The bounds of distortion: truth, meaning and efficacy in digital geographic representation

Citation for published version:
https://doi.org/10.1080/23729333.2017.1301348

Digital Object Identifier (DOI):
10.1080/23729333.2017.1301348

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
International Journal of Cartography

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The bounds of distortion: truth, meaning and efficacy in digital geographic representation

Lucas Godfrey¹ and William Mackaness¹

¹ The School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom

Abstract: Even maps that strive for a precise representation of spatial relationships use techniques of distortion to embed a geographic extent within a two dimensional plane – be it a page or a screen. The graphical design of geospatial information does however conform to a consensus around the conceptual limits of that distortion, with an overall design framework that constrains the design to ensure comprehension and the effective recognition of geographic entities and relations. Although constraints are necessary, we argue that the advent of digital technology, particularly in mobile mapping, warrants re-examination of the parameters of these distortions. Here we introduce the concept of ‘the bounds of distortion’ as a device for considering the conceptual boundaries of map design, and as a foundation for further work to investigate how these bounds may be redefined to better support map users with more effective graphical information. The focus here is navigational information, and in particular urban navigation and interaction with the graphical representation of urban geography and public transport networks.

Keywords: Spatial cognition, Wayfinding, Navigation, Mobile cartography, Map effectiveness, Cartosemiotics, Multiscale mapping, Focus maps

1. Distortion and the graphical communication of geographic space

While in many domains of knowledge and practice, distortion is seen as something counter-productive, as something to be avoided, distortion is in fact central to the effective visual communication of information. As Barbara Tversky observed: “We change the truth to tell a bigger truth…” (Tversky, 2015). Faced with either a two or three dimensional space, with limited area or volume, and a fundamental constraint of human-scale, we must make decisions as to how, through a process of abstraction and simplification, we can effectively convey the essential qualities of underlying geographic phenomena. Diagrams allow us to convey topological, temporal and spatial relationships with an immediacy that is difficult and often impossible with words alone; while cartography has formalised the use of two dimensional space for the purpose of communicating caricature and geographical relationships, with Robinson reflecting on the ambition to reduce the spatial
characteristics of a large area so as to “bring things into view…” (Robinson, 1995). Despite its negative connotations, distortion is a necessary part of the process of bringing ‘things into view’, reflected in the choice of projection and map generalisation methodologies in order to support communication of spatial relationships in ‘large-scale’ or ‘transperceptual’ space (Downs et al., 1977; Freundschuh et al., 1997).

The conceptual limits to these boundaries of distortion vary with map type and anticipated audience (for example varying according to whether the map is schematic, thematic, topological or topographic). In this paper we introduce the concept of the bounds of distortion, and consider five key approaches that have been explored in terms of seeking to redefine these bounds, before proposing opportunities for future work. The central aim is to consider how the graphical display of geospatial relationships may benefit from a reconsideration of the assumptions that underlie cartographic design. This work is particularly concerned with the efficacy of navigational information in the context of the small screen devices that have become the dominant platform of interaction for most map users.

It should be highlighted that our concern is not maps that ‘look’ distorted, but rather, we are concerned with cartographic practice that makes use of techniques which elevate the needs and capacities of the map user above the goal of displaying absolute spatial relations. Here we argue for an imperative to develop cartographic approaches which draw less sharp a distinction between metric and non-metric spaces, and linked to this, draw less distinction between the structural and functional representation of geography. This line of investigation is closely linked with the transition from ink to data and the practical implications of using ‘smart’ mobile devices as our primary means of accessing spatial information.

2. Foundations of digital geographic representation

We feel it is necessary to begin with a brief consideration of some key technical and conceptual foundations which underpin cartographic practice and are important in the framing of our arguments.

2.1 Space, maps and mappings

The most primitive spatial structure is the topological relation. Topology can describe structures in \( n \) dimensional space as the fundamental properties of topological relations can be seen as being independent of an ‘embedding space’. Euclidean space provides an embedding space for either two or three dimensional structures, with positional information being lost in the process of transforming from a Euclidean space to one that is purely topological. This issue of data loss is common to all representational processes in which the phenomenon is represented in an alternate, often lower dimensional space, moving from the large-scale to the
small-scale. Throughout the cartographic process, scale and level of detail act to govern the structure of the conceptual space in which the geographic entities and their relations are displayed. While there have been many approaches proposed to representing relative spatial relations, the prevailing map design paradigm is significantly biased toward a formalisation of the physical world that describes an absolute space – a container within which the entities lie.

