GENERALIZATION TECHNIQUE FOR 2D+SCALE DHE DATA MODEL

Hairi Karim[a], Alias Abdul Rahman[a], and Pawel Boguslawski[b]

[a]3D GIS Research Group, Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia, 81310, Johor Bahru, Johor, Malaysia - wmhairigis@gmail.com and alias@utm.my
[b]FET - Department of Architecture and the Built Environment, University of the West of England - Pawel.Boguslawski@uwe.ac.uk

KEY WORDS: scale dimension, 2D-to-3D scale generalization, extrusion techniques, data model, and the DHE data structure.

ABSTRACT

Different users or applications need different scale model especially in computer application such as game visualization and GIS modelling. Some issues has been raised on fulfilling GIS requirement of retaining the details while minimizing the redundancy of the scale datasets. Previous researchers suggested and attempted to add another dimension such as scale or/and time into a 3D model, but the implementation of scale dimension faces some problems due to the limitations and availability of data structures and data models. Nowadays, various data structures and data models have been proposed to support variety of applications and dimensionality but lack research works has been conducted in terms of supporting scale dimension. Generally, the Dual Half Edge (DHE) data structure was designed to work with any perfect 3D spatial object such as buildings. In this paper, we attempt to expand the capability of the DHE data structure toward integration with scale dimension. The description of the concept and implementation of generating 3D-scale (2D spatial + scale dimension) for the DHE data structure forms the major discussion of this paper. We strongly believed some advantages such as local modification and topological element (navigation, query and semantic information) in scale dimension could be used for the future 3D-scale applications.

1. INTRODUCTION

An ideal 3D or 4D (a spatial 3D with additional non-spatial dimension) spatial model must have capability to support most of basic analysis or/and advance analysis. There are three basic requirements that need to be preserved in order to provide basic and complex analysis in n-dimensional (where n>1) spatial model representation. The selection of data structure for model implementation should meet all GIS data requirements in order to widen the analysis functions in the 3D/4D spatial model. Navigation between spatial objects and extracting neighbourhood information are among the examples for the advance analysis available in an ideal data model. The 3D/4D spatial data model supposedly able to support geometry, topology and semantic integration within a single model.

The first motivation of this research is the recently developed DHE data structure for 3D spatial GIS data that fulfils GIS requirement for nD GIS modelling. As compared to other data structures available in market such as CAD system, the DHE able to provide the most consistent GIS elements such as geometry, topology and attribute within single data structure.

Secondly, there are two candidates of data structures that able to model higher dimensional e.g. 4D and above; namely generalised maps (G-Maps) and the DHE (Ohori et al., 2013a). The future research on this DHE data structure is widely open to be explored for the higher dimensional modelling. However, the concept and implementation of scale dimension in the DHE data structure should be started with lower dimension primitives (spatial 2D) in order to minimize the risks and upcoming problems while tackling the local cases of 3D-scale applications.

Third and lastly, the emergence of scale element for the non-spatial dimension now has taking the lead in the GIS modelling research trend for the past few years. Although some researchers from previous studies that focusing on scale dimension failed to achieve the goals, the knowledge and finding of their studies could be considered as a very valuable asset for this research improvement.

Section 2 explains the approach on scale dimension and the DHE data structure. The remaining sections describe the new approach on scale generalization, results and conclusion.

2. THE CONCEPT OF SCALE AND DATA STRUCTURE

This section is divided into two main literature; scale dimension and data structure.

2.1 Scale Dimension

GIS multidimensional may consist of one, two, three or more spatial dimensions that represent spatial objects (Gold, 2005). However, there is a wide category or definition on the third dimensional model; a 2D geometry that consists of spatial dimension of coordinate X and Y, and the third one is used to represent with other non-spatial dimension such as scale, time and etc (Oosterom, 2005).

According to previous literature, most of early works on the non-spatial modelling end up with problems and unable to solve these problems which mainly due to unsupported available data

structure. A few researches succeeded with introducing multi-scale concepts in storing line simplification such as Multi-scale Line tree (Jones and Abraham, 1986), Arc-tree (Günther, 1988), and Binary Line Generalization (BLG) (Oosterom, 1990).

