

Continuous Road Network Generalization throughout All Scales

Suba, Radan; Meijers, Martijn; van Oosterom, P.J.M.

DOI

[10.3390/ijgi5080145](https://doi.org/10.3390/ijgi5080145)

Publication date

2016

Document Version

Final published version

Published in

ISPRS International Journal of Geo-Information

Citation (APA)

Suba, R., Meijers, M., & van Oosterom, P. J. M. (2016). Continuous Road Network Generalization throughout All Scales. *ISPRS International Journal of Geo-Information*, 5(8), Article 145.
<https://doi.org/10.3390/ijgi5080145>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Article

Continuous Road Network Generalization throughout All Scales

Radan Šuba, Martijn Meijers * and Peter van Oosterom

GIS Technology, OTB—Research for the Built Environment, Faculty of Architecture and the Built Environment, Delft University of Technology, Julianalaan 134, 2628 BL Delft, The Netherlands; r.suba@tudelft.nl (R.Š.); P.J.M.vanOosterom@tudelft.nl (P.O.)

* Correspondence: B.M.Meijers@tudelft.nl; Tel.: +31-15-278-5642

Academic Editors: Monika Sester and Wolfgang Kainz

Received: 25 April 2016; Accepted: 29 July 2016; Published: 13 August 2016

Abstract: Until now, road network generalization has mainly been applied to the task of generalizing from one fixed source scale to another fixed target scale. These actions result in large differences in content and representation, e.g., a sudden change of the representation of road segments from areas to lines, which may confuse users. Therefore, we aim at the continuous generalization of a road network for the whole range, from the large scale, where roads are represented as areas, to mid- and small scales, where roads are represented progressively more frequently as lines. As a consequence of this process, there is an intermediate scale range where at the same time some roads will be represented as areas, while others will be represented as lines. We propose a new data model together with a specific data structure where for all map objects, a range of valid map scales is stored. This model is based on the integrated and explicit representation of: (1) a planar area partition; and (2) a linear road network. This enables the generalization process to include the knowledge and understanding of a linear network. This paper further discusses the actual generalization options and algorithms for populating this data structure with high quality vario-scale cartographic content.

Keywords: vario-scale; smooth; continuous map generalization; road network; scales

1. Introduction

Automated map generalization has been an important research area for years [1–3], and it is also the topic of this paper. More specifically, instead of generating a sequence of road maps with different discrete scales/Levels Of Detail (LOD) as has been common so far, we propose incremental generalization throughout the whole continuous spectrum of LODs. It is based on the idea that if features are generalized in small steps, smooth transitions between subsequent object representations can be derived [4–7]. As a consequence, it progressively leads to a simpler and simpler map. These generalization steps are stored in the vario-scale data structure called tGAP (topological Generalized Area Partition), which has been proposed in [8,9].

The vario-scale concept can potentially be applied to a broad target user group from navigation software, through desktop GIS to mobile applications. However, recent development is focusing on on-line applications with effective vector data transfer [7]. This provides a different approach to conventional discrete scale maps on the Internet. Those discrete maps have abrupt changes between scales, which can lead to disorientation, result in confusion and eventually the frustration of map users. Figure 1 shows a detailed road network in one map scale in comparison to the sparse network at the next available scale.



Figure 1. An example of map fragments at two scales (left at a larger scale, right at a smaller scale). Note that only the water and road network are displayed for these two scales, which were obtained directly from the original multi-scale database without any content modification (source: OpenStreetMap, styling: MapBox Studio).

In this paper, our working assumption is that by capturing the whole generalization process from large-scale input data, being progressively collapsed from areas (large scale) and/or merged into lines (small scale), it is possible to obtain a better representation with the following properties:

- Better preserved (road) network connectivity.
- Changes in geometry and/or classification are gradual.
- Small details are gradually eliminated (not important dead-ends (cul-de-sacs) are removed and no new dead-ends are created).
- Road network semantics are taken into account, and the overall map impression is improved.

This may give in theory a better impression to the user, e.g., while zooming in and out. Note that user testing should verify that in one of the future publications [10].

On the other hand, a side effect of the transition phase from large to small scale is that area and line representations are mixed together, e.g., second order roads can at the same map scale be represented partially by areas and partially by lines. Thus, our approach does not rely on homogeneous road segment representations by design. This is a non-trivial issue for conventional data structures and requires specific data modeling and modification of the data structure, both of which are the main goals of this paper. In previous papers on vario-scale data using the tGAP structure, it has been mentioned several times (in the “Future Work” section) that it should be possible and highly desirable to include features that have a line representation to improve the content of the map significantly. Now, for the first time, this is being realized and tested with real data.

To summarize briefly, our main goal is to introduce line feature representations in the tGAP structure to have better support and data content for smooth zooming. To accomplish such a goal, we have the following requirements based on our vario-scale concept:

1. The most important is to deal with mixed areas and line representations.
2. To generalize in small steps.
3. To use an area partition as input.
4. To preserve the meaning of road network.

The remaining part of this paper is structured as follows. Section 2 reviews related work and explains the vario-scale data structure in more detail; followed by Section 3, which explains the whole process of road network generalization in detail. Section 4 demonstrates the results. Finally, Section 5 concludes the paper and presents some open questions for our future work.

2. Related Work

A number of earlier attempts have explored ways of generalizing road networks. The roads (similarly rivers and water channels) as linear/infrastructure objects are the “backbone” of many map types. They improve the legibility of maps and help users to orient and recognize the depicted real-world situation more easily. On top of that, road network generalization forms a prerequisite for all other topographic generalization action (operators) and is thus a fundamental operation in the overall process of map and database production [11]. Therefore, road processing has been extensively studied where at a given scale, two main representations of the roads may be found: linear and area representations. On large-scale maps, for countries such as Belgium, the Czech Republic, The Netherlands or the United Kingdom, a road segment is represented by area geometries. Together, the collection of road areas represents the road network. These areas form an implicit road network graph comprised of edges and nodes. At smaller scales, the road segments are represented by line geometries. These lines correspond more directly to the edges in the road network graph. At smaller scales (with road segments directly represented as edges), the road network generalization can have the emphasis on: (1) the linear road representation itself; or (2) the areas between the roads as regions bounded by minimal road loops or cycles, e.g., containing built-up area, forest or terrain. The linear emphasis approach considers the network as a set of linear connected elements, while the regions emphasis approach concentrates on the “space between” the roads. Sometimes, the term “area partition” [12] or “mesh density” [13–15] is used. For smaller scale road network generalization, where road sections are represented as lines, both views (either linear or regions emphasis) have their own advantages and disadvantages [12].

Moreover, at the smaller scales there are two main perspectives within the approach with a focus on the linear network representation: (1) strokes; and (2) segment generalization. The stroke-based generalization groups the road segments into longer lines, which may cross without explicitly intersecting. The decision to intersect may be based on some criteria, such as geometry (angle between segments), topology (node degree of two), attribute or classification [13,16]. In the segment-based generalization, the road segments (from junction to junction, where the topological degree of end nodes is greater than two) are the smallest atomic elements for removal. The main advantage of using strokes is evident: it preserves information about the connectivity between segments. This indicates that the stroke-based approach can be a useful generalization tool. However, Turner 2007 [17] points out that segment analysis (the creation of a segment map as known in the space syntax community) can give comparable or better output compared to a stroke-based method and even produces more meaningful results, e.g., better in correlation with observed vehicular flow in reality.

Furthermore, there are indications of developments shifting towards smooth, continuous solutions in generalization, resulting in vario-scale representations. This is especially the case for the road network. Cecconi et al. investigate automated on-demand generalization for the web [18,19]. Even though they study multi-scale solutions only, some concepts are valid for our solution, as well. They assume that corresponding objects of the different LODs are linked together, and the user with the help of on-the-fly generalization could select an intermediate scale between them. In this way, the perceived generalization can be understood as an “interpolation” (or morphing) process between two different geometries. They also analyze the scope of the applicability of the generalization operators over the desired range of scales. They investigated the limits of applicability, where the “regime” of operator changes; however, they consider representation as a continuous function of scale, which has parallels with our approach.

Brewer 2007 and Touya 2013 describe an interesting tool called ScaleMaster, which supports automatic multi-scale generalization [20,21]. It is based on the model that formalizes how to generalize map features from different datasets through the whole range of targeted scales. Despite the fact that the tool focuses on generating a multi-scale/multi-representation solution, the idea of defining generalization actions for a range of map scales is similar to our approach.

Noellenburg et al. give interesting examples of smooth transition for linear features between their representations at two scales [22]. It focuses on interpolation between two scales instead of line simplification. Even the fact that it is more similar to morphing in computer graphics (animation) than map generalization (generating smooth content in our case), it still gives a nice example of representation through multiple scales.

