Comparing the vario-scale approach with a discrete multi-representation based approach for automated generalisation of topographic data

Martijn Meijers[1], Jantien Stoter[1,2] and Peter van Oosterom[1]

[1]OTB, TU Delft, The Netherlands
Email: {b.m.meijers|j.e.stoter|p.j.m.vanoosterom}@tudelft.nl

[2] Kadaster, Apeldoorn, The Netherlands

1. Introduction

Datasets at different scales can be obtained and maintained by two main approaches that differ considerably. The first option is the vario-scale approach that offers the possibility to derive a map at an arbitrary map scale (hence its name) by performing once a fully automatic generalisation process on the input data set. This sounds as the optimal situation. However, since fast and fully automated generalisation of topographic data is still a challenge, National Mapping Agencies (NMAs) employ a second option, i.e. the multi-scale approach, to produce and maintain maps at different, predefined scales. Examples are reported in the literature for IGN, France (Lecordix et al. 2007), KMS, Denmark (Foerster et al, 2010), ICC, Catalonia (Baella & Pla 2005), Germany (AdV, 2007), and The Netherlands (Stoter et al, 2009).

The tGAP data structure (van Oosterom, 1995; van Oosterom, 2005; Meijers, 2011) provides a promising implementation of the vario-scale approach in which the most detailed data is stored once, and an incremental object by object generalisation process is run and represented in a data structure, which can afterwards be used to efficiently obtain any arbitrary scale on the fly (Vermeij et al, 2003; Meijers, 2011). Figure 1 shows the concept of tGAP.

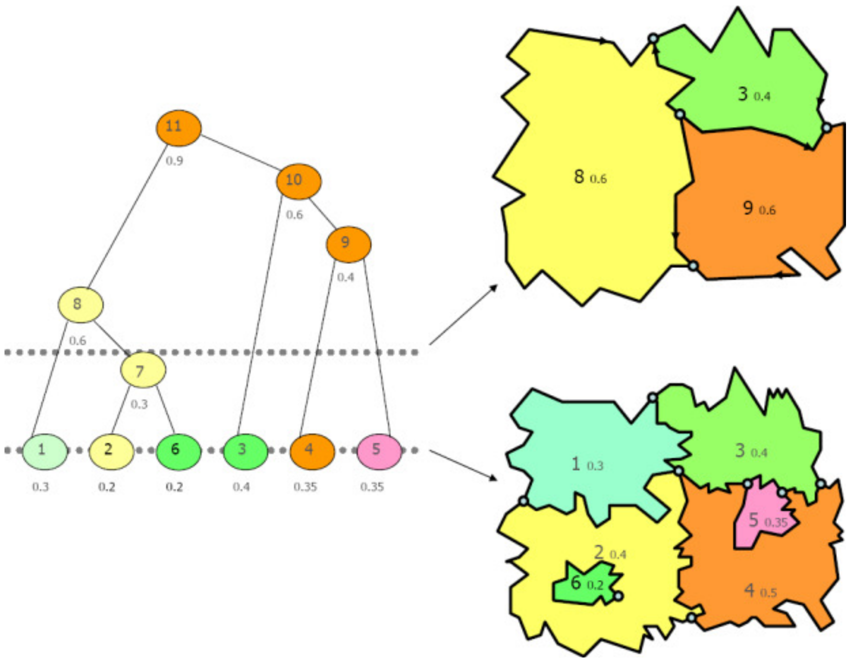
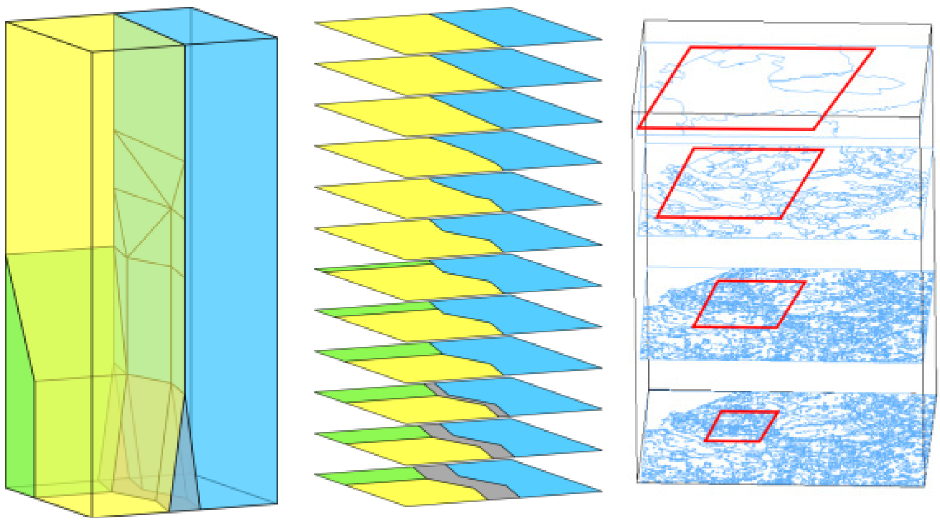


