Map Classification according to their Visual Global Properties

Catherine Domingues*, Tahar Kharchi**

* IGN, COGIT Lab., 73 avenue de Paris, F-94160 Saint-Mandé
** UPV – IRD, GRED Lab., 911, avenue Agropolis, F-34394 Montpellier

Abstract. Making on demand maps is a problem which is shared by plenty of national mapping agencies. One strategy to help a map maker to make his/her map is based on analogy. This strategy requires a large example base. Thus, the point is to explore this map base in order to offer the map maker a reasonable number of relevant examples. In this paper, a survey which aims to: i) evaluate the visual global properties of topographic map examples and ii) make groups of maps which will be offered to mapmakers in the software interface is presented. The results i.e. the map groups qualified by their visual properties are shown.

Keywords: on-demand map, map visual property, survey.

1. Introduction

Making on demand maps is a problem which is shared by plenty of national mapping agencies or cartographic companies. Indeed, the French national mapping agency (IGN) offers a web service *Carte à la carte* (*On-demand map*) by which each Internet user can specify the scale, the orientation, the title, the illustration, the center of the map to obtain his/her own customized map.

A map can be seen as the result of the application of a legend to a geographic data set. Different strategies may be used to help a map-maker to make his/her map. For example, extracting color palettes from paintings and making different legend proposals from these palettes (Christophe & Ruas 2009). Another strategy[1] is based on analogy: instead of acquiring cartographic knowledge, the map-maker can select maps among different map examples. Secondly, the corresponding legends are applied to the map-maker's data to obtain maps which are supposed to be like the original examples. This strategy requires a large example base in order to give choices to the map-maker to better meet his/her needs. Nevertheless, the larger the example data base, the more difficult it is to show all the examples at the same time. Thus, the point of this proposal is to explore this map base in order to offer the map-maker a reasonable number of relevant examples.

Previous work (Domingues 2008) defines map visual global properties, some of which are:<*property*>: *realistic, precise, original, balanced*. These properties may provide a way to structure the example data base by defining map groups based on them. The software interface helps the map-maker by guiding him/her towards relevant map groups according to their properties. In this paper, the survey which aims to evaluate the visual global properties

[1] This is the strategy which has been implemented in the IGN *Carte à la carte*.

of topographic map examples is shown in part 2. The statistical analysis of the survey allows us to identify the most exaggerated maps according to each visual property; this is presented in part 3. Results are discussed in part 4 before conclusions and future perspectives.

2. The survey

The survey aims to evaluate the properties of topographic map examples and to make groups of maps based on these properties. The ability to have a higher response rate is one challenge of this survey: the more people participating, the more reliable the survey data are. Thus it was decided to use the World Wide Web to collect data regarding the appearance of the maps.

The advantages of an online survey include rapid access to numerous potential respondents and previously hidden populations, instant data collection (results are automatically stored in a database), immediate feedback, lower cost of construction and postage. The most important point is the similarity of experimental conditions, i.e. respondents are near or in the same conditions when commenting on maps (object of the survey) or when constructing them by use of a software, the maps are shown on the screen with varied and uncontrolled context conditions.

Despite everything, online survey displays some disadvantages. The limitation of this technology in this case can be summed up as: misunderstanding some questions by people and lack of seriousness when answering them. Also, map appearance depends on the screen size as well as on the screen calibration adjusted to each respondent.

2.1. Map description

Database maps are supposed to be examples of symbolization for map-makers. Indeed, the map-maker cannot imagine the visual properties of his/her map by merely looking at the map legend. Neither can he/she imagine them by looking at the application of the legend to example data. Thus, in order to minimize the discrepancies between the visual properties of the database examples and the map which is obtained by applying the example legend to the map-maker's data, map examples are classified according to the landscape they represent. The landscape is supposed to be described by two parameters: density and relief. Density may be: *urban* or *rural* and relief: *mountainous*, *plane* or *sea area*. Consequently, the map-maker is supposed to know which landscape is predominant in his/her map to visualize relevant examples. For instance, Figure 1 and Figure 2 show map samples with two different landscapes, respectively: *rural* and *mountainous area* and *urban* and *sea area*.

