

Algorithm of automatic digital cartographic generalisation with the use of contractive self-mapping

Abstract. The research of modern cartography in the field of digital generalisation focuses on the development of such methods that would be fully automatic and give an unambiguously objective result. Devising them requires specific standards as well as unique and verifiable algorithms. In metric space, a proposal for such a method, based on contractive mapping, the Lipschitz and Cauchy conditions and the Banach theorem, using the Salishchev metric, was presented in the publication (Barańska et al., 2021). The method formulated there is dedicated to linear objects (polylines). The current work is a practical supplement to it. It presents the practical implementation of the algorithm for automatic and objective generalisation. The article describes an operational diagram of the subsequent stages of the proposed generalisation method. In the test example, a binary tree structure of an ordered polyline was created. It was simplified in two selected scales and its shape after generalisation was illustrated. The resulting polyline obtained by the fully automatic method was verified in terms of accuracy.

Keywords: digital generalisation, contractive self-mapping, minimum dimensions of Salishchev, polyline (segmented line) of binary tree structure, contraction triangles

1. Introduction

The task of cartography is to present an image of the Earth or its part on a map at an appropriate scale. In 1866, Emil von Sydow distinguished three problems to be solved in creating a map called “reefs of cartography” (Sydow, 1866). They are: mapping the reference surface, generalisation and visualisation of the three-dimensional reality on a plane. The generalisation stage, so far, has been carried out mainly with the help of the cartographer’s knowledge and skills. In (Barańska et al., 2021), the authors presented a method of fully automatic generalisation of the geometry of objects on a map, at any scale. The development of methods of generalising (simplifying) the geo-

metry of object features in digital cartography consists in improving the existing algorithms, without their validation in terms of the similarity of the geometry of features before and after the process (Salishchev, 2003), (Chrobak, 2010), (Chrobak et al., 2017), (Chrobak et al., 2019), (Courtial et al., 2020), (Kronenfeld et al., 2020) (Barańska et al., 2021). Until now, generalisation involved the use of approximate methods based on fractals, artificial intelligence and their evaluation with the method of least squares. Their main disadvantage is the ambiguity of the results. The authors of the automatic solution used an exact method. Its basis is the theory of contractive self-mapping (Barańska et al., 2021). It takes into account the requirements of the generalisation process, i.e. reduction at

scales $s < 1$. Geodetic data (points of cartographic control base) generalised in this method belong to the metric space. This ensures their continuity and unambiguity. The distances between points describing the geometry of the object are defined in the process. This makes it possible to clearly determine in the mapping the recognisability of the object at a given scale. Moreover, in contractive mapping, the transformation of an ordered polyline maintains the requirements of the binary tree system in the created triangles, which allows for preserving the Lipschitz contraction condition in the process (Barańska et al., 2021). The created triangles are unambiguously verified by recognition norms, defined by K. Salishchev (2003).

The developed generalisation algorithm evaluates the obtained results, i.e. a feature necessary for its use in automatic generalisation. Preparation for automatic data generalisation requires fully defined standards and unique and verifiable algorithms for particular groups of features. This in turn makes it possible to use these databases to draw objects directly on maps. As a result, the data collected in the database and their unique generalisation at various scales lead to standardisation of sets.

The article proposes a solution based on self-regulation of contraction (contractor for scale $s = 1$), which meets the requirements of Banach's fixed-point theorem (Barańska et al., 2021). The method of generalising object geometry features at scales $s < 1$ uses contraction, thanks to which in contractive self-mapping created for source data it constitutes one standard mapping to be used for generalising the geometry of an object at any scales $s < 1$. The article is intended for researchers working on the theory of cartography and algorithms for objective automatic simplification of object geometry.

2. Objective digital generalisation of an ordered polyline L_u at scales $0 < s < 1$

Generalisation of the geometry of an ordered polyline (L_u) at scales $0 < s < 1$ is a contractive self-mapping with an objectively justified unique result. This increases the degree of reliability of the result and eliminates iteration, which significantly reduces the costs of the generalisation process.

Contractive mapping involves base triangles (BT) formed on the existing points of the cartographic network as well as the beginning and end cartographic control points of the object. The task of triangles BT is to link the cartographic control of the object in the mapping to the existing cartographic network, which in the processes of harmonisation and interoperability of obtaining information reduces their costs (Directive, 2007). It would be advisable for the existing meta-data of the object in the geodetic and cartographic data centres to be supplemented with a cartographic control network (Krzywicka-Blum, 2017).