Chimani et al. apply a method where they remove edges of the road network map one by one. Therefore, the map is getting simpler and simpler [6]. There are almost unlimited numbers of possible orders in which edges can be removed. Therefore, they try to define the sequence of removing edges that gives the best result based on some conditions, while preserving graph connectivity. The best simplified map is the map with minimal change in connectivity. The method tries to minimize the sum of all differences for individual simplified maps, similar to the principle of the least square adjustment, where the minimized total of all changes is the optimal solution. To compute all possible permutations, they used linear programming. This optimal solution was used as a benchmark as the best quality reference. They then developed two novel heuristic optimization algorithms and compared them with the benchmark linear programming solution. They compared how well the two different algorithms approached the benchmark. It is one of the first papers to focus on global criteria during the continuous generalization; however, the quality of the generalization is still problematic. It shows that connectivity is not only a sufficient criterion for a good map and does not cover all necessary aspects of road network generalization. They faced the problem that a road segment of an important stroke can go missing or that a part of the network does not always nicely span the map extent, leaving large parts of the map empty. Furthermore, the overall impression of the map, where large rural and small urban areas should still be recognizable in a later stage of generalization, is still an issue.

Another truly continuous solution is proposed by Li and Zhou, including experiments with real data and an extensive evaluation [15]. They compare both approaches; strokes and mesh density, and combine them in their so-called integrated method to create a universal solution for road network using the advantages of both approaches. Their solution generates two separate linear and areal hierarchies. They combine them together to provide continuous multi-scale representations of a road network. This rather complex solution is based on the omission of features, with the analysis of what should be eliminated. This analysis takes place only at the beginning of the process. Changing parameters of the generalization process is not considered; e.g., rules/parameters adapted during different phases of the generalization. On the other hand, the performance test looks really promising and suggests that the approach is quite feasible and will be good for on-the-fly use.

Our approach uses a specific vario-scale data structure, which stores the results of map generalization actions; features are generalized in small steps, progressively leading to simpler and simpler maps. This process is based on repeatedly processing the least important feature, based on a global criterion. We assume that objects can be well generalized with optimized algorithms and appropriate parameter settings for use at any map scale. The intermediate representations of the objects can be defined by a sequence of generalization operations that eliminate small and less important objects to satisfy the representation constraint, while simplifying the boundaries of these objects. The vario-scale data structure, known as tGAP (topological Generalized Area Partition), which has been proposed in [8,9], captures these incremental changes with minimal redundancy. Both the detailed objects at the largest scale and the intermediate objects generated during the generalization process are stored in a set of database tables up to the most generalized representation for the smallest scale. The additional value of the vario-scale concept compared to the multi-scale approaches lies in the following aspects:

1. Redundant storage is eliminated as much as possible by avoiding duplication of features (as in multi-scale) and storing the shared boundaries between neighboring areas instead of the explicit polygons themselves, i.e., using a topological data structure composed of nodes, edges and faces.
2. More intermediate representations can be constructed, because the automated generalization process records for every topological primitive (node, edge or face) the valid range of map scales

for which this element should be shown. This can be used to meet the increasing demand from users for additional maps at intermediate scales [23].

3. The lineage of the generalization steps is stored explicitly and implies the links between generalized and original objects. Obtaining these links is regarded as a very difficult task for multi-scale databases. The main advantage of this is future updates of the features without re-generating the dataset.

In this perspective, we consider the vario-scale approach as a production of a sequence of successively more generalized maps, so that these maps go well together, similar to [6]; instead of considering each level of generalization independent. In our perspective, the sequence of generalization steps is more valuable than just the final map of specifically-identified scale. Our method has no explicit lower and/or upper bound with respect to the scale-range we target. This makes the method very generic.

Moreover, the principle of the tGAP structure can lead to smoother user interaction when zooming and panning [24,25]. It is also possible to create an explicit 3D representation of 2D features where scale is integrated as the third dimension. It allows a single real-world feature to have a single database representation, by contrast with discrete scale approaches, which not only have different representations, but where these are often maintained separately.

In the recent past, we first focused on roads in the smaller scales [26], where road objects were only represented by lines, and we stressed preserving the connectivity of the network. Later, we focused on road network features in the larger scales [27], but these investigations were limited to the theoretical concepts. This paper places both phases in an automated process where both approaches are seamlessly integrated. It deals with the problems of integration, designing an algorithm and setting up the generalization rules to preserve road connectivity. The main challenge remains; the generalization must be carried out in small steps. Later, it tests the developed method with real data.

3. Data Structures and Processing Strategy

This section introduces our proposed processing strategy for road network generalization (Section 3.3). Before doing so, it first introduces terminology in Sections 3.1 and 3.2 to be able to better explain our strategy.

3.1. Granularity

Continuous generalization requires geometric changes (big or small) between generalization steps in the process. We call the amount of the features (data) changing in one generalization step the granularity [27], and we distinguish the following levels:

- course granularity, when all features are involved at once, e.g., all roads are omitted.
- medium granularity, when all features of a certain class or subclass are processed, e.g., all local roads with speed limits are removed.
- fine granularity when one single feature is processed, e.g., one dead-end road is removed.
- finest granularity, when a part of a single feature is involved, e.g., one road segment is removed.

Since we aim at a more gradual transition without significant modification in geometry, the changes should be small. In our case, finest granularity is then optimal. This guarantees that the changes are as small as possible, aligning well with the vario-scale concept.

We apply merge/remove, collapse/split and simplification operations for these object parts, which can result in a road object composed from segments represented by mixed lines and faces; see Figure 2. From a traditional cartographic point of view this might seem less favorable, but from a vario-scale point of view, it is desirable. The differences in the representation can be compensated in the visualization by applying proper styling, when the line segment is represented by lines of the same thickness equal to the width of the adjacent area segment. Furthermore, the fact that the generalization process is performed in small steps leads to simpler problems, which are easier to

compute or implement. On top of that, the history of steps is stored explicitly, and this implies that the links between generalized and original objects are present. Often, these links are missing in multi-scale implementations.

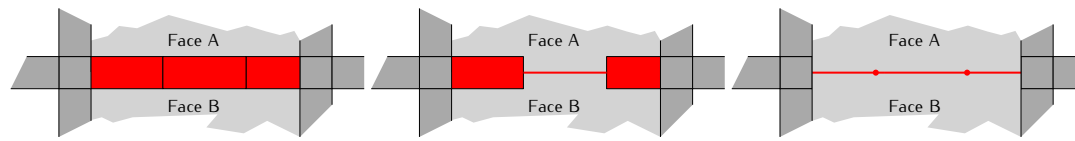


Figure 2. Side effect of the gradual transition from one scale to another (from left to right). For some reasons (perhaps different attributes; e.g., road surface type, name or speed limit), the red road consists of three parts. To achieve a gradual transition, the individual parts are generalized separately. It changes the representation from areal at the most detailed scale (left) to semi-linear at the “halfway” scale (middle) to linear at the final scale (right). Be aware of the fact that the complete road in one moment of the process is represented by both areal and linear parts at the same time. Note that one of the consequences is a topological change, where Face A and Face B become adjacent.

3.2. Level of Abstraction

The input datasets that are currently supported within the tGAP structure are modeled as a two-dimensional polygonal map, i.e., as a partition of the plane in a geometric sense, without gaps and overlaps. As a result, the data structure contains only topological primitives; nodes, edges and faces, where one area object in the map corresponds to just one topological face. The same is true for roads in the input dataset, these are represented by faces only (areas).

Besides classification and the planar area partition, there is more information/semantic information implicitly present in the large-scale input map, such as the linear networks (road infrastructure or rivers and water channels). These linear networks are implicitly in the input data. We wish to preserve their natural meaning in the target map of the smallest scale as well. Even though these features are part of a network they are often not explicitly modelled. Therefore, we will make the implicit information about the role features play inside the network explicit, even when not given as input.

Figure 3 shows a simple example of such a network in the input (large scale) (see Figure 3a) and in the target small scale (see Figure 3d). The road network can be easily derived from both figures. When two road segments are both represented by areas (large scale), then they are incident when they share an edge. Where at least one of the road segments is not represented by an area, then they are incident when they share at least a point (node). However, it is not so simple to keep track of the linear network as the map changes from scale to scale in a more gradual way; see the intermediate steps in Figure 3b, c. Therefore, the same situation is captured in Figure 4, but this time, the linear graph of the road network is depicted for better understanding of the relationships among the road objects. As mentioned, the road segments can be incident with other road segments. Depending on the number of incidences, a road segment plays the role of junction (node) or connection (edge) in the linear network graph. From Figure 3a–d, we can see that the geometric embedding of the objects in the map changes gradually, while the conceptual linear network graph stays the same. This gives us an effective tool for meaningful road network generalization throughout the scales.

Keep in mind that in the following description, the road segments of different dimensions can be in an interaction with each other, e.g., a 1D line can be in an interaction with a 2D area road segment. Therefore, the road segments can be classified (especially in gradual changing scales) based on the number of other incident road segments and are included in the data structure as follows:

- A road segment is classified as an isolated segment when it has no other road segments incident.
- A road segment is classified as a dead end if it has exactly one other road segment incident. It is represented either by a face or by an edge in the topological data structure.

- A road segment is classified as road connection of the network when it is incident with exactly two other road segments. It is represented either by a face or by an edge in the data structure.
- A road segment is classified as road junction of the network when incident to more than two other road segments. It is represented either by a face or by a node in the data structure.