Scaling dimension in GIS has gained some popularity in these recent years due to the demands from users and applications. GIS research community is now moving forward for integrating a highly formal definition of geo-data (Oosterom and Stoter, 2010) and thus focuses on designing the most efficient framework and implementation for scale and temporal dimension. In general, the scale dimension can be categorized into two categories; multi-scale and vario-scale.

2.1.1 Multi-scale Approach

The current implementation frameworks of multi-scale integration with 2D spatial data are either using the generalization techniques or storing the individual level of detail data into separate databases (Fig. 1). On the other hand, 3D spatial GIS data normally uses a generalization technique (aggregation, simplification and other approaches) or a set of pre-defined level of details (LoDs, see Fig. 2) which are topologically disconnected to represent GIS model in scale dimension.

Most previous works implemented different generalization methods in order to produce less-detailed LoDs from the higher ones. Hiding less important details or/and reducing the data storage volume are available options for an efficient 3D analysis and visualization (Baig et al., 2011).
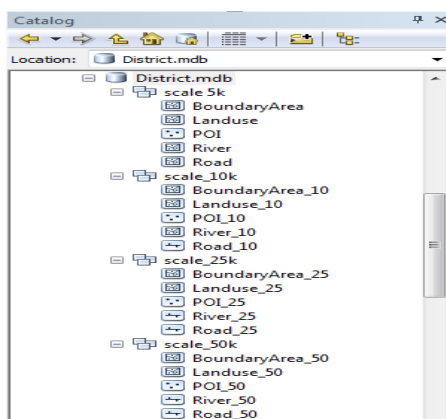

Fig. 1. A 2D redundancy problem in multi-scale approach

Fig. 2. Multi-scale concept in CityGML (Sester, 2007)

Storing the same dataset in different scale ratio (pre-defined datasets) and databases will introduce data redundancy in term of geometry and attribute. Thus, it consumes a lot of storage volume as compared to store only a single dataset (most detailed) and undergoes any generalization process for a less detailed layer. Multi-scale approach also causes a lot of

upcoming problems such as difficulty in updating process (need more works to update certain area of the data for every pre-defined scale; either geometry and attribute), easily to introduce errors if it is not update/handle properly, hardly preserve the geometry consistencies and less topological relationships between LoDs.

2.1.2 Vario-scale Approach

In vario-scale approach, each area of the map can be represented by a topological face (Meijers, 2011), uses any data structure such as Half Edge (HE) for the horizontal plane (slice) to preserve the topology and geometry of the object. The level of detail (LoD) is regarded as the additional dimension (third dimension) (Meijers, 2011). LoD level also represented as a series of horizontal slices plane and represented by concept of "importance" (Meijers, 2011). The vario-scale structure is the extension from the tGAP which use the hierarchical tree structure.

The importance (referred as '*imp*', based on Meijers' term) of the objects is highly depending on the feature classification and the size of the object. A small city is more importance than a large area of forest; a large residential area is more importance than a small water body and others. The new importance value is calculated after the process of generalization (merging) from two predecessors' objects as shown in Fig. 3 and Fig. 4.
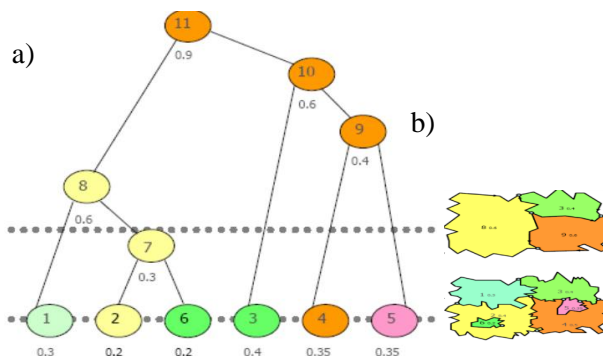

Fig. 3. Structure of variable-scale model; a) tGAP hierarchical tree for semantic information and *imp* value,
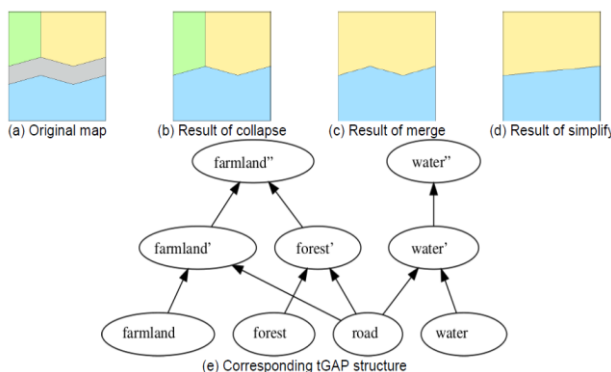b) geometry and topology in horizontal slice planes

