

Multivariate Mapping in High Quality Atlases

Huber Stefan, Sieber René, Ruegsegger Marianne, and Hurni Lorenz

Institute of Cartography

ETH Zurich, Switzerland

e-mail: {huber, sieber}@karto.baug.ethz.ch

Summary

Thematic interactive atlases are evolving into spatial knowledge systems managing huge amounts of thematic and topographical data for visualization. In fact, there are several complementary ways to visualize and interact with the richness of data. Multivariate mapping is – besides animation – one of the most promising techniques. Starting with a sound discussion of visual variables occurring in a map, a new generic approach of multivariate mapping in high quality atlases is introduced. It supports the consistent handling of different map layer types (point, line, area, raster, chart, etc.) and multiple visual variables. This leads to a non-restrictive technical solution with a consistent implementation and application, equipped with a standardized interface for all kinds of map layer types. The concept of multivariate mapping is implemented in the Atlas of Switzerland 3, which is currently under development; different types of map layer supporting multivariate mapping are presented.

Introduction

Thematic maps are multivariate per se; two dimensions of the plane and at least one visual variable encode information. There are different but complementary ways to map a larger number of data variables: the use of several visual variables on a single graphical mark, single view superimpositions of data variables [DiBiase et. al 1994], small multiples of maps, shown side by side [Bertin 1977, Tufte 2001], and charts. Traditional printed data maps – especially atlases [e.g., Atlas de la Suisse 1965-1997, Atlas DDR 1976-1981] – tend to show as much data as possible on one map. This approach mainly focuses on map reading, less attention has been paid to map analysis and interpretation [terms after Muehrcke and Muehrcke 1998]. In interactive multimedia cartography the number of possible individual maps based on a data table is huge; visual variables can be combined in different ways. Therefore different aspects can be shown on different maps as well. The main focus of this paper is on parameterizing and applying multiple visual variables in a multimedia cartography environment. We illustrate the concepts by the current prototype of version 3 of the digital Atlas of Switzerland project [Sieber and Huber 2007].

Visual variables

Maps are visual structures; thematic maps use visual variables to visualize information. A sound knowledge of the concept of visual variables introduced by Bertin [1967] is indispensable to implement interactive multimedia cartography environments. The discussion leads to a definition of multivariate maps based on visual variables.

Visual structures use graphical marks and the marks' visual variables to encode information. Visual structures are transformed (scaling, clipping, projection onto the display, etc.) to visualize information. Bertin [1967] distinguishes three elementary types of marks: Points (0D), lines (1D), and areas (2D); surfaces (2.5D) and volumes (3D) have been added by other authors. Statistical surfaces may be interpreted as representing two spatial dimensions combined with a size attribute rather than three spatial dimensions [DiBiase et al. 1994]. Thematic raster data form a regular grid; grid cells are treated mostly as areas.

The graphic sign system of Bertin [1967/1977] introduces eight visual variables, grouped in variables of the image (ordering variables) and differential variables. For Bertin, ordering variables are: 2 dimensions of the plane (2 dimensions du plan), size (taille), and value (valeur). Differential variables are: texture (grains), color (couleur), orientation (orientation), and shape (forme). The terms brightness [e.g., Wilkinson 2005] and lightness [e.g., Slocum et al. 2005] refer to value. Bertin is primarily interested in a 2D-color space, holding saturation (chroma) constant at the highest level. Although often used the term texture is ambiguous and thus rather diffuse. Wilkinson uses granularity – Bertin's original grain – for texture Slocum et al. use spacing instead.

Several authors proposed extensions to Bertin's visual variables. Spiess [1970] mentions that the meaning of orientation and shape differs for points from the meaning for lines and areas. For points, he suggests the differentiation between inner and outer orientation and shape. Morrison [1974, 1984] proposes color intensity (saturation) as additional ordering visual variable. In addition Morrison adds pattern arrangement to the list of visual variables. McCleary [1983] suggests irregular patterns; the visual variable arrangement [Slocum et al. 2005] refers to McCleary's suggestion. A more complex ordering system for texture is discussed in Caivano [1990, 1994]. Caivano [1991] introduces *Cesia*, an invented word that includes visual signs such as brilliance, glossiness, specular reflectiveness, transparency, etc. In Wilkinson [2005] transparency is one of the 11 major visual variables. MacEachren [1992] introduces focus as visual variable to represent uncertainty. Later MacEachren [1995] revises focus and proposes a tripartite approach called clarity. Clarity depends on the visual variables crispness, resolution, and transparency. Wilkinson's [2005]

visual variable blur refers to focus. More extensions have been proposed by other authors, e.g., perspective height [Slocum et al. 2005].

