Automatic selection of symbols for diagrams and choropleths in multiscale thematic mapping

Nina Yurova*, Timofey Samsonov* **

*Department of Cartography and Geoinformatics, Lomonosov MSU, Faculty of Geography, 1 Leninskiye Gory, Moscow, Russia, 119234.
** Delaunay Laboratory of Discrete and Computational Geometry, Yaroslavl State University, 14 Sovetskaya str., Yaroslavl, Russia, 150000.

Abstract. In this work the problem of the automated symbolization of multiscale thematic maps is considered. Authors offer solution for automated selection of diagram properties and formulas for transformation of their diameters. The rules of formalization graphic transformation illustrated on charts and choropleth were developed.

Keywords: multiscale maps, diagrams, choropleth maps

1. Introduction

Automation of generalization and symbolization of data on multiscale maps is one of the key issues of modern cartography. Automated generalisation has been touched in many fundamental studies (Buttenfield & McMaster 2001, Li 2007). However symbolization and graphics replacement in addition to generalization takes place during the scaling (Brewer, Buttenfield, 2007). Automation of this process is not developed enough. Usually it is necessary to symbolize every scale level individually, which is time-consuming task. Various diagrams are using to represent statistics with different levels of aggregation in online analytical processing (OLAP) systems (Stolte et al. 2003, Wang Chen Bu Yu 2010).

In this experiment we focused on symbolic transformations that are applied to charts and cartograms (choropleth maps) when scale is changing. Developed formal rules can be used for automatic translation of symbology through scale levels.

2. Methodology

Transformation of symbology in a transition from large scales to small consists of several operations: changing visualization method, changing symbols and graphical variables.

A set of formal rules for symbol transformation of diagrams and cartograms has been developed. These rules assume that several levels of detail (LoD) in database are available for every type of object and objects are linked through levels. This allows to keep track of how object geometry and attributes are changed during the transition to the next scale, and to select symbol transformation rule accordingly.

2.1. Diagrams

Diagrams are intensively used on socio-economic maps. Consider we have point objects that are symbolized using diagrams. If the map is zoomed out this type of symbol can be applied to areal territorial unit (i.e. transformation from municipality to province). On the next zooming territorial unit is switched to next superordinate hierarchical level. This process is considered as a set of separate formal transformations of symbology.

In this study we considered only the size of the diagram and not considered its inner structure (if it is a pie chart).

A good strategy is to develop a united classification that considers data values from all hierarchical levels. This step allows avoiding reclassification during scale change. For every level of detail an appropriate classes are simply selected. Classification can be made automatically (with standard classification algorithms such as Natural breaks or Quantile) or manually. In this case study we made manual classification based on traditional ranges of population.

Range of values on in every LoD is analyzed and appropriate function (linear, quadratic, cubic, logarithmic) is applied to derive diagram radius from the value (Acknowledgements

This work is supported by the Russian Government project 11.G34.31.0053 and Russian Foundation for Basic Research grant 12-05-31088-mol_a

). We used conditional scale with magnification factor set-to 1,2 - 1,6. User defines diameter for smallest class. Preferable scales of visualization for each hierarchical level are defined by analyzing statistics on minimum bounding rectangles (Freeman Shapira 1975).

Chart size varies both with the scale of resulting image, the range of values and the territorial units' size. Using of the scale-depended function to determine the diagram size ie appropriated when the territorial unit level is keeped constant. In this work fomula (1) is offered. The chart size index depends only on scales ratio. Resulting diagrams measurements are worked out (Table 1).

$$R_s = \frac{S_{k+1}}{S_k} \quad (1)$$

where R_s — scales ratio, S_{k+1} — next LoD scale, S_k — current scale.

