

Unsupervised Symmetric Polygon Mesh Mapping. The Dualism of Mesh Representation and Its Implementation for Many Layered Self-Organizing Map Architectures.

The Dualism of Mesh Representation and Its Implementation for Many Layered Self-Organizing Map Architectures

Conference Paper

Author(s):

Standfest, Matthias

Publication date:

2014

Permanent link:

<https://doi.org/10.3929/ethz-a-010248840>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Unsupervised Symmetric Polygon Mesh Mapping

The Dualism of Mesh Representation and Its Implementation for Many Layered Self-Organizing Map Architectures

Matthias Standfest

ETH Zürich, IA - Chair of Information Architecture, Switzerland

<http://www.ia.arch.ethz.ch>

standfest@arch.ethz.ch

With this paper we present a fully automated semantic shape similarity detection based on N-rings with further potential for shape synthesis in a topological correct feature space. Therefore a way of symmetric encoding of geometry, optimized for the use as feature-vector in self-organizing maps, is introduced. Furthermore we present a modified kernel for the detection of the best matching unit in self-organizing maps especially designed for a data topology differing from the default predecessor/successor structure. Finally we provide the results of a conducted experiment clustering building blocks of an area in Zürich, Switzerland.

Keywords: *Unsupervised Machine Learning, Geometry Clustering, Self-organizing Map, Mesh Synthesis, Probabilistic Modelling*

BACKGROUND

One intrinsic problem of the architectural design process is the issue of how to compare different designs and how to interpolate between them. As a necessity for any synthesis driven design approach this wicked problem has yet to be fully automatized, especially in consideration that existing methods fail in practicability because of expensive manual data pre-processing. So far in the field of architecture those methods can be divided into two main groups, corresponding to the particular kinds of compared data: either the comparison of shape determining rules instead of a shape itself (cf. Stiny, Mitchell 1978; cf. Fezer 2009; cf. Eastman 1972; cf. Rittel 1977; cf. Alexander et al. 1977; cf. Dave et al. 1994) or the comparison of key figures computed as arbitrary shape

features of an underlying architecture (cf. Langenhan et al. 2011; cf. Dillenburger 2010; cf. Derix et al. 2012). Clearly, for both approaches pre-processing of data in different ways is necessary. While the first group is focusing on the analysis of decision trees leading to a shape, in order to allow sustaining a limited ability for synthesizing interpolated shapes, exactly this ability rapidly decreases with the increasing complexity of the solution-space; up to a point where finally generalizable approaches like the highly abstract Case Based Reasoning (cf. Dave et al. 1994) need to fully sacrifice the ability for shape-synthesis in order to maintain a certain kind of interpretable and general ontology. Additionally, when comparing shape determining rules a larger solution space is mainly achieved by increasing the complexity of

the underlying rule sets. This problem can be partly avoided by comparing a set of designs according to arbitrary figures (cf. Dillenburger 2010; cf. Derix et al. 2012) or topology informed labels (cf. Langenhan et al. 2011) rather than the decision trees describing the shapes preceding design processes. Such highly abstract and generalized models for computing key figures provide easy applicable semantics for a much larger solution space, because they minimize the costs for a useful comparison by reducing the amount of semantic information needed. While these methods render only a handful basic architectural elements required for being labelled during the data pre-processing stage, the figures used for comparison lack the potential of being translated unambiguously into architectural shapes. Thus we identified that the potential for shape-synthesis on the one hand and the generalizability of the labelling method on the other hand tend to be mutually exclusive. However, for a seamless integration of a procedure of this kind into any architectural design process it is a necessity to provide both analysis and synthesis capabilities. In order to do so and to provide exactly those discussed processes this paper investigates, inspired by actual developments in the field of computer graphics, the potential of a method based on a purely geometric representation of architecture. The goal was to work with the latent semantics hidden underneath the topological *contexts* of the mesh geometries.