Fig. 4. The process of generalization and hierarchical tGAP structure (Oosterom and Meijers, 2011b)

The tGAP model generally capable to overcome the data redundancy problem via topological connector since it was constructed starting with the most detailed dataset to the coarser

dataset. It also preserves the topology of the horizontal plane and semantic within the scale axis. However, the tGAP structure (linking appropriate 2D objects at different levels of detail) is not solving the arisen problems entirely. Among the highlighted problems are; the solution only fixed number of representations, no link between the same object a different representation and the information on geometry or topology is stored multiple times which caused a waste of memory (Ohori et al., 2013b).

The 3D vario-scale tGAP model are generated via Space-Scale Cube (SSC) which can be classified as a data model for the tGAP structure. For the 2D dataset, vario-scale SSC can be obtained by slicing the SSC model in the horizontal plane; 3D-to-2D concept of vario-scale. Via SSC model (Fig. 5), the topology of the vario-model in vertical axis is highly dependent on the hierarchical b-tree structure which only can be navigated using the procedure of attribute query in the database.
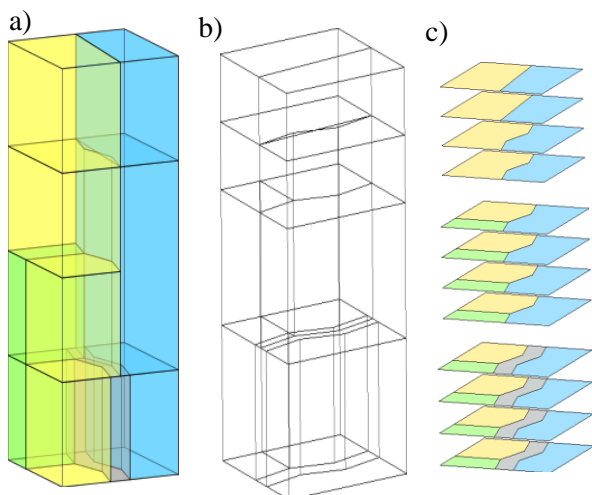

Fig. 5. A classic tGAP model (Meijers, 2011); a) Classic SSC, b) wireframe model, c) slicing planes for several LoDs

The first phase of this research is to review the existing conceptual models in implementing scale dimension as the third dimension. Further, the best concept or approach is selected and will be implemented within the DHE data structure for scale implementation. Thus, our proposed model will be able to preserve the advantages of both DHE data structure and the selected conceptual model.

2.2 Data Structure

A data structure is defined as a specialized format or approach for organization and storage of digital geographic dataset in a particular spatial data model. It refers to the approach on how to encapsulate the geometry, topology and semantic information of the GIS data systematically. The main purpose of defining the data structure in the data model is to enable any process of analyzing the world with its objects and phenomena within the digital model (Boguslawski, 2011). Different data structures can be implemented in a single data model (Ledoux, 2006).

A data model however is a conceptual model of the real world (Wright et. al., 2008). It can be defined as a method or logical or mathematical way of visualizing and representing data (informational needs) in an information system. From GIS point of view, a spatial data model is needed to simplify the sketch of real world and achieving disperse model in geographical way of representation within a computer (Liangchen et al., 2008).

Generally 3D data models can be divided into four main groups: 3D geometric models, 3D topological and graph models, 3D city models and 3D CAD models (Boguslawski, 2011; Lee and Zlatanova, 2008). The spatial models are used in accordance with a particular application, for example using Constructive Solid Geometry (CSG) or Boundary Representation (B-rep) (Boguslawski, 2011) in 2D or 3D environment of CAD system.

The selection of an appropriate data structure to be implemented inside a data model is made based on the targeted applications and dimensionality structure (Boguslawski, 2011; Gold, 2005). In fact, it is impossible to design a general-purpose data structure in which capable to run in all kind of situations (Meijers, 2011).