2.2 Truth and meaning in cartography

Distortion is both real and imagined. It is ‘imagined’ in the sense that people do not ‘see’ maps completely objectively but are influenced by a number of cognitive quirks that lead to a distorted understanding even if the distortion is not present in the map itself. A simple example is our tendency to see vertical lines as being shorter than horizontal lines of the same length (Held, 1971). Another, more far reaching example is the issue that perception does not happen in isolation but is a process that occurs in parallel with [mental] projection, which has a ‘looser coupling’ to stimuli and allows us to ‘see’ what could be there, rather than what is actually presented to us (Kirsh, 2009).

Beyond these issues of veracity however, is the communication and perception of meaning, something that is often of far greater significance to spatial decision making. From the perspective of semiotic analysis, the meaning of a map is a function of the ‘triadic structure’, in which the underlying geography is the ‘object’, the map is the ‘sign’, and the map user, the ‘interpreter’. Within this construct the map is usually considered to be an ‘icon’, or a sign that is considered to display the quality of ‘likeness’ to the object. While cartosemiotics supports a more detailed analysis of meaning within the context of geospatial information (Schlichtmann, 2009), the key overall issue here is recognition. Whatever form a map takes, the interpreter must be able to recognise which underlying object a graphical feature represents. If two buildings are generalised into a single form at lower levels of detail, this is acceptable if the map user recognises that they are now looking at a generalised urban area rather than a single large building. These transition points at which we find ourselves looking at a fundamentally different phenomenon remain a challenge for automated map generalisation. Thus we might argue that this is a ‘semantic distortion’. Similarly, there are occasions where a distortion of the angle or length of a connecting edge is necessary in order to convey the essential character of the geography being visualised. For example, we seek to preserve the gently flexing path of the railway line because this is a defining property that differentiates it from other phenomena. We suffer other distortions in order to preserve the conveyance of this distinguishing property. Overall then, we see trade-offs among different distortions in order to convey various characteristic properties and ultimately, to communicate meaning to the user.
3. A consensus on the conceptual limits of distortion in map design

While digital, and particularly mobile devices, have changed how we interact with geographic information, the underlying design framework remains little changed. The consensus around map design, that was reached during the development of paper-based maps, still dictates the fundamental approach to geographic representation, despite the transition from ink to data.

While approaches to generalisation specify the exact process by which spatial data is displayed on a device, our concern here is the aspects of map design that may support wider levels of distortion in cartographic representation. So beyond, say, the specific algorithm used to determine the method of displaying a road’s width, what is the more fundamental approach to distortion that is common to all modern maps? In our analysis of this issue, we will distinguish between the representation of metric and non-metric spaces, and the associated approach to distortion in each.

In a geographical context, metric space describes a space in which the Euclidean distance between all geographic objects is defined by a distance function on the overall set. This provides a consistent mapping from reality to the conceptual model, and on to the target structure. A topographic map would be one such example. Non-metric space in a geographical context includes maps that relax this approach, usually with the purpose of highlighting the functional, as opposed to structural character of the geography. In other words, maps that do not adhere to metric spatial relations make this departure with a view to increasing the saliency of key elements of the overall spatial structure that help to ‘specify action’ (Klippel et al., 2005). When considering the bounds of distortion, it is necessary to distinguish between these two high level cases, and also to consider how the two are integrated.