METHOD

Already known from different methods for architecture or urban planning (cf. Stadler et al. 2007), most approaches in computer graphics need to pre-process the purely geometrical information provided for enriching it with certain semantic or topological information first. The main focus hereby lies on the segmentation into semantically tagged shapes (cf. Chaudhuri et al. 2011) which is necessary to compare and cluster shape elements according to different geometrical or topological features (cf. Funkhouser et al. 2004) or to automatically label identical functions

(cf. van Kaick et al. 2013). Others even established a workflow where this information is used to combine existing shapes into meaningful new combinations (cf. Chaudhuri et al. 2011). And thereby they provide methods for both analysis and synthesis of polygon meshes - the missing link for the architectural design process. While the main problem of those techniques is the costly supervised or even manual segmentation, a problem also leading to a growing demand for pre-segmented data repositories (cf. van Kaick et al. 2013), this paper presents a way of substituting the necessary semantic information by a purely contextual one, analogue to unsupervised image (cf. Bengio 2009) and text processing (cf. Bellegarda 2004).

Most methods in clustering or probabilistic modelling of geometry are either based on arbitrary feature descriptors for the mesh segments like angles, diameters and much more complicated ones (cf. Sidi et al. 2011) or topology informed semantic labeling (cf. Kalogerakis et al. 2012). We already described those tactics when previously mentioning the ways of comparing designs in the field of *CAAD*. The two problems arising are yet again the need for large labeled (pre-segmented) data sets and, regarding the feature descriptors, the difficulty of how to decide which of the vast amount of possible descriptors are relevant for clustering in general or if not a contextual approach would be preferable to the intrinsic one (cf. Standfest et al. 2013). Both problems are shown to be avoidable by using unsupervised learning methods for creating high-level features from only unlabeled data (cf. Le et al. 2011). And undoubtedly this is also part of the observable trend towards minimizing the amount of semantic information needed for state of the art data analysis. To contribute to this development, this paper is focusing on the problem of comparing unlabelled geometry data, by using both many layered neural networks and n-gram inspired mesh processing. We show how emergent descriptors allow clustering of architecture comparable to state of the art approaches in geometry clustering while the same method still manages to maintain synthesizing abilities (figure 1).

Figure 1
result of the
conducted
experiment on
clustering 48
randomly chosen
building blocks
according to the
latent semantics of
the unlabelled
mesh geometry

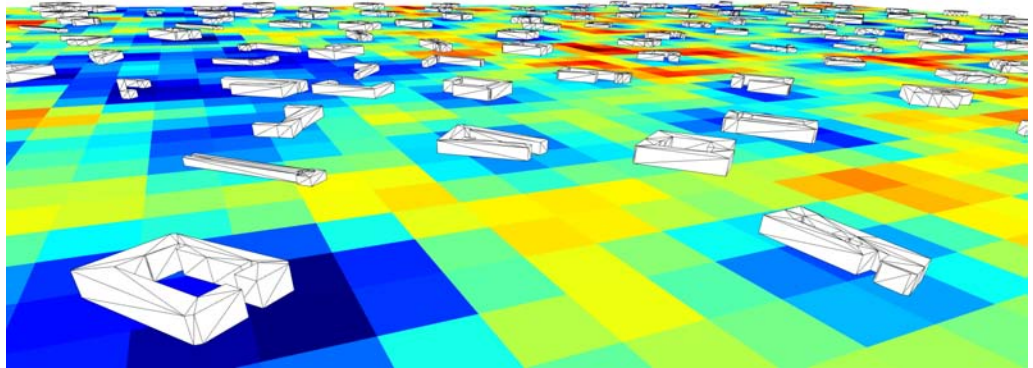


Figure 2
constructing a
2-ring (right) as a
collection of the
outer leaves of the
nested 1-rings
(middle) with the
corresponding root
1-ring (left)



