SHAPE ANALYSIS OF MULTI-SCALE BUILDING FEATURES

Melih Basaraner and Sinan Cetinkaya

Yildiz Technical University, Department of Geomatic Engineering, Division of Cartography, 34220 Esenler, Istanbul Turkey, Tel.:+90 2123835342-45, E-mail: {mbasaran, cetinkaya}@yildiz.edu.tr

Abstract
Buildings can generally be represented as individual polygons until 1:100 000 in topographic maps. As scale decreases, the level of detail of the features is reduced with generalization to preserve the readability on map. In this context, shape analysis is important because it can be used in generalization operator and parameter selection, generalization quality evaluation as well as scale inference for VGI and remotely sensed data. Shape analysis of polygons are often carried out with various shape indices. A shape index usually measures a specific shape characteristic of a polygon. In this paper, it is aimed at investigating the response of different shape indices to the level of detail variation in building features depending on scale. For this purpose, five shape indices were used and evaluated for measuring shape complexity of individual building features at six different scales ranging from 1:1 000 to 1:100 000.

Keywords: Shape analysis, Shape indices, Multi-scale, Cartographic generalisation, Building features

INTRODUCTION
Shape is a significant characteristic that needs to be analysed in order to understand spatial phenomena better (Wentz, 1998). Shape has both geometric and semantic nature. Geometric properties/elements of a spatial feature’s own geometry and auxiliary geometries are usually employed to quantify the shape while the semantic definition of the feature is utilised to interpret the shape. Geometric properties/elements include area and perimeter etc., and the auxiliary geometries include convex hull, minimum bounding rectangle (MBR) etc. while the semantic definition corresponds to the feature types (classes) which have usually specific shape characteristics. There are many methods for quantifying shape. Among them, shape indices are widely used in spatial applications. They usually measures a single shape characteristic of a polygon. In this paper, it is aimed at investigating the response of different shape indices to the level of detail variation in building features depending on scale. For this purpose, five shape indices were used and evaluated for measuring shape complexity of individual building features at six different scales ranging from 1:1 000 to 1:100 000.

Spatial features are scale-dependent because different spatial applications usually involve analysing and visualising spatial phenomena at different levels of detail (Li, 2007; Mackaness et al., 2014). To be specific, as the scale of a spatial dataset decreases, the level of detail of the feature is reduced. Transitions from larger to smaller scales are performed with generalization. When dealing with maps, map features are subject to cartographic generalisation during scale transition. Cartographic generalisation is guided by geographic meaning, i.e. geometric, structural and semantic characteristics of and relationships among map features, and graphic limits, i.e. minimum discernible sizes of and distances between map features. In other words, during the transfer of the features from source to target scale, both aspects have to be taken into account. Extracting geographic meaning involves applying structure recognition techniques to the spatial dataset used. Spatial structure can be differentiated depending on the analysis level. When it is at the feature level, it corresponds to the shape of a feature. On the other hand, it corresponds to the relationships among a group of features belonging to same and/or different feature classes when it is at inter-features level. In this respect, being a structural characteristic of a feature, shape is a significant characteristic about geographic meaning that needs to be analysed before, during and after generalisation. Apart from generalization, shape analysis can contribute to scale inference for volunteered geographic information (VGI) and remotely sensed data.

There are very few studies on scale-based shape analysis. Paszto et al. (2015) use shape metrics originally developed for landscape ecology to assess their response to the generalised geometries of building footprints. In this context, they create 22 levels of detail by applying a simplification algorithm to four buildings using different tolerance values. In their study, there is no connection with map scale and it seems that they experimentally choose the tolerance values. Touya and Reimer (2015) present a study to assign a scale to individual OpenStreetMap (OSM) features. They use several parameters such as vertex density, median edge length and smallest edge. However, they do not use any shape index in their processes. Remaining studies focus solely on shape analysis and they do not make any connection with scale or level of detail. MacEachren (1985) compares, categorises and evaluates various compactness indices. Medda et al. (1998) extend Boyce and Clark’s shape index to recognise and classify urban shapes. Wentz (1998) examines the need for a shape analysis capability in GIS. Zhang and Lu (2004) classify and review significant shape representation
and description techniques in the literature. Yang et al. (2008) shortly describe and compare shape-based feature extraction and representation methods. Basaraner and Cetinkaya (2017) investigate the performance of 20 shape indices and proper shape index-classification scheme pairs to characterise shape complexity of building footprints in GIS. Among these indices, two of them are new.