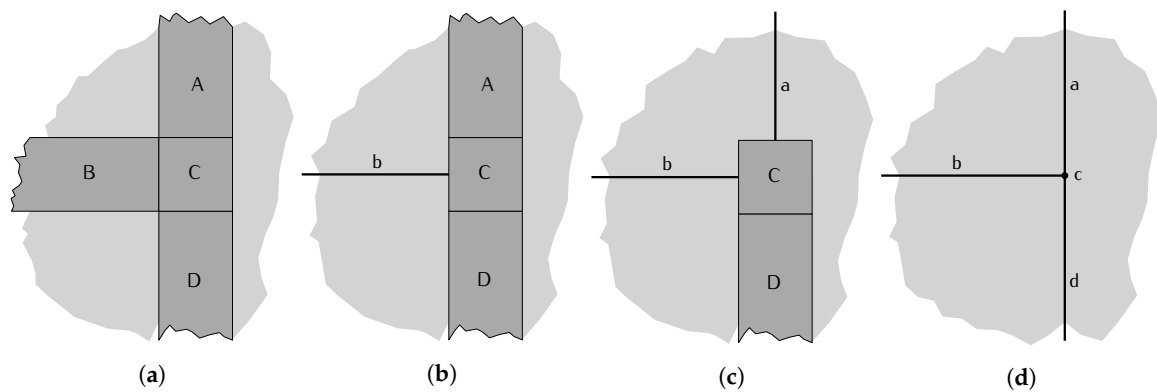


Figure 3. Example of road network generalization from the large scale (a), through intermediate step (b,c), to the final scale (d). Geometrical representation of the road segment/junction changes from 2D areas to 1D segments/0D point, but “feature” (semantic), and their role in linear network remains the same.

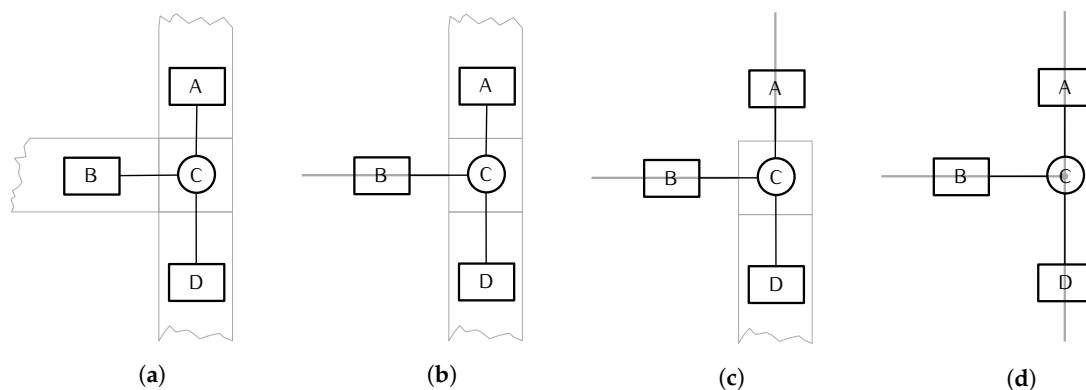


Figure 4. (a–d) The linear network graph represents the same situation as Figure 3. It captures the topological relationships of road objects where a rectangle indicates a road connection and a circle a road junction.

In this way, we can associate any selected feature in the map at any stage of the process. The assumption related to the input data is that connections and junctions of a road network are well defined in the input dataset; see Figure 5a. However, when the input data do not conform to this assumption (see for an example Figure 5b), an additional pre-processing step must be taken. This can be done by applying the constrained Delaunay triangulation to obtain properly-classified (connection or junction) area road segments, as proposed by Uitermark et al. [28].

3.3. Strategy for One Generalization Step

The generalization process to generate content for the vario-scale data structure is based on the tGAP-principle (find the least important object and merge it with a most compatible neighbor) extended using linear network knowledge. Note that the principle is designed in a very generic way,

such that it is possible to mix different generalization operations. It is also integrated in the sense that all features for which the operations are performed are treated together (all features are geometrically integrated in the same planar area partition that is used as input). There are many design decisions in the development of this process, and we will label these as design decision i (with i a sequence number). Often, there are several alternative options, but based on our experience and some limited testing, we present our initial “best guesses” for these decisions.

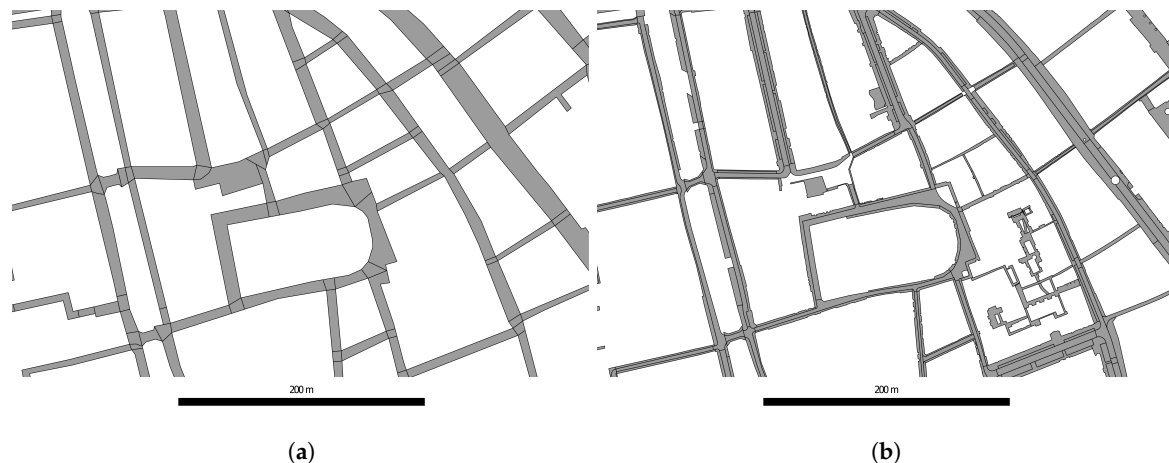


Figure 5. Two map fragments of different possible input showing the city center of Leiden, The Netherlands: first (a), the topographic map (TOP10NL) intended for use at a 1:10,000 map scale; second (b), the BGT base map (in Dutch: Basisregistratie Grootschalige Topografie) intended for use at a 1:500–1:5000 map scale. Note that only road network features are displayed.

Design Decision 1: We distinguish in the creation of the vario-scale content for the data structure three classes of objects only: roads (sub-classified as either junction or connection), water and other objects. This will drop the number of object types during the creation of the vario-scale to three, which makes decisions more transparent. Besides road network processing, this also allows us to treat water differently from other non-road classes. Note that the original classification of the other classes is kept and used later on in visualization (but not during the creation of the vario-scale structure). Alternative processing design decisions here could have been made: just two classes (road, other: even simpler), three classes (same, but more refined, with two subclasses for water: junction, connection), more classes (as currently present on the topographic base maps, where the classification of these other classes is used; e.g., for selecting the most compatible neighbor to merge with).

At the beginning of the process, every face in the structure gets an importance value based on the type and the size (area) of the feature (in the initial large-scale map, there are only area features). Note that the computation of the importance value can be refined; see Section 5. Based on the importance value for every face, the process starts picking one face after another and performs specific actions based on the type of the chosen face. The face with the lowest importance value is processed first (**Design Decision 2:** as in the integrated data structure, both area and line representations of roads (and other features) are possible; an alternative to having just faces in the importance queue is also having line or node features in the importance queue.) Depending on the type of a face, there are the following processing options; see Figure 6.

- The selected face is a road junction and will be either merged with the adjacent road junction or preserved until all adjacent road connections are collapsed. If so, then also the face itself can be collapsed. If not, then the importance is raised and the face is put back in the queue (and will be processed later) (**Design Decision 3:** Instead of postponing the processing of the junction, it would be possible to directly collapse it to a node (even if not all incident connections are collapsed)).

- The selected face is a road connection and will be merged with the adjacent road connection. If there is no such face, then it is collapsed to a line.
- The selected face is water and will be merged with another adjacent water face. If there is no such face, then it is collapsed to a line.
- The selected face is the other object and will be merged with an adjacent other object, if there is any; otherwise, the face is collapsed (**Design Decision 4:** If no adjacent face of type other is present, an alternative design decision instead of collapse would be to raise the importance and put it back in the queue. Later on, when one (or more) of the neighbor road faces have been collapsed to a line, then the selected face might have an adjacent other face). Adjacent other objects with no collapsed roads lying between are the most optimal to merge with. If there is no such adjacent option, another object with the least important collapsed road (edge) between is selected. When the collapsed road lies between, the faces will be merged (and the collapsed road will be removed).

This recipe guarantees to generalize the roads in a meaningful way and is continuous for all faces in the structure. Roughly speaking the following happens with the roads: Initially, the area road segments are collapsed, and later on, the merging of the other areas with each other takes place. To which neighbor the other area is merged depends on the edge between. If there is no collapsed (line) road segment in between, then this has preference. If all edges represent collapsed road segments, then the least important one is selected, and this decides with which neighbor area it should be merged. With this type of merge, the unimportant linear road segments are “automatically” removed, as well.

It is important that the least important collapsed road (line) is determined by looking at its classification, the local configuration (connectivity) and length. Specifically, for every potentially collapsed road, we look at its classification first. If this gives a “winner”, we pick this road as the least important collapsed road. If this results in a tie (collapsed roads having the same classification), we compute for every collapsed road a connectivity value. The connectivity value for one road is defined by how many routes between other roads go via this road, with more passing routes meaning higher connectivity. The least important collapsed road is the one with the lowest connectivity value. As a last resort, in the case of roads also having the same connectivity value, the shortest of the collapsed roads is selected as the least important one. An in-depth description is given in [26], where this has been extensively tested.