2.2.1 Dual Half Edge (DHE) Data Structure

For the higher dimension modelling, there are several data structures that are able to represent models in basic four or more dimensions (Ohori et al., 2013a). They are: Nef polyhedral (Bieri and Nef, 1988), polytopal meshes (Sohanpanah, 1989), simplex-based ones (Paoluzzi et al., 1993; Shewchuk, 2000), and decompositions of polytopes (Bulbul et al., 2009). However, despite the fact that they are able to maintain various topological relationships, none of them can provide efficient access to the dual graph of a model for the navigation purposes.

Ohori et al. (2013) had identified two candidates of data structures that able to model higher dimensionality. One of them is the Dual Half Edge (DHE). The DHE uses two structures; the dual and the primal that are presented as entities connected to form a graph (Boguslawski, 2011). The dual structure conforms to the rules of the 3D Poincaré duality (Boguslawski, 2011). A cell in primal structure presented as a single dual vertex; a face as a dual edge, an edge as a face and meanwhile, a vertex is presented as a cell Fig. 6 illustrates the concept of 3D Poincaré duality applied in DHE data structure.
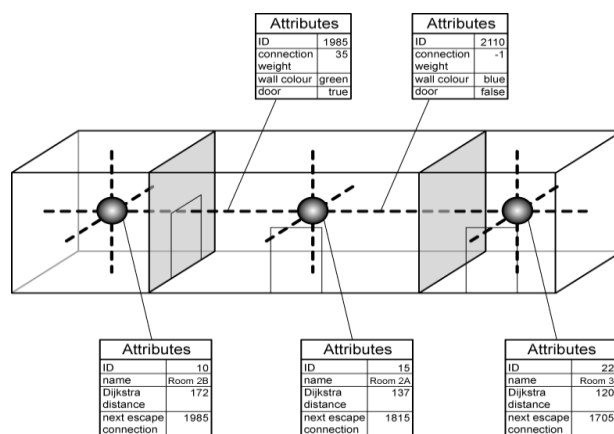

Fig. 6. Poincaré duality with the primal space to support topology and the geometry of 3D cell (Boguslawski, 2011).

Navigation in DHE data structure is possible because all of half-edges in the model are connected via pointers. Half-edge is the concept uses to represent edges in the model in which all edges are split into two directed halves. The navigation is performed from half edge to half edge, each half edge consist of directed half edge with dual half edge. There are four basic set of operator pointers available for navigation; *Sym, Nextv, nextf*, and the dual for every primitives such in Fig. 7.
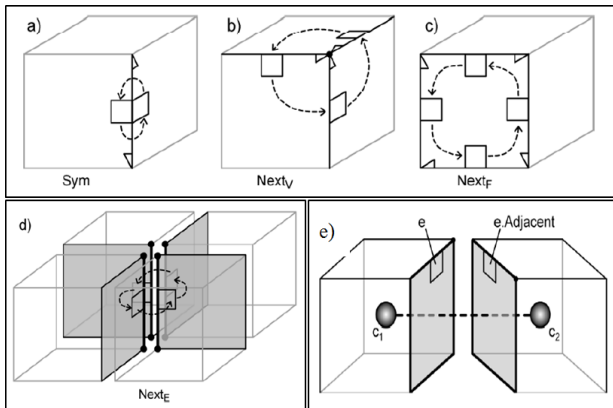
Fig. 7. Navigation of Dual Half Edge (Boguslawski, 2011).

2.2.2 DHE Vario-scale Data Model: Concept

The implementation of scale dimension within the DHE data structure can be considered as a new data model according the data model definition. The DHE data structure have the capability to support the geometry, topology and attribute and with additional scale (let's say, l) dimension to replace Z axis, the semantic element in DHE is being explored exhaustively to minimize the data redundancy in scaling datasets. The concept of DHE vario-scale model is illustrated in Fig. 8.
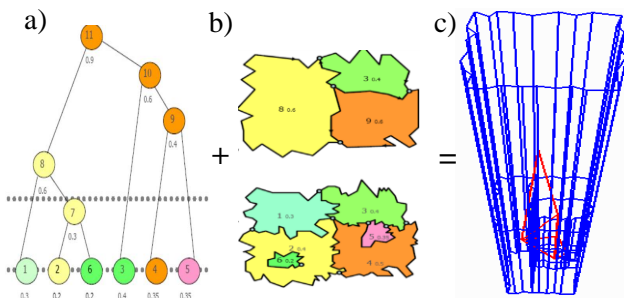

Fig.8. The idea on implementing the DHE vario-scale data model in a single model

The DHE can model the given tGAP dataset efficiently. For a special case such as a polygon with holes, the Triangulated Irregular Network (TIN) model is required to model this situation. But, cost-wise, it is quite expensive in term of storage requirement to store the scale model. Table 1 shows the result comparison of the basic and triangulated planar model for half edge and DHE data structures.