There is not enough study that analyses shapes of multi-scale features. Therefore, the aim of this study is to investigate the response of shape indices to the level of detail variations of spatial features during scale transitions. For this purpose, five shape indices were selected from the literature and used for measuring shape complexity of individual building features at six different scales ranging from 1:1 000 (1:1K) to 1:100 000 (1:100K).

**METHODOLOGY**

This study consists of two main phases: 1) Deriving multi-scale building features through generalisation, 2) Applying shape indices to quantify the shapes of building features at multiple scales.

**Multi-scale building features through generalisation**

In the first phase, buildings are simplified with an algorithm specifically designed for buildings which are usually of rectangular shapes (Wang and Lee, 2000). The simplification tolerance is set according to the graphic limits specific to buildings (after Basaraner and Selcuk, 2008) (Figure 1). To be specific, the minimum granularity threshold value is used as the simplification tolerance. In fact, the threshold values about graphic limits are those that have to be satisfied on a map at target scale, so these values are multiplied with target scale factor to find the parameters that are applied during simplification because a GIS works with real world coordinates. In this way, graphic limits are converted to the threshold values about graphic resolution.

**Shape analysis of multi-scale building features**

In the second phase, shape index values are automatically computed in GIS through scripts. In this context, five shape indices are used given in Table 1. The geometric properties/elements of the polygon’s own geometry and auxiliary geometries form the basis of shape index computation. The geometric properties/elements include area, perimeter, length/distance and coordinates and the polygon’s own geometry and auxiliary geometries include vertices, interpolated boundary points (IBPs), centroid, radial lines, equal-area circle (EAC), convex hull (CH), minimum bounding rectangle (MBR), minimum area bounding rectangle (MABR) and equal-area rectangle (EAR). They are illustrated in Figure 2. These auxiliary geometries are either virtually created or mathematically utilised in formulations. Clearly speaking, it is not required to create an EAC and an EAR as it can be seen from the related equations.

The shape indices tested in this study get values approximately between 0 and 1. To be specific, the minimum value is greater than 0 for all of the indices, while the maximum value is 1 except for Equivalent Rectangular Index (ERI). This index gets 1.128 for a circle as the maximum value. As shape complexity decreases, the index values increase. For Roughness Index (RI), extra points are interpolated along the boundary of a polygon (Basaraner and Cetinkaya, 2017).
Shape complexity of building features mainly depends on (1) the number of the edges, (2) the homogeneity/the heterogeneity of the angles in the corners, (3) the regularity/the irregularity of the edge lengths, (4) the size and/or the number of intrusions and/or protrusions along the boundary relative to the size of the polygon, and (5) the symmetry/asymmetry in the geometry (after Basaraner and Cetinkaya, 2017).

