A Metric for Polygon Comparison and Building Extraction Evaluation

Janja Avbelj, Rupert Müller, and Richard Bamler *Fellow, IEEE*

***Abstract*—Standardization of evaluation techniques for building extraction is an unresolved issue in the fields of remote sensing, photogrammetry, and computer vision. In this paper, we propose a metric with a working title 'PoLiS metric' to compare two polygons. The PoLiS metric is a positive definite and symmetric function that satisfies a triangle inequality. It accounts for shape and accuracy differences between the polygons, is straightforward to apply, and requires no thresholds. We show through an example that the PoLiS metric between two polygons changes approximately linearly with respect to small translation, rotation, and scale changes. Furthermore, we compare building polygons extracted from a digital surface model to the reference building polygons by computing PoLiS, Hausdorff and Chamfer distances. The results show that quantification by PoLiS distance of dissimilarity between polygons is consistent with visual perception. What is more, Hausdorff and Chamfer distances overrate the dissimilarity when one polygon has more vertices than the other. We propose an approach towards standardizing building extraction evaluation, which may also have broader applications in the field of shape similarity.**

***Index Terms*—Building extraction, Metric, Polygon comparison, Quality assessment, Shape similarity**

I. Introduction

OBJECT extraction and modeling from images has been an active research area in computer vision and remote sensing in the past few decades. The increasing spatial resolution of satellite and aerial imagery together with ongoing developments of methods enable accurate and (semi) automatic object detection. The extracted objects are then represented in vector or raster format. The latter is usually a result of classification methods, in which each pixel is labelled, whereas objects in vector format are represented by points, lines, and polygons. When focusing on building extraction rather than on general classification, many proposed methods obtain 2D building footprints represented by 2D polygons in vector format. Hence, a performance evaluation is required for comparing methods either among each other and to the reference. If both, extracted and reference data, are in vector format, a measure is needed that fulfills the following requirements:

- compares polygons, not only point sets, with different number of vertices
- is insensitive to additional points on polygon(s)
- is monotonic and also has linear response to small changes in translation, rotation, and scale
- is a metric in the mathematical sense.

J. Avbelj, R. Müller, and R. Bamler are with the German Aerospace Center, Earth Observation Center, Oberpfaffenhofen, 82234 Wessling, Germany. E-mail: (Janja.Avbelj, Rupert.Mueller, Richard.Bamler)@dlr.de see http://www.dlr.de/eoc/

Manuscript received February XX, 2014; revised May xx, 2014.

A. State of the art

Several authors tackled shortage of standard evaluation techniques for building polygon extraction by proposing, assessing, and comparing these techniques [1]–[10]. Nevertheless, commonly accepted evaluation indices and metrics for building extraction evaluation were not yet agreed on in the community. [5] Lists several indices for evaluating building extraction and [3] observes quantity differences between per pixel and per object evaluation methods, as well as a significant difference among per object based indices. Some of the commonly used indices are correlated, e.g. perimeter ratio and few image moments or completeness and correctness. Thus, [6] joined and decorrelated several indices in a hierarchical evaluation system, resulting in a single index.

Per pixel evaluation of detected buildings is analogous to classification accuracy assessment using *matched rates* [6], e.g. completeness, correctness, and quality rates. These three rates are insensitive to additional points on polygons, and respond linearly when small changes in translation, rotation and scaling occur. The major issue with *per pixel* evaluation for polygon comparison is that it requires rasterization of vector data [4], [7] and consequently influences the accuracy. However, such evaluation is straightforwardly applicable, requires no thresholds and can serve for a quick assessment [5]. The matched rates can handle vector data indirectly, if instead of a number of pixels the areas of polygons are considered [5], [9]. Furthermore, when indices are computed *per object*, a new definition problem rises of true/false detected objects and so a threshold must be set [2], [3], [10].

Shape similarity measures quantify *per object* similarity between a reference and an extracted object in raster or vector format. Their desired properties depend on the application [11]. For instance, the affine invariance is desired for some object recognition tasks, but not for evaluating building footprints, because two footprints rotated, translated, and/or scaled to each other should be recognised as different. Several shape similarity measures on *boundary level* are being used for building polygon extraction [5], [6], [10]. For instance, turning angle function and Fourier descriptors are both translation, rotation and scale invariant, whereas area and perimeter ratio are only translation and rotation invariant.