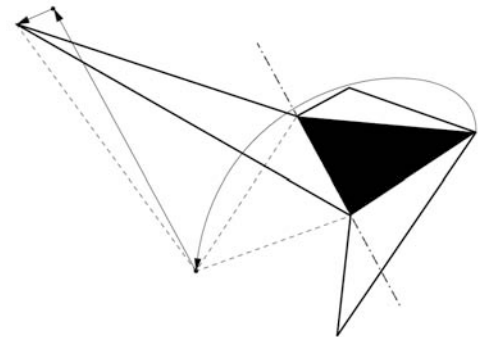
Nested N-Rings

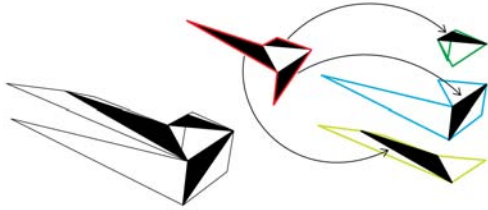
As an alternative to the popular goal of achieving *semantically correct* sub graphs (cf. Kalogerakis et al. 2012) the segmentation of polygon meshes may also aim for same sized patches. This can be done either via intersecting the mesh with r radius balls (cf. Mitra et al. 2006) or as in our case by creating patches of similar underlying topology. Like N-grams in language processing (Shannon 2001 [1948]), small topologic configurations instead of predefined subdivisions now carry latent semantics. Therefore we chose to compare each triangle according to its topologic neighbourhood, a format we further call N-ring (figure 2).

These N-rings of each triangle are strongly related to N-grams and are used to replace, from natural language strongly influenced, standard semantics with its statistical model (cf. Bellegarda 2004). To achieve a fully contextual encoding without intrinsic semantics (cf. Deleuze et al. 2008), we build upon the dualism in mesh representation already used for polygon mesh compression (cf. Rossignac

1999). Adopted for constrained Delaunay triangulation (cf. Chew 1987) this means we work with the triangle relations instead of the triangles themselves. A shift easily accomplished by constructing the folding transformation for each pair of neighbours (figure 3). We define this fold by one rotation angle α and one pair of relative uv transformations. Due to the fact that certain angles appear much more often than others (e.g. 180° where the mesh is planar or 90° as the most common edge) this design is especially optimized for architectural or urban geometries and supports thereby subsequent clustering efforts (as does minimizing the number of triangles by using constrained Delaunay triangulation).

Figure 3
triangle folding
process, showing
the rotation axis as
well as the u- and
v-transformation



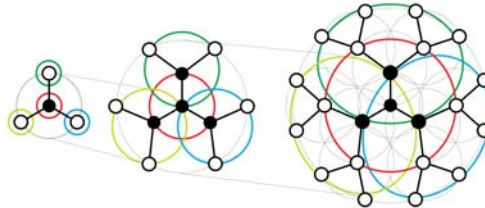


Analogue to this contextual encoding paradigm, N-rings with distance greater than one edge may be composed out of smaller rings as alternative to constructing them from scratch, allowing the hierarchical processing of geometry (figure 4). For robustness reasons we designed those nested rings to create overlapping branches, resulting in the composing of a N-ring of size i out of four N-rings of size $i - 1$ or 4^{i-1} N-rings of size 1. This way we ensure a fixed data topology for the whole recursive process. The presented kind of segmentation is fully unsupervised and works with any unlabelled triangular 2-manifold polygon mesh without boundary. We optimize the meshes before converting them into training vectors (as basis for the obligatory Delaunay triangulation planar straight-line graphs are detected first after reducing some noise like welding vertices or flattening of peaks). Nevertheless the underlying data encoding topology differs fundamentally from well investigated predecessor-successor patterns of data streams (like text or audio signals) causing consequences for the further below described best matching unit (BMU) kernel of the used self-organizing map (SOM).

Mesh-SOM

The presented mesh segmentation allows working with a fixed data topology for every nested N-ring of arbitrary size in form of a tree graph with one root node and three leaf nodes, thus a star S_3 with 3 edges what we further call a *claw*. Every node of the graph is again a nested claw itself (with at least one overlapping edge for connection purposes) and may be decomposed until the lowest order of unring representation is reached (figure 5). Because the SOM is trained layer by layer the conceptual complex-

ity of each level increases and therefore bigger trees may be interpreted as more abstract features than smaller ones - a common method when compared to multi-level usage of Gabor filters (cf. Gabor 1946) or N-grams which are essential for deep believe networks (cf. Bengio 2009). This means that we compare the triangles not only by its N-rings but according to their immediate but increasingly detailed neighbourhood. As a result overlapping edges (a third of the branches is oriented towards the root claw) improve the robustness of our encoding significantly.



Replacing predecessor/successor-pairs with claws leads to a further change in the feature vector managing of our process. The clearly identifiable root ring and the order of the claw's leaves determined by the normal of the initial triangle result in only one remaining ambiguity: the index of the initial leaf. When converted into a vector the values of the root node are always the first few dimensions, but the following topology of the three neighbouring rings (the leaves) forms a 3×3 circulant matrix C . For synthesis each nested claw (subring) needs to be rotated for finding the matching fold of the root claw (figure 6), because of this uncertainty. On the other hand, for the analysis phase the BMU kernel is adopted to find the best of three variations (the most frequent initial claw) of a feature vector and is thereby increasing the quality of its SOM.