On the points of the cartographic network of base triangles BT, contraction triangles (CT) of polyline L_u are created, which are the matrix for its generalisations. Triangles CT preserve the necessary contraction condition in contractive mapping of a polyline. Their bases are greater than their heights, which was proved in the work (Barańska et al., 2021). The set of triangles CT of polyline L_u for $s = 1$ is the standard of contractive self-mapping and generalisations for the scale $0 < s < 1$. The result of the self-mapping of the data of the polyline is verified by the source data of the ordered polyline L_u as presented in the paper (Barańska et al., 2021). Triangles CT are formed on the legs of triangles BT. The number of triangles BT and CT depends on the degree of compilation of the geometry of the polyline and its length, following the rule "from the general to the specific". The process of creating a CT is as follows: on the vertices of a polyline, triangles are created in the binary tree system. In the sections of the generalised polyline L_u its point function is regular except for some areas where the so-called singular points have been identified (Barańska et al., 2021). The lengths of bases and heights of the triangles are verified by the K. Salishchev norm (Salishchev, 2003). After the generalisation of the polyline L_u with contractive mapping, the lengths of the sides and heights of triangles CT are compared with the K. Salishchev norm.

3. Implementation of the polyline generalisation algorithm

The polyline generalisation algorithm is implemented in several stages (fig. 1). The first two stages concern the preparation and ordering

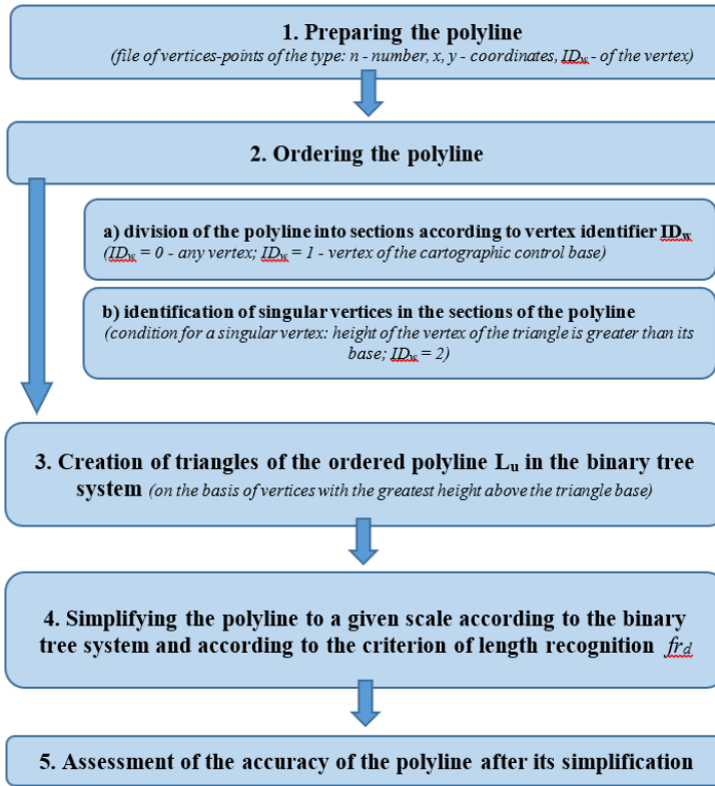


Fig. 1. Stages of polyline generalisation

of the polyline. An open polyline, consisting of n -vertices, is defined by X, Y planar coordinates of the vertices in a given planar coordinate system. Among the vertices, fixed vertices should be distinguished, which cannot be removed from the polyline, e.g. as a result simplifying it, and singular vertices. Fixed vertices are cartographic control points, e.g. both ends of the polyline, and other vertices which are cartographic control points (and form a base triangle BT). In order to distinguish the vertices of a polyline, each vertex, apart from its coordinates, should be assigned an identifier (ID_W) with the following values: $ID_W = 0$ – any vertex; $ID_W = 1$ – vertex of the cartographic control base; $ID_W = 2$ – singular vertex.