Next to the matched rates and shape similarity measures, also the *positional accuracy* of extracted building footprints can be computed. For instance, the root mean square error (RMSE) between reference and extracted points of a building polygon [5], [6] or Euclidean distance between centroids of the objects [1]. Both these indices use vertices of extracted building footprints without accounting for the edges. However, [8] accounts for the reference edges by computing the RMSE between extracted points and nearest points on the corresponding reference building polygon.

For naming various *indices* or *measures* for building extraction evaluation, also the term *metric* is used as a synonym, i.e. in its broader sense [3], [6]. In strictly mathematical terminology a metric or a distance function defines the distance between elements of a set and is a positive definite and symmetric function that satisfies a triangle inequality [12], as used in [10], [11]. So, a RMSE is not a metric in contrast to the Euclidean distance between two centres of mass.

In this paper, we define a metric for comparison of **po**lygons and **li**ne **s**egments, hereafter referred to as PoLiS metric. It accounts for positional and shape differences by considering polygons as a sequence of connected edges instead of only point sets, and fulfills the mathematical conditions for a metric (Sec. II). Moreover, the PoLiS metric is straightforward to implement and responds approximately linearly to changes in translation, rotation, and scale. We recognize the importance of existing matched rates, similarity, and positional accuracy measures for building footprint evaluation, e.g. [1]–[7], [9], [10]. Thus, the proposed PoLiS metric should be considered as an alternative approach for *per object* building extraction evaluation [5], [6], which may also have broader applications in the files of shape similarity.

II. Metric

A. Metric definition

Let S be any set of objects. A metric d on a set S is a distance function $d : S \times S \mapsto \mathbb{R}$. For all $x, y, z \in S$, the function d must satisfy the following conditions [12], [13]

$$d(x,y) \geq 0 \text{ and}$$
$$d(x,y) = 0 \Leftrightarrow x = y \qquad \text{positive definiteness} \qquad (1)$$
$$d(x,y) = d(y,x) \qquad \text{symmetry} \qquad (2)$$
$$d(x,y) + d(x,z) \geq d(y,z) \qquad \text{triangle inequality.} \qquad (3)$$

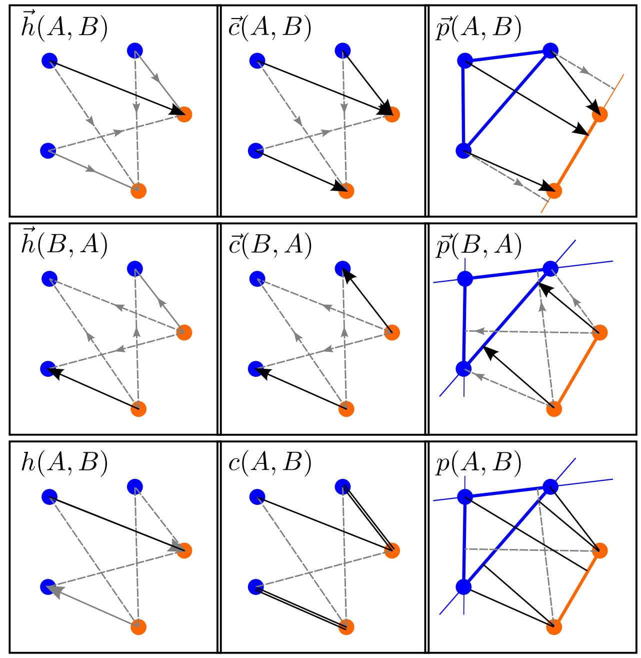
Many shape similarity measures are based on distances between points [11]. For example, the Euclidean distance denoted $\|\cdot\|$ is defined by a function $e : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$, where $e(x,y) = \|x-y\| = (\sum_{i=1}^{n}(x_i - y_i)^2)^{\frac{1}{2}}$, $x, y \in \mathbb{R}^n$.

Now, let us define A and B as two sets of points, with elements $a_j \in A$, $j = 1, \ldots, q$ and $b_k \in B$, $k = 1, \ldots, r$, respectively. The Euclidean distance $e(a_j, b_k)$ between any two points of the sets A and B can be computed, if correspondences between points are known. For applications like stereo-matching or comparing generalized to more detailed shapes, distances are needed that allow a different size of the sets $q \neq r$, e.g. Hausdorff or Chamfer distance (Subsec. II-B).