3.1 The bounds of distortion in metric geographic representation

The core design convention in this case is that a metric space defines the absolute location of all elements of the overall set. A key property of this approach is that changes of scale do not affect relative distances. Limits to visual acuity require cartographic distortions via the process of generalisation (e.g. exaggerated road widths and simplification of geometric form). Ultimately scale (or level of detail) governs the limit to these graphical treatments; map real estate constraints mean that as we zoom out, each phenomenon will reach a ‘conceptual cusp’ (Muller 1991, Mackaness et al., 2005) or ‘generalisation point’ (Ratajski, 1967) – a point at which it is no longer feasible or meaningful to visualise that phenomenon. This links with the notion of ‘map capacity’ (Ratajski, 1967), whereby map content is optimised for the given area relative to scale and thematic focus. This reflects the hierarchical view of geography that is formalised in the data models which support geographic representation within metric spaces.
3.2 The bounds of distortion in non-metric geographic representation

In non-metric space, the design convention is focused around schematisation, often with topological relations maintained. Within this conceptual framework, cardinal bearing is maintained where possible, but is not strict. From a semiotic perspective, there is a consistent graphical treatment of iconic elements with emphasis on network structure and the broader overarching caricature of the phenomena being represented. In transportation themed maps, this approach supports a functional view of the geography, focused on supporting the user in a route orientation strategy (Skagerlund et al., 2012). The challenge becomes one of supporting both ‘turn-by-turn’ approaches to conceptualising space together with a survey strategy, in which an overview of the geography is comprehended. The failure to convey the latter is why people arrive at a destination but don’t know how they got there.

Integration of the two approaches – metric and non-metric, is often achieved by means of nesting one within the other, with a hard boundary between the two. In the case of integrating the two (i.e. on the same ‘page’), the bounds of distortion are defined as the area contained within the ‘box’ corresponding with the area ‘underneath’ the box, to the extent that coherence is maintained between the content ‘inside’ and ‘outside’. A variety of graphical techniques are used to both warn the user of these two different spaces and of the connections between them. For example, using continuity principles to show how different graphical elements in each representation are recognised as being the same road.

4. Five design approaches that explore the bounds of distortion

As part of our investigation into how distortion in cartographic practice may be reframed to focus more on the contextual needs of the user, we reflect on five approaches.

4.1 Hard Boundary

The simplest approach is that of a ‘Hard Boundary’ - an approach that supports the display of two or more spatial constructs within a single ‘page’. Usually this approach is used to display a metric and non-metric space in such a way as a user may be able to quickly switch between the different views and orientate themselves in either, relative to the other. The classic example of this Hard Boundary between two views is a hub and spoke or ‘spider’ map, in which a higher level of detail is shown near the centre of the extent, with highly schematised routes that cover a larger area being displayed around the topographically correct region. Figure 1, a Lothian Buses map of the bus network in the City of Edinburgh, is an example of this. In the example in Figure 1, the main central street and the area immediately surrounding it are shown in a higher level of detail, with the city’s bus
routes shown in terms of their entry and exit points in the context of this central hub area. Despite the fact that schematisation of the larger scale map has been designed specifically to provide consistency with the roads in the smaller scale view, this design still fits within the Hard Boundary approach, with a clear boundary separating the two different representations of space.

4.2 Global Transformation

Metric space in the context of cartography and GIS is synonymous with a transformation that preserves certain spatial aspects of the underlying geography – for example minimising the distortion of angles between locations (conformal). This is however a specific and narrow interpretation of metric space, with a metric space more generally being any space where a distance function is applied to all members of the set. Global Transformation is an approach that can be used to reinterpret the representation of the entire extent to communicate a derivative property of the underlying spatial structure. It refers then to metric and metric-like approaches which do not conform to traditional representational objectives. While there is no upper limit on the variety of applications of this approach, it is most often used as a way of representing spatial relationships in terms of time geography (Hägerstrand, 1985). A map in this case, is used to communicate time taken to travel as a derivative of the lived experience of the geography, as opposed to the absolute length of the distances within the given scale. Simple applications of this approach include the emphasis of denser urban areas which typically take longer to traverse than more suburban areas, despite the absolute distances being shorter.
An interesting example of this approach is Tom Carden’s ‘Travel Time Tube Map’, that applies a temporal transformation to the schematic representation of the London Underground train network. The nature of this network is such that overall distance, even in a highly schematised map design, is a poor indicator of the actual time needed for a journey; despite its experimental nature, this is a very interesting concept when considering the efficacy of spatial information.

Another example of Global Transformation is Benedikt Gross’s ‘Metrography – London Tube Map to large scale collective mental map’ (Figure 3). It uses the structure of a non-metric space to transform the metric representation of the city.

