AN ALGORITHM DETERMINING THE FRONT EDGE OF BUILDINGS TO PLACE POIS

Fatih Gülgen, Batuhan Kılıç, Özge Öztürk

Associate Professor Dr. Fatih Gülgen;
Yildiz Technical University, Geomatic Engineering Department;
Istanbul, Turkey;
+902123835443, fgulgen@yildiz.edu.tr;

PhD student Batuhan Kılıç;
Yildiz Technical University, Geomatic Engineering Department;
Istanbul, Turkey;
+902123835326, batuhank@yildiz.edu.tr;

MSc student Özge Öztürk;
Yildiz Technical University, Geomatic Engineering Department;
Istanbul, Turkey;
+902123835340, ozgeo@yildiz.edu.tr;

Abstract
Points of interests (POIs) describe a geographic entity that users are focused on such as a school. The different types of POIs are represented by the cartographic symbols. Its positional accuracy on a map is usually considered good if the symbols are inside a polygon feature. However, the correct location for the POIs is the place close to the front edge of building footprints. In this study, we aim to develop an algorithm which identifies the possible front edges of buildings by considering the geometry of the network and building blocks in urban areas. The building blocks are derived from amalgamating the footprints near to each other in the partitions obtained from the street network. The each line segment is extracted from the blocks, and its probability of being on the front edge is calculated. Finally, the results are compared by the real statuses obtained from the street perspectives.

Keywords: VGI, POI, Geocoding, Navigation, Map symbol, Building footprint,

INTRODUCTION

Points of Interest (POI) are zero-dimensional features which refer to specific locations or real-world entities in geographical space, such as historical sites, landmarks, public services, shops, restaurants, or bars etc. (De Tré, 2013; Jonietz and Zipf, 2016). The POIs are the basis for most of the data supporting location-based applications. The different types of them are represented by the cartographic symbols which are designed to illustrate actual entities. They easily crowdsourced by citizens, and are often a major part of the volunteered geographic information (VGI) projects supported by open or commercial organizations. In the data-rich environment provided by VGI projects, one of the fundamental issues is the quality of POIs. There are several studies focused on the quality of POI datasets based on the different assessment methods such as the measuring completeness index (Mashadi et al. 2015), the logical consistency (Corcoran et al. 2010), the actuality, the thematic accuracy (Fan et al. 2014) and the positional accuracy (Hochmair and Zielstra, 2012; Arsanjani, et al 2013).

In most cases, the quality of the crowdsourced data is assessed by a reference dataset obtained from authoritative or commercial sources. For instance, Girres and Touya (2010), made a comparison between OSM POIs and IGN BD topo maps to determine the positional accuracy, the semantic accuracy and the currentness of datasets. Typical indicators for the positional accuracy are Euclidean distances between co-referent points, distance deviations on the X- and Y-axis, or the evaluation whether, and in case of line or polygon features how much of, a feature from a crowdsourced dataset is located within a certain buffer zone computed around a reference feature (Jonietz and Zipf, 2016). Touya et al. (2017) examined the spatial relationships between the POIs and its expected surrounded features to determine the positional accuracy. They evaluated that the quality of a POI is poor, if there are impossible or improbable spatial relations within
the features, e.g., a shop POI is outside a building. This use of spatial relations to guide the quality assessment were also implemented in Touya et al. (2013).

In such case, the main problem is that POIs are points that represent a geographic entity that has a geographic extent often much larger than the point (Touya et al. 2017). Commonly, the volunteer citizens commonly place them directly on top of the specific footprint or directly in front of it (Hart and Zandbergen 2013). These are valid and acceptable representations. The positional accuracy is usually considered good if the symbols are usually inside a building footprint that is used as the reference spatial data by capturing the edge of the roofs from high resolution georeferenced satellite images. There are two ways used to place the cartographic symbols on the building footprint: put the POI at the center of the footprint, or at the entrance of the building (Touya et al. 2017). However, the correct location for the POIs is the place close to the front edge of the building footprints. The front edge is determined by considering the actual status of the building's main entrance door. This is usually not considered by the citizens. Also, it is very important issue that needs to be taken into consideration in the consistency of geocoding and navigation applications for the online map services such as OSM, Google, Bing, etc. In this study, we aim to develop an algorithm which identifies the possible front edges of buildings by considering the geometry of the network and building blocks in urban areas. Accordingly, the characteristic points where a volunteer can place the cartographic symbol to the entrance of the building will be presented as suggestions. The user will be able to place the symbol on the building footprint properly by snapping the most appropriate between the suggestions.

METHODOLOGY AND DATASET

The algorithm proposed in this study depends on the assumption that the front edge of the footprint is the building's main entrance. The front edge is defined according to the building and street features in the topographic dataset. The characteristic points for the main entrance are determined throughout the front the building’s edge.