B. Hausdorff and Chamfer distances

A directed Hausdorff distance $\vec{h}(A,B)$ between the sets A and B is defined as the maximum distance (Fig. 1a) between each point $a_j \in A$ and its closest point $b_k \in B$

$$\vec{h}(A,B) = \max_{a_j \in A} \min_{b_k \in B} \|a - b\|. \qquad (4)$$



(a) Hausdorff metric (b) Chamfer metric (c) PoLiS metric

Fig. 1. Distances (black solid lines) between two sets of points, i.e. A (orange) and B (blue). Directed $\vec{\cdot}(A,B)$ (first row), $\vec{\cdot}(B,A)$ (second row) and symmetric Hausdorff h (Fig. 1a), Chamfer c (Fig. 1b) and PoLiS metric p (Fig. 1c) between A and B is shown. Arrows represent direction in which the distance is computed, grey (solid or dashed) connections between points show an intermediate step in computing a distance, i.e. an underlain Euclidean distance between points. The PoLiS metric is defined for polygons and not for point sets, so connections between the points (blue and orange lines) are established (Fig. 1c).

A directed Chamfer distance $\vec{c}(A,B)$ between the sets A and B is defined as the sum of the distances (Fig. 1b) between each point $a_j \in A$ and its closest point $b_k \in B$

$$\vec{c}(A,B) = \sum_{a_j \in A}^{q} \min_{b_k \in B} \|a - b\|. \qquad (5)$$

Both, directed Hausdorff $\vec{h}$ and directed Chamfer distance $\vec{c}$, fail to fulfill the condition of the symmetry eq. (2), and are therefore not a metric in the mathematical sense. Thus, to fulfill eq. (1)–(3), $\vec{h}$ is symmetrized by computing the maximum of directed Hausdorff distances [11], [14], [15]

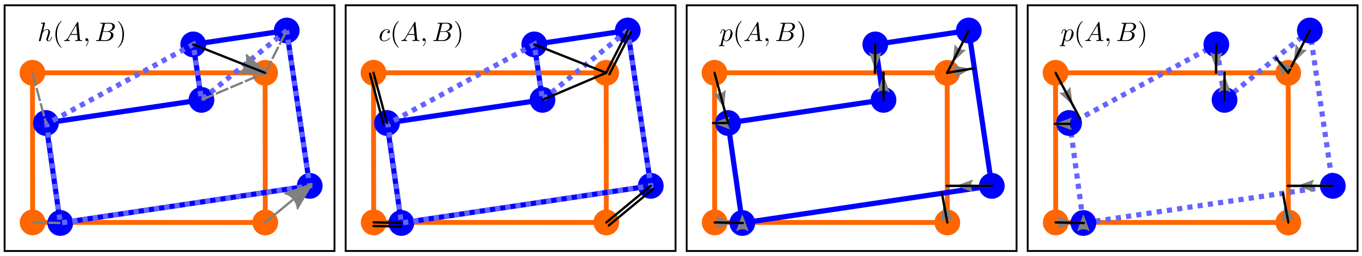
$$h(A,B) = \max\{\vec{h}(A,B), \vec{h}(B,A)\}. \qquad (6)$$

In analogy with $\vec{h}$, $\vec{c}$ is symmetrized by summing the normalized directed Chamfer distances

$$c(A,B) = \frac{1}{2q}\vec{c}(A,B) + \frac{1}{2r}\vec{c}(B,A) \qquad (7)$$

or by computing the average or median between $\vec{c}(A,B)$ and $\vec{c}(B,A)$ [15]. The Hausdorff and the Chamfer distances in eq. (6), (7) are defined with an underlying Euclidean distance, thus any other distance could be underlain.

The Hausdorff distance is a measure for the highest dissimilarity between the two point sets, and is therefore sensitive to outliers, whereas the normalized Chamfer distance quantifies the overall average dissimilarity [14]. Moreover, they are both



(a) Hausdorff distance (b) Chamfer distance (c) PoLiS distance, example 1 (d) PoLiS distance, example 2

Fig. 2. Distance (solid black lines) between extracted building footprint A (orange) and a reference building footprint B (blue), marked with solid black lines. The PoLiS metric (Fig. 2c, 2d) accounts for the shape, whereas the Hausdorff (Fig. 2a) and the Chamfer metric (Fig. 2b) are independent of the connections between the point sets (solid and dotted blue lines). The dotted light blue lines demonstrate one alternative way to connect the point set B into a polygon. Annotations analogue to description under Fig. 1.

sensitive to additional points on polygon edges (Fig. 3), and have a non-monotonic response. Thus, a different measure must be found for building polygons.