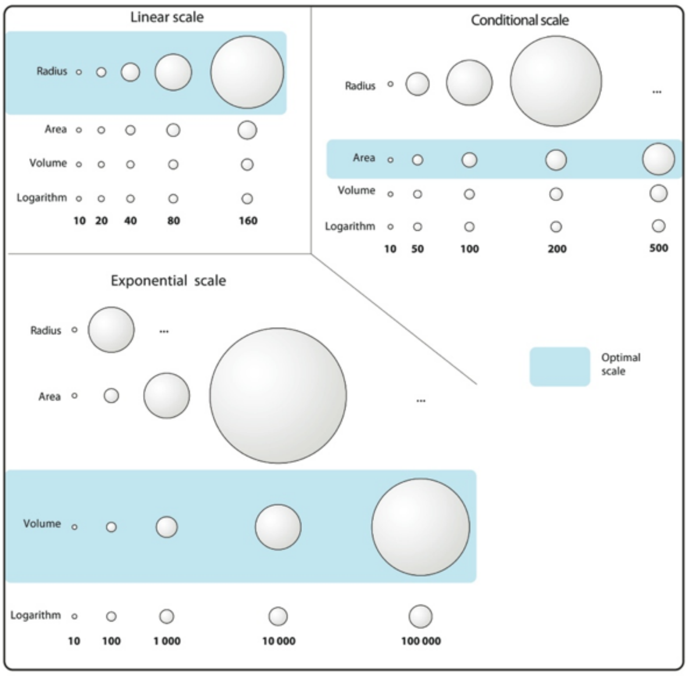
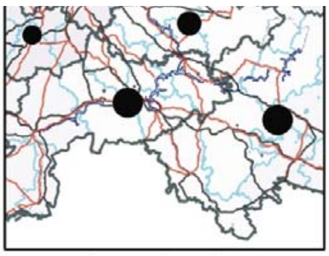
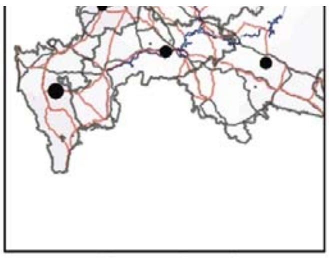


Figure 1. Different symbol scales

The range ... ɿes if ter-
ritorial un ... of abso-
lute indica ... l and ap-
propriate ... from the
value and ... ıulas (2)
and (3). T ... ıits' area
size relatio ... and val-
ue both. B ... ters were
obtained (



Economic regions
(scale 1:24 000 000)



Economic regions
(scale 1:48 000 000)

Figure **2**, (scale 1:24 000 000) (scale 1:48 000 000)

Figure **3**).

$$k = \frac{\sqrt{R_a}}{R_s} \quad (2)$$

$$k = \frac{\sqrt{R_v R_a}}{R_s} \quad (3),$$

where R_s — scales ratio , R_a – logarithmic areas ratio, R_v – logarithmic values ratio.

$$R_v = \log_n \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{j=1}^{m_i}\frac{V_i}{v_{ij}} \quad (4),$$

$$R_a = \log_n \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{j=1}^{m_i}\frac{A_i}{a_{ij}} \quad (5),$$

where n is number of superordinate units, mi is a number of current units inside i-th superordinate unit, A_i — area of i-th superordinate unit, V_i — value of $_i$-th superordinate unit, a_{ij} — area of j-th unit inside i-th superordinate unit, v_{ij} — value of j-th unit inside i-th superordinate unit.

Level	Municipalities		Provinces		Economic regions	
VALUE	1:1,5M	1:3M	1:6M	1:12M	1:24M	1:48M
<	12	6	4	2	1,30	0,65
10,0	16	8	5	3	2	1
50,0	20	10	7	4	2	1
100,0	26	13	9	5	3	1
500,0	34	17	12	6	4	2
1000,0	45	22	16	8	5	2
5000,0	58	29	20	10	6	3
10000,0	75	38	26	13	8	4
20000,0	98	49	34	17	11	5

>	127	64	45	22	14	7
Ratio	1,3	1,3	1,3	1,3	1,3	1,3

Table 1. Diagrams sizes (mm) by formula (2)

The methodology was assessed on population data from official Russian census (Figures 2-3). Formula 3 was considered to give better results, diameter is well adapted to scale and a gradual decrease of mean radius is achieved during scale transitions.