Note that another (non-road) area is collapsed if and only if it is completely surrounded by areas of roads, e.g., a face of grass is between two faces of road (grass strip between lanes of a highway). This is a rare case where the collapse of a non-road feature is the most favorable. Another option would be to return the non-road face back into the queue and wait until at least one road nearby is collapsed (as indicated in Design Decision 4). However, the collapse operation is preferred in this case, because it makes sense to assign the parts of this unimportant face to the neighboring road faces, and it will reduce the number of faces by one.

Design Decision 5: In the beginning of this section, we defined road junctions and road connections. During the generalization process the configuration changes, which gives two options: (1) faces keep their original road (junction/connection) subclassification even if it is in contradiction with our definitions; or (2) faces are reclassified when needed in order to remain consistent with our definitions for junctions and connections. We opt for the first option, because this results in a slightly better vario-scale cartographic quality according to our visual inspection; cf. Section 4.3.

The above-described iterative generalization proceeds in steps. Normally, there is one face fewer after every step, and the number of faces never increases. There is some delay when a road junction area still has a road connection area as a neighbor, causing the processing of the junction to be postponed. However, the neighboring road connection areas will all be collapsed to lines at a certain moment in the process, and after that, also the delayed processing of the road junction areas can take place. This proves that our algorithm will not enter an infinite loop.

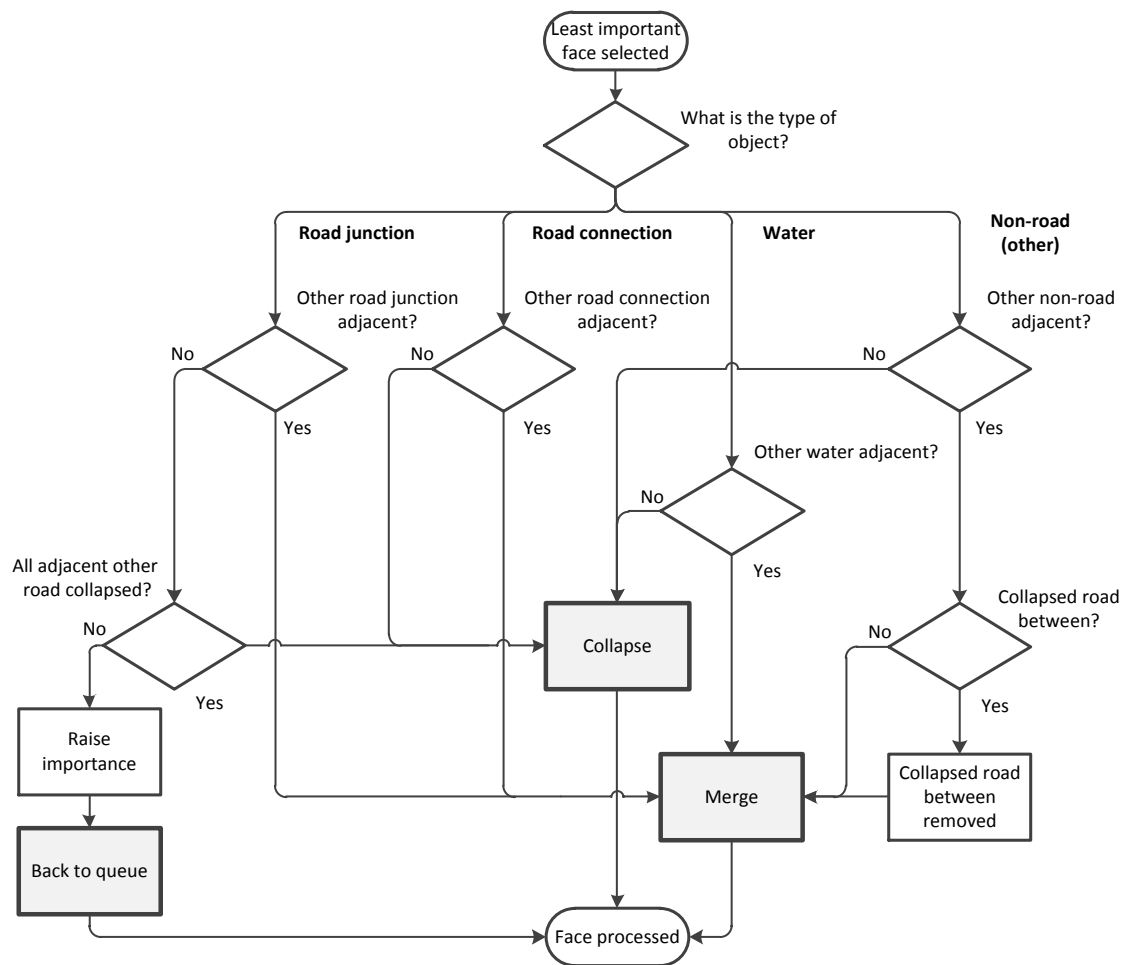


Figure 6. The way one generalization step is performed. Note that the road area will never merge with a non-road (other) area.

4. Results

This section addresses the cartographic quality (Section 4.1), the quantitative analysis carried out (Section 4.2) and additional discussion items (Section 4.3).

4.1. Cartographic Quality

For our experiments, we loaded a subset of the Dutch topographic map (TOP10NL) intended for use at 1:10,000. The two following regions were used:

- A rural region, area of 7 km × 7 km with 11,300 faces as input; see Figures 7 and 8.
- A city center (of the city of Leiden). It is a 1 km × 1 km region with 19,400 faces as the input; see Figure 9.

These datasets are provided as simple feature polygons, where terrain, water and road network layers together form a complete planar area partition. The road segments are present in the dataset with geometries for road junctions and connections. The layer of buildings lies on top of those layers. First, we “fused” all data layers together and created a planar partition as input that we converted to a topological data structure with the help of Feature Manipulation Engine (FME). At this stage, all objects are represented by areal geometries. Then, this input is processed into a vario-scale structure with the help of merge/remove and spit/collapse [29] generalization operations. Line simplification could be included in the process; see Figure 10c, d. However, this is not the case at this stage of the research process in order to see the features’ geometry without any additional effects.