III. PoLiS METRIC

In this section we propose a PoLiS metric for comparing **po**lygons and **li**ne **s**egments. Let us assume points a_j of a set A (Subsec. II-A) represent salient points of a shape, e.g. a building footprint. Thus, they can be connected in a closed polygon in $\mathbb{R}^2$ (Fig. 2). We denote with the same capital letter, e.g. A, point set and polygon with the same points and vertices respectively. Then, the points a_j, $j = 1, \dots, q$ of the set A represent the vertices of the closed polygon A, where the first and the last vertex coincide $a_1 = a_{q+1}$, $j = 1, \dots, q+1$. A boundary ∂A consists of $q+1$ vertices a_j of the closed polygon A, q edges, and points that lie on the boundary. We refer to point of a polygon, which has defined coordinates, as vertex, even if it is not a corner point of a polygon and lies on the polygons' boundary. Any point, e.g. $a \in A$, without subscript can be either a vertex or a point without explicitly defined coordinates. Analogically to the point set A, the point set B can be considered as a closed polygon B with $k = 1, \dots, r+1$ vertices.

A PoLiS distance $\vec{p}(A, B)$ between polygons A and B (Fig. 1, 2c, 2d) is defined as the average of the distances between each vertex $a_j \in A$, $j = 1, \dots, q$ of A and its closest point $b \in \partial B$ (not necessary a vertex) on the polygon B

$$\vec{p}(A, B) = \frac{1}{q} \sum_{a_j \in A} \min_{b \in \partial B} \|a_j - b\|. \tag{8}$$

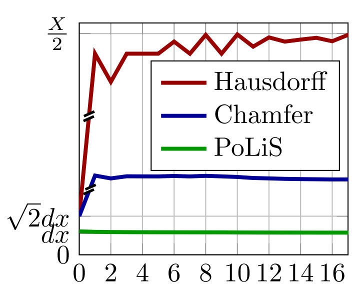
The directed PoLiS distance $\vec{p}$ is symmetrized, similar as in eq. (6) and (7) by summing up the directional distances

$$\begin{aligned} p(A, B) \quad &= \frac{1}{2q} \sum_{a_j \in A} \min_{b \in \partial B} \|a_j - b\| + \\ &+ \frac{1}{2r} \sum_{b_k \in B} \min_{a \in \partial A} \|b_k - a\|. \end{aligned} \tag{9}$$

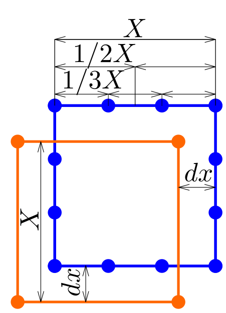
The normalization factors $\frac{1}{2q}$ and $\frac{1}{2r}$ are needed to quantify the overall average dissimilarity per point, same as for the Chamfer distance, eq. (7). The units of a PoLiS distance (Fig. 2c, 2d) are the same as the units of the polygon vertices.

IV. EXPERIMENT

We carry out two experiments on a synthetic and a real dataset. In both experiments the performance of the proposed PoLiS metric (Sec. III eq. (9)) is compared to the Hausdorff eq. (6) and the Chamfer eq. (7) distance (Subsec. II-B). Unit of all metrics are in [m].