Comparing two current categorizations of visual variables for point, line, and area [Slocum et al. 2005, based on MacEachren 1994 and Wilkinson 2005], there is consent regarding the color variables (ignoring terminological differences). However, categorization of texture related variables and form variables for points vary. Wilkinson isolates almost all texture related variables while Slocum et al. retain Bertin's approach. Wilkinson's texture pattern combines Slocum's texture related variables size, shape, and arrangement. Andrienko and Andrienko [2006] separate Wilkinson's approach for texture related visual variables and distinguish five texture related visual variables: arrangement, density (granularity), size, shape, and orientation.

A generic approach to applying visual variables [after Wilkinson 2005] in a digital environment requires among other things that:

- Each visual variable addresses a distinct graphic feature: Exterior form and interior texture pattern have to be separated.
- A visual variable is capable of representing both continuous and categorical data.
- A minimum number of different values of a visual variable should be distinguishable.
- When two visual variables are combined, they should remain visually separable.

In his classification, Wilkinson distinguishes form (including *size*, *shape*, and *rotation*), color (*brightness*, *hue*, *saturation*), texture (*granularity*, *pattern*, *orientation*), and optics (*blur*, *transparency*). This approach is directly applicable to computer cartography.

For cartographic purposes location (position in the plane) is inherent as well as shape and rotation for lines and areas, and size for areas in 2D space. In 3D space statistical surfaces may depend on area size. But there is no rule without exception: Cartograms that warp space to visualize quantitative data use location, size and – depending on the type of cartogram – shape to encode information. Bertin [1967] divides visual variables into two groups: (1) dimensions and (2) variables “of the third dimension”, often called retinal variables. In cartography, dimensions are used to place marks, while retinal variables define the appearance of marks. Concerning visual variables, (1) a map is defined as *visually* multivariate if more than one retinal variable is used to visualize information and (2) a map is defined multivariate *in terms of data input* if more than one data variable is visualized.

A Generic Multivariate Approach for Atlas Cartography

In high quality multimedia atlases, there are several complementary ways to deal with visual variables. Data variables can be visualized in a strictly analytical sense, applying a single visual variable to one specific map symbolization type. Mapping population data is a good example: The number of people is usually simply depicted by means of the variable size, using a single shape and a single color, typically resulting in proportional dot maps. These rather simple maps ignore the combinative possibilities that retinal variables offer. They leave space to visualize related data, using a combination of retinal variables and interactive techniques, e.g., juxtaposition (split screen), superimposition (overlay), or adjustable transparency.

While simple maps are usually easy to grasp, cognitive research indicates that people are most attracted by visualizations of a well-balanced medium complexity [Tufté 2001, Dansereau 2005]. To transfer these findings to a multimedia atlas environment, mapmakers should use multivariate symbolization techniques more frequently.

The Adaptive Map Concept [Sieber and Huber 2007] already supports various facets of multimedia cartography, such as map layer management, multivariate symbolization, automatic adaptive zooming, and self-adapting legends and analysis tools. This concept for high-quality atlases is now extended by a new approach of *Generic Multivariate Mapping*. The term “generic” refers to the technical configuration and graphical visualization by means of data and a set of instructional parameter. The intentions of this retinal-variable-based approach are manifold:

- Minimize programming and production efforts, by generating different maps from the same database.
- Ensure a consistent map description language that adapts to various types of maps.
- Facilitate a flexible map design.
- Allow for the exploration of mapped data variables with interactive tools.

The Generic Multivariate Mapping approach handles different *map layer types* (point, line, area, raster, chart, etc. for 2D and 2.5D maps) and *multiple visual variables*. A combination of heterogeneous map layer types share an almost homogeneous set of functionality. This allows e.g., to overlay data variables in a flexible way.