	Basic model	Triangulated planar model
Half Edge (Primal)		
Dual Half Edge		
Number of Edges	2,864 edges (1,432 for HE)	7,076 edges (3,538 for HE)
Storage required	248kB	504 kB

Table 1. Storage requirement of the model

In order to minimize the storage volume required in storing the same dataset, TIN is not an efficient option to proceed with.

Thus, working with the highest details of 2D dataset with generalization approach is last option that should be tested.

3. TOPOLOGICAL 2D-TO-3D GENERALIZATION TECHNIQUE

The available 2D-to-2D generalization or 3D-to-3D generalization of spatial dataset are not applicable for this scenario because of several factors:

- Topology is encrypted in the data structure
- DHE comfort Poincaré duality rule that need a synchronize generalization for both geometries
- Scale dimension requirements/parameters

Thus, the normal visualization techniques and generalization algorithms such as based on emphasis, displacement, simplification, aggregation, individual or group of features and others are not applicable for this research.

Hence, a new generalization technique needs to be derived to support the scale implementation. This 2D-to-3D generalization technique basically derived to transform the 2D spatial dataset to 3D-scale (XYl) with topology implementation; in DHE environment, comforts Poincaré duality rules.

The proposed generalization technique comprising three main sections or steps:

- Extrusion: process of transforming 2D-to-3D scale
- Merge by Face: topological and geometry correction
- Topology Simplification: selection of tree structure

3.1 Extrusion

This research introduces three extrusion techniques as a concept and option of acquiring 3D-scale model from the 2D scale dataset. Fig. 9 shows a sample of scale dataset to demonstrate the extrusion techniques (Fig. 10-12).

Fig. 9. 2D scale dataset for demonstration purpose

3.1.1 Extrusion by Polygon (Normal extrusion)

Extrusion by polygon or face is a normal technique to generate a 3D model out of 2D faces such as from 2D as-built drawing and plan. This technique creates a cell (3D box) which uses the same 2D coordinate from the selected 2D polygon by adding the third dimension as normally refers to Z-dimension (height). (Ohari and Ledoux, 2013) further describes the process of extrusion by this technique to generate higher dimensional of GIS spatial dataset (3D).

Fig. 10. Example of Normal Extrusion technique (polygon).

3.1.2 Extrusion by Line

The use of a single line or polyline as based reference to extrude the other edges. It is sased on *winning and losing* concept; an object is gradually losing until it is completely disappear and another one is gaining the space to replace the old one.
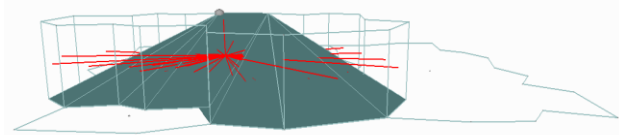

Fig. 11. Example of Line Extrusion technique

3.1.3 Extrusion by Point

The use of centre point with extrusion height value as the based reference to extrude the other edges. This extrusion is still based on *winning and losing* concept. Extrusion by point can be used for cases of A completely inside B (hole or island), or the point touching the boundary of B and others.
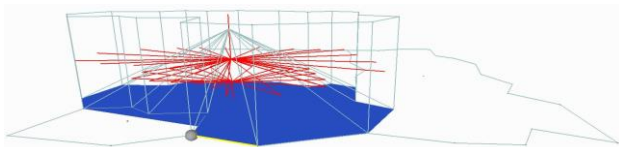

Fig. 12. Example of Point Extrusion technique

3.2 Merge by Face

Merge by face is one of the basic Euler Operators for modification/updating purposes. 'Merge by Face' operator is basically allows for merging of two cells joined by a shared face into one cell – the shared face is removed from between the cells and then deleted. This operation changes only occur in the internal cells, thus it can be used inside a cell complex where the external cell is not directly accessible (Boguslawski, 2011).