(a) Number of added vertices per side

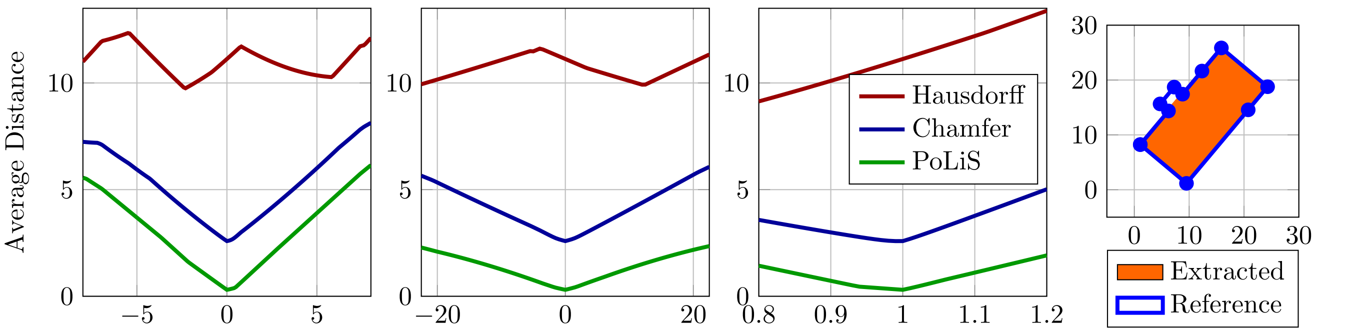


(b) Polygons

Fig. 3. Sensitivity of the Hausdorff (3a, red), the Chamfer (3a, dark blue) and PoLiS (3a, green) distance to additional vertices on the edges of the reference polygon (blue). The extracted (3b, orange) and the reference polygon (3b blue) have both side length of X and are translated for dx, where $dx << X$. Vertices added to edges of the reference polygon are equally distributed, so that the distance between two subsequent vertices is always equal.

```
procedure PoLiSMetric(A, B)
    p1, p2 ← 0
    for j = 1, ..., q do                    ▷ for every point a_j ∈ A
        p1 ← p1 + minDistPt2Poly(a_j, B)
    end for
    for k = 1, ..., r do                    ▷ for every point b_k ∈ B
        p2 ← p2 + minDistPt2Poly(b_k, A)
    end for
    p ← p1/(2q) + p2/(2r)
    return p                                ▷ PoLiS distance value
end procedure
```

Fig. 4. Pseudocode for computing PoLiS metric between two closed polygons, A and B. The procedure MINDISTPT2POLY computes the shortest distance between a point and a polygon given by its vertices, and can be for a 3D case replaced by computing a minimal distance between a 3D point and a polyhedron.



(a) Translation [m] (b) Rotation [°] (c) Scale [ ] (d)

Fig. 5. Hausdorff, Chamfer and PoLiS metric as a function of translation, rotation and scale. A polygon under consideration (orange filled polygon) is compared to the reference (blue polygon). The area of polygons is equal. Hence, the extracted polygon is generalized representation of the reference polygon, i.e. without division in two units and additional structure.

A. Synthetic data

We define two polygons of equal area, the reference polygon (Fig. 5d, blue), which has two points on the edges and a small structure, and the extracted polygon (Fig. 5d, orange), which is a rectangle. This is a typical building polygon detection scenario: two additional points dividing a building into two building units are not detectable from remote sensing images and the small structure is not distinguishable due to the spatial resolution of the image.

The extracted building is translated (Fig. 5a), rotated (Fig. 5b) and scaled (Fig. 5c) according to the centroid of the reference polygon. The Hausdorff distance is not appropriate for quantification of dissimilarity between two polygons, because it has no minimum at an initial position of the extracted polygon and has a non-monotonic response. In contrast, the Chamfer and the PoLiS distances have a minimum at the initial position of the extracted polygon and the impact of changes in translation, rotation, and scale can be approximated by a linear function. Moreover, georeferencing accuracy in remote sensing is normally in much smaller ranges than in Fig. 5a–5c, e.g. $< \pm 8$ m, $< \pm 22.5°$ and $< 1 \pm 0.2$. The graphs (Fig. 5a–5c) of the Chamfer and the PoLiS distance have different slopes. This is one reason, why the numerical values of the metrics can be compared only relative to each other, next to the different minimum values (Subsec. IV-A) and different definitions of the metrics (eq. (6), (7), (9)).

B. Real data

For every pair of extracted and reference building footprints (Fig. 6), the Hausdorff (Fig. 7a), Chamfer (Fig. 7b), and PoLiS (Fig. 7c) distances are computed. The building footprints are extracted from a digital surface model (DSM) with the method described in [16] (Fig. 6, grey areas). The DSM with 1 m spatial resolution is resampled from a LiDAR point cloud with an average density of 1.69 points/m^2. The reference building footprints (Fig. 6, blue), provided by the City of Munich, are detailed cadastral data.