In a highly interactive atlas environment a project-specific selection and aggregation of these variables is recommended. Each map layer type has its particular set of *aggregated visual variables*. An aggregated visual variable is defined to base on one or more visual variables and to provide an interface with parameters and project specific defaults. Each aggregated visual variable conveys a distinct graphic feature to be mapped (one degree of

freedom). Hence, the number of aggregated variables determines the degree of freedom of the marks. The state of an aggregated visual variable is either constant or variable (data driven). The redundant use of visual variables is possible and leads to visually multivariate maps.

The symbolization criteria for each type of map layer are merely based on generation and visualization issues: a specific layer type should allow for every cartographic variation. For example, common chart types [see Schnabel 2007] as pie, wing, or bar chart are representations of interval graphs or small multiples of interval graphs [Wilkinson 2005], e.g., population pyramid can be understood as dual bar chart. Likewise, the construction of e.g., a pie and a divided area chart or a wing and a bar chart is nearly the same: Although the graphical appearance differ, only the coordinate system changes from polar to rectangular.

The multivariate symbolization process applies on two levels: First, multiple visual variables (shape, size, rotation, brightness, etc.) can be applied to marks like colored circles or rotated rectangles. To improve overall map readability an attribute can be assigned to more than one visual variable, as it is the case with proportional dot symbols in the Atlas of Switzerland 2 [2004], where data variables are mapped with a redundant combination of size and color. On a second level, charts – categorized as interval graphs or small multiples of intervals – containing different thematic variables, like e.g., gender (split up in male/female) or age classes, are built up by a combination of marks with their visual variables.

In general, this multivariate approach treats each visual variable equally. Consequently, all visual variables work on all measurement scales. For each visual variable there is an extended set of parameters including useful project-related cartographic and aesthetic default values. In a specific atlas project, the selection of these multiple visual variables is often steered by design and usability considerations.

This generic multivariate approach leads to an open and non-restrictive technical solution for many kinds of different map layers. The symbolization can be configured with a flexible parameterization, based on a standardized XML-interface.

Implementation

The design, development and implementation of the Atlas of Switzerland 3 is an iterative and incremental process. A multitude of maps, available data, various degrees of user interaction, and technical constrains, etc., lead to the current expandable implementation. The base elements of the atlas are the maps [fig. 1]. Each map consists of several layers

that are loosely coupled. The layer's task is on the one hand to link data tables, graphical data, and visual variables, and on the other hand to visualize itself in the map and to create and manage its interactive tools. An XML-based interface called Map Description File serves as container for the first task and holds all the visual variable related settings. Great importance has been attached to a clearly arranged and concise XML-language in order to facilitate reassigning visual variables or data. A capable mathematical expression parser evaluates mathematical expressions that may contain variables and functions and manages data access accordingly.