3.3 Topology Simplification

The original simplified version of the 3D DHE data structure works by removing the dual structure entirely from models. The removed dual structures are not useful for advanced analysis and could save some storage space. The volume of the primal (dual nodes) and connections between cells (dual edges) are no longer present in the model post removal, thus attributes cannot be assigned to these entities (Boguslawski, 2011).

For storing topology of 3D-scale DHE data model, further simplification is crucial in order to minimize the storage space of the original simplified 3D DHE for cell complexes. This new simplification need to be integrated with the concept of hierarchical tree structure which can be assigned only once in the result of merged by face dual edge cells. The usual components of dual primitives such as nodes, edges and faces are not useful for this dimension and can be removed completely for minimizing the space for storage and improve the efficiency of query duration.
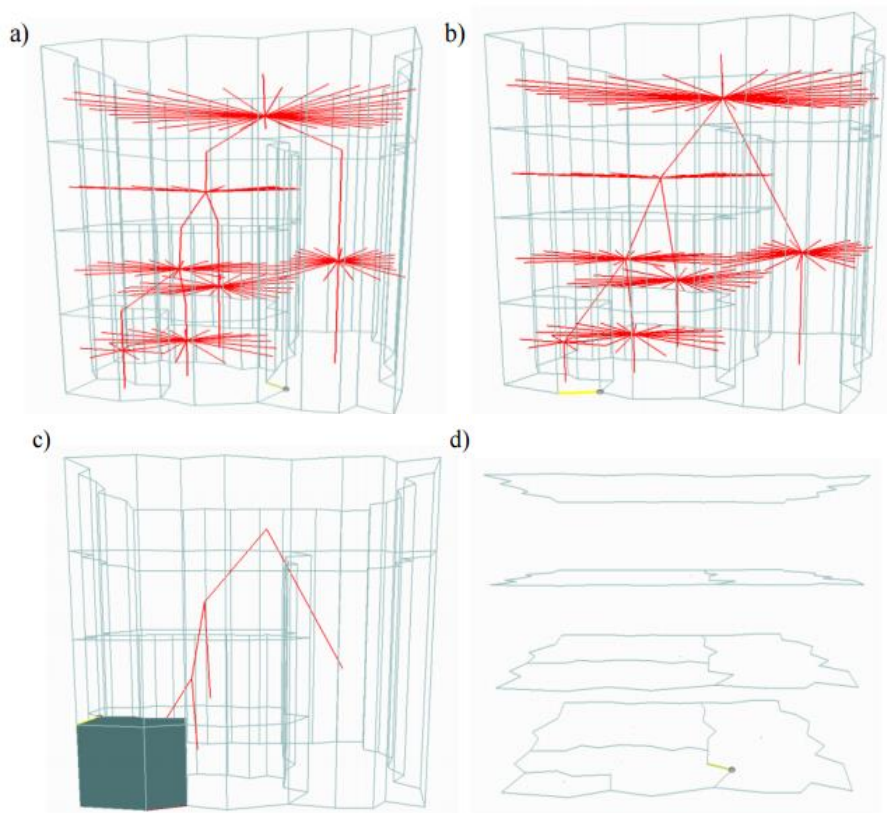

Fig. 13. Overall process of generalization; a) extrusion technique, b) Merge by Face operator, c) Simplified version of the dual (for navigation and query across scale level), d) corresponding 2D maps with different LoDs.

4. EXPERIMENT

Fig. 13 shows the overall process of 2D-to-3D scale generalization to produce n-LoD 2D maps (where n is the number of objects). The result shows the utilization of normal generation technique (extrusion by polygon).

3D-scale DHE data model without the 2D-to-3D generalization (especially extrusion technique) will produce a mess environment for the dual edge (topology) such as in Fig. 14. This model is not valid and produces a false or disturbance in the navigation between LoDs.

The topological extrusion is quite complex since it involves some rules to preserve the topology structure. Ledoux and Meijers (2010) discussed the construction of 3D objects based on extrusion technique to preserve the topology consistency in a case of hole/island.