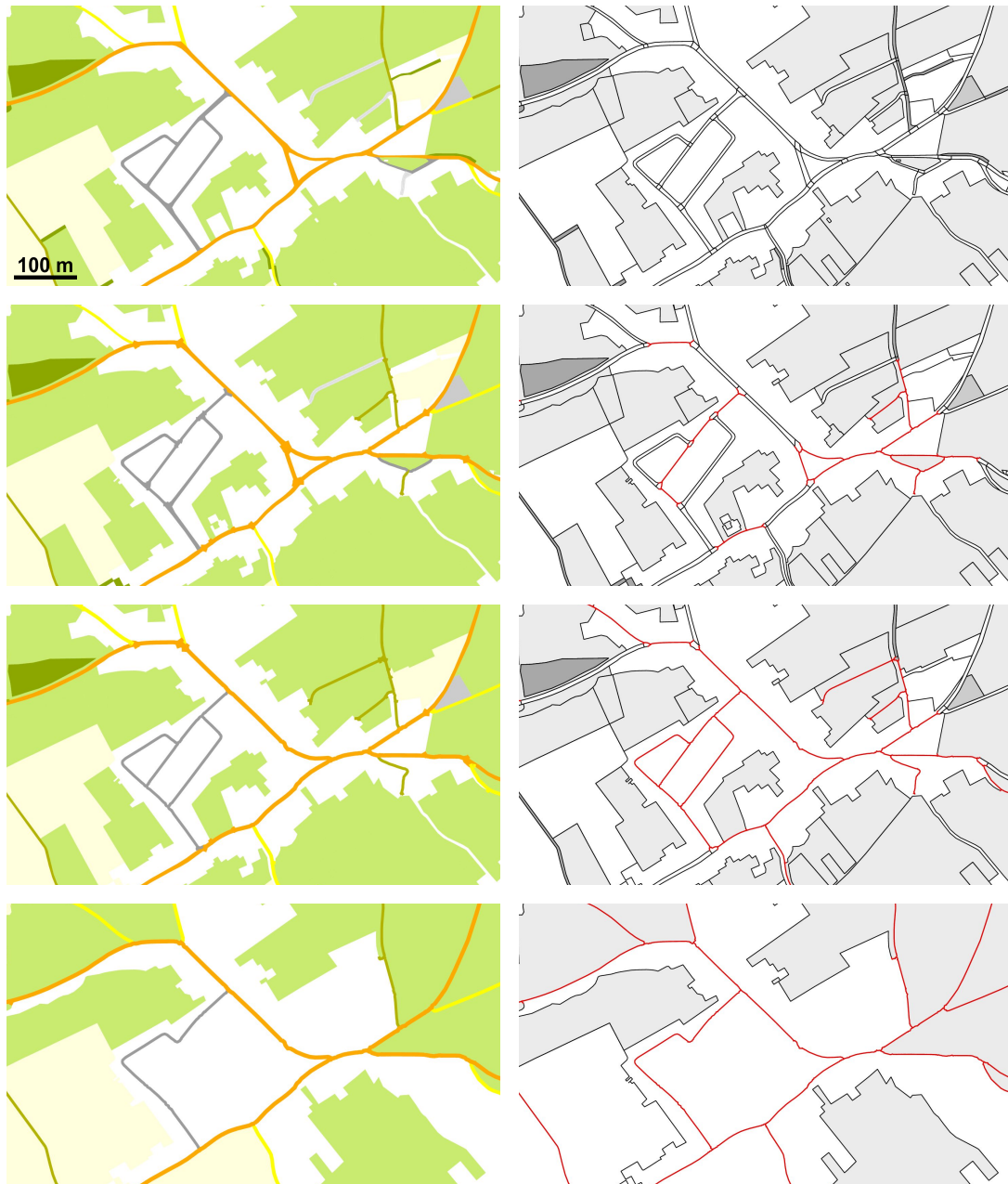


Figure 7. Detailed situation throughout the scales. All maps are at the same scale (and the exact scale is not very relevant). It illustrates how the structure evolves. The top maps are at the input scale, where all objects are represented by areas. The “halfway” scales are in the middle, where the representation of roads is mixed (by areas and by lines) and generalized scales are at the bottom, where roads are represented only by lines. The styled map (**left**) corresponds to the situation in the tGAP structure (**right**). Red lines (right) indicate collapsed road edges. Note: all map fragments are displayed at the same scale to clearly show the effect of the generalization process (in reality, the bottom map fragments should be shown at smaller scales).

Figure 7 shows the impact of our method on real data. It demonstrates that by making small generalization steps, we got incrementally a simpler map. At intermediate scales two representations for road objects (areas and lines) are mixed.

Figure 8 presents a sequence of maps for a small part of the rural region retrieved from the tGAP structure as generated by our algorithm. It demonstrates the outcome of the algorithm (vario-scale content), but it does not correspond to the correct user impression as the scale is fixed in this figure and vario-scale map use should be an interactive experience during zooming. At least it gives an

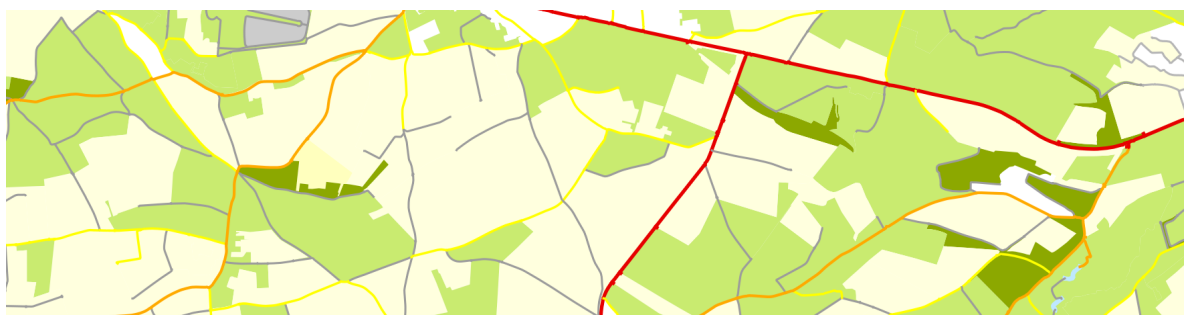
impression of the content of the different map scales. Figure 9 shows a proportional re-sized map sequence to give a better impression of how a user should perceive the derived maps.



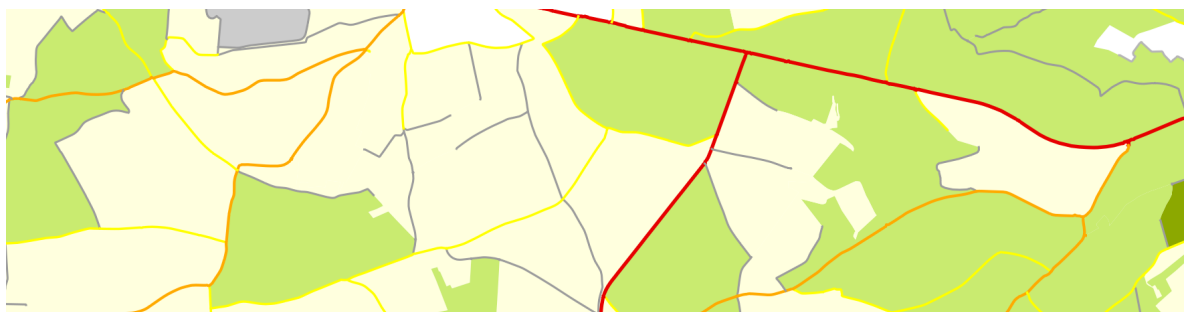
(a)



(b)



(c)



(d)

Figure 8. An example of the generalization process in the rural region (all at the same scale). The top figure (a) shows the input. The other figures (b), (c) and (d) show map fragments after 86.9%, 94.4% and 98.2% of the process, where 0% is input (a) with 11,300 faces and 100% is one face.



Figure 9. (a–f) An example of the generalization process for the city center dataset. (a) shows the input; The other subfigures (b–f) represent 60.6%, 68.7%, 80.8%, 88.9% and 95.0% of the process, where 0% is the input with 19,400 faces and 100% is one face. Note: the map fragments are displayed at the intended target scales; (a) 1:10 k; (b) 1:16 k; (c) 1:18 k; (d) 1:23 k; (e) 1:30 k; (f) 1:45 k.

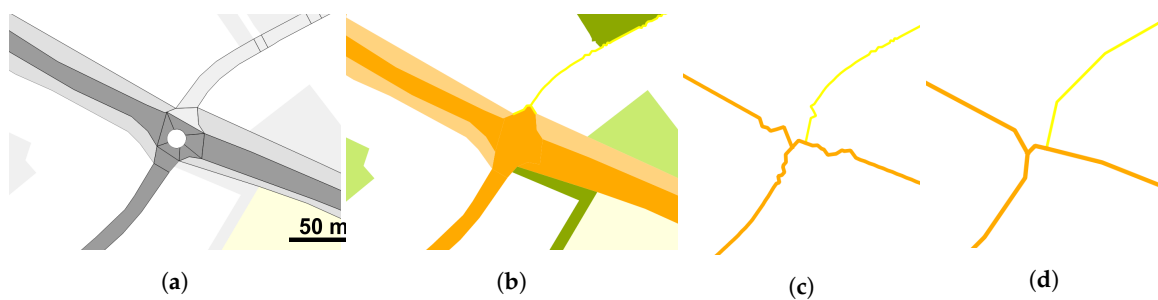


Figure 10. Example of the process for a roundabout. It changes the representation from areal in the initial partition in (a), partial areal and partial linear in (b) to linear (c). (d) demonstrates the option when line simplification is used.

Figure 10 presents a detailed situation of a simple roundabout. Note that these types of infrastructure objects are not present as separate entities, nor classified as such; their road segments are dealt with individually.

4.2. Quantitative Analysis

Due to the nature of illustrations on paper, only specific intermediate map scales can be shown. The intention is to use the vario-scale data in combination with interactive zooming and panning operations over the map. However, there are only a few existing measures to evaluate continuous map generalization in general. Thus, we used visual inspection and compared our results to previous developments. To give a better quantitative notion of the process, we also generated some graphs, which give better insights into the whole generalization process.

First, Figure 11 shows the proportion of feature classes throughout the generalization process. For every generalization step, we count how many objects in the tGAP structure there are for a certain feature class. Then, those feature classes are grouped into “water”, “terrain”, “buildings”, “roads” and “other” super classes. This graph corresponds to the example depicted in Figure 8. Similarly, Figure 12 shows the area these objects cover in the structure. Note that a collapsed road has no area (even if it is still a map object) and that measured in the area of the roads has a smaller share than when expressed proportionally (%). Near the end of the graph in Figure 12, road objects do not occupy any area, while there are still road line segments. The water bodies are small and occupy only a small portion of dataset; therefore, they do not survive long in the process.

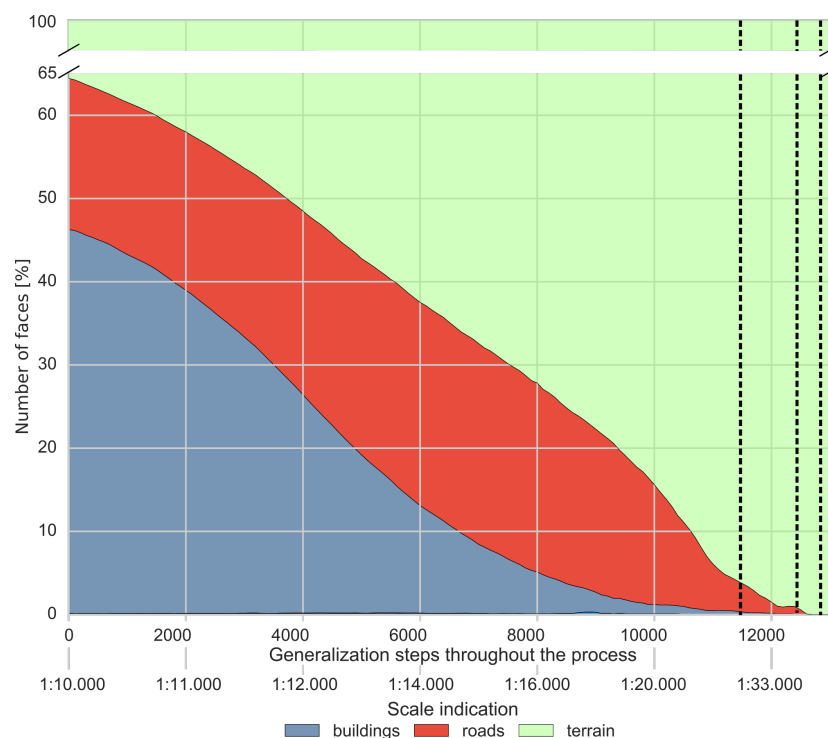


Figure 11. Ratio of feature classes throughout the generalization process in topological Generalized Area Partition (tGAP). The numbers are relative (number of faces per feature class divided by all faces present in the map). Note that the total amount of objects decrease throughout the process. The vertical dashed lines indicate the map scales shown in Figure 8b–d.