### Table 1. Shape indices used in this study (adapted from Basaraner and Cetinkaya, 2017)

<table>
<thead>
<tr>
<th>SHAPE INDEX</th>
<th>EQUATION</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circularity (CI)</td>
<td>$CI = \frac{A_{PN}}{A_{EPC}} = \frac{4\pi A_{PN}}{P_{PN}^2}$</td>
<td>Circularity measures compactness of a shape by measuring the areal deviation of a polygon (PN) from its equal-perimeter circle (EPC).</td>
</tr>
<tr>
<td>Convexity (CNV)</td>
<td>$CNV = \frac{A_{PN}}{A_{CH}}$</td>
<td>Convexity measures a polygon’s degree of being curved inward or outward by measuring the areal deviation of a polygon from its convex hull (CH).</td>
</tr>
<tr>
<td>Rectangularity (REC)</td>
<td>$REC = \frac{A_{PN}}{A_{MABR}}$</td>
<td>Rectangularity measures the areal deviation of a polygon from its minimum area bounding rectangle (MABR).</td>
</tr>
<tr>
<td>Roughness Index (RI)</td>
<td>$RI = \frac{\mu_{ibp}^2}{A_{PN}} + \frac{P_{ibp}^2}{P_{PN}^2} \times 42.62$</td>
<td>Roughness Index is used as a measure of compactness and calculated through the area, the perimeter and the radial distances ($r_{ibp}$) between interpolated boundary points and the centroid $(x_0,y_0)$ of a polygon. It has two advantages against the indices based on area-perimeter ratio such as circularity: (1) less sensitive to the elongation (i.e. the aspect ratio of a polygon) and (2) more responsive to the roughness (i.e. intrusions and protrusions along the boundary of a polygon).</td>
</tr>
<tr>
<td>Equivalent Rectangular Index (ERI)</td>
<td>$ERI = \frac{P_{EAB}}{P_{PN}} = \sqrt{\frac{A_{PN}}{A_{MABR}} \times \frac{P_{MABR}}{P_{PN}}}$</td>
<td>Equivalent Rectangular Index measures perimeter deviation of a polygon from its equivalent (equal-area) rectangle.</td>
</tr>
</tbody>
</table>

**Figure 2. A polygon’s own geometry and auxiliary geometries (adapted from Basaraner and Cetinkaya, 2017)**
EXPERIMENTAL STUDY

For the experimental study, 25 building features were selected from a topographic map dataset at 1:1K (Figure 3). In the generalisation phase, these features were simplified in ArcGIS software according to the graphic limits of multiple scales. Thus, 1:5K, 1:10K, 1:25K, 1:50K and 1:100K scales were derived from 1:1K. For this purpose, the tolerance parameters given in Table 2 were used in the simplification process. In this context, if the buildings did not provide the area threshold, they were enlarged in accordance with cartographic generalisation principles. The resulting geometries were then interactively postprocessed to eliminate generalisation errors.

![Figure 3. Original building features at 1:1K scale](image)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Granularity (m)</th>
<th>Minimum area (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1K</td>
<td>0.3</td>
<td>0.25</td>
</tr>
<tr>
<td>1:5K</td>
<td>1.5</td>
<td>6.25</td>
</tr>
<tr>
<td>1:10K</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>1:25K</td>
<td>7.5</td>
<td>156.25</td>
</tr>
<tr>
<td>1:50K</td>
<td>15</td>
<td>625</td>
</tr>
<tr>
<td>1:100K</td>
<td>30</td>
<td>2500</td>
</tr>
</tbody>
</table>

Table 2. Parameters used in the simplification process

In the shape analysis phase, the five shape indices were automatically computed in GIS using the original and the derived datasets. For the Roughness index, the number of interpolated boundary points (IBPs) was set as 300. Scale-based variations of building features and their corresponding shape index values were given in Figure 4.
Figure 4. Sample multi-scale building features and their shape index values based on scale.
RESULTS AND DISCUSSION

