

# Conceptual framework for enhancing visual change point analysis in generalized multi-temporal displays

Jochen Schiewe<sup>a</sup>\*

<sup>a</sup> HafenCity University Hamburg, Lab for Geoinformatics and Geovisualization (g2lab), jochen.schiewe@hcu-hamburg.de

\* Corresponding author

**Abstract**: Large amounts of point data are being generated and depicted in web maps, in particular promoted by the availability of Volunteered Geographic Information (VGI). This data volume causes usability problems for the visual presentation and exploration, especially by decreasing rendering performance and increasing geometric or semantic point clutter. Typically, generalization techniques are applied in order to overcome these clutter problems. Appropriate constraint-based approaches have been developed for static and single displays.

Due to the increasing amount and relevance of multi-temporal point data, the aim of this contribution is to present a so far missing conceptual framework that consists of constraints related to change point analysis. These constraints describe the effects of the generalization on preservation and legibility of change information. Here, different types of change point analysis are differentiated, namely: existential changes, changes of semantic properties, and changes of spatial properties. For this purpose, difference frames between two given frames of the multi-temporal data set are used as a model of the mental representation of change information.

The resulting constraints can be used to describe and to compare the suitability of various generalization operations for specific tasks. However, in a later stage they can also trigger the generalization processes as such. This could lead to a new way of thinking, since the optimization of generalization processes should work on difference frames rather than on single views.

Keywords: change analysis, VGI, multi-temporal analysis, generalization, constraints

# 1. Introduction

More and more point data representing points of interest (POIs), other qualities (such as accident locations) or quantities (such as noise levels) is being generated. This is mainly due to the increasing volume and relevance of Volunteered Geographic Information (VGI). With a variety of services and software solutions, especially online map-based mashups with Application Programming Interfaces (APIs), both expert and nonexpert users can collect and present data over the internet (Huang and Gartner, 2012). As VGI point data is acquired over longer periods of time or even in real time, multitemporal visualizations become more and more relevant. At this point it is of interest to visually detect and analyze changes in multi-temporal point displays.

Especially VGI point data typically has an enormous volume as well as semantic and temporal heterogeneity. Both aspects can drastically reduce usability in visual presentation and exploration, leading to

- a decline in rendering performance;
- the effect of *geometric point clutter* (fig.1-left), i.e., a very dense or even overlapping display of markers or symbols resulting in unsatisfactory visual appearance and affecting recognition and exploration performance;



Figure 1. Left: Geometric point clutter (McDonald's stores in Hamburg. Right: Semantic point clutter (hotel prices in Hamburg); source: Google Maps)

• the effect of limited overview and slow readability due to too many different (or even non-comparable) values or categories (hereafter referred to as *semantic point clutter*, fig.1-right).

Typically, generalization techniques are applied in order to overcome these clutter problems. While these processes improve legibility, they might also disturb relevant change information. In order to describe effects of the generalization on legibility and preservation of change information, constraints and related measures are introduced (Stoter et al., 2014). They can be used to assess the (possibly contradictory) effects of (different) generalization workflows with respect to a required change analysis task.

So far, constraints were only defined for static displays. In light of the aforementioned relevance of multi-temporal data, the aim of this contribution is to outline a conceptual framework consisting of constraints related to given types of change point analysis. These constraints are not just designed for description and comparison purposes, but also for triggering generalization processes as such in a later stage.

The remainder of this paper is organized as follows: Section 2 describes related work on point generalization, especially the constraint-based approach. The general framework for visual change point detection and analysis is set out in section 3, while section 4 gives a detailed description of selected preservation and legibility constraints as a function of change types. Finally, section 5 gives a brief discussion and outline of future work.