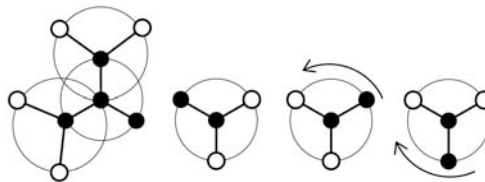


Figure 4
composing a higher level ring out of four lower level rings (red: root, dark green: leaf A, cyan: leaf B, light green: leaf C)

Figure 5
nested hierarchical data topology (red: root, dark green: leaf A, cyan: leaf B, light green: leaf C) of level 1 (not nested), level 2 (biring) and level 3 (3-ring)

Figure 6
circular data topology leaves 3 possible index leaves per claw

This limitation to three possible combinations per vector per SOM level allows to approach the underlying combinatorial problem in a nested recursive manner (figure 7) which proves to be much faster than proceeding just the direct way. Only the initial learning level has twice the amount of variations to be compared. The simple reason is that we try to further densify the solution space of the SOM by the possibility of flipping the normal of each triangle (changing the order of its neighbours). Finally comparing the different variations of each vector is done by adopting the used GPU SOM BMU kernel (cf. Wittek 2013) to not only find the best matching unit, but the best matching unit possible out of n variations - we call this kernel further *BMUofN*. This modification is the very reason why we are able to produce necessary findings otherwise impossible to compute in a time sensitive manner (table 1).

The rest of the process is basically following the 2-grams based *web-SOM* method for uncovering latent semantics of texts (cf. Ritter et al. 1989): For each N-ring level one SOM is first sorting the concepts (the different kinds of triangles) according to its contexts in the form as described. For better training results those feature vectors are also standardized but not normalized (cf. Blayo 1992). Subsequently a so called domain map for each polygon mesh is created. This is done by creating the domain feature vectors on basis of histograms over the triangle map which again is smoothed by a Gaussian convolution kernel (cf. Kaski et al. 1998). As a result we get SOMs arranging the compared meshes according to lower or higher level emergent features which establishes a process of strictly unsupervised clustering of unlabelled architectural geometries. Summing up, we present a way of algorithmic modelling (cf. Breiman 2001) geometrical semantics as an alternative to the intrinsic ones provided by semantically enriched pre-processed data.

RESULTS

For a first test of the method we compared a set of 48 building blocks (the level of detail is 1) randomly chosen from the area of Zürich Altstetten, leading to the comparison of 3064 triangles in total. The resulting U-matrices of the triangle maps (figure 8) clearly show significant clustering of similar triangle contexts. Although this kind of semantics may differ from the kind the human mind would create. Further research needs to be conducted in how far this algorithmic modelled one is able to be used to substitute previously missing semantic classes (cf. Sidi et al. 2011). In addition to the emergent clustering we observed that our modified BMU kernel leads to much clearer results, e.g. at the initial training level (figure 8). Finally the rendering of the triangle map of level 3 reveals the lack of necessary additional data to do high quality high level abstraction. A problem which is in accordance to the existing studies showing the necessity of big data sets for deep learning algorithms (cf. Le et al. 2011). This underdetermina-

Figure 7
computational
costs due to
possible tree
combinations

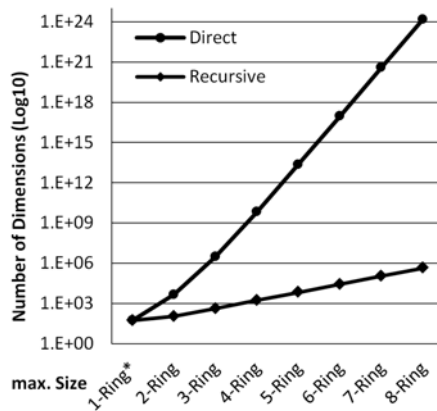


Table 1
relation between
recursive and direct
data processing

Max. Ring	1-Rings	Faces	Dimensions	Recursive : Direct
1-Ring	1	4	9	1:1
2-Ring	4	10	36	1:42
3-Ring	16	22	144	1:6783
4-Ring	64	46	576	1:3.75E+06
5-Ring	256	94	2304	1:3.24E+09
6-Ring	1024	190	9216	1:3.17E+12
7-Ring	4096	382	36864	1:3.21E+15
8-Ring	16384	766	147456	1:3.28E+18