Singular vertices are searched for among vertices with the identifier $ID_W = 0$. For this purpose, triangles CT are sequentially created from the following vertices of the polyline: W_i , W_{i+1} and W_{i+2} ; the created triangles have the

base $p_{i,i+2} = W_i W_{i+2}$ and legs $p_{i,i+1} = W_i W_{i+1}$, $p_{i+1,i+2} = W_{i+1} W_{i+2}$. In these triangles, the following are calculated: height h_{i+1} of the vertex W_{i+1} above the base $p_{i,i+2}$. A vertex is singular if it meets the condition: $h_{i+1} \geq p_{i,i+2}$ (Barańska et al., 2021). The singular vertex is assigned an identifier $ID_W = 2$. After identifying singular vertices, the polyline takes the form of an ordered polyline L_u .

3.1. Creating triangles in the binary tree structure

The binary tree of a polyline is a structure of interconnected and branching contraction triangles CT formed from the vertices of polyline L_u . This tree is created according to the principle “from the general to the particular”. The initial branches of the tree are polyline sections delimited by fixed vertices $ID_W = 1$. For a given section, the length of the triangle

base included between the vertices which are limits of the section is calculated. A given section of the polyline is divided into successive branches (segments) if it consists of more than two vertices. A section consisting of only two vertices: the beginning and the end vertex (i.e. the side of the original polyline) should be understood as a triangle with a vertex with “zero height”. For all vertices of a given section, their heights above the base are calculated, and then the vertex with the greatest height is selected. The vertex with the greatest height and the base form a triangle of the polyline. The vertex with the greatest height divides the given polyline section into two consecutive branches (sections). The branching process is then repeated on each newly created branch of the tree. As a result, we obtain a binary tree of triangles, the end branches of which are triangles of “zero height”, constituting any sides W_iW_{i+1} of the original polyline. Figure 2a shows an example of polyline $W_1W_{10}W_{20}$, with the vertices of the cartographic control base: W_1 , W_{20} (the ends of the polyline) and W_{10} . Table 1 shows the X, Y coordinates of all vertices of the polyline with their ID_W identifiers.

Two singular vertices were identified in this polyline: W_4 and W_{18} . The ordered polyline is shown in figure 2a. Figure 3 shows the structure of the binary tree of triangles for this ordered polyline L_U . Each block of the structure constitutes a branch of the tree, i.e. a triangle with the given base W_iW_j and its length (p). The description of the triangle in a given block of the structure is completed by the vertex (W_h) with the greatest height (h_W) above the base

Tab. 1. Coordinates of the vertices of polyline $W_1W_{10}W_{20}$ with singular vertices W_4 and W_{18}

NW	X	Y	ID _W	NW	X	Y	ID _W
1	0.0	0.0	1	11	85.1	101.8	0
2	15.5	4.1	0	12	78.0	111.9	0
3	23.5	16.1	0	13	65.0	110.1	0
4	30.0	50.0	2	14	59.0	124.8	0
5	49.1	21.8	0	15	46.1	121.2	0
6	55.2	39.5	0	16	39.9	131.9	0
7	64.8	48.1	0	17	30.5	131.2	0
8	71.6	70.1	0	18	20.0	100.0	2
9	84.2	76.8	0	19	13.9	129.0	0
10	98.8	93.8	1	20	0.0	135.0	1

(fig. 3). The last branches of the tree are the sides (W_iW_{i+1}), or “zero-height triangles” ($h = 0.0$).

The formation of the structure of a binary tree of triangles begins in the sections between the vertices of the cartographic control base (fig. 3). For example, for the section W_1W_{10} of the polyline, the vertex W_5 was selected because its height above the base $p_{1-10}=W_1W_{10}$ is the greatest and amounts to 18.0 (fig. 3). Due to the selection of the vertex W_5 , the legs W_1W_5 and W_5W_{10} of the triangle $W_1W_5W_{10}$ constitute successive branches (sections) of the polyline. On the first of them (W_1W_5), a triangle $W_1W_4W_5$ with the highest vertex W_4 ($h_4 = 33.5$) and