The color bar for each metric in Fig. 6 is scaled from the best extracted building footprint (dark green) to the worst (red), i.e. from the minimum to the maximum distance value. The rectangular and elongated L-shaped buildings are well

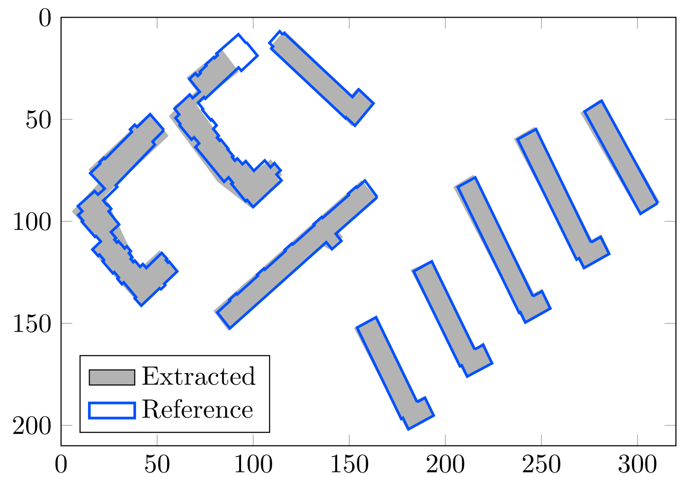
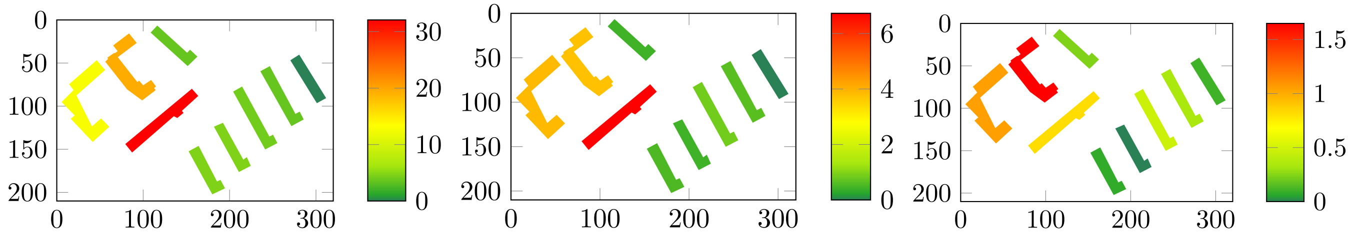


Fig. 6. Extracted (grey filled polygons) and reference (blue polygons) building footprints. Evaluation of this extracted building footprints with different metrics is shown in Fig. 7.

estimated (green), yet some differences between the metrics occur. Moreover, the worst estimated building footprint, according to the h and c distances (Fig. 7a, 7b, red), is the elongated building with several small structures. On the contrary, according to the PoLiS distance the worst estimated building (Fig. 7c, red) is the one where a part was not extracted Fig. 6, due to the vegetation on this part of the roof. Thus, the PoLiS metric penalizes missing estimated areas more than generalization of the boundary. This is what we require for the application at hand.

C. Discussion

The PoLiS metric is a metric in mathematical sense, straightforward to implement, and apply (Fig. 4). We showed that it changes approximately linearly to expected small changes between the extracted and the reference polygon. It is, like the RMSE defined in [8], robust towards partitioning of the polygon by adding vertices on the polygon edges (Fig. 5, 3). However, RMSE is not a metric (is not symmetric) and has a parabola-like response to small changes in translation, rotation and scale.



(a) Hausdorff metric (b) Chamfer metric (c) PoLiS metric

Fig. 7. Evaluation of building footprint extraction (Fig. 6) with Hausdorff (7a), Chamfer (7b) and building (7c) metric, all in [m]. The values of the metrics can be compared relative to each other, but not numerically. Thus, for easier visual interpretation, a colour bar for each metric is scaled from the worst (red) to the best extracted building footprint (dark green).