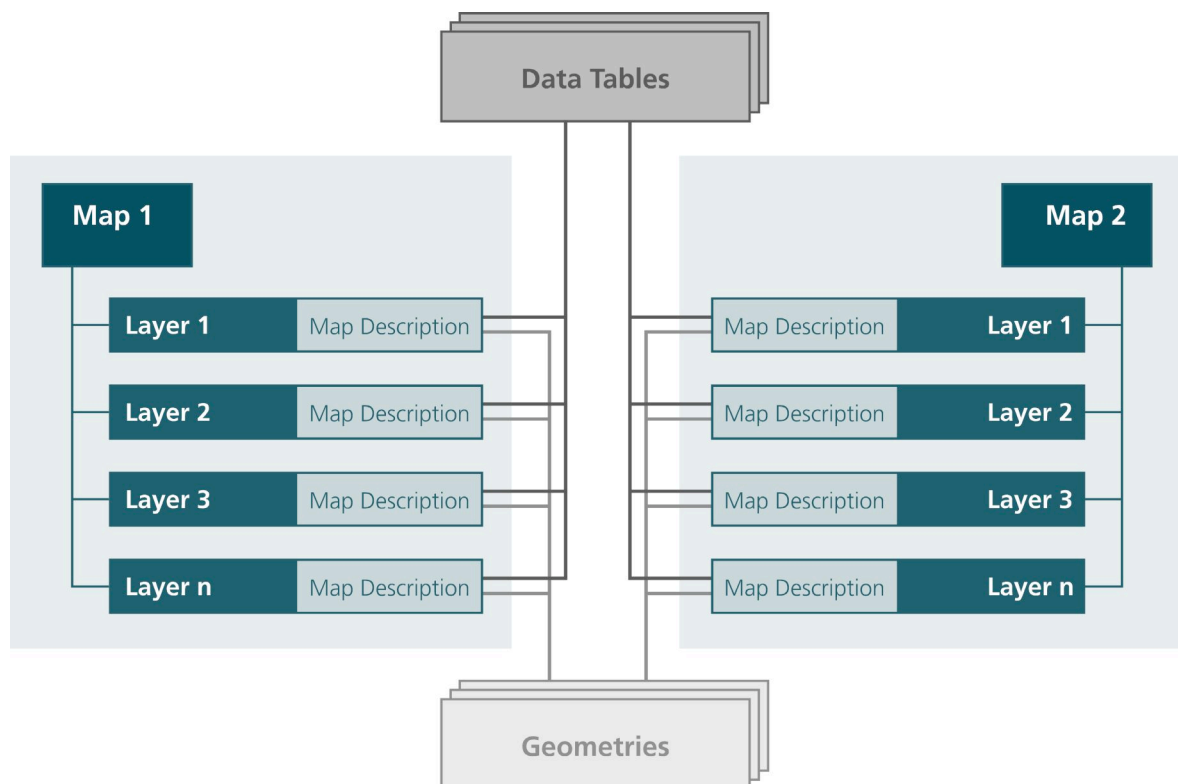


Fig. 1 Atlas of Switzerland 3: Linking map layers and data

We classify layers by their types of marks: an *area* layer contains polygons for e.g., choropleths, a *line* layer contains polylines for e.g., isopleths, and networks, and a *point* layer contains points for symbols or charts. For technical reasons, a *raster* layer is a separate type. In practice this classification helps to organize the map vertically and to support comprehensible interactive tools.

The software engine of the Atlas of Switzerland 3 currently supports several aggregated visual variables [tab. 1]. The benefit of aggregated visual variables will be illustrated by the aggregated variable color/transparency. Brightness is the only independent component of color. For instance a change of fully saturated hue results in a change of brightness. Color schemes used in cartography often affect more than one color component [see e.g., Brewer et al. 2003]. And transparency heavily interacts with color on a higher level. In the

digital realm, color/transparency can be simply set by color schemes with an alpha channel. With two color/transparency variables more ambitious colorization is possible [fig. 2a]. The number of available aggregated variables will be limited by restrictions of the layer type, e.g., shape of areas and lines are inherent. The layer types area and point support outlines: Regarding visual variables, outlines will be treated as lines of the layer type line.

Map Layer Type	Aggregated Visual Variables in the Atlas of Switzerland 3				
	Color/Transparency 1 2	Shape	Size	Rotation	Texture (Pattern/Granularity)
Area	● ●		(● 3D space)		
Line	● ●		●		●
Point	● ●	●	●	●	
Chart	● ●		●		

Tab. 1 Atlas of Switzerland 3: Currently supported aggregated visual variables

A few samples from the Atlas of Switzerland illustrate the potential of the multivariate approach. Figure 2a shows a choropleth map that uses two color/transparency variables. Switzerland is a multilingual country but each language is clearly dominant in its own area. The prevailing language is linked to hue, the percentage of resident population with the prevailing language as main language (how dominant is the language) mainly to saturation. In figure 2b the voting results of the Alpine Transit Initiative are depicted, using dot size and hue to show the number of votes and percentage of Yes votes. Considering the choropleth background information of cumulative Yes votes of the Cantons (bipolar color scheme), it clearly shows communes with coherent (e.g., red dots on red area) and divergent voting results.