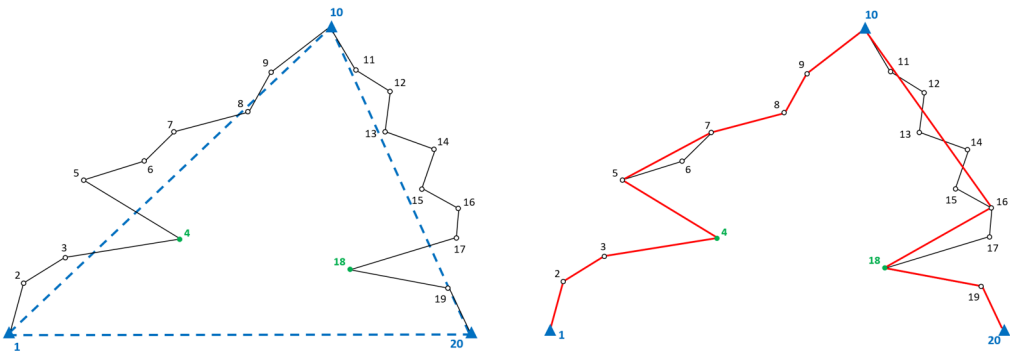


Fig. 2. Ordered polyline L_U with cartographic control points W_1 , W_{10} , W_{20} : a) base triangle BT (blue) and identified singular vertices W_4 and W_{18} (green); b) polyline, simplified to a scale of 1:20,000 (red)

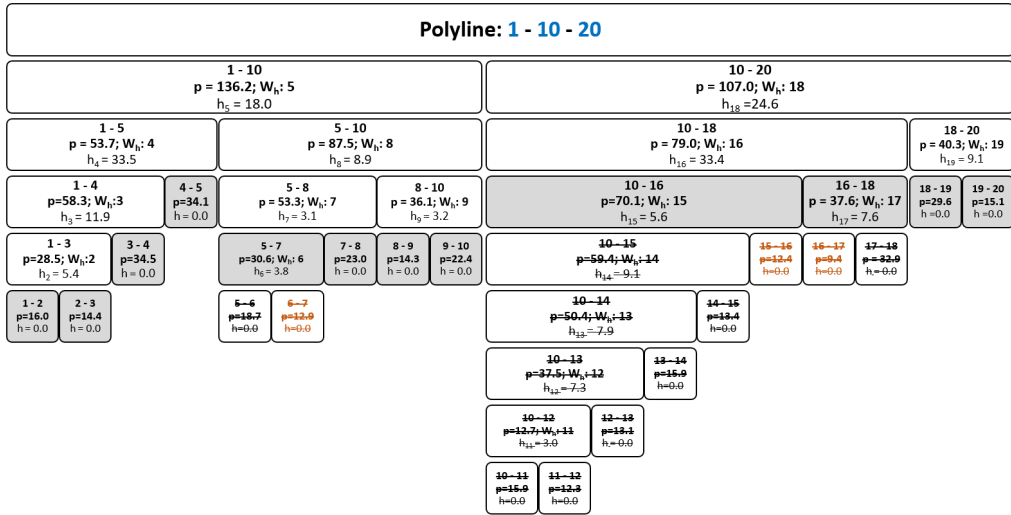


Fig. 4. Structure of the binary tree of ordered polyline L_U after simplification from scale 1:10,000 to scale 1:20,000 (in brown – triangles that do not meet the fr_d condition; crossed out – triangles rejected as a result of simplification; shaded blocks – the polyline remaining after simplification)

As a consequence of simplifying the polyline to scale 1:20,000, the ordered polyline L_U was reduced to the form: $W_1W_2W_3W_4W_5W_7W_8W_9W_{10}W_{16}W_{18}W_{20}$ (fig. 4). Its shape is shown in figure 2b. Simplification of the polyline resulted in the removal of 7 vertices: $W_6, W_{11}, W_{12}, W_{13},$

$W_{14}, W_{15}, W_{17}, W_{19}$. The simplification of the polyline in any other scale will have a similar course. For example, at a scale of 1:50,000, where $fr_d = 35.0$, the simplified polyline will consist of only 7 vertices: $W_1W_5W_8W_{10}W_{16}W_{18}W_{20}$ (fig. 5)