Figure 1: The working of the current tGAP structure. Based on the importance value of the objects, objects are aggregated at smaller scales and the data structure can be queried at any arbitrary scale.

The idea of vario-scale was even taken further in van Oosterom and Meijers (2011) who introduced a 3D data structure to represent 2D + scale data in a space partition resulting in a true vario-scale structure for 2D maps: i.e. a delta in scale leads to a delta in the map (and smaller scale deltas lead to smaller map deltas until and including the infinitesimal small delta) for all scales. The structure is called smooth tGAP and its integrated 2D space and scale representation is stored as a single 3D geometric data structure. The polygonal area objects are mapped to polyhedral representations in this space-scale cube, see Figure 2. The polyhedral primitive is integrating all scale representations of a single 2D area object. Together all polyhedral primitives form a partition of the space-scale cube: gaps and overlaps are not allowed (not in space or scale). Obtaining a map at a single scale is deriving an horizontal slice through the structure. The structure can be used to implement smooth zoom in an animation or morphing style by ascending or descending along the scale dimension.
Note that in Figure 2 a and b the two objects in the DLM (Digital Landscape Model as represented in the space-scale cube) only shrink before respectively disappearing and collapsing to a line. However, when the content of this DLM is depicted and the various objects are 'styled' (for polygon both the interior is coloured and the boundary/ casing is coloured, potentially with same colour as interior), then in the 'virtual DCM' (Digital Cartographic Model, which is not explicitly stored) this would have the effect of been enlarged first (to increase their visibility), before respectively disappearing or switching to the line based representation. See also the discussion in Subsection 2.1.



(a) (b) (c)

Figure 2. Examples of the integration of the 1D scale dimension and 2D space into a 3D datastructure (i.e. DLM). (a) Every map object, 4 in this case, is represented as one polyhedron. (b) 2D Maps are slices (cross sections) of this 3D model. (c) For interactive use, apart from taking a slice, also a bounding box filter should be applied.

Although the key principle of both approaches (vario-scale and multi-scale) overlap (i.e. support consistency and data reduction), the current implementations are driven by different motivations and therefore the solutions differ. Current vario-scale implementations have been engineered by focusing on object consistency (w.r.t thematic, geometric and topological aspects within a scale and between scales) and maintaining this consistency throughout the automated generalisation process. Another key feature of the current vario-scale implementation is supporting smooth zooming functionalities, i.e. avoiding shocks when zooming in or out, and efficient data streaming within Internet environments possible (because redundancy is avoided by the explicit topological data structures). On the contrary, the multi-scale approach is originally motivated by maintaining maps at predefined and independent map scales, created by cartographers at NMAs. Redundant data storage is considered to be a lesser problem and also smooth changes may not be a concern. For example the reduction of streets by removing all minor roads in one step is common practice when going from a large-scale map to a midscale map. In addition humans may use different concepts for representation in different scales and cartography has followed the graphical representations of these concepts. This may not be in line with the vario-scale concept that aims mainly at reducing data in a gradual manner.

Considering these differences, we can say that vario-scale aims at assuring 100% consistency between different scales as well as reducing redundancy to improve efficiency with a focus on continuous changes, while the multi-scale approach aims at controlling the redundancy in a discrete fashion.