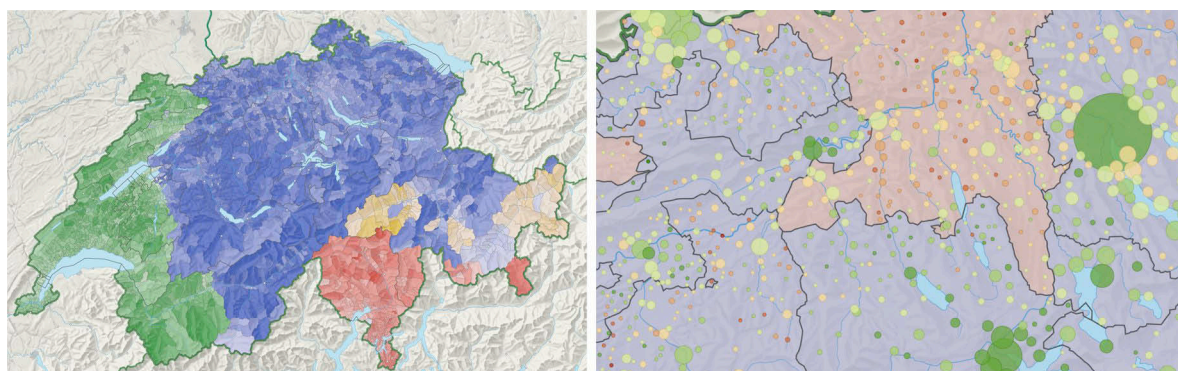


Fig. 2 a) Prevailing language (hue) and percentage of main language on resident population (saturation); b) Voting results (bipolar color scheme), number of voters (size) and overall results of the Swiss Cantons (hue)

Figure 3a shows the passenger traffic on the Swiss rail network. While size represents the total number of passengers, hue stands for single-track (orange) or double-track railways (red). Depicting commuter activity is a rather difficult task (fig. 3b). The example informs about the commuting process, where size is used to indicate the number of daily commuters and texture (dotted lines) and color to differ commuter directions.

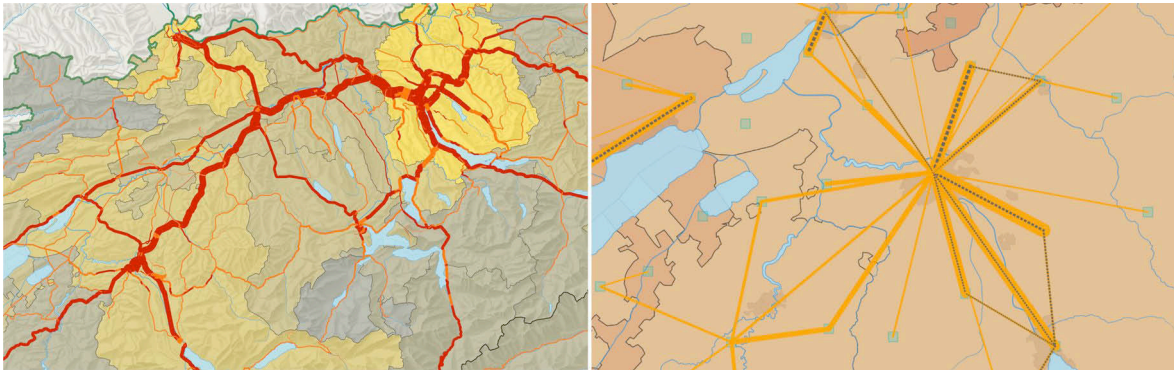


Fig. 3 a) Passenger traffic by rail (size) and number of tracks (hue); b) Commuting between districts: Number of commuters (size), direction (color, texture)

The symbols [fig. 4a] represent the types of airports using the visual variables hue, shape, and size. The rotation of the symbol corresponds to orientation of the main runway. Airspace classes are visualized by colored areas. Figure 4b presents a statistical 2.5D surface based on raster data. Size represents population density and color the atmospheric pollution by sulfur dioxide. The correlation seems to be evident.