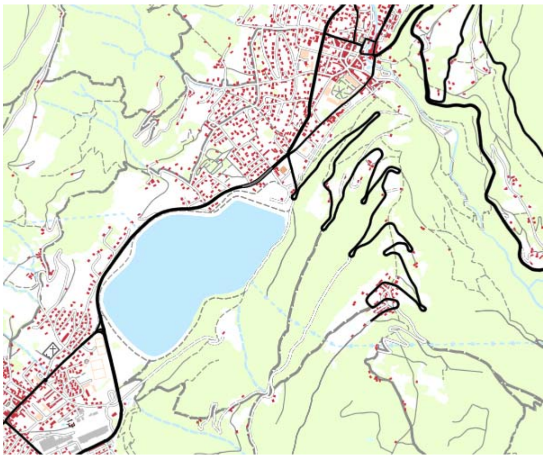

Figure 1. Faroe Islands-
density: *rural*, relief: *mountainous area*.

Figure 2. Ukraine,
density: *urban*, relief: *sea area*.

Legends have been made in the course of previous work. Several are deduced from the national mapping agencies' legends (Jolivet 2009): for example, the legend of the map in Figure 1 is a transposition of the Faroe Islands to IGN data. Others are made into colors extracted from qualified color palettes (Dominguès & Bucher 2006): for example, Figure 3 and Figure 4 show maps whose legend colors are respectively extracted from the *calm* and the *soft* color palette.

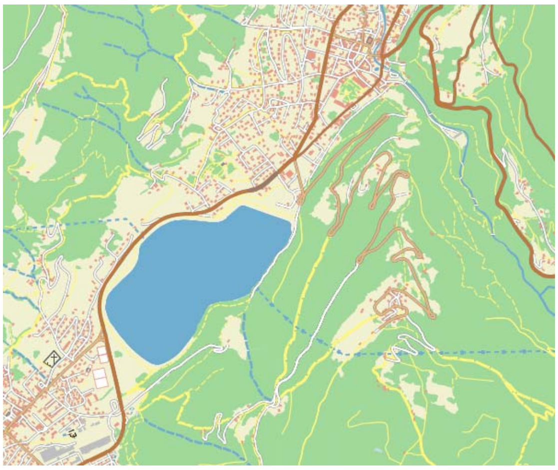

Figure 3. Legend with *calm* color palette - density: *rural*, relief: *mountainous area*.

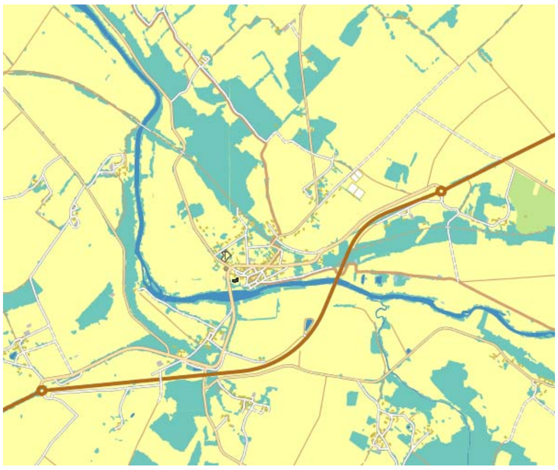

Figure 4. Legend with *soft* color palette - density: *rural*, relief: *plane area*.

Lastly, 48 maps have been selected. They have been made from 16 color palettes applied to 3 types of landscapes: *rural* and *plane area*, *urban* and *sea area* and *rural* and *mountainous area*. All of them are topographic, based on IGN data.

2.2. Survey characteristics

The construction of the questionnaire can be illustrated as follows:

- 48 topographic maps;
- one closed question about 4 properties repeated for every map: *Do you think this map is Realistic? Precise? Original? Balanced? …*;
- 6 response choices were given for each property: *don't agree at all, disagree, neither agree nor disagree, agree, strongly agree* and *no opinion*. (see Figure 5, screen capture of a page of the survey web site);
- property definitions in footnote page:
  - **Realistic** map: does the symbolization chosen refer to reality?
  - **Precise** map: does the map give the right information about the situation on the ground at first glance?
  - **Original** map: is the map distinguished from the habitual mapping uses?
  - **Balanced** map: are the different elements in the map represented in a way that none predominate the others?
- and the last page of the questionnaire permits to gather information about job, sex of the respondents and their comments (see Figure 6: screen capture of last page of the survey web site, is an illustration).