# 2. Previous work

Due to geometric or semantic point clutter frames, i.e. individual images of the entire animation sequence, have to undergo generalization processes. In cartography, different sets of fundamental generalization operations have been defined by various authors (e.g., McMaster and Shea, 1992; Hake et al., 2002; Slocum et al., 2009). From an information visualization perspective, Ellis and Dix (2007) proposed a taxonomy of techniques for the purpose of point clutter reduction. Korpi and Ahonen-Rainio (2013) presented a synthesis of cartographic and information visualization operations and demonstrated a significant overlap. With the focus on the problem of point clutter, the following generalization operations (or combinations of these) are of major importance: aggregation, simplification, selection, displacement, spatial distortion, data classification.

These generalization operations form the building blocks for an overall workflow. Rather than looking at a holistic approach or isolated generalization operations only, this contribution focuses on optimizing generalization workflows designed for specific visual change detection and analysis tasks (such as detecting loss of objects or existence of hot spots). Modelling and optimizing these workflows is often done using a constraint-based approach. Beard (1999) classified constraints (i.e., requirements that shall be fulfilled) into aspects related to position, topology, shape, structural, functional, and legibility. For practical and evaluation purposes measures for these constraints have to be introduced. Stoter et al. (2014) and Zhang (2012) provided comprehensive reviews of measures. Mackaness and Ruas (2007) grouped these into internal vs. external as well as micro, meso vs. macro categories.

Another grouping relates to the purpose of the underlying task: First, despite the abstraction, the generalized data still has to inherit the existing change information – leading to the need of *preservation constraints*. These are often described in terms of geometrical object-specific measures (e.g., change in line lengths before and after generalization). Second, change information must be readable by users (*legibility constraints*). Stigmar and

Harrie (2011) developed analytical measures for topographical maps. Based on an empirical study, best correspondence between perceived legibility and calculated measures could be observed for the following criteria: number of vertices, object line length, local density, proximity indicator, and degree of overlap. Sadahiro et al. (1997) used local density and relative local density as measures for describing the perception of clusters. Rosenholtz et al. (2007) proposed feature congestion as measure for visual clutter in order to describe user perception of map complexity. It is obvious that not only there is a relationship between perception and preservation of geometrical, semantical or topological properties, but also a close connection to Gestalt theory (laws of proximity, similarity, closeness, common regions, synchronicity, etc.).

Focusing on the visual interpretation of point patterns, Ellis and Dix (2007) presented a systematic analysis of cluster reduction techniques, resulting in the definition of eight high-level clutter reduction preservation constraints (e.g., avoiding overlap, keeping spatial information, or enabling localization). Korpi and Ahonen-Rainio (2013) developed their synthesized set of criteria and applied it to an example. Qian et al. (2006) introduced a general strategy of an automatic cartographic generalization chain that also considers the quality of point clusters. However, an objective comparison between methods or even an improvement due to changed parameter settings is not documented.

Despite of these single developments, complex measures for evaluating synoptic (high-level) interpretation tasks based on generalized visualizations are still difficult to define (Stoter et al., 2014). Harrie and Weibel (2007) argued that "the main limitation of constraint based techniques is the limitations of the constraints themselves", especially constraints for groups of objects.

Looking at techniques for implementing constraint-based generalizations and optimizations, there is considerable work on optimizing single generalization methods (Sester, 2000; Ware and Jones, 1998; Burghardt, 2000). For controlling complex processes, Harrie and Weibel (2007) argued that agent-based modelling is the most powerful modelling method in terms of applicability. Ruas and Duchéne (2007) described this technique in detail. Harrie and Weibel (2007) also noted that the success of this strategy relies on the complete and precise definition of constraints and the proper formalization of plans to guide the agent-based process.

All in all, so far developing constraints and implementing constraint-based generalizations refer to single and static views. Consequently, the motivation for this paper is to extend the previous work to multi-temporal views in order to optimize visual change point analysis in generalized views.

# 3. General framework considerations

First, the task of multi-temporal change detection and analysis is reduced to a bi-temporal setting. Changes are



Figure 2. Overall concept of bi-temporal change detection based on original data (top) and generalized data (bottom)

explored by comparing two subsequent frames (F1, F2). In the following, the difference frame (DF) is considered a good approximation for the mental representation of change information (figure 2, top row). At this point, we hypothesize that a multi-temporal change analysis can be understood as a series of bi-temporal comparisons. For example, change results over time can be stored as an evolution function (Bard, 2004).