Derivation of the Main Entrance Candidates for POIs

An urban area where the POIs are located consists of a large number of adjacent buildings touching each other. Since the front entrances are certainly on the street sides, the edges of them are principally to be detected. However, none of the back holes and adjacent sides of the building are a candidate edge. Firstly, a large block is derived from the amalgamated footprints adjacent to each other in order to determine the correct edges where the POIs can be placed (Figure 1a). The outer boundary of the large block (indicated in red lines) includes the candidate edges of each building. The inner polygon (painted in blue) is a hole which is far from the main entrance of buildings. The common boundaries between the footprints (indicated in green lines) corresponds to two adjacent buildings. Any POIs placed close to the common boundaries causes confusion.

![Figure 1. Amalgamated buildings with the candidate edges](image)

Secondly, two feature classes in line geometry are derived from the outer boundary of the large block. The first feature class contains the arcs generated by splitting the polygon boundaries to polylines at their intersections. Each polyline refers to a front edge of a specific building. The second feature class contains the straight line segments generated by...
splitting the boundaries of blocks at their vertices. Each line segment can be considered as potentially to correspond to the main entrance of the buildings.

Weighting the Line Segments

In this study, six measures are used to determine the most suitable line segments for POIs. They are extracted from the geometries of objects in the two feature classes mentioned above and the street network. These measures are:

- **The length of polyline**: The measure shows the length of each building's edge which is stored in the first feature class and which is close to the street(s). For the two buildings painted in pink and orange in Figure 2, the values are based on the lengths of purple and blue polylines respectively. If there are two adjacent buildings on either side of a building (painted in pink), the edge is only associated with one street (indicated in brown polyline). The number of associated streets for a corner building (painted in orange) can be more than one (indicated in brown and gray polylines).

- **The sinuosity of polyline**: This is a measure which shows the straightness of each edge. It is calculated by dividing the length of the straight line connecting the two end-points of a polyline by the length of the polyline. The low sinuosity value approaching to zero indicates that the building edge has more than two vertices and that the straightness is decreasing. If an edge has only two endpoints without any vertices, the sinuosity is equal to one. The sinuosity values for discrete buildings are equal zero.

- **The length of line segment**: The measure shows the segment lengths of each building's edge in the second feature class. The short line segments usually show the little details on the building edge. The longer segments can be considered to be more important than the shorter ones.

- **The closest distance to street**: This measure is based on the distance between a line segment and its the closest street feature in the network. The main entrance of a building is usually located near the street feature. The building's edge, which is closer to the street, is more suitable for placing the POIs than the farther ones.

- **The length of street**: This measure is based on a length value obtained from the street geometry closest to the line segments in the second feature class. It can be viewed as an indicator revealing the importance of streets, if there are alternative conditions in which the building's edge correlates with different streets (e.g. the brown and gray lines for the orange building in Figure 2).

- **The parallelity**: This is a measure indicating the state of line segment, which is expected to be parallel to the closest street feature. If the two lines (i.e. the edge segment of a building and its the closest street feature) are as parallel as possible to each other, the segment is suitable for placing the related POI. Otherwise, (i.e. if the two lines are perpendicular to each other), it is not suitable.

Study Area

The methodology is elaborated in an experimental testing region: Cihangir neighborhood which reflects an urban character on the European side of Istanbul province in Turkey (Figure 3). The OSM POIs for the Cihangir region are used with the source GIS data that are produced by the GIS directorate of Istanbul Municipality in testing study (Figure 3).
4). The GIS data covers an area of approximately 11.2 ha and contains 473 buildings (indicated in orange polygons) with a total area of 6.3 ha. The total length of road network (indicated in black lines) is 5.4 km in that region. While there are major shortcomings in Istanbul’s OSM data, the region is quite complete because of the activities of the OSM community.

*Figure 3. Location of the study area*

*Figure 4. OSM data overlaid with the GIS data in Cihangir, the borders were drawn by red*

**Case Study**

Each measure of the line segments was calculated and then stored in the attributes of the second feature class (Figure 5). SEG_ID contains the unique numbers that identify each line segment. PLY_ID contains the identification number for polylines in the second feature class. There is one to many relationships between the first and the second feature classes.
A polyline can consist of multiple line segments. BUILD_ID contains the identification numbers for each building footprint. One to many relationships between the line segments and the building footprints is also valid. SEG_LEN contains the lengths of each line segment. PLY_LEN contains the lengths of each polyline. SINUOSITY contains the sinuosity values of each polyline. NEAR_STR_ID contains the identification numbers of each street which is close to the associated line segment. STR_LEN contains the lengths of each streets. CLO_DIST contains the closest distances between the line segment and its closest street. PARALLELITY contains the angle values between the line segment and its associated street. An angle value close to zero means that the line segment and the street are parallel to each other. The values close to ninety degrees mean that the line is perpendicular to the street. Then, the values of these attributes are normalized in accordance with each building feature (Figure 6).

