbances are rejecting, since they mostly correspond either to vertical transitions (e.g. stairs or elevators) or to outliers (e.g. high wind velocities).

This approach is equivalent to:

\[
\sigma = \sqrt{\kappa}
\]

where:

\[
\kappa = \frac{1}{N-1} \left( q - \frac{s^2}{N} \right)
\]

where:

\[
q = \sum_{i=1}^{N} x_i^2 \quad \text{and} \quad s = \sum_{i=1}^{N} x_i
\]

Finally, the results are fused following the majority approach.
TABLE II: The data set used for evaluation. As can be seen data collected from five different days are mixed to five new data sets which simulated limitations such as not visited floors, same floor visited different days and different floors visited by different people.

<table>
<thead>
<tr>
<th>Floor</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>1</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>2</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>3</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>4</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

where \( x_i \) is the instance of the input signal and \( N \) the number of elements. A sample of this process is visualized in the figure 8, where as can be seen pressure readings (blue line) which correspond to vertical transitions, hence high standard deviation (red line), have been rejected.

Fig. 5: The STD estimated from the filtered pressure sensor. As can be seen the segments correspond to high disturbances, cause due to vertical transitions, have been successfully filtered out.

E. Altitude Estimation

The altitude is estimated by (2) as follows:

\[
h = \left( \frac{P_0}{P_i} \right)^{\frac{1}{\gamma}} - 1 \cdot T_b + 273.15 \frac{T_0}{0.0069} \tag{3}
\]

where \( P_0 \) is the reference pressure extracted from the OITransition, \( P_i \) is the current pressure value and \( T_b \) is the temperature value, which is extracted via online openly available weather stations.

F. Data Aggregation

Data aggregation is essential for identifying all floors inside a building to aggregate multiple user data, since not all users are expected to visit all floors. In the data aggregation module, multiple recorded data are fused. Grouped by a GPS radius, the data always correspond to the same building. More specific, altitudes estimated from multiple users, via their unique users identifier (UUID) and shorted by their timestamp and fused together for the classification phase. Since the reference pressure for the altitude estimation is extracted by the same device with the one used for estimating it, and approximately the same location with all the users, due to the novel approach for reference altitude extraction via the outdoor indoor transition discovery, there is no need of calibrating any sensor along different phones. In this paper we assume that every entrance of a building is at the same altitude, although the variant altitude case is answered in the future work chapter.

G. Number of Floor Estimation

Since the number of floors as well as the label of every floor (i.e. the corresponding altitude) are unknown, for classification we used a classifier able to cope with unlabeled data. The classifier K-Means is selected because of its popularity and its relatively low processing demand. For estimating \( K \) the Elbow method is selected.

The classification process can be divided into two main steps. The first step is the identification of \( K \), which corresponds to the real number of floors. In the second step the center of each cluster is recognized, which correspond to the altitude of every floor.

Identification of \( k \)

As the number of floors is unknown, in the first step, the estimation of \( K \) has to be made. For this purpose the elbow method \([20]\) is chosen. The elbow method is designed to help finding the appropriate number of clusters in a data set. It is a method of interpretation and validation of consistency within cluster analysis. As can also be seen in the algorithm 1, the elbow method looks at the percentage of variance explained as a function of the number of clusters: The optimum number of clusters is revealed, since adding another cluster doesn’t give much better modeling of the data. If we plot the variance against the number of clusters, the first clusters will add much information but at some point the marginal gain will drop, giving an angle to the graph. As a result, the number of clusters is chosen at this point.

Identifying \( K \) is essential, since it corresponds to the number of floors. As a result a wrong estimation of \( K \) can lead to an enormous error which does not reflect the real number of floors as well as the real altitude of every floor.

The Center of Clusters

After the \( K \) is identified the classification is made using K-Means, since the cluster label (i.e. each floor altitude) is unknown. The input to the algorithm is the computed vector of filtered pressure data and the \( K \) estimated above. The algorithm’s output is then a vector with the assigned classes for every input point and the cluster centroids.