Moreover, a web site was designed. The online survey tools available do not permit such a high number of pictures (48 maps) and their databases are generally limited to 100 responses, beyond which the survey tool shifts automatically to its paid version. So, the alternative is to construct our own web site and connect it to a database. Firstly, it is free; secondly, the administrator is allowed to provide his personalized graphical user interface (GUI) adapted to the questionnaire features. Besides the design look, the database in which responses, comments and all information about the questionnaire are stored, can easily be organized which constitutes a reusable information database.

Technically, the web site was developed with:

- **HTML** (Hyper Text Markup Language): this is a web markup language used to create web pages that can be displayed in a browser (or web navigator).HTML is used to present content (text, image) in a webpage. It provides the tools with which the content of a document can be structured or annotated according to different types of metadata and indications of rendering and display;
- **CSS** (Cascading Style Sheets): this is a standardized syntax that gives web designers extensive control over the presentation of their web pages and is an essential component of web design;
- **PHP** (Hypertext Preprocessor): this is one of the most famous languages widely used to create dynamic and interactive websites. PHP is generally associated to database in order to facilitate storing of data stream they receive;
- **MySQL** is a RDBMS (Relational Database Management System). Its role is to store data in an organized way to easily handle them later. Thanks to MySQL, responses of surveyed people, their job and even their comments will be treated and requested. Language used in this field is called SQL (Structured Query Language): It is standard language designed for managing data held in relational database management systems (RDBMS).

Finally, the website has been designed and tested locally via the server *WampServeur* to verify receipt of responses in the survey database. The site was hosted by *FranceServ*, an Internet hosting service company that allows individuals and organizations to make their website accessible via the World Wide Web.

The survey web site was sent by email to all our contacts, inside and outside IGN (Geographical National Institute). The idea was to reach different profiles, experts and non-experts in map-making.

The survey was also posted in GIS and mapping forums such as *GeoRezo*[2], *GIS Forum*[3], *Plaisance Practice*[4], *Geomatics-Enit*[5]. It was available from 27th June to 11th July 2012. Not only is this strategy useful to reach a huge number of people with different profiles, but it provides discussion topics around the survey as well.