The Triangulated Irregular Network (TIN) is also an alternative method to this generalization tehnique. TIN is used to model the scale dataset (with hole case or A inside B). The process of extrusion and tetrahedralization are also discussed by Ohari and Leduox (2013). However, it requires more effort to verify the model and it is found, this technique requires a lot of storage space than the generalization technique.
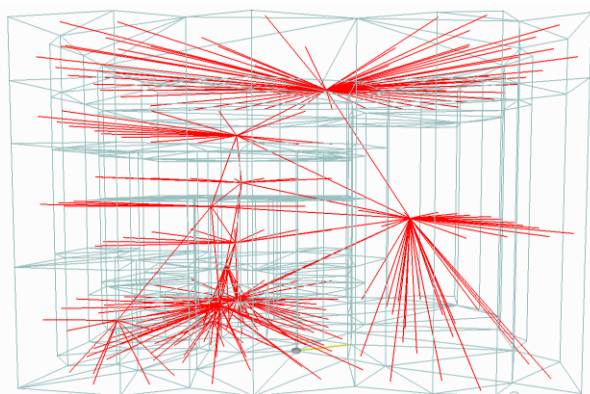

Fig. 14. The result of simply transform the 3D model into DHE environment.

Extrude by Point is the new contribution as far as the extrusion techniques for scale dimension is concerned. This technique reduces the number of edges, faces and the cost of storage for the model compared to TIN; especially to handle the case of polygon within polygon (A within B; an island is bounded by the sea or water).

The significant contribution of introducing *Extrude by Line/Polyline* technique is generalization based on boundary preservation. Some applications need their boundary to be preserved more than other aspects such as cartographic element and area. For example, the shared cadastral boundary which is a very sensitive issue in many countries and to preserve the river boundaries while the model undergo the generalization process. The most accurate term for this technique is Line/Polyline Weighted parameter; selecting the most important line segment to be preserved.

The use of vario-scale approach in deriving scale dimension as the 3rd dimension (after spatial 2D dataset) capable to minimize the data redundancy problem for any given application. The

implementation within the DHE data structure increases the potential and specific demand for topological navigation between LoDs. The introduced extrusion techniques are introduced mainly to help the transformation of scale data from 2D-to-3D and at the same time preserving the validation of the DHE model. The techniques also reduce the storage cost and broad up the application domain for point, line and polygon extrusion.

5. CONCLUSION

There are two possible approaches in implementing scale as a new dimension for spatial 2D, they are - multi-scale and vario-scale. The vario-scale approach is chosen due to its capability to support minimum topological structure for integrating various LoDs in the DHE data structure.

This paper introduced 2D-to-3D generalization technique for topological structure in supporting scale dimension. The approaches for the generalization technique are - "extrusion" (polygon, line and point extrusion), "merge-by-face" (or glue in Euler operators) and "topology simplification".

The experiment of the simple land cover dataset shows the implementation of the generalization technique using the extrusion by polygon proved that the input dataset (i.e. land cover) is capable of producing various LoDs to represent the scale dimension. The merge-by-face operator benefits the primal and the dual graph of DHE model from producing redundant faces. The topology simplification reduces the number of redundant edges thus the storage and query performance could be optimized.

The generalization technique with the three approaches (i.e. extrusion, merge-by-face and topology simplification) gives a flexible generalization methods for supporting topological vario-scale environment. For instance, it could be applied in tGAP structure, half edge and other supported topological data structures. We would like to extend this piece of work toward an automated 3D-scale generalization with 3D DHE database capability.

REFERENCES

Aarnes, J., Kippe, V., Lie, K.A. and Rustad, A.B., 2007. Modelling of Multiscale Structures in Flow Simulations for Petroleum Reservoirs, Geometric Modelling, Numerical Simulation and Optimization. Springer, pp. 307-360.

Baig, S.U., Hassan, M.I. and Rahman, A.A., 2011. Automatic Generalization of 3D Building Models – A Review, 10th International Symposium & Exhibition on Geoinformation (ISG 2011). ISPRS, Shah Alam Convention Centre(SACC), Malaysia.

Boguslawski, P., 2011. Modelling and Analysing 3D Building Interiors with the Dual Half-Edge Data Structure, University of Glamorgan, UK, 1-134 pp.

Gold, C. M. (2005, September). Data structures for dynamic and multidimensional GIS. In Proceedings of the 4th ISPRS Workshop on Dynamic and Multi-dimensional GIS (pp. 36-41).

Günther, O., 1988. Efficient Structure for Geometric Data Management. Number 337 in Lecture Notes in Computer Science. Springer-Verlag, Berlin. Springer Book.