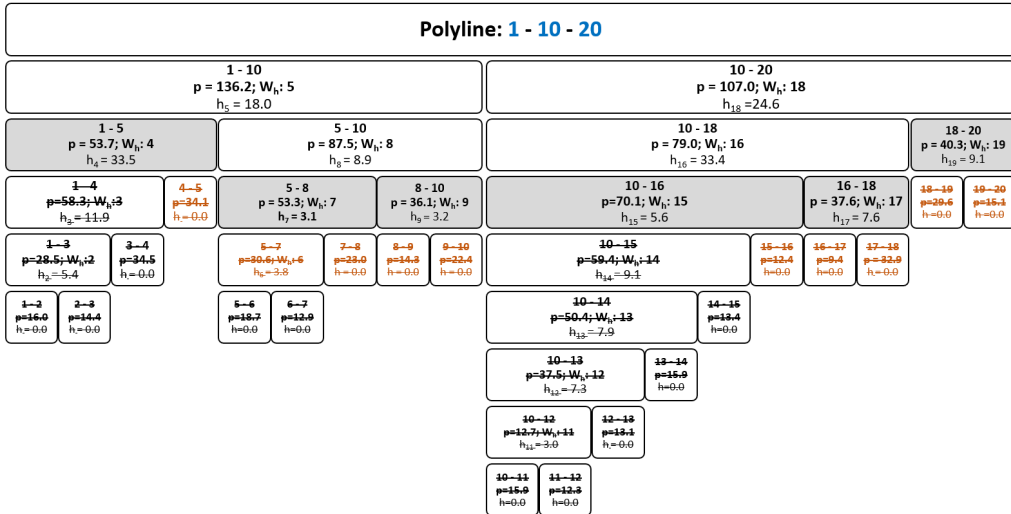


Fig. 5. Structure of the binary tree of ordered polyline L_U after simplification from scale 1:10,000 to scale 1:50,000 (in brown – triangles that do not meet the fr_d condition; crossed out – triangles rejected as a result of simplification; shaded blocks – the polyline remaining after simplification)

4. Verification of the measures of the generalised polyline L_u with contractive mapping to triangles CT according to the metric of K. Salishchev

The verification consists in comparing the source data vertices with the sum of the vertices remaining after the simplification process and the rejected vertices. Table 2 shows the number of vertices before and after the simplification process.

The result (number of vertices) obtained for the subsequent scales (tab. 2) in contractive self-mapping of polyline L_u is identical and fulfils the Banach theorem concerning the existence of one solution of generalised polyline in each scale $s < 1$.

4.1. Calculation of standard deviation

The standard deviation for polyline L_u determined after its self-mapping at scale $s = 1$ is the standard for its simplifications at any scale $s < 1$. And the result after simplifying the polyline at a given scale is verified by the K. Salishchev norm applied to the bases and heights of triangles created through the mapping. As a result of polyline generalisation with contractive self-mapping to the selected scale, some of its vertices (and thus sides) are rejected. Figure 6 shows a fragment of polyline W_1-W_{i-n} , in which, due to generalisation, the vertices W_{i+1} and W_{i+2} were rejected, leaving the vertices W_i and W_{i+3} . The remaining vertices form a new side of the polyline W_i-W_{i+3} .

For all vertices of the source polyline, their heights (h_i) in relation to the sides created after the generalisation were calculated. The heights

Tab. 2. Verification of contractive mapping on the basis of the number of vertices of a generalised ordered polyline L_u

Generalization to 1:M scale	The number of vertices of ordered polyline L_u		
	original	after simplification	rejected
1:20,000	20	13	7
1:50,000	20	7	13
1:75,000	20	4	16

of the rejected vertices of the polyline will be greater than zero, and the heights of the remaining vertices will be zero. These heights are used to calculate standard deviation (δ_h), which describes the discrepancy between the generalised polyline and the original polyline, according to the equation (2):

$$\delta_h = \sqrt{\frac{\sum_{i=1}^n h_i}{n}} \quad (2)$$

where: n – number of vertices of the original polyline.

4.2. Examining the recognisability of geometry of the generalised polyline according to K. Salishchev's metric

The equation (2) does not take into account the geometry of the vertices of base triangles BT, as the recognisability evaluation concerns the change of the scale of the generalised