Due to geometric or semantic point clutter, these frames have to undergo generalization processes. As mentioned before, various operations (such as aggregation, simplification, selection, displacement, spatial distortion, data classification) and workflows (e.g., agent-based modeling of constraints) have already been proposed for this purpose. However, for developing the conceptual framework within this paper it is not crucial which specific operations and workflows are actually used. At this point, just the existence of generalized frames (GF1, GF2) – and thus the generalized difference frame (GDF) – is of interest (fig. 2, bottom row).

The key question now is to what extent were specific change information preserved in the generalized difference frame (GDF) compared to the original difference frame (DF)?

In order to evaluate the usability and to compare with other generalization procedures, several preservation constraints can be introduced (section 4.1). Beyond this, legibility constraints must to be used to assess readability of changes (section 4.2).

The definition and selection of both types of constraints depend on the actual type of change analysis. In the following, the general typology of spatial changes described by Andrienko et al. (2003) will be used to structure related exploration constraints. This typology consists of the following change types:

- *Existential changes* describe whether objects are kept unchanged, deleted or added between two given frames. It is assumed that neither semantic nor spatial properties of points change.
- *Changes of semantic properties* describe the change in attribute values (being either classified

or metric) between two given frames. It is assumed that neither spatial properties nor the existence of points change.

• *Changes of spatial properties* describe the change in position, elevation, size, shape, orientation, volume etc. between two given frames. It is assumed that neither semantic properties nor the existence of points change.

Obviously, a more detailed classification of changes could be applied (e.g., by also considering abrupt vs. gradual changes). For the purpose of this framework, however, the general categorization seems sufficient.

# 4. Constraints

As mentioned before, the following constraints, describing either preservation (Section 4.1) or legibility (4.2) of change information, rely on the model of a bi-temporal visualization, which can be extended to a multi-temporal model by storing the sequence of bi-temporal differences.

# 4.1 Preservation constraints

#### 4.1.1 Existential changes

Given two frames of two dates (F1, F2), displayed points may represent the presence of any qualitative or quantitative feature. If *existential changes* are of interest, the difference frame (DF) stores the three possible levels of description (fig. 3, top row): point unchanged between two dates (1/1), point deleted (1/0), or point added (0/1).

Due to geometric point clutter, these frames need to undergo generalization (e.g. by aggregation). The generalized difference frame (GDF) is computed from the generalized views (i.e., GF1 - GF2).

At this point it may happen that, for example, from F1 to GF1 an aggregation must be applied due to a large point density. Since some of the original points are deleted in F2, DF2 no longer requires aggregation. Of course, the difference frame GDF now has to show both change cases (fig. 3; points marked with dashed ellipse). This can lead



Figure 3. Example of existential change detection based on bi-temporal scenes (top: original data, bottom: generalized data)

to new point clutter if there are points in the vicinity of different change classes.

As mentioned earlier, the question now is to what extent were change patterns preserved in the generalized difference layer (GDF) compared to the original difference layer (DF)? In order to evaluate the fitness-for-use several measures can be introduced:

- Preservation of global relative frequency of specific change levels: This measure that does not take space into account calculates the ratio of the relative number of specific change levels (i.e., 1/0, 0/1, and 1/1) before and after generalization.
  Example (refer to fig. 3): In the case of aggregation, which is performed separately for each change level, this measure should amount to 1. In the example the relative number of deleted points (1/0) in DF amounts to 7/20 and 7/20 in GDF, thus the ratio is 1.0. If points disappear due to generalization (e.g., through simplification or selection), this measure will become smaller or larger than 1.
- Preservation of local density of specific change levels: This measure calculates the density of each change level before and after generalization. For this, one counts the number of points within a given cell of a certain tessellation (such as regular grid, Voronoi) in DF and GDF, and calculates the ratio of these numbers. Not only in the case of removal operations, but also in the case of aggregation, densities might change due to a shift of the marker cluster symbol into a another cell.
  Example (refer to fig. 3): Taking the upper right cell (grey background) into account, local

density of deleted points is increased from 3/6 in DF to 4/7 in GDF due to the placement of the cluster marker symbol.