Second, Figures 13 and 14 provide another indication of the same example from Figure 8. With proper styling and color schema, it is not obvious which roads are still areas and which have been already collapsed. Therefore, these graphs present absolute numbers of road objects. Figure 13 captures only roads objects represented by areas. Figure 14 shows the number of edges in the structure

(representing the collapsed road objects). It shows that the process collapses the majority of road areas first (creating new road lines). In part b, the road lines between two merged faces are removed. This corresponds to our designed process.

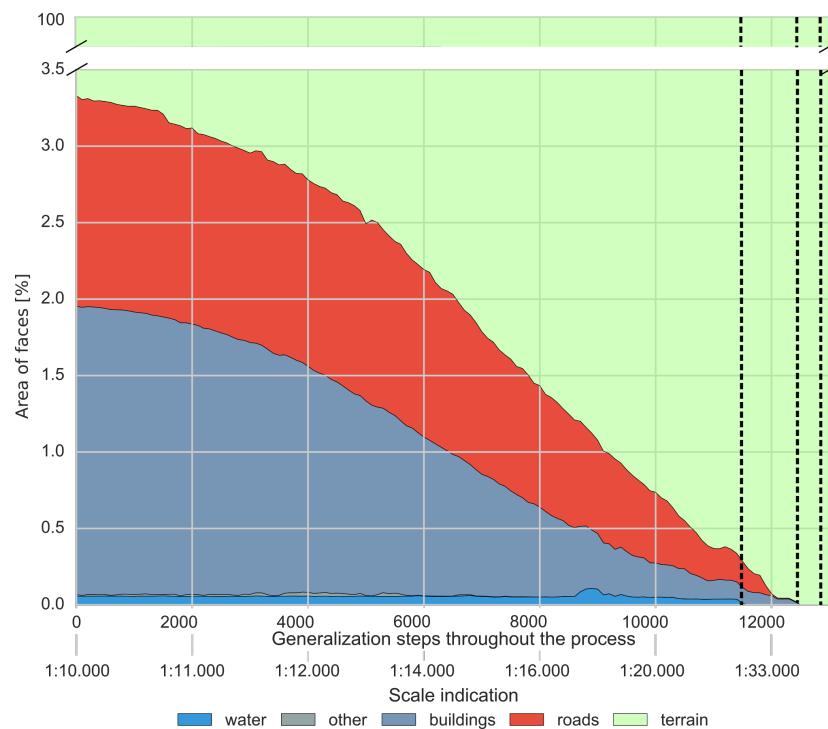


Figure 12. Ratio between the covered area in tGAP structure for feature classes throughout the process. The graph relates to the data of Figure 8. The vertical dashed lines indicate the map scale of Figure 8b–d. The numbers are relative.

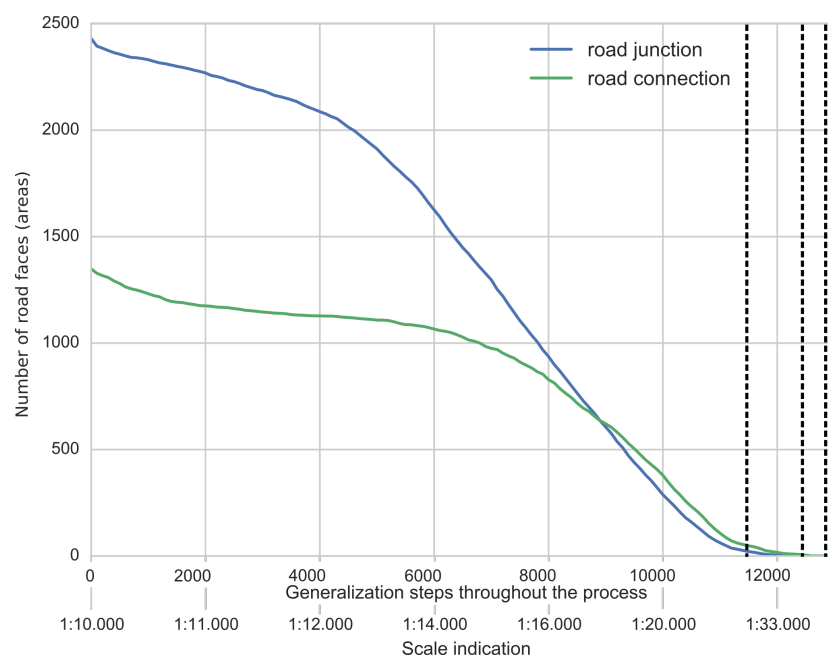


Figure 13. Number of road faces represented by areas throughout the process in Figure 8.

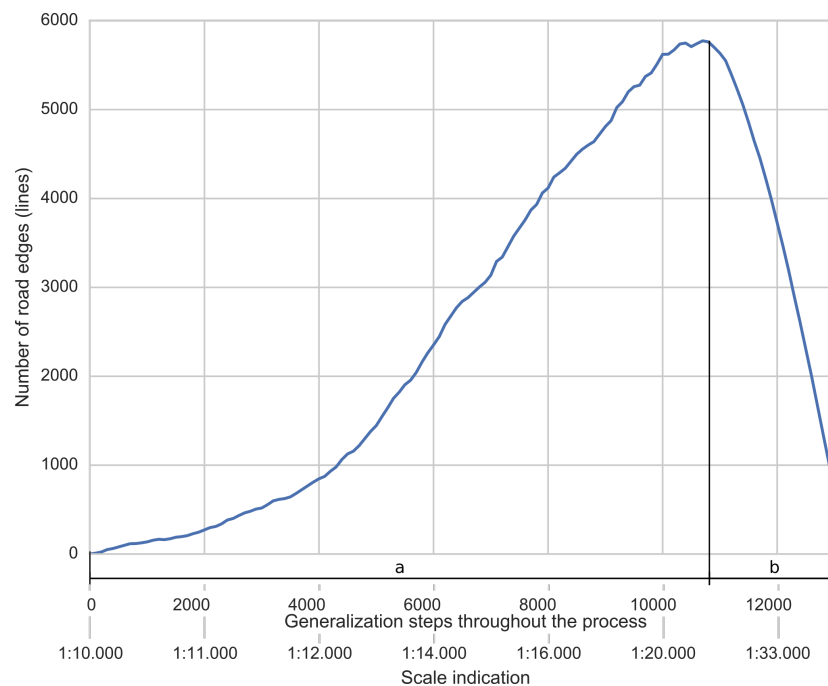


Figure 14. Number of road line edges in the structure (collapsed roads), related to Figure 8. Graph part a, where the number of collapsed road connections increases, represents the initial stage of the process, where roads are split/collapsed. Part b, where the graph decreases, indicates a situation where two area objects merge, intentionally removing the collapsed road lines between.

Graphs presented so far have shown only small value changes (second derivative). It indicates a gradual process that corresponds to our goal presented earlier. It also suggests that our generalization rules in the overall strategy were quite reasonable.

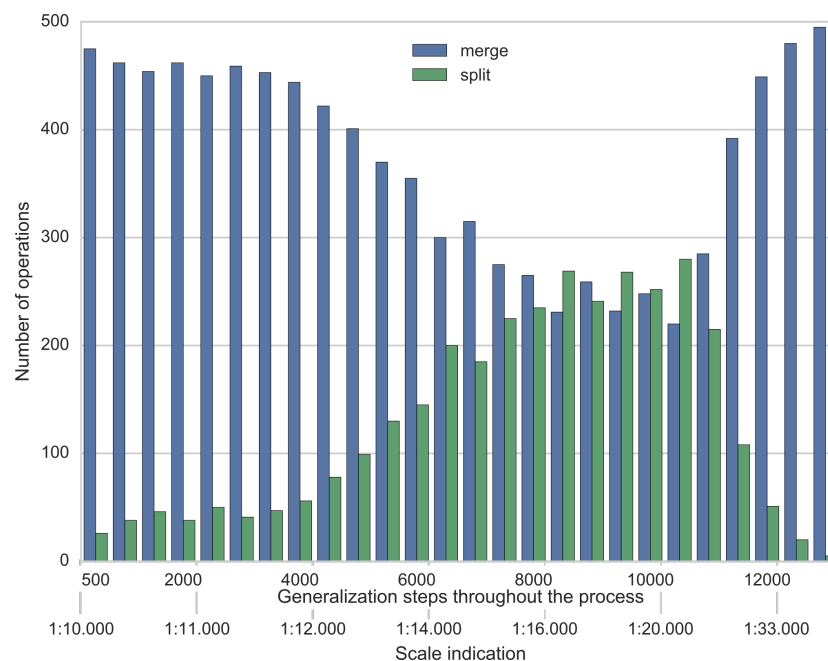


Figure 15. Application of the merge/remove and the split/collapse generalization operator throughout the generalization process. It is related to the example in Figure 8.