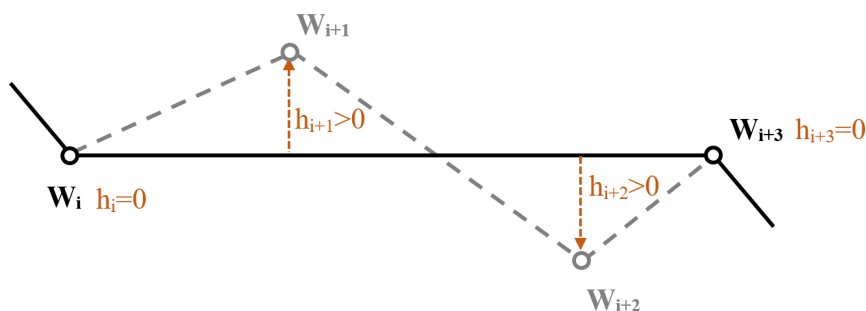


Fig. 6. Fragment of the original and generalised polyline (solid black line – polyline after generalisation; dashed grey line – rejected sides of the original polyline)

polyline according to the K. Salishchev metric. The evaluation at a scale of simplified geometry of the polyline, made on the basis of the sum of its sequences in appropriate intervals, is carried out after the simplification process. The evaluation of the simplified polyline depends on the sum of the heights of the rejected vertices in the intervals. This is due to the fact that the data of the source polyline belong to the metric space of the polyline L_u , as demonstrated in the work (Barańska et al., 2021). The cartographic control points and the singular points of the polyline have an impact on the evaluation of contractive mapping if they are not excluded from the generalisation process. Their exclusion would result in failure to meet the Lipschitz condition (Barańska et al., 2021). Retaining

these points (tab. 3 items 1, 2, 3, 4, 5, 7, 13, 14, 16, 17, 18, 21, 24, 25, 27, 29, 31) allows for an automatic generalisation process. This leads to a single and objective result of the generalised polyline, verified by K. Salishchev metric, because:

- polyline data belong to the metric space,
- contractive self-mapping with the use of the binary tree system fulfils the Lipschitz condition and the Banach theorem,
- the result of contractive mapping of the polyline meets the recognition metric of K. Salishchev.

Cartographic control points (W_1, W_{10}, W_{20} - bold numbers) and singular points (W_4, W_{18} - shaded numbers) (tab. 3: items 1, 2, 3, 4, 5, 7, 8, 9, 13, 14, 15, 16, 17, 18, 19, 20, 21, 24, 25,

Tab. 3. Verifying the K. Salishchev metric in the process of generalisation of the ordered polyline L_u with the use of contractive self-mapping (for scale 1:10,000, $fr_d = 4.0$)

Polyline fragment 1 - 10					Polyline fragment 10 - 20				
No.	Triangles CT				No.	Triangles CT			
	base	height	Points			base	height	Points	
			singular	cartographic control				singular	cartographic control
1	1-10	18.0	0	2	14	10 - 20	24.6	0	2
2	1-5	33.5	0	1	15	10 - 18	33.4; h _s =0	1	1
3	5-10	8.9	0	1	16	18 - 20	9.1; h _s =0	1	1
4	1-4	11.9; h _s =0	1	1	17	10 - 16	5.6	0	1
5	4-5	0.0; h _s =0	1	0	18	16 - 18	7.6; h _s =0	1	0
6	5-8	3.1	0	0	19	18 - 19	0.0; h _s =0	1	0
7	8-10	3.2	0	1	20	19 - 20	0.0	0	1
8	1-3	5.4	0	1	21	10 - 15	9.1	0	1
9	3-4	0.0; h _s =0	1	0	22	15 - 16	0.0	0	0
10	5-7	3.8	0	0	23	16 - 17	0.0	0	0
11	7-8	0.0	0	0	24	17 - 18	0.0; h _s =0	1	0
12	8-9	0.0	0	0	25	10 - 14	7.9	0	1
13	9-10	0.0	0	1	26	14 - 15	0.0	0	0
					27	10 - 13	7.3	0	1
					28	13 - 14	0.0	0	0
					29	10 - 12	3.0	0	1
					30	12 - 13	0.0	0	0
					31	10 - 11	0.0	0	1
					32	11 - 12	0.0	0	0

27, 29, 31) do not preserve the measures of the K. Salishchev metric, which justifies their exclusion from the mapping. These results indicate that cartographic control points should not participate in the evaluation of the generalisation of the object. Their task is to orientate the geometry of the cartographic control base, not of the object. In contractive self-mapping, the sequence of points of the polyline fulfils the measures of the K. Salishchev metric (tab. 3, items 6, 7, 10). The exclusion of singular points and cartographic control points from contractive mapping is justified, as it makes it possible to verify the existence of one objective solution of the object geometry, fulfilling the Banach theorem, in contractive self-mapping.