• Preservation of local clusters of specific change types: The comparison of clusters of change levels between DF and GDF implies taking into account the position and shape of clusters. This is a fairly complex task that can be solved in principle (Grubesic et al., 2014). However, this holistic approach is rather time consuming, contradicting the ultimate purpose of rapid assessment and triggering of generalization operations in multi-temporal displays. Previous work on static displays also avoided the holistic approach and used other measures as substitutes (such as local density; Stigmar and Harrie, 2011).

### 4.1.2 Changes of semantic properties

Given two frames of two dates (F1, F2), displayed points represent attribute values given at categorical or cardinal scale. The difference frame (DF) stores either possible value differences (with many levels of description) or just the result of a simple comparison (with three levels: increase, decrease, constant). The latter case is treated like the aforementioned case of existential changes, which also shows three levels.

In order to tackle geometric and/or semantic point clutter, generalized frames GF1 and GF2 and the respective difference frame GDF are created (fig. 4). Again, the key question is – now related to attribute value changes – to what extent were change patterns preserved in the generalized difference frame (GDF) compared to the original difference layer (DF)?

<sup>29</sup>th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan. This contribution underwent single-blind peer review based on submitted abstracts. https://doi.org/10.5194/ica-proc-2-114-2019 | © Authors 2019. CC BY 4.0 License.



Figure 4. Example of existential change detection based on bi-temporal scenes (top: original data, bottom: generalized data)

In the case of *classified data* the preservation measures for existential changes as outlined before can be applied; except that now more than three levels of description are observed (in the case of n classes, a maximum of  $n^2$ -1 levels is possible; the subtrahend ensures neglecting the change of "no value" to "no value").

For *metric* data (with a theoretical infinite number of levels) other measures need to be defined, for example:

• *Preservation of global central tendency of differences* (i.e., preserving global mean or median difference values)

 $\bigcirc$  Example (refer to figure 4): Mean difference values remain unchanged between DF (1/6) and GDF (1/6), if one takes class difference values into account for GDF.

• *Preservation of global variation of differences* (i.e., preserving standard deviation of difference values)

⇒ Example (refer to figure 4): Standard deviation of difference values is reduced from 1.17 (in DF) to 0.75 (in GDF) – a typical effect due to combining original values to classes.

• *Preservation of global extreme difference values* (i.e., the largest maximum and minimum differences in the entire data set)

 $\bigcirc$  Example (refer to figure 4): The global maximum difference value in DF (+2; marked with an arrow) is not clearly detectable in GDF where two points shows the same maximum value (+1).

• *Preservation of local extreme difference values* (i.e., differences that are surrounded by smaller or larger differences only in a specified neighborhood).

➡ Example (refer to figure 4): The local maximum difference value (i.e., a difference that is surrounded by smaller differences only) is lost after generalization (in this example global and local maximum values are identical).

• Preservation of local hot/cold spots of difference values (which can be determined, for example, with the *Getis-Ord Index*; Getis and Ord, 1992).

Data classification is an important generalizing method to reduce semantic cluttering. Schiewe (2018) describes further constraints (such as preservation of within-class or between-class homogeneity) for this case.

#### 4.1.3 Changes of spatial properties

Changes of spatial properties describe the change in position, elevation, size, shape, orientation, volume etc. between two given dates. In the following, the focus will be on positional changes only (i.e., the displacement of points).

Given two frames of two dates (F1, F2), displayed points represent any values given at categorical or cardinal scale. The difference frame (DF) stores either displacement values on a metric scale or just a simple comparison (with two levels: displacement, no displacement). The latter case can be treated like the aforementioned case of existential changes.