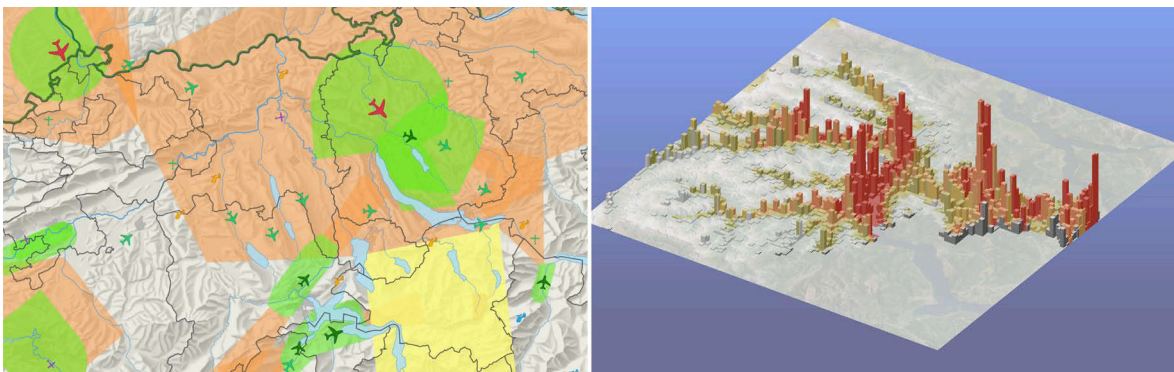


Fig. 4 a) Types of airports (hue, shape, size) and orientation of runway (rotation), and airspace classes; b) Population density (size) and SO₂ concentration (color)

The Atlas of Switzerland 3 currently supports several types of charts: bar chart, divergent bar chart, pie chart, divided pie sector, divided area chart, wing chart, and polar area chart. Net electricity production by fuel is represented by pie charts (size) [fig. 5a]. The choropleths emphasizes the percentage of nuclear energy already shown as red sectors (brightness). Figure 5b differentiates the labour market sectors (size) above and below

mean value (saturation). Using not only constant color per pie sector, the values above mean are recognizable at a glance.

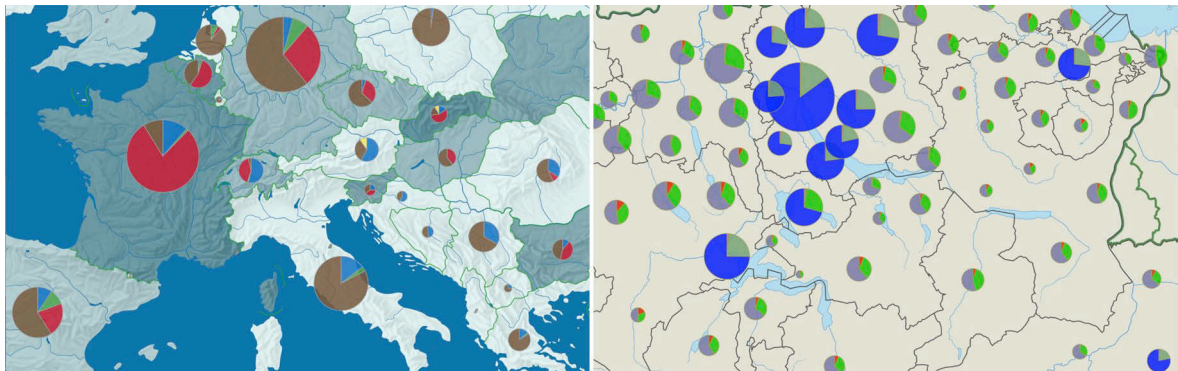


Fig. 5 a) Net electricity production by fuel (size); b) Labour market sectors, values below mean are dimmed (size, saturation)

Conclusions and Outlook

The generic multivariate approach presented shows one of the possibilities to solve the “triangle problem” between graphical data, statistical data and symbolization by a bottom-up, tree-like procedure. Because of its open structure, nearly any combination of visual variables and map layer types is theoretically possible. Practical experience with the prototype of the Atlas of Switzerland 3 – as proof of concept – shows the potential of the approach.

For future applications, the generic multivariate approach is very promising: it is transferable to the fields of animation and 3D maps and even combinations of time and space.