---

[2]http://georezo.net/forum/viewtopic.php?id=80179

[3]http://www.forumsig.org/archive/index.php/t-34622.html

[4]http://www.plaisance-pratique.com/

[5]http://www.sitealytics.com/geomatics-enit.com/

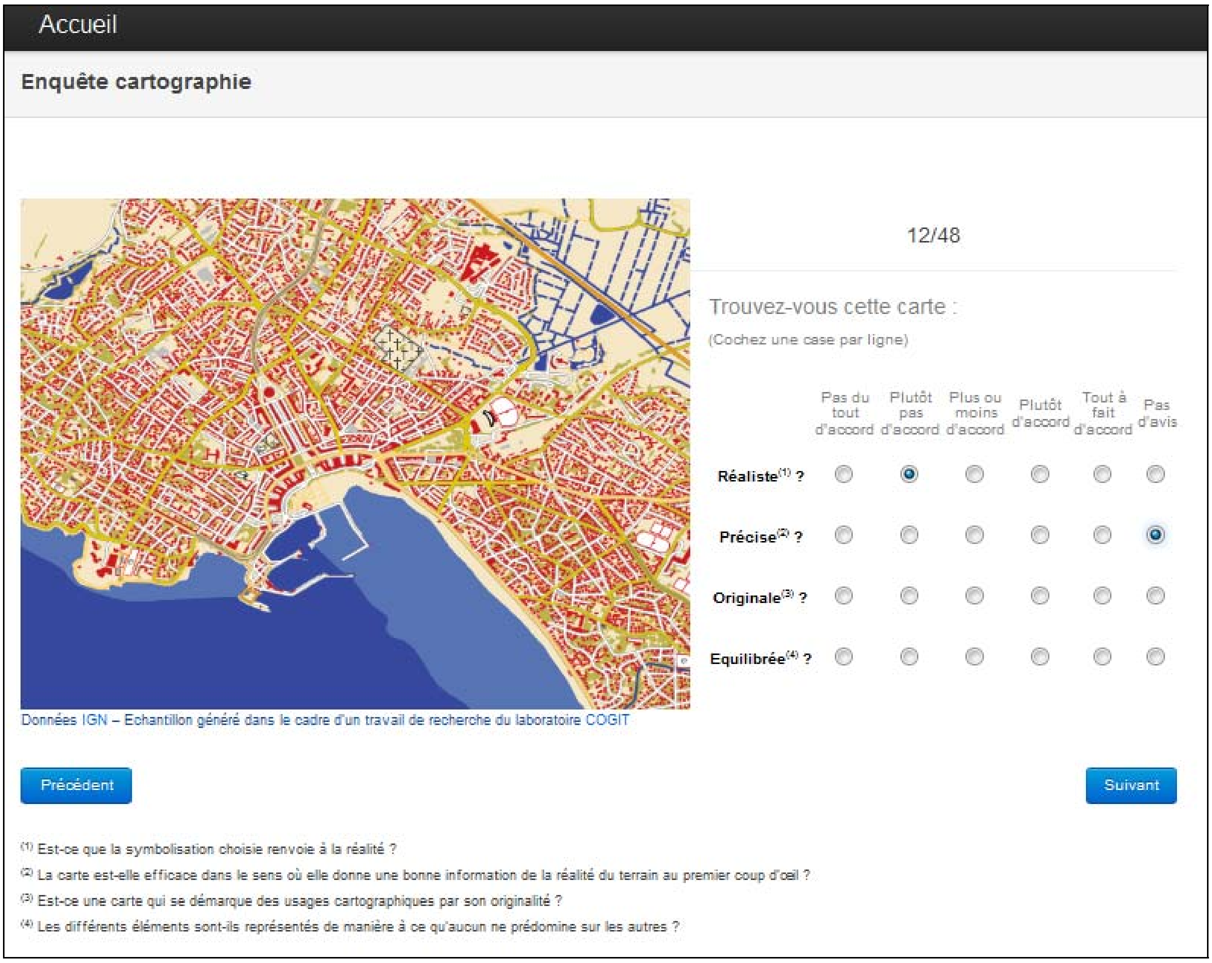

Figure 5. Screen capture of a page of the survey web site.

Figure 6. Screen capture of last page of the survey web site.

2.3. Statistical analysis

Lastly, the survey database contains more than 270 responses. Respondent profiles are very varied (see Figure 7). It is sufficient to carry out a statistical analysis for each property. The purpose is to make map groups based on their visual properties and to identify the maps which are the most <*property*> or the least <*property*>, in other words, maps which may be a caricature of a visual property. For each property, the statistical analysis (Bressoux 2010) allows us to identify the most exaggerated maps i.e. maps whose mark is the most different from the average according to the standard deviation.