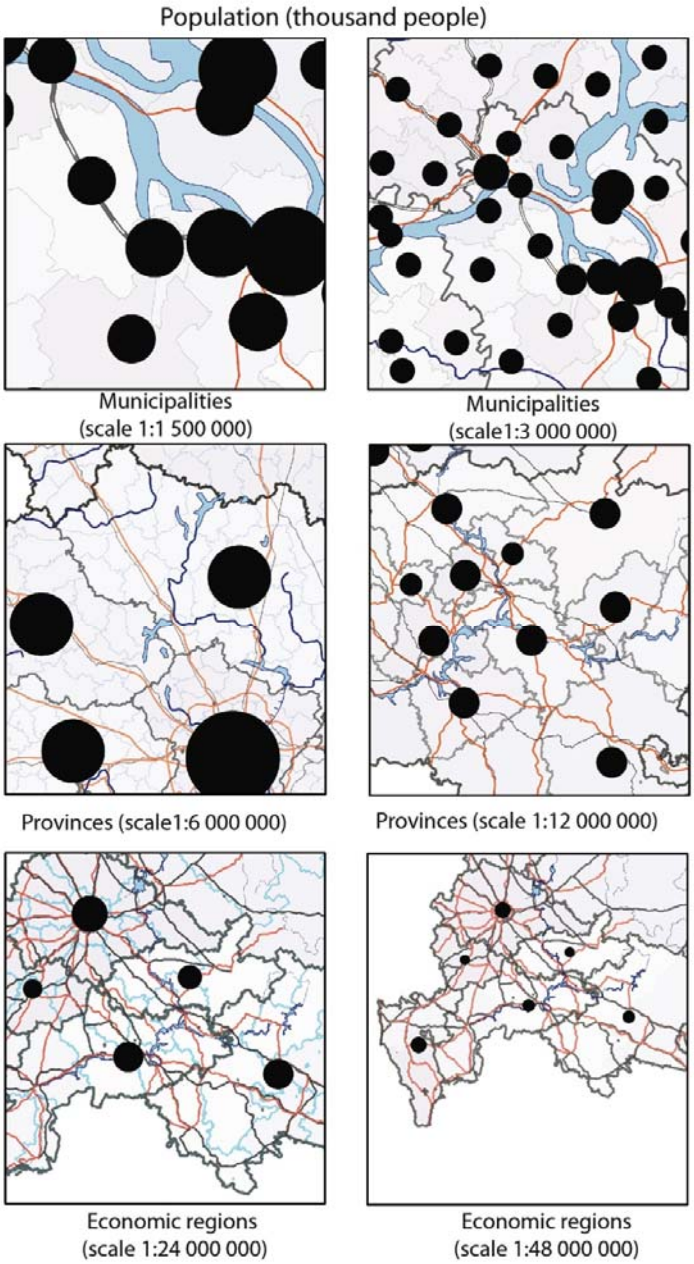


Figure 2. Fragments of the population map created using formula (2)