Shape indices yielded different values at a specific scale and the ranges of their values also became different (Figure 5). These values had a tendency to increase across the scales as a response to the simplified geometries of the building features. The boxplots given in Figure 5 showed that some outliers were produced by some of the indices. These outliers correspond to the most complex shaped building features. At 1:1K scale, only Roughness Index (RI) produced two outliers. This stems from its sensitivity to the intrusions and protrusions along a polygon’s geometry. This trend continued at successive scales except for 1:25K in this index. At 1:50K scale, Convexity (CNV) and Equivalent Rectangular Index (ERI) were also produced outliers. This showed that the reduction in the level of detail in the geometries of most building features are more apparent at that scale. Equivalent Rectangular Index (ERI) compared to Rectangularity (REC) was more sensitive to the intrusions and protrusions. Rectangularity (REC) did not reflect small changes in the level of detail in some cases. Even, its index values decreased slightly while the opposite was expected when the scale was reduced. Circularity (CI) usually responded well to the scale variations. Since the most of the buildings have rectangular shapes with different elongations at 1:100K scale, the range of the compactness was wider than the others because it was quite sensitive to the elongated shapes. Furthermore, the shapes were quite simple so the other indices tended to produce values closer to the upper limits. The ranges of the indices also considerably decreased at 1:100K scale. This means that the shapes become quite similar as scale decreases. Figure 6 showed the largely coherent average response of the shape indices to the scale variations. Only Convexity (CNV) differed from the others at 1:100K since the most of the shapes became convex. Although some scale-dependent patterns can be observed, the variety in the shape complexity of the building features poses a challenge for scale-based evaluations. In other words, some buildings have simple shapes and hence they are less affected from the generalisation while some of the others have quite complex shapes and scale variations make more substantial changes in their geometry and thus in their index values. The minimum variation in the shape of the features was observed between 1:1K and 1:10K scales. The maximum variation was observed between 1:50K and 1:100K scales. Roughness Index (RI) seems the most sensitive index to the scale-based variations. It should be noted that Convexity (CNV) yields one for any convex shape such as square, rectangle and hexagon, Rectangularity (REC) and Equivalent Rectangular Index (ERI) yield one for square and rectangle, Circularity (CI) and Roughness Index (RI) yield one for circle. Hence, the value ranges became different among some of the indices.

Scale-based shape analysis can contribute to generalisation. Before and during generalisation, shape analysis can guide the generalisation process especially when it is applied in a recursive manner. In other words, best candidate among many can be decided by utilising from the shape index values along with other criteria. After generalisation, it can assist in the quality evaluation. One of the other possible applications can be scale inference for volunteered geographic information and remotely sensed data. Both often present fairly heterogeneous levels of detail about the features. Shape analysis may be used for comparing these kinds of datasets with the authoritative datasets to identify their scales. This can be performed at feature level and/or dataset level.
CONCLUSION

This study investigated the response of shape indices to the scale variations. In this respect, five shape indices (i.e. Compactness, Convexity, Rectangularity, Roughness Index and Equivalent Rectangular Index) were used for building features at six different scales (1:1K, 1:5K, 1:10K, 1:25K, 1:50K and 1:100K) derived through generalisation. Evaluation of scale-based shape index value changes were given. The shapes of buildings vary from regular and simple to irregular and complex. The former is higher in numbers while its shape variety is low while the latter is less in numbers while its shape variety is high. The multitude of the irregular and complex shape types poses a challenge in terms of the performance of shape indices. Due to this fact, it is not so easy to make a strong connection between the index values and multi-scale feature geometries. Nevertheless, shape analysis shows potential to play a contributing role for scale-driven problems such as generalisation. For this purpose, various shape indices can be used in a collaborative manner. Furthermore, analysis of larger datasets might further elucidate their interaction.

REFERENCES


BIOGRAPHY

Melih Basaraner is an associate professor of Cartography and GIS at the Division of Cartography, Department of Geomatic Engineering, Faculty of Civil Engineering, Yildiz Technical University (YTU), Istanbul, Turkey. He holds BSc, MSc and PhD degrees in Geomatic Engineering from YTU. He spent 1-year postdoctoral research stay in COGIT Laboratory, IGN France. His research interests include multi-representation spatial databases, automated map generalisation, structure and shape recognition, and spatial analysis, GIS and cartographic applications in various fields.

Sinan Cetinkaya is an assistant professor of Cartography and GIS at the Division of Cartography, Department of Geomatic Engineering, Faculty of Civil Engineering, Yildiz Technical University (YTU), Istanbul, Turkey. His research interests include structure and shape recognition and spatial analysis, GIS and cartographic applications in various fields.