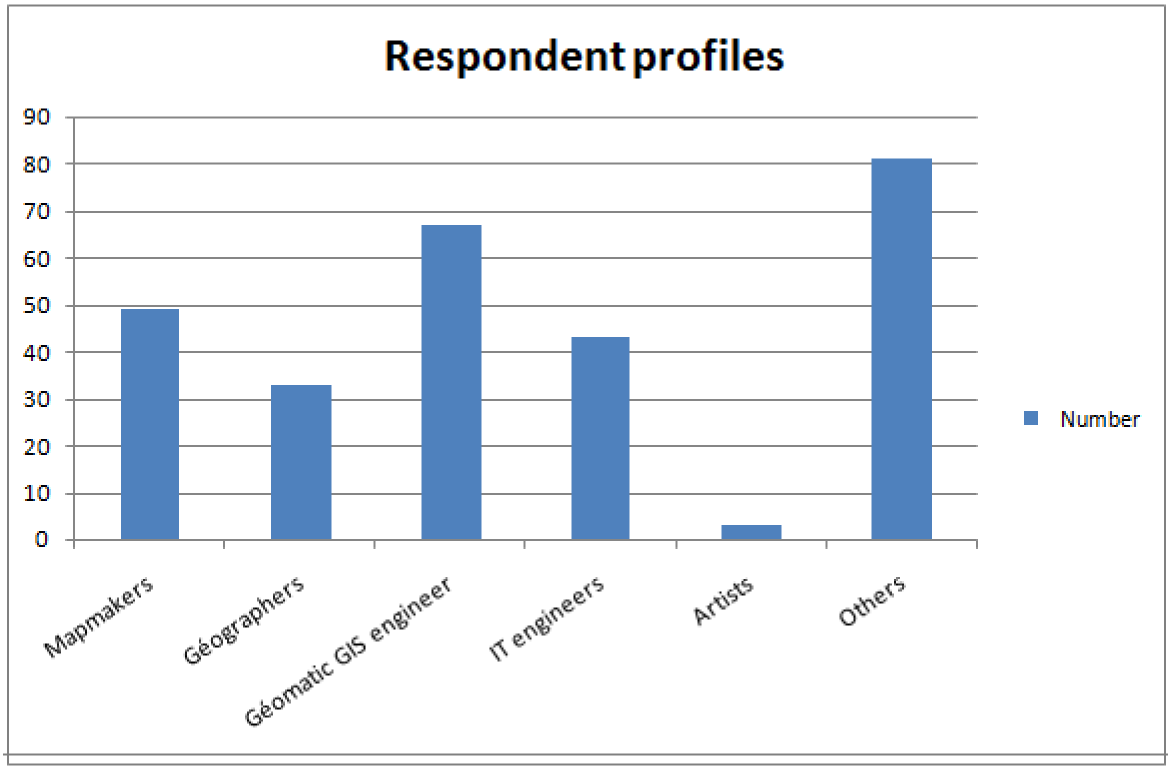

Figure 7. Respondent profiles.

The ordered qualitative choices have been translated into quantitative values to calculate average and standard deviation, for each map and each property. So, firstly, for each map and each property, the mark could have been calculated (the average of all the marks given in the survey). Secondly, for each property, the average and the standard deviation have been calculated (see Table 1).

	Average	Standard deviation
Realistic	-0,10	0,45
Precise	-0,14	0,39
Original	0,09	0,22
Balanced	-0,35	0,47

Table 1. Average and standard deviation for each visual property.

The graph below (see Figure 9) shows value distribution of the *realistic* variable. Two groups of maps can be distinguished at the two extremities: the *most realistic* and the *least realistic* maps.

We can read on the positive x-axis: the more distant from the average the *realistic* value is, the more *realistic* the map is. In the negative part: the more distant from the average the *realistic* value is, the least *realistic* the map is.

Indeed, maps with values greater than the standard deviation are very realistic. On the contrary, in the negative part, maps with values less than the value of the average minus the standard deviation, are considered unrealistic.