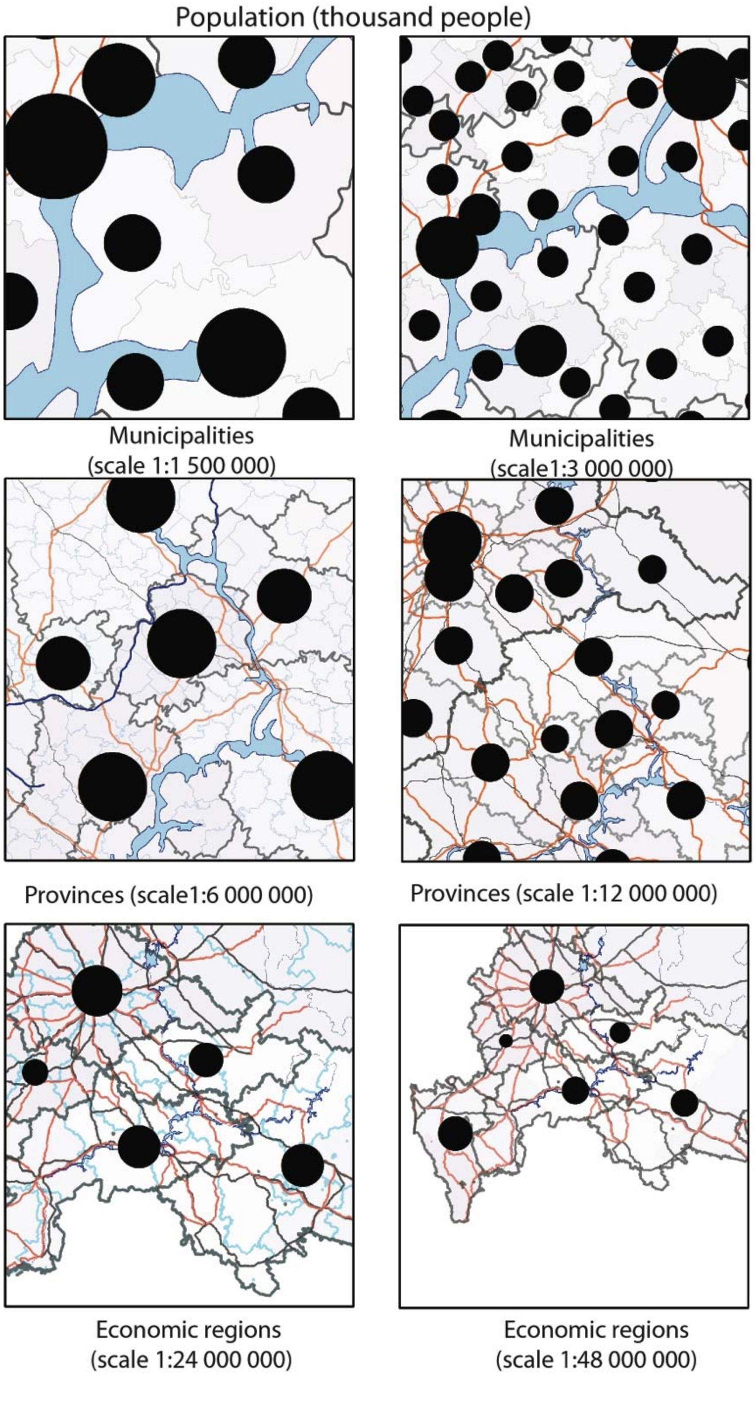


Figure 3. Fragments of the population map created using formula (3)

2.2. Choropleths

Choropleth maps are usually used for display of relative parameters per territorial units. All automatic classification methods were analyzed for suitability for using in mapping (Slocum, McMaster, Kessler, & Howard, 2008). In our case study (population density) quantile was selected as the most appropriative method for the data classification. The differences in relative indexes between territorial units of different hierarchical level are usually not so dramatic if compared to absolute values. In many cases the same classification and color scale can be used in all levels of detail, but some classes can be merged if they contain too little entities. We followed this strategy in our investigation.

More complicated part of the research was an attempt to derive a formal principle of color gradient construction, from which colors for classes should be selected. The analysis of histograms of an attribute for different LoDs of data allows creation of uniform color scale that covers distribution ranges for all levels of detail (Brewer, 2005). The minimum and maximum values and skewness are found for the entire data range including data from all LoDs. A zero skew value indicates that the values are relatively evenly distributed on both sides of the mean and data typically has a symmetric distribution (**Figure 4**).

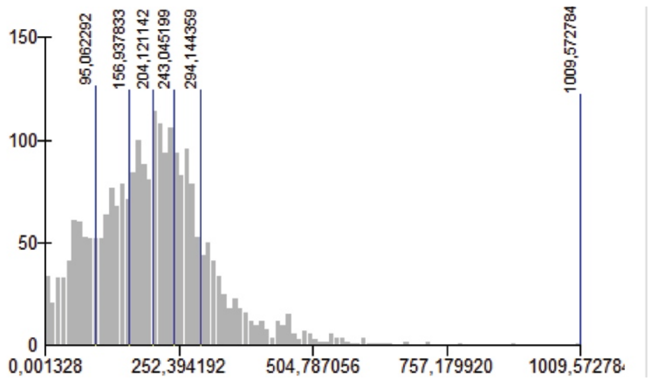


Figure 4. Symmetrical distribution of the river network density data

In this case standard color scheme is useful - representing ordered values by variation of hue and saturation (**Figure 5**).

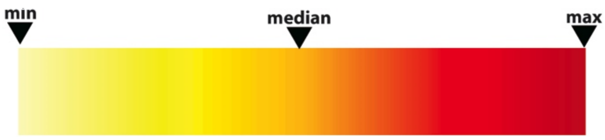


Figure 5. Color ramp using for display symmetrically distributed data

Positive or negative skew values indicate that the bulk of the values lie to the left or to the right of the mean (

Figure **6**). So the standard color schema is transformed to the multicolor one (Figure 7). This principle can be conceptualized as a color slider that drags away the color of maximum or minimum if the distribution is asymmetric. The color of maximum or minimum is replaced by a new one, because it is anomalous value.