Until now, vario-scale implementations have considered the generalisation of topographic data only to a limited extent and therefore the current vario-scale solutions need to be extended with more generalization operations and intelligence ('more semantics and context aware') to cover all issues of automated generalisation of topographic maps. In theory the vario-scale structure can hold (represent) any generalization and thus be used to efficiently generate the content of the discrete scales of a multi-representation approach (but with less redundancy and more integration between scales). Since the multi-scale generalisation approach is cumbersome, redundant and risky for inconsistencies, it is however relevant to study if and how the vario-scale approach could be applied for automated generalisation challenges of NMAs in practise (and not only in theory). The answers to these questions may yield new research questions for generalization research.

These questions are the motivation of this paper, in which we discuss the suitability of vario-scale for automated generalisation of topographic maps, including a set of open research questions to improve this suitability. In section 2 we describe these issues both in 2D and 3D. The questions will give focus to the STW research projects 'vario-scale (2D+scale) geo-information' and '5D (3D+scale+time) modeling', which started in July, 2011. Both researches implement scale as a separate dimension in spatial models and therefore for both research projects it is relevant to better understand how geo-information behaves when going up or down the scale axis.

We close this paper with a discussion in Section 3, in which we explain through a number of use cases why it would be a good idea to perform integration of scale as an additional geometric dimension during data modelling for topographic information, instead of as separate levels of detail.

The main motivation to write this discussion paper, is that until now vario-scale has been mainly an academic research topic. An next step for vario-scale research is to test it on practicality. To support this, this paper analyses the fundamental differences between vario-scale on the one hand and generalisation of topographic maps employed in practice (discrete multi-representation) on the other hand. The paper also identifies overlap and sees how the two approaches can be linked for better results. This knowledge can be used to formulate new research questions to further advance vario-scale approaches in generalisation, both in 2D and 3D.

2. Issues when applying vario-scale concepts to generalisation of topographic data

This section presents the issues when applying vario-scale concepts to the generalisation of topographic data. Section 2.1 focuses on 2D data. In theory, the vario-scale approach also works for 3D space+scale, where data is stored integrated in one 4D hypercube. The question of how vario-scale applies to 3D modelling overlaps with the issues in 2D. Therefore Section 2.2 discusses how the vario-scale concept applies to 3D city models.

2.1 Vario-scale for automated generalisation of topographic maps

To apply the vario-scale approach for solving 2D+scale problems in the map generalisation domain, several issues require further research.

At first, the three generalisation operations that are currently supported (i.e. split, smooth and simplify) implement the principle of smooth data reduction, which may not be completely in line with cartographic principles, e.g. small objects may be enlarged at smaller scales until they suddenly (and not smoothly) disappear at a certain scale level and become part of the neighbouring object. Also symbolised objects may cover larger map areas than their geometric counterparts, which requires operations not only resulting in data reduction, e.g. enlargement, displacement and other operations reported in (Regnauld, 2008).

Furthermore, to implement topography-specific concepts in vario-scale requires better understanding of how geo-information changes at a scale transition. In this respect the work of Dilo et al. (2009) is relevant. This work implements a smooth transition between two existing topographic data sets collected independently for two different scales, by which the most detailed data set is generalised to the given representation at the smaller scale.

When considering a vario-scale approach for topographic maps, one should realise that the current automatic generalisation process for vario-scale solves best the generalisation problem of mono-thematic maps (legend consisting one class), such as land use maps which also need to be a planar partition. The vario-scale implementation is therefore at this moment not directly suitable to generalise the heterogeneous set of classes of topographic maps. In those maps different types of objects with different geometrical characteristics are common (and also do not always form a planar partition). However, all geographic objects (also when having different dimensions and not forming a planar partition) could be embedded in one data structure, so that the adjacency relationships between all objects is explicitly available to the generalization algorithms. Examples of class-specific constraints, which are currently not supported by our generalisation algorithms to create the vario-scale structure are preservation of the rectangular outline of buildings, and preservation of networks, for road and water objects.