![Travel Time Tube Map](carden2011)

This is not a map we could use for day-to-day wayfinding, but we argue it is a mapping of people’s sense of place insofar that millions of Londoner’s organise their thinking around the distorted space of the London Underground schematic.

Such a representation reminds us of Anchor Point Hypothesis posited by Couclelis et al. (1987). The Anchor-Point Hypothesis describes the phenomena in which people’s cognitive conception of space is distorted by their variable level of familiarity and association with certain parts of a geographical environment, and that this leads to a unique, personal understanding of the structure of that environment. Technology now enables the modelling of an individual’s familiarity and associations thus enabling cognitive and behavioural aspects to take a more central role in the design of personalised maps.

This idea has been further explored through Gross’s ‘MapMap Vauxhall’ project, in which Gross created spatial transformations of a region of London based on research participant’s mental conception of the area expressed through the recording of individual sketch maps (Figure 4). While this process was manual rather than automated, it is nevertheless an interesting example of a Global
Transformation methodology that redefines the bounds of distortion in order to fit more closely with an individual’s natural perception of space.

Fig 3. Metrography - London tube map to large scale collective mental map (Gross, 2012)

Fig 4. MapMap Vauxhall - Mashup Mental Maps and OpenStreetMap (Gross, [online])

4.3 Focus Maps

The third approach is through the use of ‘focus maps’, with early investigations into the efficacy of maps on mobile devices using this approach (Zipf et al., 2002). Focus maps use computational and graphical techniques to draw the user’s attention to certain aspects of the geography.
Harrie et al.’s variable-scale mapping for small scale cartography is an early example of integrating multiple scales into single map views with mobile devices in mind, however a high level of visible distortion is apparent in this approach. ‘Focus, Glue and Context’, is a term coined by Yamamoto et al. (2009). Instead of the distortion being spread across the central region (Figure 6), it refers to the idea of containing the distortion in a narrow band around the focused part of the map. This ring is the ‘glue’ between the focus area and the surrounding context. Furnas noted that: “The fundamental motivation of a fisheye strategy is to provide a balance of local detail and global context...” (Furnas, 1986). Focus, Glue and Context however supports a much broader range of objectives by allowing for an intermediate area of transition between a focal region and its surrounding.

Focus, Glue and Context can be seen as a way of integrating different spatial representations (different levels of detail) in scenarios where it is also necessary to maintain a low level of schematisation. In other words, map content that describes context is not lost in the process of focusing on a specific region or geographic entity. A user may find fish eye distortion disconcerting. By containing the distortion in the ring (the focus, glue, context approach), it makes the distortion less apparent – conforming more to expected norms.

A mixing of these ideas is reflected in Haunert et al.’s use of focus in the context of representing road networks. Here we see a more subtle use of Focus, Glue and Context that supports map legibility through the use of ‘locally valid scale factors’
4.4 Variable Scale Route

The fourth approach is one pioneered by Agrawala and Stotle (2001). The Variable Scale Route approach makes use of highly schematised representations to emphasise journey segments that are best displayed with varying levels of detail. The LineDrive project adopted this idea for long car journeys. The issue was long journey segments (often with the least need for guidance), consumed the largest portion of space on the device’s screen. The solution was to reduce the space devoted to long journey segments in order to provide more space for detailed segments. Here the distortion varies as a function of the complexity of decision making.

As with other approaches to the automation of schematic maps for mobile devices (e.g. Anand et al., 2004, Anand et al., 2006), maps that use the Variable Scale Route approach conceal the nature of this distortion by suppressing structural and contextualising information that would otherwise expose the distortion - the focus being on the conveyance of functional information almost exclusively (Klippel et al., 2005).
4.5 Variable Level of Detail

The prevailing design paradigm for digital maps makes use of variable levels of detail in so far as the level of detail is changed after each generalisation point is crossed (Ratajski, 1967), i.e. a transition across ‘conceptual cusps’ based on either zooming in or zooming out. What is outside of the current bounds of distortion however is a variable level of detail within each scale, based on the user task. Thus Variable Level of Detail in the context of redefining the bounds of distortion can be seen as another form of focus map, with functional information given precedence but without completely suppressing the broader structural content of the map, as with the Variable Scale Route. In the Chorematic Focus Map method (Klippel at al., 2004) and other approaches within this category, the objective is to tease out the salient information but to leave an intermediary level of detail that supports navigation and goes beyond a simple A-B route representation.