When no additional points on the edges of polygons are present, the c takes similar values as the PoLiS distance. Yet, the Chamfer and also the Hausdorff metric are very sensitive to the additional points on edges (Fig. 3). What is more, these two metrics compare only the point sets. On the contrary, the PoLiS distance considers the shapes of the polygons by computing distances to the polygon edges. If one of the polygons has much larger number of vertices than the other, the numerical value of the PoLiS metric underestimates the actual dissimilarity, because of normalization factors. However, under the assumption of small translation, rotation and scale, the relations between values of PoLiS distances are consistent relative to each other, in contrast to the values of Hausdorff and the Chamfer distances (Fig. 7).

V. Conclusion

We propose a new PoLiS metric for comparing polygons that quantifies overall average dissimilarity per polygon vertex. The metric can be used to assess the quality of extracted building footprints, when reference data are available. The performance of the metric is tested on synthetic and real data examples, i.e. evaluating extracted building footprints from the remote sensing image. The PoLiS distance estimates similarity between polygons with different number of vertices better than the Hausdorff and the Chamfer distance and comparable, if the number of vertices is similar. The proposed metric changes approximately linearly, when small changes in translation, rotation, and scale between the polygons occur. Moreover, it is a combined measure, which takes into account positional accuracy and shape differences between the polygons.

The PoLiS metric can be straightforwardly extended to a 3D PoLiS metric (Fig. 4). Moreover, a potential of the maximum of the PoLiS distance could be exploited as a measure for the highest dissimilarity between polygons. The metric is suitable for any polygon comparison, nevertheless our first direct goal is its application in the field of building extraction evaluation.

Acknowledgment

The authors would like to thank the city of Munich, Department of Environment and Health for providing the reference building footprints.

References

[1] Q. Zhan, M. Molenaar, K. Tempfli, and W. Shi, “Quality assessment for geo-spatial objects derived from remotely sensed data,” *International Journal of Remote Sensing*, vol. 26, no. 14, p. 29532974, 2005.

[2] M. Awrangjeb, M. Ravanbakhsh, and C. S. Fraser, “Automatic detection of residential buildings using LIDAR data and multispectral imagery,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 5, pp. 457–467, Sep. 2010.

[3] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, “A comparison of evaluation techniques for building extraction from airborne laser scanning,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 11–20, 2009.

[4] J. Shufelt, “Performance evaluation and analysis of monocular building extraction from aerial imagery,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 311–326, 1999.

[5] W. Song and T. Haithcoat, “Development of comprehensive accuracy assessment indexes for building footprint extraction,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 402–404, 2005.

[6] C. Zeng, J. Wang, and B. Lehrbass, “An evaluation system for building footprint extraction from remotely sensed data,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 3, pp. 1640–1652, 2013.

[7] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, “Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 2, pp. 135–149, Jun. 2007.

[8] F. Rottensteiner, G. Sohn, M. Gerke, J. D. Wegner, U. Breitkopf, and J. Jung, “Results of the ISPRS benchmark on urban object detection and 3D building reconstruction,” *ISPRS Journal of Photogrammetry and Remote Sensing*, in press.

[9] J. Shan and S. D. Lee, “Quality of building extraction from IKONOS imagery,” *Journal of surveying engineering*, vol. 131, no. 1, p. 2732, 2005.

[10] L. Ragia and S. Winter, “Contributions to a quality description of areal objects in spatial data sets,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 55, no. 3, pp. 201–213, Sep. 2000.

[11] R. Veltkamp, “Shape matching: similarity measures and algorithms,” in *Shape Modeling and Applications, SMI 2001 International Conference on.*, 2001, pp. 188–197.

[12] E. T. Copson, *Metric Spaces*. Cambridge; New York: Cambridge University Press, 1988.

[13] R. C. Veltkamp and M. Hagedoorn, “State-of-the-art in shape matching,” Utrecht University, Departement of Computer Science, The Netherlands, Tech. Rep., 1999.

[14] J. Zhang, R. Collins, and Y. Liu, “Representation and matching of articulated shapes,” in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*, vol. 2, 2004, pp. II–342–II–349 Vol.2.

[15] A. Tsai, C. F. Westin, A. Hero, and A. Willsky, “Fiber tract clustering on manifolds with dual rooted-graphs,” in *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR ’07*, 2007, pp. 1–6.

[16] J. Avbelj, D. Iwaszczuk, R. Müller, P. Reinartz, and U. Stilla, “Coregistration refinement of hyperspectral images and DSM: An object-based approach using spectral information,” *ISPRS Journal of Photogrammetry and Remote Sensing*, in press.