Although the current tGAP data structure has ways to treat various classes differently by assigning importance and priority measures, the possibility to apply different generalisation operators depending on specific classes has never been turned in an implementation. Also relevant is, whether it is useful and possible to always create a smooth transition in the scale dimension: e.g. do not show all second order road network at once, but let it gradually appear? It relates to the following question: Can there always exist a strict ordering (sequence) of applying generalisation operations on objects (e.g. merge and split) such that objects can be sorted and scheduled for being generalised, leading to a gradual and smooth change of the map over map scale? In the current tGAP implementations this was introduced in order to proof the correctness of the various generalization operations (and avoid interference of simulations nearby generalization actions). However, from the geo-information quality point of view, it may even be preferred for some of 'object by object' sequenced generalization steps to be performed in parallel (but still all smooth); e.g. an merge of two objects in the north can be done in parallel with a merge of two other objects in the south (with similar importance of the disappearing objects) as these will never interfere; see section of 4.3 of van Oosterom and Meijers (2011).

There are also interesting questions[1] related to the cartographic principles when visualizing the content of the tGAP structure (both in static situations and also in dynamic smooth zoom operations). For example: when a road is an area on the largest scale and in the tGAP structure the road area is collapsed to a road centreline via a completely smooth transition, then in the visualization a shock might appear if the road area is displayed by colouring the area (which becomes infinitely thin before it is represented by a line. The line is also infinitely thin, but displayed with line symbology, to make it visible. One possible approach to avoid this to display 'shock', is not simply colouring the road area, but also provide the proper casing of this area, which might even be of the same color. Cartographic usability tests will have to result in guidelines for appropriate vario-scale symbology. Another interesting cartographic question is: assuming that as in Figure 2 the forest (green) has to be removed and the area may be taken by the neighbour field (yellow). Is the 'best' gradual transition indeed gradually reducing the size of the forest and enlarging the geometry of the field or is another gradual transition to be preferred; e.g. gradually changing the colour green into yellow. This latter option does also support the 'delta scale results in a delta map principle'. Cartographic usability tests will have to make clear which option is to be preferred (in principle the tGAP structure is capable to represent both alternatives). The initial proposition of the authors is that shrinking/enlarging will provide better interaction experiences (as colour mixing might result in an infinite range of mixed colours, to which it is difficult to match a legend to figure out the classification of the different objects).

2.2 Vario-scale for automated generalisation of 3D topographic data

For the question of how the vario-scale concept would apply to generalisation of 3D city models, we focus on the scale concept currently implemented in CityGML, since this is the international standard for 3D topographic information. CityGML is the OGC standard for modelling and exchanging 3D city and landscape models (CityGML, 2012). Similar to the maps at a series of scales of NMA's (e.g. 1:25k, 1:50k, 1:250k), CityGML supports the concept of separate scales or levels of detail. That is, CityGML includes five predefined independent levels of detail (LODs), ranging from only the terrain, via block models, to more 3D details and finally to the interior of buildings including furniture. These different LODs of 3D city models coexist and individual objects are not explicitly linked together.

Similar to the multi-scale approach of 2D topographic data, this multi-LOD approach has several limitations. First, it is particularly difficult to query through different LODs and to keep different LODs consistent after updating. The different LODs are poorly connected, and on-the-fly derivation of lower LODs from a higher LOD is not supported. Therefore, consistency between different LODs cannot be assured. A specific problem of the different LODs in the 3D city model is that the

[1] Some of the questions were raised by Bettina Speckmann (TU/e) after the vario-scale geo-information presentation by the last author of this paper at the 8th Dutch Computational Geometry Day, 19 January 2012 (Utrecht University, The Netherlands).

different LODs refer to individual objects only, i.e. aggregation is not supported, and higher LODs cannot consist of parts from a lower LOD. Related to this problem is the lack of a notion of semantic change at a scale transition, for example, the concept that individual trees at a higher LOD may change to a forest at a lower LOD is not supported. For users of CityGML with LODs as reference data to which own application data should be added, the question arises to which LOD to link-to? Given the fact that the LODs are independent, the same application data may need to be linked several times.

A vario-scale approach for 3D models would offer the possibility to continuously zoom-in and out across levels of detail, without jumping to another discrete representation (as in CityGML), because the LODs are integrated in the 4D data structure itself. In addition it allows the continuous representation of a city model, i.e. not restricted to the arbitrary five fixed LODs (in the case of CityGML). Slicing the 4D data cube permits us to obtain a 3D city model at any given LOD.