H. Implementation in CityGML

In our research, we concentrate on the derivation of the floor number and their height. This does not allow us to create a complete LoD4 model. As a result, we enhance LoD2...
model geometry with the hull geometry for each floor. For this purpose, we introduce LoD2+, as visualized in Figure 6.

In LOD2 and higher LODs the outer facade of a building can be modeled semantically by the _BoundarySurface. _BoundarySurface is a part of the building’s exterior shell with an assigned function like wall WallSurface, roof RoofSurface, ground plate GroundSurface, outer floor OuterFloorSurface, outer ceiling OuterCeilingSurface or ClosureSurface. For indoor modelling FloorSurface, InteriorWallSurface, and CeilingSurface can be used [21]. In [10] the authors enhance the CityGML scheme with a new feature class Storey which has five attributes: class, function, usage, storeyHeightAboveGround and storeyVolume.

To model the indoor geometry, we keep LoD2 representation using _BoundarySurface and add indoor geometry for each storey using FloorSurface, InteriorWallSurface, and CeilingSurface as well as feature class Storey introduced by [10]. In addition, we propose a further attribute of the feature class Storey storeyAltitude. This attribute is necessary for our application, as the output of a navigation device is an altitude and not the height above the ground. This extension is not included in the current version of the CityGML specification, however we suggest to include it in the next release.

For the dynamic generation of the CityGML model citygml4j [22] is used. It is an open source library for Java which binds the XML Schema definitions of CityGML to a Java object model.

V. EVALUATION

This chapter presents the evaluation of the proposed method for the dynamic vertical mapping from user smart-phone data. In [V-A] in the evaluation of various phone poses, [V-B] presents the evaluation for the human velocity robustness, [V-C] OITransition recognition and [V-D] floor number identification and altitude estimation on various weather conditions. As visualized in the table [1] twelve data sets are obtained from three different buildings, during four different days with different ambient pressure and temperature conditions. Additionally, five data sets were obtained with five different phone poses and the same weather conditions and routes, for testing the robustness of the sensor itself on different phone poses. Data were also collected from three different human walking velocities for the evaluation of the stair removal algorithm.

A. Influence of Different Phone Poses

A pre-study is conducted where the influence of different phone poses is estimated. Those poses are: (1) phone on hand, (2) phone in jacket pocket, (3) phone in a random position (i.e. multiple poses are involved), (4) phone in the swinging hand and (5) phone in trouser pocket. It is found that the influence on pressure due to different smart-phone poses is minor, as can be seen in figure 7. Considering the fact that the maximum error along 600 unique pressure readings each of the five tested poses does not exceeds 30 cm maximum vertical displacement.

B. Evaluation of Stair Removal

For evaluating the efficiency of the stair removal method, the data are collected from the same route in three different visits and walking velocities, approximately 1x, 1.5x and 2x.

For evaluating the efficiency of the stair removal phase, we recorded data with three different walking velocities, approximately 1x, 1.5x and 2x, as can be seen in [8] while climbing five pairs of stairs on a building. As demonstrated in
the results, the algorithm scores a precision of 93.8%, recall 95.5% and F-Score 94.6%, on correctly identifying the stairs, with the same sliding window length for all data sets.

<table>
<thead>
<tr>
<th>TABLE III: Confusion Matrix of Stair Removal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Floors</td>
</tr>
<tr>
<td>Stairs</td>
</tr>
</tbody>
</table>

C. Evaluation of Reference Pressure Area

As described in the chapter V-C, the reference pressure value which is used for the altitude estimation with the barometric formula, corresponds to the location that follows the entrance of a building. For the identification of the building entrance the outdoor-to-indoor-transition is taken into consideration, where the transition is identified similar to [18], as detailed described in [V-C] in our scenario, the ambient light and the GPS uncertainty are taken into consideration. The approach is evaluated in three different buildings with four collected data sets for each building, during day and night. The collected data and the algorithm used for the evaluation are open sourced and can be found here [16].