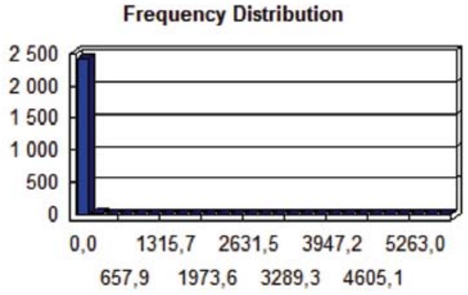


Figure 6. Asymmetrical distribution of data as an example of the population density in municipalities of Russia

Figure 7. Color ramp using for display asymmetrically distributed data

Using of the two and three key colors allows ranging data into more classes and focusing on the critical or anomalous values.

At reduction of scale there is a changeover to larger territorial units. Thus from a uniform color scale the interval corresponding to range of values at this level is selected (Figure 8).

3. Conclusion

The developed rules can be used in systems of drawing multiscale maps as follows: if the symbology for the most detailed mapping level is set, it can be translated to less detailed levels automatically. Formulas were derived for automatic translation of diagram radius between scales. A principle of color scale selection for choropleth map using color slider is introduced.

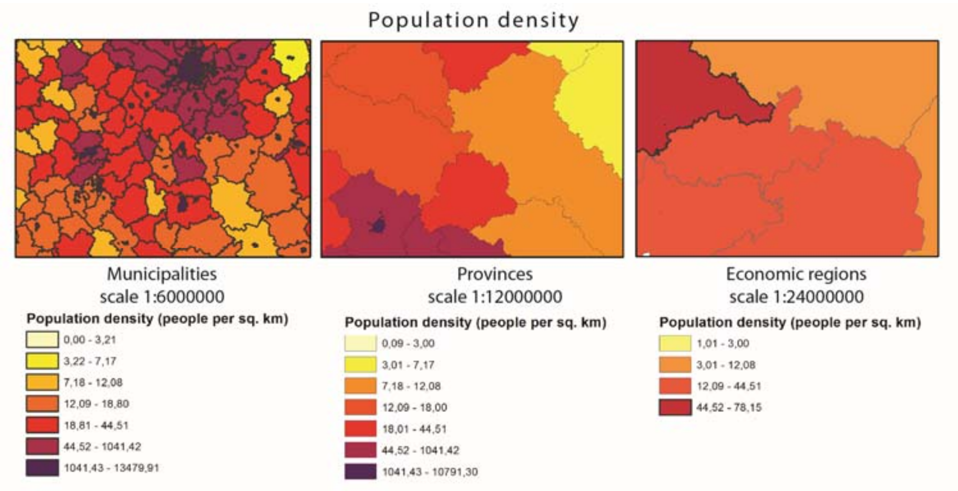


Figure 8. Choropleth map of population density

Acknowledgements

This work is supported by the Russian Government project 11.G34.31.0053 and Russian Foundation for Basic Research grant 12-05-31088-mol_a

References

Brewer CA (2005) Designing better maps. A guide for GIS users (First edit.). Redlands: ESRI Press, 2005.

Brewer CA, Buttenfield BP (2007) Framing Guidelines for Multi-Scale Map Design Using Databases at Multiple Resolutions // Cartography and Geographic Information Science, 2007, No34(1), pp. 3-15

Buttenfield BP and McMaster RB (1991) Eds., Map Generalization: Making Rules for Knowledge Representation, Longman Scientific and Technical, London, 1991.

Freeman H, Shapira R. (1975) Determining the minimum-area enclosing rectangle for arbitrary closed curve, Communication of ACM, 18, 409–413

Li Z. (2007) Algorithmic Foundation of multi-scale spatial representation. CRC Press, 310 p.

Slocum TA, McMaster RB, Kessler FC, & Howard HH (2008) Thematic Cartography and Geovisualization (3rd Editio., p. 576). New York: Prentice Hall

Stolte C, Tang D, and Hanrahan P. (2003) Multiscale visualization using data cubes. IEEE Trans. on Visualization and Computer Graphics, 2003, 9:176–187.

Wang B, Chen G, Bu J (2011) Multiscale visualization of relational databases using layered zoom trees and partial data cubes // Proceedings of IVAPP 2010, 11 p.