Finally, Figure 15 shows the usage of the different generalization operators throughout the process. Exactly one operator is applied in every generalization step, either the merge/remove or the split/collapse. The graph summarizes what happened every 500 steps and shows the ratio between the operators. One can observe that a lot of merge/remove operations happen at the beginning when tiny faces are merged. Those faces are mainly slivers from the preprocessing step when layers of buildings were “fused” together with the other layers. Since they are small, they have low importance and are processed first, but this could have been an additional preprocessing/cleaning step.

Later in the process, the split/collapse operator is more dominant because road objects are processed. Finally, the merge operator becomes more significant again because most of the roads are collapsed, and other objects are then merged together, removing any collapsed roads between.

4.3. Additional Discussion Points

The results above have shown some reasonably good outcomes in automated and continuous map generalization. However, there are still quite a number of design decisions (as mentioned above) and some additional issues, which have been encountered during the design and implementation. In most cases, the best solution is not yet known, and further research is needed. The list of additional issues includes:

4.3.1. Road Classification

Road objects were classified as road junction or road connection, based on the number of incident road segments, where a junction should have more than two road neighbors. There are two options when this classification can take place (Design Decision 5). In the first option, the objects are classified in a pre-processing step, and then, the same knowledge is used throughout the whole process. In the second option, more dynamic, the objects are reclassified during the process, when needed.

Figure 16 shows the processing sequence for both approaches, starting from an initial configuration where humans would recognize the strip of grass between two roads running in parallel. We can see that static classification (on top in the figure) identifies two junctions and two connections at the beginning of the process. Then, connections are merged together and then collapsed. On the other hand, dynamic classification (at the bottom) recognizes everything as junction objects (all objects have more than two incident roads). Then, two road junctions are merged, and a new classification identifies one new road connection, which is collapsed later. The last step is the merge of the two remaining road junctions.

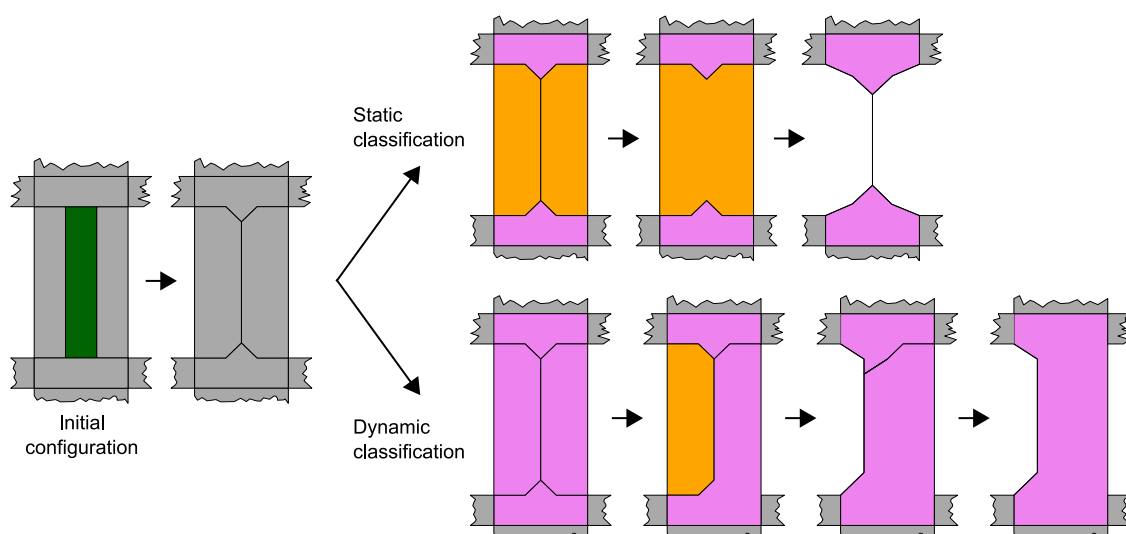


Figure 16. Alternate road object classification during the generalization process. It starts with collapsing the strip of grass between roads. Road junctions in pink and road connections in orange.

The generalization process would continue and remaining junction(s) would be collapsed. Note that for dynamic classification this would happen much later because the final junction is larger (higher importance) and would lead to a different geometry of collapsed roads.

Besides technical aspects, such as memory use or time complexity, a static approach may lead to incorrect classification (of junctions and connections according to our definition) during the process, but it gives a slightly better overall cartographic impression during visualization. Therefore, we used it in our implementation.

4.3.2. Reclassification after Split/Collapse

Design Decision 6: When the face of the road object is collapsed, the newly-created edges should carry a correct classification; see Figure 17. The face is transformed based on the skeletonization to a set of edges. Most of the time, one “main” branch corresponds to the shape of the original face. However, the most appropriate reclassification for new edges is not so clear. Should all branches receive the same classification or should only the main branch be classified?

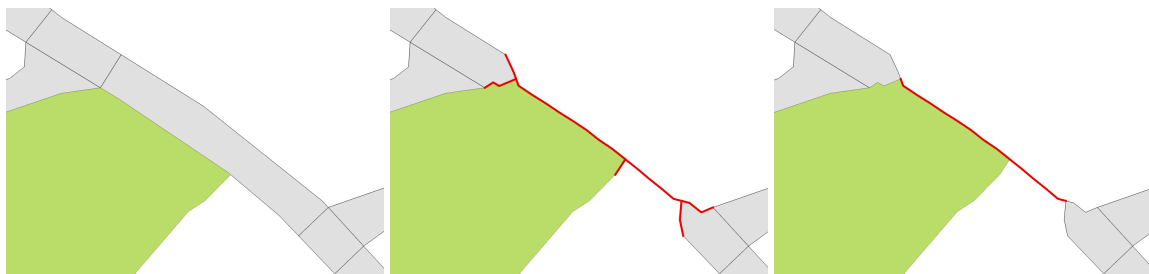


Figure 17. Reclassification after split/collapsed operation. One road connection object from the initial configuration (**left**) can be reclassified in two ways. Either all branches (in red) carry road classification (**middle**) or only the main branch carries information (**right**).

New classification for all edges guarantees good connectivity, because it is more easily detected in the topology. The implementation is simpler and more obvious for further processing. Therefore, we used it in our approach. Nevertheless, it slightly deforms the original network. The network is “spreading”; see the middle in Figure 17 and the detail in Figure 18. On the other hand, another option is to classify only the main branch of the collapsed object. This way, the road network is prevented from unwanted spreading. However, the whole network in the domain becomes more and more shattered throughout the process and connectivity analysis will be more difficult.

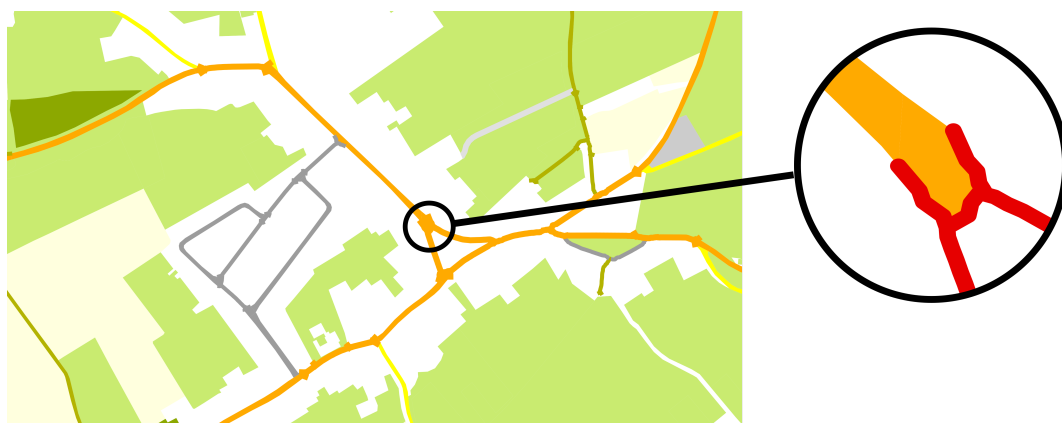


Figure 18. Side effect of all branches reclassification. This detail relates to Figure 7.

4.3.3. Crossing of Multiple Networks

The input map is a projection of 3D space into a 2D map. This means that the information about linear networks crossing each other at multiple height levels should be preserved somehow in the map, allowing us to use that knowledge in the processing. However, what should we do if there is no such knowledge available? Obtaining data is one thing (e.g., if it is available in TOP10NL), but how do we keep that knowledge during the process when objects change their representation? Additionally, how should the priority of individual networks be set? Another not really clear aspect is the solution of the case where more important networks cross each other, e.g., a road network crossing rivers and water channels. What to do in such a case is not really clear at this moment.