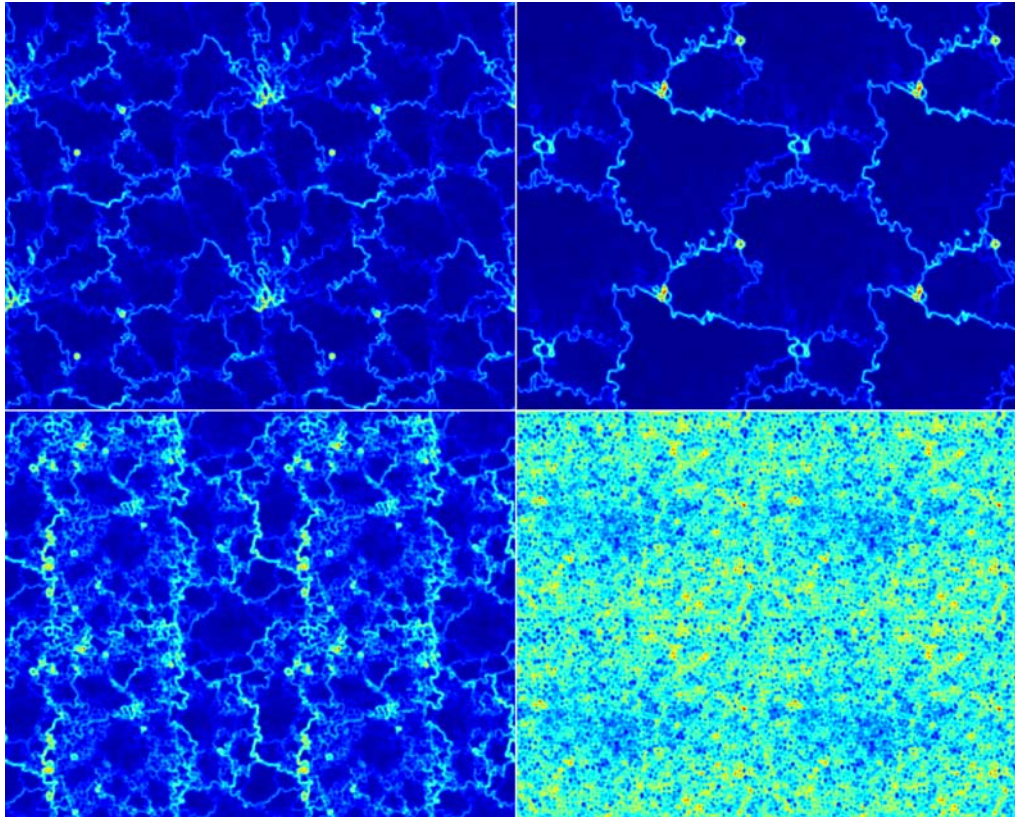


Figure 8
triangle maps with
toroid topology,
t.l.t.b.r.: level 1
(default kernel),
level 1 (modified
kernel), level 2
(modified kernel),
level 3 (modified
kernel). The color
scale reflects the
cluster borders,
blue cells are close
neighbours while
bright ones are
further apart.

tion of the achieved clustering as result of deficient data is closely related to the concept of algorithmic modeling itself (cf. Lyre 2004).

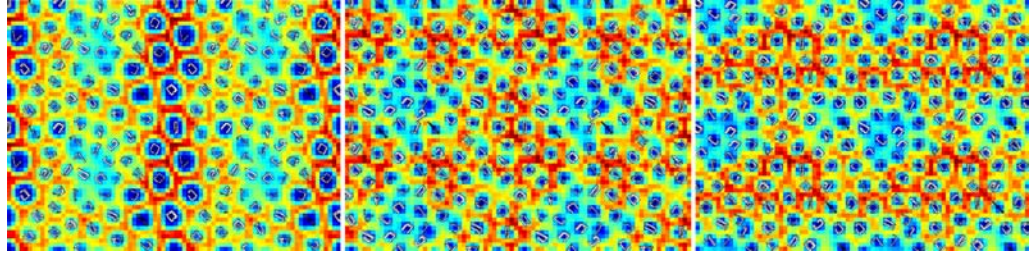
Level	Map Size	Dimensions	T-Error	Q-Error	QC-Error
Level 1*	120x150	1x9	0.32735	0.00581	0.00286
Level 1	120x150	6x9	0.06560	0.00141	0.00064
Level 2	120x150	3x36	0.28949	0.04733	0.00885
Level 3	120x150	3x144	0.37794	0.14427	0.01230