Jones, C.B. and Abraham, I.M., 1986. Design Considerations for a Scale Dependent Cartographic Database. 384-398.

Jones, R.R., McCaffrey, K.J.W., Clegg, P., Wilson, R.W., Holliman, N.S., Holdsworth, R.E., Imber, J. and Waggott, S., 2009. Integration of Regional to Outcrop Digital Data: 3D visualisation of Multi-scale Geological Models. Computers & Geosciences, Volume 35(1): 4-18.

Ledoux, H., 2006. Modelling Three-dimensional Fields in Geoscience with the Voronoi Diagram and its Dual. PhD Thesis. Certificate of Research, University of Glamorgan.

Ledoux, H. and Gold, C.M., 2007. Simultaneous storage of primal and dual three-dimensional subdivisions. Computers, Environment and Urban Systems, 31(4): 393-408.

Ledoux, H. and Meijers, M., 2010. Topologically consistent 3D city models obtained by extrusion. International Journal of Geographical Information Science, 25(4):557–574, 2010. pp. 4-7.

Lee, J. and Zlatanova, S., 2008. A 3D Data Model and Topological Analyses for Emergency Response in Urban Areas. Geospatial Information Technology for Emergency Response. Taylor & Francis.

Lindeberg, T., 1994. Scale-space Theory: A Basic Tool for Analysing Structures at Different Scales. Journal of Applied Statistic, Volume 21(2): 225-270.

Liangchen, Z., Guonian, L., Yehua, S., Hangbo, X., & Haixia, W. (2008). A 3D GIS spatial data model based on cell complex. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII (Part B2).

Meijers, M., 2011. Variable-Scale Geo-Information, PhD Thesis. Technische Universiteit Delft, Netherlands.

Ohori, K.A., Boguslawski, P. and Ledoux, H., 2013a. Representing the Dual of Objects in a Four-dimensional GIS, International Workshop on Geoinformation Advances. 1-12, pp. 4.

Ohori, K.A., Ledoux, H. and Stoter, J., 2013b. Modelling Higher Dimensional Data for GIS Using Generalised Maps, 13th International Conference on Computational Science and Its Applications, Ho Chi Minh City, Vietnam.

Ohori, K.A. and Ledoux, H., 2013. Using Extrusion to GenerateHigher-dimensional GIS Datasets. Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, November 2013, pp. 398–401. ISBN: 978-1-4503-2521-9.

Oosterom, P.v., 1990. Reactive Data Structure for Geographic Information Systems, PhD Thesis. Leiden University.

Oosterom, P.V., 2005. Variable-scale Topological Data Structures Suitable for Progressive Data Transfer: The GAP-face Tree and GAP-edge Forest. Cartography and Geographic Information Science(32): 331–346.

Oosterom, P.V. and Stoter, J., 2010. 5D Data Modelling: Full Integration of 2D/3D Space, Time and Scale Dimensions, 6th international conference on Geo- graphic information science. Springer- Verlag, Berlin, Heidelberg, pp. pages 310–324.

Peuquet, D.J., 2001. Making Space for Time: Issues in Space-time Data Representation. Geoinformatica, 5: 11-32.

Raper, J., 2000. Multidimensional geographic information science. Taylor & Francis, London.

Sester, M., 2007. 3D Visualization and Generalization. Institude of Cartography and Geoinformatics. Leibniz University of Hannover, Online presentation. http://www.ifp.uni-stuttgart.de/phowo/2007/presentations/320sester.pdf (pp. 10)

Vervoort, J.M., Rutting, L., Kok, K., Hermans, F.L.P., Veldkamp, T., Bregt, A.K. and Lammeren, R.v., 2012. Exploring Dimensions, Scales, and Cross-scale Dynamics from the Perspectives of Change Agents in Social–ecological Systems. Ecology and Social, 17(4):24.

Worboys, M.F., 1994. A unified model for spatial and temporal information. The Computer Journal, 37(1): 26-34.

Wright, D., Knoche, J. and Kugler, T., (2008). GEO 580: GIS Data Models and Data Structures (Online Instruction Report). (2). Retrieved April 2016, from University of California. http://dusk.geo.orst.edu/buffgis/Arc9Labs/Lab2/lab2.html