As in 2D, also for 3D, it is a relevant question how continuously zooming in and out supports the users' requirements of 3D applications. We can take a house with a balcony as example. Should the balcony appear at once, or should it gradually be morphed to its final look and feel, moving as you go? If far away enough, all changes become mapped to one pixel on the screen, when would a user perceive these changes as rendering artefacts or would it always be useful to show changes in this smooth way? Furthermore, has this gradual change meaning in the transition, or does only the end result of such a change carry meaning? One could argue that for a certain representation LODn does not contain enough information and LOD(n+1) contains too much information to be optimal in information content.

Also the implementation of a vario-scale approach for 3D city models, requires the addition of generalisation operations that enable contextual generalisation of 3D topographic information and the inclusion of semantic concepts in the 4D data structure. For example aggregation of single buildings to building blocks when going from LOD2 to LOD1, as studied in Guercke et al (2011) and Zhu (2010). Currently the CityGML LODs are mainly defined for buildings. These LODs form the targets or 'anchor-points' for creating a vario-scale structure. Tests will have to be conducted to further explore the behaviour of the representations at an arbitrary scale between two fixed LODs ('horizontal slice with hyper plane in the 4D hypercube resulting a 3D city model at the requested scale'). Perhaps more interesting is to use the 4D hypercube to generate perspective 3D scenes with more detail close to the viewer and gradually less detail further away (without any topology problems, gaps of overlaps, that fixed LOD-based solutions have). A pragmatic solution to take into account perspective viewing and generating a generalized 3D model was developed by Mao et al., 2011, 2012. However, in this work no attention is paid to topological consistency of the result, which may be disturbing of there are unwanted intersections.

For other types of objects than buildings, it is also important that the LODs are well defined. Recently, fixed LODs were defined for tunnels (Breunig et al., 2012). It will be interesting to explore what is the behaviour of the vario-scale representation of such a tunnel object, which meets at the specific scale the well-defined LODs.

To further support the change of semantic concepts at scale steps in 3D, an a solution similar to the constrained tGAP approach as applied in 2D (Dilo et al. 2009) is needed. That is, if different LODs are available, we generalise between them to obtain a smooth transition, but ensure that the resulting object at the given LOD is the same as the existing object at that LOD. This lays down an explicit link between existing LODs. This has been studied in Bédard and Bernier (2002).

3. Discussion

In this paper we considered the vario-scale approach for automated generalisation of topographic data as currently employed by NMAs (i.e. to produce maps at predefined scales).

The multi-scale approach of NMAs could be seen as an intermediate solution awaiting the successful ultimate vario-scale solution in which you only store the most detailed version of the map with additional structure obtained from an automatic generalisation process, enabling to efficiently ('on the fly') produce a representation at any required scale. However in this abstract we showed that the specific application areas differ and therefore algorithms that populate the current state-of-the-art vario-scale structure can only be used to solve certain parts of the generalisation of topographic maps, and so far the vario-scale solution alone is not suitable to solve all the issues of multi-scale maps. The current automated processes of NMAs contain a lot of cartographic knowledge that still has to be implemented in the generalization process to populate good quality vario-scale data structures, but there are no know issues that might make it impossible to represent this in the tGAP structure. In addition new challenges for topographic data can be better met by a vario-scale approach (although the gradual and smooth scale transitions of vario-scale also bring their own challenges):

- Vario-scale data will be 100% backwards compatible with multi-representation databases: the discrete maps at discrete scale steps can be defined as views on the vario-scale data cube. A tuned vario-scale solution (where good map generalisation has been performed and 'recorded' in the tGAP structure) can thus well be used to generate database representations at different scales that can be the source of cartographic representations. For example, the portal that implements the Dutch SDI (i.e. Publieke Dienstverlening op de Kaart, i.e. Public Map Services) needs a topographic background map at fourteen different map scales. On a similar note: Sometimes you need less-detailed data sets to be able to perform spatial analyses on a national level. These can be derived by taking the right slice from a vario-scale structure.
- The automated generalisation process for creating vario-scale datasets benefits from structuring and cleaning the data explicitly in a topological data format, as this allows the generalisation process to be expressed more concisely and the cleaning follows the mantra of garbage in, garbage out. Integrity checking has become a first class citizen in our implementation: explicit structure is better than implicit, because if coded explicitly in the data structure (and update/generalisation operations are introducing no errors while modifying the structure), then there is no need for additional rules (or it can be independently checked, similar to topology checking, e.g. validation in Oracle).
- To make continuous zooming a reality becomes easier with vario-scale data. Users can then benefit from smooth interactions and better preserve their mental model, as Nordic (2007) showed.
- Streaming, in which already sent ordered geometry is re-used, will only be possible with an approach where redundancy is a concern from the start. Data consistency for such an approach is a must have; from technological point of view it will not be possible to perform this on unstructured data and guaranteeing the result (you will have to rely on that what you send earlier matches/connects with what you receive later). Data streaming becomes more important in distributed environments (Internet/client-server) with very large datasets: having a coarse overview first, adding gradually more detail, makes it possible to stop