It is interesting to note a possible gain in the connectivity of two linear networks: roads and water. At the larger scale and with a single classification per face, it is not possible to model the fact that both the road network and the river network are properly connected. However, at the smaller scales, when the road and water segments are collapsed to lines and nodes, both networks can become connected again (and better represent the nature of both networks).

5. Discussion and Future Work

Up to now, our vario-scale method could only be used to represent area features. We have shown that line features in our current vario-scale solution may be introduced. We have designed an algorithm which provides the fully-automated generalization process that preserves a road network throughout all scales, which gives reasonable results. The algorithm maintains knowledge about a road network even in situations where roads are partially represented by lines, as well as partially represented by areas. We presented the necessary modifications that have to be applied to accommodate such an algorithm and still follow the idea of small generalization steps. For this, we have presented six design decisions. We also have presented that it is sufficient to use a large-scale planar partition with only area objects and their classification as input. We have validated our approach on a test dataset together with some quantitative measurements. With the generated results, we now have an opportunity to conduct user testing of the vario-scale principle. However, there are still the following open questions that need to be addressed:

- Objective validation of the cartographic quality of the results is difficult. Direct comparison of the vario-scale map with existing multi-scale solutions is not adequate. Both approaches are different in principle and not really comparable. Therefore, user acceptance testing may provide additional validation (in addition to quantitative metrics).
- In our current approach, every generalization step starts by picking the least important face for which a generalization operation is performed. This selection could be improved. Instead of only faces (area presentation of feature), also the selection of edges (line representation of features) should be possible. Picking objects based on a more advanced importance function, using their size, classification, the connectivity parameter, semantic or traffic information could be considered.
- Other options of the various design decisions as mentioned in this paper could be further analyzed.
- The advanced treatment for water network (also junctions and connections), rail networks, buildings or other feature types could be included. It might be very well the case that during the generalization process additional knowledge or different treatment is needed for these types.
- The actual classification of the other feature types could be used. This may enable better, more sensitive importance and compatibility functions, resulting in better cartography.
- Line simplification in the generalization process could be included to create the vario-scale data structure.
- Besides Top10NL data, the proposed approach should also be applied to other datasets; e.g., Corine (smaller starting scale) or Dutch BGT (Basisregistratie Grootschalige Topografie) (larger starting scale).

Acknowledgments: This research is supported by the Dutch Technology Foundation STW (Project Number 11185), which is part of the Netherlands Organisation for Scientific Research (NWO) and which is partly funded by the Ministry of Economic Affairs.

Author Contributions: All three authors have contributed to this paper with discussions and writing. Radan Šuba wrote most of the initial manuscript and implemented, with the help of Martijn Meijers, the workflow proposed. All authors have revised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Buttenfield, B.P.; McMaster, R.B. *Map Generalization: Making Rules for Knowledge Representation*; Longman Scientific & Technical: Harlow, UK, 1991.
2. Weibel, R. Generalization of spatial data: Principles and selected algorithms. In *Algorithmic Foundations of Geographic Information Systems*; van Kreveld, M., Nievergelt, J., Roos, T., Widmayer, P., Eds.; Springer: Berlin, Germany, 1997; pp. 99–152.
3. Mackaness, W.A.; Ruas, A.; Sarjakoski, L.T. *Generalisation of Geographic Information: Cartographic Modelling and Applications*; Elsevier: Oxford, UK, 2007.
4. Sester, M.; Brenner, C. Continuous generalization for visualization on small mobile devices. In *Developments in Spatial Data Handling*; Springer: Berlin, Germany, 2005; pp. 355–368.
5. Midtbø, T.; Nordvik, T. Effects of animations in zooming and panning operations on web maps: A web-based experiment. *Cartogr. J.* **2007**, *44*, 292–303.
6. Chimani, M.; van Dijk, T.C.; Haunert, J.H. How to eat a graph: Computing selection sequences for the continuous generalization of road networks. In Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Dallas/Fort Worth, TX, USA, 4–7 November 2014.
7. Huang, L.; Meijers, M.; Šuba, R.; van Oosterom, P. Engineering web maps with gradual content zoom based on streaming vector data. *J. Photogramm. Remote Sens.* **2016**, *114*, 274–293.
8. van Oosterom, P. Variable-scale topological data structures suitable for progressive data transfer: The GAP-face tree and GAP-edge forest. *Cartogr. Geogr. Inf. Sci.* **2005**, *32*, 331–346.
9. van Oosterom, P.; Meijers, M.; Stoter, J.; Šuba, R. Data structures for continuous generalisation: tGAP and SSC. In *Abstracting Geographic Information in a Data Rich World*; Burghardt, D., Duchêne, C., Mackaness, W., Eds.; Springer International Publishing: Berlin, Germany, 2014; pp. 83–117.
10. Šuba, R.; Driel, M.; Meijers, M.; van Oosterom, P.; Eisemann, E. Usability test plan for truly vario-scale maps. In Proceedings of the 19th ICA Workshop on Generalisation and Multiple Representation, Helsinki, Finland, 14 June 2016.
11. Weiss, R.; Weibel, R. Road network selection for small-scale maps using an improved centrality-based algorithm. *J. Spat. Inf. Sci.* **2014**, *2014*, 71–99.
12. Edwardes, A.; Mackaness, W. Intelligent road network simplification in urban areas. In Proceedings of the GIS Research UK 2000 Conference (GISRUK 2000), Pittsburgh, PA, USA, 4–8 April 2000.
13. Thomson, R.C.; Richardson, D.E. The “Good Continuation” principle of perceptual organization applied to the generalization of road networks. In Proceedings of the 19th International Cartographic Conference, Ottawa, ON, Canada, 14–21 August 1999; pp. 1215–1225.
14. Chen, J.; Hu, Y.; Li, Z.; Zhao, R.; Meng, L. Selective omission of road features based on mesh density for automatic map generalization. *Int. J. Geogr. Inf. Sci.* **2009**, *23*, 1013–1032.
15. Li, Z.; Zhou, Q. Integration of linear and areal hierarchies for continuous multi-scale representation of road networks. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 855–880.
16. Zhou, Q.; Li, Z. A comparative study of various strategies to concatenate road segments into strokes for map generalization. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 691–715.
17. Turner, A. From axial to road-center lines: A new representation for space syntax and a new model of route choice for transport network analysis. *Environ. Plan. B Plan. Design* **2007**, *34*, 539–555.
18. Cecconi, A.; Weibel, R.; Barrault, M. Improving automated generalisation for on-demand web mapping by multiscale databases. In *Advances in Spatial Data Handling*; Springer: Berlin, Germany, 2002; pp. 515–531.

19. Cecconi, A. Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping. Ph.D. Thesis, University of Zurich, Zurich, Switzerland, 2003.
20. Brewer, C.A.; Buttenfield, B.P. Framing guidelines for multi-scale map design using databases at multiple resolutions. *Cartogr. Geogr. Inf. Sci.* **2007**, *34*, 3–15.
21. Touya, G.; Girres, J.F. ScaleMaster 2.0: A ScaleMaster extension to monitor automatic multi-scales generalizations. *Cartogr. Geogr. Inf. Sci.* **2013**, *40*, 192–200.
22. Nöllenburg, M.; Merrick, D.; Wolff, A.; Benkert, M. Morphing polylines: A step towards continuous generalization. *Comput. Environ. Urban Syst.* **2008**, *32*, 248–260.
23. Dumont, M.; Touya, G.; Duchêne, C. Automated generalisation of intermediate levels in a multi-scale pyramid. In Proceedings of the 18th ICA Workshop on Generalisation and Multiple Representation, Rio de Janeiro, Brazil, 21–22 August 2015.
24. Meijers, M.; van Oosterom, P. The space-scale cube: An integrated model for 2D polygonal areas and scale. In Proceedings of the 28th Urban Data Management Symposium, Delft, The Netherlands, 28–30 September 2011; Volume 38, pp. 95–102.
25. Van Oosterom, P.; Meijers, M. Towards a true vario-scale structure supporting smooth-zoom. In Proceedings of the 14th ICA/ISPRS Workshop on Generalisation and Multiple Representation, Paris, France, 30 June–1 July 2011; pp. 1–19.
26. Šuba, R.; Meijers, M.; Huang, L.; van Oosterom, P. Continuous road network generalisation. In Proceedings of the 17th ICA Workshop on Generalisation and Multiple Representation, Vienna, Austria, 23 September 2014; pp. 1–12.
27. Šuba, R.; Meijers, M.; van Oosterom, P. Large scale road network generalization for vario-scale map. In Proceedings of the 18th ICA Workshop on Generalisation and Multiple Representation, Rio de Janeiro, Brazil, 21–22 August 2015; pp. 1–10.
28. Uitermark, H.; Vogels, A.; van Oosterom, P. Semantic and geometric aspects of integrating road networks. In Proceedings of the 2nd International conference on Interoperating Geographic Information Systems, Zurich, Switzerland, 10–12 March 1999; pp. 177–188.
29. Meijers, M.; Savino, S.; van Oosterom, P. SPLITAREA: An algorithm for weighted splitting of faces in the context of a planar partition. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 1522–1551.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).