To quantify the quality of the SOMs regarding the clustering for analysis purposes as well as the accuracy for synthesis purposes we used well established figures like the topographic error and the quantiza-

tion error (cf. Uriarte et al. 2008). As a modification of the last one mentioned we calculated the average Euclidian distance not only over the whole feature vector but instead over the single three dimensional components each ring consists of. Thus the *component quantization error* (QC -Error) is more revealing in this context and is better suited for comparing the different SOM levels on a unified basis (table 2). The computed statistics of the test scenario emphasize the impressions we already got from U-matrix renderings: first of all the significant differences between the default BMU kernel and our modified one which is densifying the solution space. Then

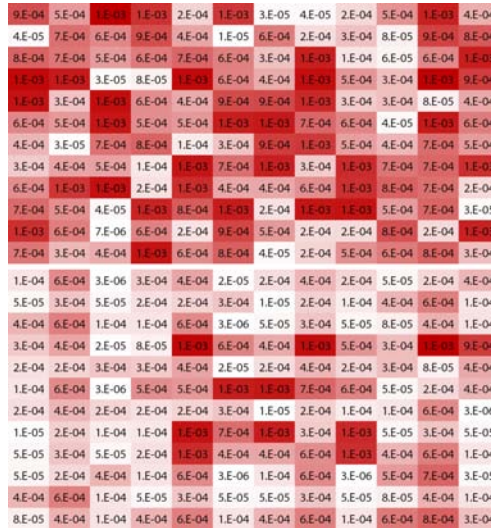
Table 2
SOM statistics (level
1* is calculated with
default BMU kernel
method)

Figure 9
domain maps of
level 1 (based on
the distribution of
unirings), 2 (based
on the collection of
birings) and 3
(based on further
nested 3-rings)



we observe the lack of data resulting in a rather great topographic error within our most abstract SOM. And finally we get positive results of our second level triangle map regarding a convincing synthesizability.

Figure 10
sample deviation of
trained level-3
feature vector to its
original before (top)
and after (bottom)
leveling the error by
using the
underlying
redundancy



Despite the still limited dataset the three resulting domain maps already show reasonable clustering and arrangements of the evaluated building blocks (figure 9). Regardless of the significant quality differences between the underlying triangle maps of level two and three the two corresponding domain maps of these levels show surprisingly few differences. This is partly caused by the Gaussian convolution filter we

applied to smoothen the histograms and partly result of a probable limit of reasonable N-ring sizes - a limit research in phrase-based language processing as already suggested (cf. Zollmann et al. 2008).

Finally our tests confirm the self-stabilizing effect of our robust nested encoding method (figure 10). When re-translating, the produced QC-errors of level 3 feature vectors are between 30% and 40% smaller than without redundancy and appear especially small in close neighborhood to the root triangle.

CONCLUSION

The dominant outcome of this paper is a fully automated shape similarity detection with the potential for shape synthesis in a topological correct feature space. Therefore a way of symmetric encoding of geometry, optimized for the use as feature-vector in self-organizing maps, is introduced. The conducted experiments further illustrate how different unlabelled polygon meshes can be aligned according to latent semantics. In future this could have the potential of rendering expensive, manually edited, semantically enriched geometry repositories obsolete. The successful implementation of the described techniques is directly connected to a vast number of design related theories. One can interpret the synthesized scenarios as *concretization of possibilities* (cf. Flusser 1994) while the nodes of the SOM itself can be viewed from an actor-network theory perspective (cf. Wassermann 2010). The underlying duality of the mesh encoding is discussed as *double articulation* (cf. Deleuze et al. 2008) and the appli-

cation of the method enforces the opinion that design is rather a *redesign* than *design ex nihilo* (cf. Latour 2009). Of course, this also has an effect on the discussions of authorship in architecture (cf. Carpo 2011) and on the role of geometric representations in the design process and its communication (cf. Evans 2011).