**Figure 5. Attributes of line segments**

<table>
<thead>
<tr>
<th>Seg_ID</th>
<th>Ply_ID</th>
<th>Build_ID</th>
<th>Seg.Len</th>
<th>Ply.Len</th>
<th>Sinuosity</th>
<th>Near_Str_ID</th>
<th>Str.Len</th>
<th>Clo.Dist</th>
<th>Parallelity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>455</td>
<td>11.12263</td>
<td>38.410226</td>
<td>0.205768</td>
<td>76</td>
<td>217.47456</td>
<td>3.047724</td>
<td>0.279263</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>466</td>
<td>4.5765748</td>
<td>38.410226</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>467</td>
<td>10.53193</td>
<td>38.410226</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>468</td>
<td>8.7513125</td>
<td>38.410226</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>469</td>
<td>38.956682</td>
<td>38.410226</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>167</td>
<td>6.5411477</td>
<td>6.821447</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>167</td>
<td>5.9813425</td>
<td>6.821447</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>167</td>
<td>0.8163271</td>
<td>6.821447</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>57</td>
<td>0.164265</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>57</td>
<td>2.9314249</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>57</td>
<td>0.5736862</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>57</td>
<td>0.59329</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>57</td>
<td>0.8383271</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>57</td>
<td>0.340513</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>57</td>
<td>1.150647</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>57</td>
<td>0.3550261</td>
<td>6.8794534</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>176</td>
<td>6.2914282</td>
<td>20.65913</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>176</td>
<td>2.3714282</td>
<td>20.65913</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>176</td>
<td>1.365459</td>
<td>20.65913</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>176</td>
<td>2.322398</td>
<td>20.65913</td>
<td>0.205768</td>
<td>76</td>
<td>227.47245</td>
<td>3.056611</td>
<td>0.279263</td>
</tr>
</tbody>
</table>

**Figure 6. Normalized attributes of line segments**
In order to calculate the priority values of the line segment for a building, the attributes are weighted. The values of weights are selected according to existing experiences: the closest distance to street is 0.35; the parallelity is 0.35; the length of line segment is 0.10; the length of street is 0.10; the sinuosity of polyline is 0.05; the length of polyline is 0.05. The highlighted line segments with the highest priority value for each building in the study area are shown in Figure 7.

Figure 7. The most important line segments for placing POIs

Figure 8 shows the POIs derived from the two street perspectives on the highlighted streets in the test area. They are shown in the red points. Additionally, The POIs collected from the OSM database are shown in the blue points. A total of 62 POIs have been identified from the perspective views. The eleven is located on the street intersection. They are known as the corner buildings. Only one of them is on the most important line segments. The remaining ten are on the second important line segments. 51 points are on the attached buildings on a single street. 45 of these are on the most important line segment. Only 6 points are on the second important line segment.

Ten of the eleven POIs in the corner buildings are visible on the OSM. One of them has not been marked on the database yet. The 44 of 51 POIs found in the attached buildings are also available in the OSM database. The 7 POIs has not been marked on the OSM yet. When the POIs derived from the street perspectives are used as references, the mean deviation distance for a total numbers of 54 POIs in the OSM database are shown in Table 1 together with the mean deviation of the line segments that are found to be important. The results show that the first-order line segments, identified by the proposed method, are very close to the front edge of the buildings, which are located on a single street. However, it may not be suitable for the corner buildings, because they have at least two facades to be used as the entrance. In this case, the second-order line segments may be suitable for the main entrance.
Figure 8. The POIs collected from street perspectives (red) and OSM (blue) on the two streets (highlighted)

Table 1. The mean deviation distance between the references and the derived data

<table>
<thead>
<tr>
<th>Buildings located on</th>
<th>Numbers</th>
<th>Distances between the referenced POIs and the line segments</th>
<th>Distances between the referenced POIs and the OSM data</th>
</tr>
</thead>
<tbody>
<tr>
<td>junction</td>
<td>10</td>
<td>5.6</td>
<td>5.7</td>
</tr>
<tr>
<td>a single street</td>
<td>44</td>
<td>0.2</td>
<td>3.6</td>
</tr>
</tbody>
</table>

CONCLUSION

POIs are one of the most important components contributing to the development of digital map services. The correct location for POIs is the place close to the front edge of the building footprints. The front edge is determined by considering the actual status of the main entrance. In this study, a method, which identifies the possible front edges of buildings by considering the geometry of the street network and building blocks in urban areas, is proposed. The first proposed line segment, one of the possible parts of the front edge of a building, is generally close to the entrance, if the building's facade is looking at a single street. For the corner buildings which have at least two facades to be used as the entrance, the second proposed line segment can be located close to the entrance of a POI. In order to solve this problem, it is necessary to evaluate the buildings in two different conditions according to the number of facades looking at the streets.

REFERENCES


Jonietz, D. and Zipf, A. (2016), Defining fitness-for-use for crowdsourced points of interest (POI). ISPRS Int. J. Geo-Inf., 5, 149.


**BIOGRAPHY**

Fatih Gülgen is Associate Professor of Geomatics Engineering at Yildiz Technical University, Turkey. His main research interests focus on cartographic generalization, digital terrain modeling and algorithmic foundations of geographic information systems.

Batuhan KILIÇ is Geomatic Engineering researcher and Ph.D. student at Yildiz Technical University, Turkey. His master thesis focused on online geocoding approaches.

Özge Öztürk is Geomatic Engineering researcher and MSc student at Istanbul Technical University, Turkey. Her master thesis focused on space syntax.