As can be seen in the figure 9, the OITransition (red dots) is successfully identified in all trials of the particular data set, while the location of the entrance can also be approximated in 1.6 m in latitude and 5.5 m in longitude.

![Fig. 9: Locations which correspond to the detection of the OITransition. The figure includes nine different determined locations for the entrance to the building (red dots).](image)

For better evaluation of this component, data were combined, as visualized in the Table II to five new incomplete data sets for testing the robustness of our algorithm in the following two challenges. The first, is the problem of obtaining five data sets from five people walking on five different floors during five different days. The second, is the problem of obtaining data from five people walking on the same floor during five different days. An example of the constructed data set follows: the data set a consist of the 0 and 1st floor on the 1st day, the 0, 1st and the 2nd floor on the second day, the 0, 1st and 3rd floor on the third day, the 0, 1st and 4th floor on the fourth day and the 0, 1st and 5th floor on the fifth day.

![Fig. 10: The result of the elbow method for five constructed data sets. The data set used for this evaluation as well as the algorithm can be found here [16]. Those data are from different days and floors, hence for identifying the correct number of floors there is only the option of aggregation of data.](image)

The number of floor estimation process is divided into two main steps. In the first step, the number of classes is estimated, as explained in detail in section IV-G. In the second step the altitude of each floor is estimated. For the first step, the elbow method is used. As can be seen in figure 10, the distortion threshold for the elbow method is decided at 99.2% for all buildings, while K is altered from 2 to 10.

As can be seen, the uniformity of the data in each cluster, expressed by the percentage of their variance, does not further altered after 5 (K = K_temp + 1 = 6, because of the way the algorithm works [20]). The algorithm is successful on identifying the correct number of clusters (floors) in all data sets.

<table>
<thead>
<tr>
<th>TABLE IV: Building I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>indoor</td>
</tr>
<tr>
<td>outdoor</td>
</tr>
</tbody>
</table>

After the number of clusters (floors) is identified by the elbow method, the altitude of each floor is estimated. As can be seen in the figure 11, the ground floor is defined at 0 altitude, the 1st floor is predicted at 5.2 m the 2nd floor at 10.36 m, the
3rd floor at 14.47 m, the 4th floor at 18.7 m and last the 5th floor is predicted at 23.37 m. The absolute maximum error is at 1.4 m, with a standard deviation of 0.5 m.

![Floor altitude vs. Clusters](image)

**Fig. 11:** Floor height estimation for all data sets that correspond at the same building. The building consist of 5 floors.

**Conclusions**

To conclude, this paper describes our developed and evaluated framework for the dynamic mapping of the vertical characteristics of a building. For achieving such a goal we had to manage uncertain sensor data, collected via crowdsourcing. We estimated the altitude of each of these collected data sets via the barometric formula. For achieving this, we introduced a novel approach for the extraction of reference pressure during the outdoor-to-indoor-transition of the user inside a building, which is identified through sensor fusion. Additionally, we faced an unsupervised classification problem, where the number of floors as well as the altitude of each floor are unknown. Finally, we proposed a way to map those characteristics enhancing the standards of the CityGML.

**Future Work**

In the future we will put the focus on generalizing our model to enable it to dynamically vertical map the indoor of any type of building, by including mechanisms able to classify between pressure changes due to vertical displacements and wind changes, as well as able of enhancing existing models such as the CityGML, with longitude, latitude and altitude information extracted through crowd-sourcing via the GPS sensor and our OITransition model. In this way the coordinate system model will become easier convertible to the global positioning system coordinates while becoming robust against entrances located at different altitudes. By applying architectural knowledge, the robustness over the lack of data can be addressed (e.g. usual floor height). A more robust model for the reference point selection which might take into consideration intermediate floors can also be considered.

**Acknowledgment**

This work is part of the TUM Living Lab Connected Mobility project and has been funded by the Bayerisches Staatsministerium für Wirtschaft und Medien, Energie und Technologie.

**References**