Again, generalized frames GF1 and GF2 and the respective difference frame GDF are created in order to tackle geometric and/or semantic point clutter. The key question is – now related to displacement values – to what extent were change patterns preserved in the generalized difference frame (GDF) compared to the original difference frame (DF)?

If displacement values are classified into *distance classes*, the aforementioned preservation measures for existential changes can be applied (see also 4.1.2).

If *metric displacement values* are taken into account the constraint measures described for semantic changes can be simply be transferred as follows:

- *Preservation of global central tendency of displacement* (i.e., preserving global mean or median displacement)
- Preservation of global variation of displacement (i.e., preserving standard deviation of displacement)
- Preservation of global extreme value of displacement
- Preservation of local extreme values of displacement
- Preservation of local hot/cold spots of displacement

#### 4.2 Legibility constraints

As stated in section 2, legibility measures have been defined by a couple of authors. These measures that refer to single and static layers and describe the degree of geometric point clutter (such as the degree of overlap of points; Stigmar and Harrie, 2011) can also be transferred to single frames of multi-temporal visualizations.

In addition, aspects of *cognitive overload* originating from the sequential display of scenes should be considered. One option is to determine the relative number of changed points compared to the total number of points within each difference frame. Introducing also the different levels of descriptions, one can derive an entropy measure. However, this does not consider two contradicting aspects: On the one hand, it is difficult to find a very small number of changes in a layer consisting of many points (i.e., overload of non-changed points). On the other hand, it is also challenging to detect a large number changes during the short occurrence of a frame (i.e., overload of changed points). The latter problem should be even further differentiated as exploration becomes more difficult the more levels of change descriptions occur (e.g., detecting existential changes 1/0 and 0/1 between two frames is more challenging compared to detecting 1/0 changes only).

Obviously, a differentiation and further description of overload constraints is necessary, including empirical user tests designed to describe the overload thresholds for these conflicting constraint measures as a function of the display time for each frame.

# 5. Discussion and further work

The aim of this paper was to outline a general framework for enhanced visual change point analysis in generalized multi-temporal displays. Since this framework has only a conceptual character so far, further studies must be carried out.

There are two basic assumptions for this framework (section 3):

- First, the difference frame (DF) is used as an approximation for the mental representation of change information.
- Second, a multi-temporal change analysis can be treated as a series of bi-temporal comparisons.

Of course, empirical studies have to confirm these hypotheses and adjust constraints accordingly.

With regard to the central elements of the framework, the list of constraints and relates measures makes no claim to completeness yet. On the other hand it must be reduced to an operational minimum. In the case of preservation constraints, this will be done primarily by analytical testing of various data sets and applying correlation or principal component analysis. In terms of legibility constraints, empirical studies have to be carried out in order to derive the best match between perceived legibility and calculated measures. One important and interesting aspect relates to the cognitive overload constraints that require adaptive legibility constraints as a function of the number of frames a display time for each frame.

As already mentioned, the constraints are not only intended for comparison purposes, but - in a later stage also for triggering the generalization processes as such. This could lead to a different way of thinking: In order to



Figure 5. Concept of simultaneously generalizing frames of bi-temporal data set

consider preservation and legibility in the difference frames, the optimization of generalization should work on the difference frame rather than on single frames (fig. 5). This again leads to the requirement of respective implementation and testing.