data transfer when enough detail is received (e.g. when using a 3D globe): compare to the mantra of Schneidermann (1996) in data visualization: 'overview first, more details on demand'.

- As mentioned several times before, it should be noted that current streaming approaches only consider geometric aspects by adding/removing additional vertices or segments. They do not consider thematic multi-scale aspects, i.e. a changing legend over a scale range: e.g. built-up area is not represented at highest level of detail, only at mid-scale range. Another example is a roundabout: in the Dutch topographic dataset at the highest level of detail only the separate road parts are represented, but not the fact that these together form a roundabout. With a vario-scale approach the gradual transition from individual road parts to a roundabout can be stored, when this notion is incorporated in the data structure.
- Mixed-scale slices are possible (close more detail, far away less) without having to perform an additional gluing step as indicated in section 2.2, which would be necessary with an approach with discrete levels of detail, see Figure 3.

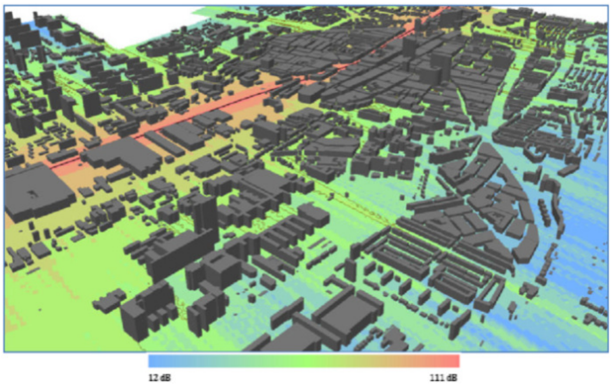


Figure 3. Noise modelling in 3D (caused by a railway in downtown Delft) would benefit from having more detail available close to the source of the noise, while further away less detail is needed. Note that this figure in not trying to show a nice looking perspective view image, but that fact that also computations (such as simulations) may be most effective when using vario-scale representations.

Although this paper focused on how vario-scale in its current state can support multi-scale generalisation, also the reversed question is relevant, i.e. what can a multi-scale approach bring to vario-scale representations. A multi-scale approach allows the production of a representation as required by the user (in a constrained tGAP fashion). This knowledge can be captured and (partly) embedded in vario-scale structure, in the end to make the implementation better suitable for end users-maps.

Acknowledgements

This research is supported in part by the Dutch Technology Foundation STW (project numbers 11300 and 11185), which is part of the Netherlands Organisation for Scientific Research (NWO) and partly funded by the Ministry of Economic Affairs, Agriculture and Innovation.

References

AdV. 2007. Produktblaetter Arbeitskreis Geotopographie, Arbeitsgemeinschaft der Vermessungsverwaltungen der Laender der Bundesrepublik Deutschland.

Baella, B. & Pla, M.. 2005. Reorganizing the topographic databases of the Institut Cartographic de Catalunya applying generalization. In Proceedings of the 8th ICA workshop on Generalisation and Multiple Representation. A Coruña, Spain.

Bédard, Y. & E. Bernier, 2002. Supporting Multiple Representations with Spatial View Management and the Concept of "VUEL". Joint Workshop on Multi-Scale Representations of Spatial Data, ISPRS WG IV/3, ICA Com. on Map Generalization, July, 7-8, Ottawa, Canada.

Breunig, M., A. Borrmann, S. Hinz, E. Rank and M. Schilcher, 2012.Towards 3D Geoinformatics and Computational Engineering Support for Cooperative Tracks Planning. In proceedings FIG Working Week 2012, 6-10 May 2012, Rome, Italy.