Further consequences for the design process are at least as manifold as the theoretic implications: in future a generalizable analysis of architectural and urban structures is a necessity not only for synthesis driven design approaches but also for any kind of architectural impact assessment. Especially in consideration of *Big Data* and algorithmic modelling, this process is able to form a geometric data stream so that other streams can be mapped onto, so data mining processes can finally re-manifest into geometric output. As a next step in the development of this methodology we are working on modified point set registration (cf. Gelfand et al. 2005; Wang et al. 2008) for re-translating the feature vectors of the domain maps back into polygon meshes to better visualize the output of the trained probabilistic shape space.

REFERENCES

- Alexander, C, Ishikawa, S and Silverstein, M 1977, *A pattern Language*, Oxford University Press, New York
- Bellegarda, JR 2004, 'Latent Semantic Language Modeling for Speech Recognition', in Johnson, M (eds) 2004, *Mathematical foundations of speech and language processing*, Springer Science + Business Media, New York, pp. 73-103
- Bengio, Y 2009, *Learning deep architectures for AI*, Now Publishers, Hanover, Mass
- Blayo, F 1992, 'Kohonen Self-Organizing Maps: Is the Normalization Necessary?', *Complex Systems*, 6(6), pp. 105-123
- Breiman, L 2001, 'Statistical modeling: The two cultures', *Statistical Science*, 16(3), pp. 199-231
- Carpo, M 2011, *The alphabet and the algorithm*, MIT Press, Cambridge, Mass
- Chaudhuri, S, Kalogerakis, E, Guibas, L and Koltun, V 2011, 'Probabilistic reasoning for assembly-based 3D modeling', *ACM Transactions on Graphics*, 4, pp. 35-43
- Chew, LP 1987 'Constrained Delaunay Triangulation', *Proceedings of the third annual symposium on Computational geometry*, pp. 215-222
- Dave, B, Schmitt, G, Faltings, B and Smith, I 1994 'Case-based design in Architecture', *Artificial intelligence in design*, pp. 145-162
- Deleuze, G, Guattari, F and Massumi, B 2008, *A thousand plateaus*, Continuum, London
- Derix, C, Gamlesæter, Å, Miranda, P, Helme, L and Kropf, K 2012, 'Simulation Heuristics for Urban Design', in Müller Arisona, S, Aschwanden, G, Halatsch, J and Wonka, P (eds) 2012, *Digital urban modeling and simulation*, Springer, Berlin, Heidelberg, pp. 159-180
- Dillenburger, B 2010 'Space Index: A retrieval-system for building-plots', *Proceedings of eCAADe 2010*, Zürich, p. 893-899
- Eastman, CM 1972, 'Preliminary report on a system for general space planning', *Commun. ACM*, 15(2), pp. 76-87
- Evans, R 2011, 'Durch Papier Sehen', in Voorhoeve, J (eds) 2011, *Welten schaffen*, Diaphanes, Zürich, p. 157-194
- Fezer, J 2009, 'A Non-Sentimental Argument', in Gethmann, D and Hauser, S (eds) 2009, *Kulturtechnik Entwerfen*, Transcript, Bielefeld, p. 287-304
- Flusser, V 1994, 'Vom Subjekt zum Projekt', in Bollmann, S and Flusser, E (eds) 1994, *Schriften*, Bollmann, Bensheim, pp. 9-160
- Funkhouser, T, Kazhdan, M, Shilane, P, Min, P, Kiefer, W, Tal, A and Rusinkiewicz, S 2004, 'Modeling by example', *ACM Transactions on Graphics (TOG)*, 23(3), pp. 652-663
- Gabor, D 1946, 'Theory of communication. Part 1: The analysis of information', *Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26), pp. 429-441
- Gelfand, N, Mitra, NJ, Guibas, L and Pottmann, H 2005 'Robust Global Registration', *Eurographics Symposium on Geometry Processing 2005*
- van Kaick, O, Xu, K, Zhang, H, Wang, Y, Sun, S, Shamir, A and Cohen-Or, D 2013, 'Co-hierarchical analysis of shape structures', *ACM Transactions on Graphics*, 32(4), pp. 69:1-69:10
- Kalogerakis, E, Chaudhuri, S, Koller, D and Koltun, V 2012, 'A probabilistic model for component-based shape synthesis', *ACM Transactions on Graphics*, 31(4), p. 55
- Kaski, S, Honkela, T, Lagus, K and Kohonen, T 1998, 'WEBSOM—self-organizing maps of document collections', *Neurocomputing*, 21(1), pp. 101-117
- Langenhan, C, Haß, S, Weber, M, Petzold, F, Liwicki, M and Dengel, A 2011 'Investigating research strategies for accessing knowledge stored in semantic