5. Conclusion

The article is a response to the tasks set out in the INSPIRE Directive of 2007 (Directive, 2007) and concerns spatial data in terms of the direct or indirect location of a geographical area and the needs of users. Particular attention in the article is drawn to the optimal use of data when employing computer software, as well as the accuracy of data processing. For this purpose, contractive mapping along with the principle of "from the general to the specific" were employed, as well as the Lipschitz condition for strengthening the continuity of a uniform function, and the Banach fixed-point theorem and a solution resulting in the maximum degree of reliability of the result. The article presents a solution to the problem of cartographic generalisation, which is one of the main problems of cartography to this day (Sydow, 1866). In a previously published article (Barańska et al., 2021), which was the basis for the current work, the theoretical problem of generalisation was presented. This article describes an algorithm that verifies the existence of an objective

result of the data simplification process. Unlike many others, this algorithm does not rely on an iterative process. The presented algorithm is illustrated with an example of a polyline with vertices of the cartographic control base and singular vertices. The results obtained using the presented method of generalisation lead to the following conclusions:

a) The polylines were examined using the example of first degree curves. The next stages of research should concern the algorithm of generalisation of curves of higher degrees for maps at scales 1:10,000 and smaller.

b) In contractive self-mapping of a polyline belonging to the metric space L_u , its single result verifies the fulfilment of the Banach theorem.

c) The simplification of an ordered polyline L_u by contractive self-mapping to the scale $s < 1$ has one objective solution.

d) The result of the generalisation of polyline L_u also takes into account the heights of the triangles, which are compared with the recognition norm of K. Salishchev.

e) In the study of the geometry of the polyline after contractive self-mapping, full compatibility with the recognition norm of K. Salishchev was obtained.

f) Cartographic control points of the polyline change the orientation, not its geometry, in the process of its generalisation.

g) The process of polyline generalisation should be carried out separately from the transformation between the coordinate systems, because transformation changes the angular-linear measures.

Acknowledgements

The article was prepared under the research subvention of AGH University of Science and Technology No. 16.16.150.545.

Literature

- Barańska, A., Bac-Bronowicz, J., Dejaniak, D., Lewiński, S., Krawczyk, A., Chrobak, T. (2021). A Unified Methodology for the Generalisation of the Geometry of Features. *SPRS International Journal of Geo-Information*, 10 (3), art. no. 107, 1-25. <https://doi.org/10.3390/ijgi10030107>
- Blum-Krzywicka, E. (2017). Elements of Maps Contents with (0D) Point Reference Units. In: *Map Functions*; Springer Geography Switzerland: Cham, Switzerland, 41-84. TWO:10.1007 / 978-3-319-47358-1_2
- Chrobak, T. (2010). The role of least image dimensions in generalized of object in spatial databases. *Geodesy and Cartography*, 59, 99-120, DOI: 10.2478/v10277-012-0004-y
- Chrobak, T., Szombara, S., Kozioł, K., Lupa, M. (2017). A method for assessing generalized data accuracy

- with linear object resolution verification. *Geocarto International*, 32, 238-256. <https://doi.org/10.1080/10106049.2015.1133721>
- Chrobak, T., Lupa, M., Szombara, S., Dejnia, D. (2019). The use of cartographic control points in the harmonization and revision of MRDBs. *Geocarto International*, 34, 260-275. <https://doi.org/10.1080/10106049.2017.1386721>
- Courtial, A., ElAyedi, A., Touya, G., Zhang, X. (2020). Exploring the Potential of Deep Learning Segmentation for Mountain Roads Generalisation. *ISPRS International Journal of Geo-Information*, 9(5), 338. <https://doi.org/10.3390/ijgi9050338>
- Directive, (2007). <https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2007L0002:20070515:PL:HTML>
- Kronenfeld, B., Buttenfield, B., Stanislawski, L. Map generalisation for the Future. *ISPRS International Journal of Geo-Information*, 9(8), 468. <https://doi.org/10.3390/ijgi9080468>
- Salishchev, K. (2003). Kartografia ogólna. Wydawnictwo Naukowe PWN: Warszawa, Poland.
- Sydow E. (1866). Drei Karten-Klippen – Geo-kartographische Betrachtung. *Geographisches Jahrbuch*, 1, 348-361.