References

- Andrienko, N., and Andrienko, G. (2006): *Exploratory analysis of spatial and temporal data: a systematic approach*. Berlin: Springer.
- Atlas de la Suisse (1965-1997): *Atlas der Schweiz/Atlas de la Suisse/Atlante della Svizzera*. Wabern: Eidgenössische Landestopographie.
- Atlas of Switzerland 2 (2004) [DVD / 2 CD-ROM]: *Atlas der Schweiz/Atlas de la Suisse/Atlante della Svizzera/Atlas of Switzerland 2*. Wabern: Bundesamt für Landestopografie.
- Atlas DDR (1976-1981): *Atlas Deutsche Demokratische Republik*. Gotha/Leipzig: Haack.
- Bertin, J. (1967): *Sémiologie graphique: les diagrammes, les réseaux, les cartes*. Paris: Mouton. (English ed. 1983: *Semiology of graphics: diagrams, networks, maps*. Madison WI: University of Wisconsin Press.)
- Bertin, J. (1977): *La graphique et le traitement graphique de l'information*. Paris: Flammarion. (English ed. 1981: *Graphics and graphic information-processing*. Berlin: De Gruyter.)

- Brewer, C. A., Hatchard, G. W., and Harrower, M. A. (2003): ColorBrewer in print: a catalog of color schemes for maps. *Cartography and Geographic Information Science* 30(1), pp. 5-32.
- Caivano, J. L. (1990): Visual texture as a semiotic system. *Semiotica* 80(3/4), pp. 239-252.
- Caivano, J. L. (1991): Cesia: A system of visual signs complementing color. *Color Research and Application* 16(4), pp. 258-268.
- Caivano, J. L. (1994): Towards an order system for visual texture. *Languages of design* 2, pp. 59-84.
- Dansereau, D. F. (2005): Node-link mapping principles for visualizing knowledge and information. In: Tergan, S. O. & Keller, T. (eds.): *Knowledge and information visualization. Searching for synergies*. Berlin: Springer, pp. 61-81.
- DiBiase, D., Reeves, C., MacEachren, A.M., von Wyss, M., Krygier, J. B., Sloan, J. L., and Detweiller, M. C. (1994): Multivariate display of geographic data applications in earth system science. In: MacEachren, A. M. and Frasor Taylor, D. R. (eds.): *Visualization in modern cartography*. Oxford: Pergamon, pp. 287-312.
- MacEachren, A. M., (1992): Visualizing uncertain information. *Cartographic Perspectives* 13, pp. 10-19.
- MacEachren, A. M., (1994): *Some truth with maps: A primer on symbolization & design*. Washington DC: Association of American Geographers.
- MacEachren, A. M. (1995): *How maps work: representation, visualization, and design*. New York NY: Guilford.
- McCleary, G. F. (1983): An effective graphic "vocabulary". *IEEE Computer Graphics & Applications* 3(2), pp. 46-53.
- Morrison, J. L. (1974): A theoretical framework for cartographic generalization with emphasis on the process of symbolization. *International Yearbook of Cartography*, 14, pp. 115-127.
- Morrison, J. L. (1984): Applied cartographic communication: Map symbolization for atlases. *Cartographica* 21(1), pp. 44-84.
- Muehrcke P. C., and Muehrcke, J. O. (1998): *Map use: reading, analysis, and interpretation*. 4th ed. Madison WI: JP Publications.
- Schnabel, O. (2007): *Benutzerdefinierte Diagrammsignaturen in Karten: Konzepte, Formalisierung und Implementation*. Zürich: ETH (Diss. ETH)
- Sieber, R., and Huber, S. (2007): Atlas of Switzerland 2 – A highly interactive thematic national atlas. In: Cartwright, W., Peterson, M. P., and Gartner, G. (eds.): *Multimedia cartography*. 2nd ed. Berlin: Springer, pp. 161-182.
- Slocum, T. A., McMaster, R. B., Kessler, F. C., and Howard, H. H. (2005): *Thematic cartography and geographic visualization*. 2nd ed. Upper Saddle River NJ: Pearson-Prentice Hall.
- Spiess, E. (1970): Eigenschaften von Kombinationen graphischer Variablen. In: Arnberger, E. (ed.): *Grundsatzfragen der Kartographie*. Wien: Österreichische Geographische Gesellschaft.
- Tufte, E. R. (2001): *The visual display of quantitative information*. 2nd ed. Cheshire CT: Graphics Press.
- Wilkinson, L. (2005): *The grammar of graphics*. 2nd ed. New York NY: Springer.