- models', *Proceedings of eCAADe 2011*, Ljubljana, pp. 403-411
- Latour, B 2009 'A Cautious Prometheus? A Few Steps Toward a Philosophy of Design', *Proceedings of the 2008 Annual International Conference of the Design History Society*, pp. 2-10
- Le, QV, Ranzato, MA, Monga, R, Devin, M, Chen, K, Corrado, GS, Dean, J and Ng, AY 2012 'Building high-level features using large scale unsupervised learning', *Proceedings of the 29th International Conference on Machine Learning*, Edinburgh, Scotland, UK
- Lyre, H 2004, *Lokale Symmetrien und Wirklichkeit*, Mentis, Paderborn
- Mattausch, O, Panozzo, D, Mura, C, Sorkine-Hornung, O and Pajarola, R 2014, 'Object Detection and Classification from Large-Scale Cluttered Indoor Scans', *Computer Graphics Forum*, 33(2), p. to appear
- Mitra, NJ, Guibas, L, Giesen, J and Pauly, M 2006 'Probabilistic fingerprints for shapes', *Proceedings of Eurographics Symposium on Geometry Processing*, pp. 121-130
- Rittel, HWJ 1977, *On the Planning Crisis: Systems Analysis of the first and Second Generations.*, Institut für Grundlagen der Planung IA, Universität Stuttgart, Stuttgart
- Ritter, H and Kohonen, T 1989, 'Self-organizing semantic maps', *Biological Cybernetics*, 61(4), pp. 241-254
- Rossignac, J 1999, 'Edgebreaker: Connectivity compression for triangle meshes', *Visualization and Computer Graphics*, 5(1), pp. 47-61
- Shannon, CE 2001 [1948], 'A mathematical theory of communication', *Mob. Comput. Commun. Rev. (ACM SIGMOBILE Mobile Computing and Communications Review)*, 5(1), pp. 3-55
- Sidi, O, van Kaick, O, Kleiman, Y, Zhang, H and Cohen-Or, D 2011, 'Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering', *ACM Transactions on Graphics*, 30(6), p. 126
- Stadler, A and Kolbe, TH 2007 'Spatio-semantic coherence in the integration of 3D city models', *Proceedings of the 5th International ISPRS Symposium on Spatial Data Quality*, Enschede
- Standfest, M and Moosavi, S 2013 'Fuzzy Modelling with Self-Organizing Maps', *Proceedings of Design Modelling Symposium 2013*, Berlin, p. 243-256
- Stiny, G and Mitchell, WJ 1978, 'The palladian grammar', *Environment and planning B*, 5(1), pp. 5-18
- Uriarte, EA and Martin, FD 2008, 'Topology Preservation in SOM', *International Journal of Applied Mathematics and Computer Sciences*, 2(9), pp. 874-877
- Wang, F, Vemuri, BC, Rangarajan, A and Eisenschenk, SJ 2008, 'Simultaneous nonrigid registration of multiple point sets and atlas construction', *IEEE Trans Pattern Anal Mach Intell (IEEE transactions on pattern analysis and machine intelligence)*, 30(11), pp. 2011-2022
- Wassermann, K 2010 'SOMcity: Networks, Probability, the City, and its Context', *Proceedings of eCAADe 2010*, Zürich, pp. 197-205
- Wittek, P 2013, 'Somoclu: An Efficient Distributed Library for Self-Organizing Maps', *Preprint - <http://arxiv.org/pdf/1305.1422>*, tba, p. tba
- Zollmann, A, Venugopal, A, Och, F and Ponte, J 2008 'A systematic comparison of phrase-based, hierarchical and syntax-augmented statistical MT', *Proceedings of the 22nd International Conference on Computational Linguistics*, Manchester, pp. 1145-1152