CityGML, 2012, OpenGIS® City Geography Markup Language (CityGML) Encoding Standard, version 2.0, http://www.opengeospatial.org/standards/citygml.

Dilo, A. and P. van Oosterom and A. Hofman, 2009. Constrained tGAP for generalization between scales: The case of Dutch topographic data. Computers, Environment and Urban Systems, Volume 33(5), pp. 388–402.

Foerster,T., J.E. Stoter and M Kraak. 2010. Challenges for Automated Generalisation at European Mapping Agencies: A Qualitative and Quantitative Analysis, In: The Cartographic Journal, Vol. 47(1), pp. 41–54.

Guercke, R. and T. Götzelmann, C. Brenner, M. Sester, 2011. Aggregation of LoD 1 building models as an optimization problem, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 66(2), March 2011, pp. 209–222.

Lecordix, F., Gallic, J.L., Gondol, L. & Braun, A.. 2007. Development of a new generalization flowline for topographic maps. In 10th ICA workshop on Generalisation and Multiple Representation. Moscow.

Mao B., Ban, Y., and Harrie, L., 2011. A framework for visualization of city models based on CityGML and X3D. ISPRS Journal, Vol. 66, pp. 198–208.

Mao B., Harrie L. and Ban, Y., 2012. A Dynamic Typification Method of 3D City Models using Minimum Spanning Tree. Computer, Environment and Urban Systems, Vol. 36, pp. 233–244, doi:10.1016/j.compenvurbsys.2011.10.001.

Meijers, M., Variable-scale Geo-information, PhD thesis Delft University of Technology, December 2011, 235 p. Published by Netherlands Geodetic Commission, Publications on Geodesy 77, Delft, 2011.

Midtbø T. and Nordic T., Effects of Animations in Zooming and Panning Operations on Web maps: A Web-based Experiment, The Cartographic Journal, Vol. 44 (4), pp. 292–303, doi:10.1179/000870407X241845.

van Oosterom, P. and Martijn Meijers, 2011, Towards a true vario-scale structure supporting smooth-zoom In: Proceedings of the 14th Workshop of the ICA Commission on Generalisation and Multiple Representation & the ISPRS Commission II/2 Working Group on Multiscale Representation of Spatial Data, 2011, Paris, 19 p.

van Oosterom, P., 2005. Variable-scale topological data structures suitable for progressive data transfer: The gap-face tree and gap-edge forest. Cartography and Geographic Information Science, Vol. 32, pp. 331–346.

van Oosterom, P., 1995, 'The GAP-tree, an approach to "On-the-Fly" Map Generalization of an Area Partitioning.' GISDATA Specialist Meeting on Generalization, Compienge, France, 15-19 December 1993. Chapter 9 in: GIS and Generalization, Methodology and Practice. Editors J.C. Mueller, J.P. Lagrange and R. Weibel. Taylor & Francis, London, pages 120–132, 1995.

Regnauld, N. and R.B. McMaster, 2007, A Synoptic View of Genealisation Operators, Chapter 3, in: Generalisation of Geographic Information: Cartographic Modelling and Applications (International Cartographic Association); William A. Mackaness (Editor), Anne Ruas (Editor), L. Tiina Sarjakoski (Editor)

Schneidermann, B. (1996), The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In Proceedings of the IEEE Symposium on Visual Languages, pages 336-343, Washington. IEEE Computer Society Press, 1996.

Stoter, J.E., J. van Smaalen, N. Bakker, P. Hardy. 2009. Specifying map requirements for automated generalisation of topographic data, The Cartographic Journal Vol. 46(3) pp. 214–227 August 2009

Vermeij, M., van Oosterom, P., Quak, W., and Tijssen, T. (2003). Storing and using scale-less topological data efficiently in a client-server dbms environment. In GeoComputation 2003, University of Southampton, Southampton, UK.

Zhu, Q., Junqiao Zhaoa, Zhiqiang Dua, Yeting Zhanga, 2010, Quantitative analysis of discrete 3D geometrical detail levels based on perceptual metric, Computers & Graphics Vol. 34(1), February 2010, pp. 55–65