In Mackaness et al. (2011), multimodal journeys that cross transport networks are shown with varying levels of detail dependent on sub task. For example a higher level of detail would be shown where the task is on foot as opposed to via public transport (see Figure 9.). In a similar approach, Schmid et al. (2010) proposed emphasising detail at the start and end of the journey (as is often required in prototypical journeys), however this was proposed within the context of ‘route aware’ maps that embed emphasised route information within the broader structural map content, which can from there be further simplified to provide map views that convey ‘disambiguated’ start and end environments (as in illustration ‘c’ in Figure 10).
5. Key opportunities for future work

By considering the strengths and weaknesses of these five design approaches we propose a research agenda that seeks to improve map efficacy through a reworking of these bounds of distortion in the context of digital and personalised mapping.
5.1 Integrating structural and functional information: variable forms of spatial representation in single map views

A key area that is yet to be effectively designed and implemented is the presentation of multiple forms of space in single map views. The aim here would be to develop Global Transformation operations such that we can derive and visualise the optimal form of visualisation for each stage of the journey – i.e. for each ‘sub-task’ (Mackaness et al., 2011). While Schmid et al.’s route-aware maps go some way to achieving this, route-aware maps only address Variable Level of Detail, and do not allow for the integration of varying underlying spatial representations. While Variable Scale Route maps can integrate different forms of space, they do this at the expense of geographical context, so the ambition here is to find an approach to automation that effectively satisfies both route and survey orientation strategies – i.e. provides the structural content while also showing the spatiality of the actions that are needed to complete the task, but in a simplified form.

5.2 The application of non-Euclidean spatial concepts

Mixing metric and non-metric structures in single map views requires an altogether different conceptualisation of space, and we suggest the exploration of non-Euclidean approaches to formalising space is another area for further investigation. To give an example: the bounds of distortion in contemporary mapping products is closely tied to the constraints of a two dimensional Euclidean plane, however the reality of the underlying geography and our personal experience of it, is closer to a manifold in a higher dimensional space. So a framework for specifying this manifold such that it could be queried to produce a spatial structure that could be embedded in the two dimensions of the screen would be an example of an area for further research.

5.3 Developing and validating mobile solutions

As illustrated in this paper, there is a historic body of research inspired by the imperative to more effectively display geospatial information on small screen devices. The number of research projects that have explored potential approaches to distortion in this context, that actually included the development and validation of mobile prototypes with users, is very low, and is highlighted as a key gap for further work.

6. Concluding comments

We have introduced the bounds of distortion as a conceptual device that supports the analysis of established map design, with a particular aim to lay the foundation for further work to redefine these bounds in the context of the efficacy of maps for navigation. While truth is a key issue in cartographic representation, we have argued
that the recognition of meaning is of greater influence on the effectiveness of navigational information, and that making better use of the limited space on the screens of our ubiquitous mobile devices means that there is considerable potential value in extending the work already done to redefine the way distortion is applied in cartographic design. We reviewed five high level approaches, and have highlighted their strengths and weaknesses in the context of users’ core orientation strategies.

To date generalisation techniques have focused on algorithms that respond to legibility constraints within narrow changes of scale, with a focus on topographic mapping (with emphasis on locational accuracy). We would propose a broader and more ambitious remit that considers wider scale transitions across a broader range of thematic and schematic maps. And even beyond this, a remit that explores a significant loosening of the boundaries of distortion, one that facilitates the optimal embedding of multi scale spaces, and ultimately one that chimes more closely our conceptualisation of, and familiarity with place.

Acknowledgements

We would like to thank the Ordnance Survey and the United Kingdom Engineering and Physical Sciences Research Council for their support.

References


Gross, B. (2012). Metrography – London Tube Map to large scale collective mental map [online] Available at: http://scale-collectibenedikt-gross.de/log/2012/02/metrography-london-tube-map-to-large-scale-collective-mental-map


Tversky, B. (2015). How Diagrams Communicate. ISVIS [online] Available at: https://www.youtube.com/watch?v=B_X0PGgJjTs