# 6. References

- Andrienko, N., Andrienko, G. and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. Journal of Visual Languages and Computing & Computing. 14: 503–541.
- Bard, S. (2004). Quality Assessment of Cartographic Generalisation, Transactions in GIS. 8(1): 63-81.
- Beard, K. (1999). Constraints on rule formation. In: Buttenfield, B. and McMaster, R. (Eds.): Map Generalization: Making Rules for Knowledge Representation, London: Longman: 121–135.
- Burghardt, D. (2000). Controlled Line Smoothing by Snakes. Geoinformatica. 9(3): 237-252.
- Ellis, G. and Dix, A. (2007). A Taxonomy of Clutter Reduction for Information Visualisation. IEEE transactions on visualization and computer graphics. 13. 1216-23. DOI: 10.1109/TVCG.2007.70535.
- Getis, A. and Ord, J.K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis. 24: 189-206.
- Grubesic, T.H., Wei, R. and Murray, A.T. (2014). Spatial Clustering Overview and Comparison: Accuracy, Sensitivity, and Computational Expense, Annals of the Association of American Geographers, 104:6, 1134-1156, DOI: 10.1080/00045608.2014.958389
- Hake, G., Grünreich, D. and Meng, L. (2002). Kartographie. Verlag De Gruyter, 8th edition.
- Harrie, L. and Weibel, R. (2007). Modelling the Overall Process of Generalisation. In: W. Mackaness et al. (Eds.): Generalisation of Geographic Information: Cartographic Modelling and Applications, Elsevier: 67-87.
- Huang, H. and Gartner, G. (2012). A Technical Survey on Decluttering of Icons in Online Map-based Mashups. In M.P. Peterson (Ed.): Online Maps with APIs and WebServices, Springer: 157-175, DOI: 10.1007/978-3-642-27485-5 11.
- Korpi, J. and Ahonen-Rainio, P. (2013). Clutter Reduction Methods for Point Symbols in Map Mashups. The Cartographical Journal, 50(3): 257-265.
- Mackaness, W.A. and Ruas, A. (2007). Evaluation in the map generalisation process. In: Mackaness, W.A. et al. (Eds.): Generalisation of Geographic Information: Cartographic Modelling and Applications; Series of International Cartographic Association, Elsevier Science: Amsterdam, The Netherlands: 89–111.
- McMaster, R.B. and Shea, K.S. (1992). Generalization in Digital Cartography, Washington, DC: Association of American Geographers.
- Qian, H., Meng, L., Wu, F. and Wang, J. (2006). The generalization of point clusters and its quality assessment

based on a polarization approach. Mapping and Image Sciences. 4: 44-63.

- Rosenholtz, R., Yuanzhen, L., and Nakano, L. (2007). Measuring visual clutter. Journal of Vision. 7: 1–22. DOI:10.1167/7.2.17
- Ruas, A. and Duchêne, C. (2007). A prototype generalisation system based on the multi-agent system paradigm. In: Mackaness, W.A. et al. (Eds.): Generalisation of Geographic Information: Cartographic Modelling and Applications; Series of International Cartographic Association, Elsevier Science: Amsterdam, The Netherlands: 269–284.
- Sadahiro, Y. (1997). Cluster perception in the distribution of point objects. Cartographica. 34(1): 49-61.
- Schiewe, J. (2018). Task-Oriented Visualization Approaches for Landscape and Urban Change Analysis. ISPRS International Journal of Geo-Information, 7(8): 288-303.
- Sester, M. (2000). Generalization Based on Least-squares Adjustment. International Archives of Photogrammetry and Remote Sensing. 33(B4): 931-938.
- Slocum, T.A., McMaster, R.B., Kessler, F.C. and Howard, H.H. (2009). Thematic Cartography and Geovisualization, 3rd edition. Prentice Hall.
- Stigmar H. and Harrie L. (2011). Evaluation of analytical methods to describe map readability. The Cartographic Journal. 48(1):41–53.
- Stoter, J., Zhang, X., Stigmar, H. and Harrie, L. (2014). Evaluation in Generalisation. In: Burghardt D., Duchêne C., Mackaness W. (Ed.): Abstracting Geographic Information in a Data Rich World. Lecture Notes in Geoinformation and Cartography. Springer, Cham: 258-298.
- Ware J.M. and C.B. Jones (1998). Conflict Reduction in Map Generalisation Using Iterative Improvement. GeoInformatica. 2(4): 383-407.
- Zhang, X. (2012). Automated evaluation of generalised topographic maps. PhD dissertation, ITC, University of Twente.

<sup>29</sup>th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan. This contribution underwent single-blind peer review based on submitted abstracts. https://doi.org/10.5194/ica-proc-2-114-2019 | © Authors 2019. CC BY 4.0 License.