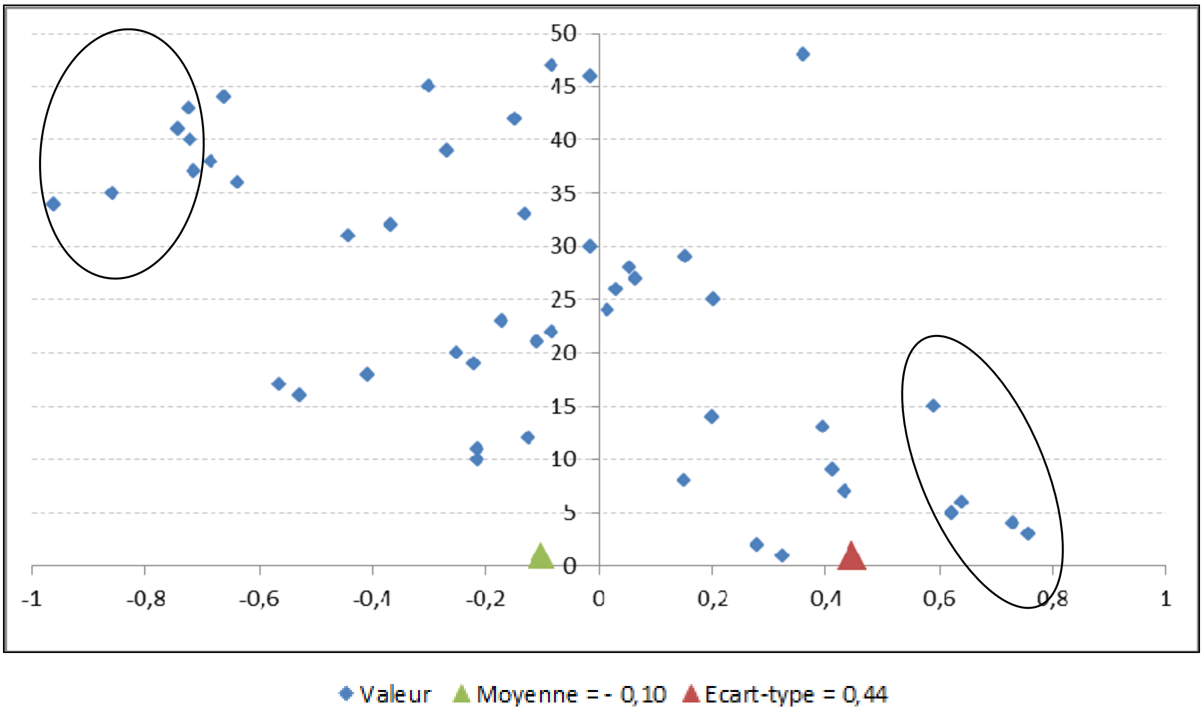

Figure 9.Value distribution graph of the *realistic* property.

Therefore, the two groups related to the *realistic* property can be sorted. The examples below give an idea about the two most realistic maps (Figures 10 and 11) and the two least realistic maps (Figures 12 and 13):

Two most *realistic* maps

Figure 10. Germany legend - density: *urban*, relief: *sea area*.

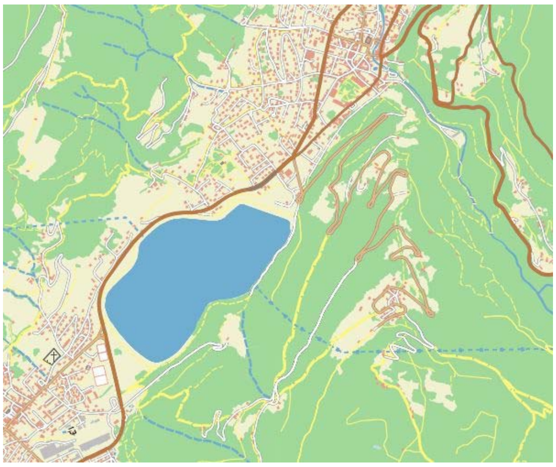

Figure 11. Legend with *calm* color palette - density: *rural*, relief: *mountainous area*.

Two least *realistic* maps

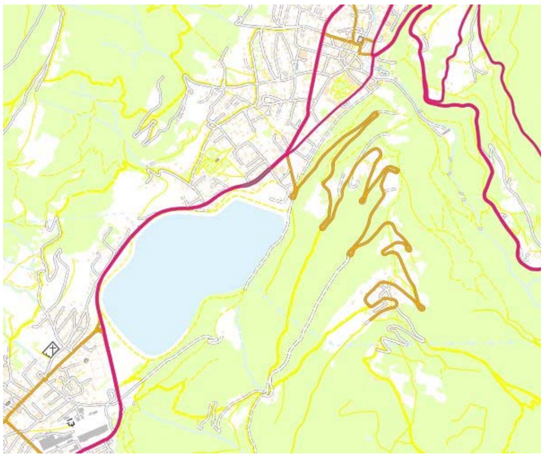

Figure 12. UK legend - density: *rural*, relief: *mountainous area*.

Figure 13. UK legend - density: *rural*, relief: *plane area*.

3. Results and discussion

The statistical classification approach, in Figure 7, has been generalized by the other properties: *precise*, *original*, and *balanced*. Eight groups of maps are constructed; they are summed up in Table 2:

Map property	Intensity	Legend
Realistic	*Most*	*Germany/urban/sea area* *Calm* palette/*mountainous area* *Calm/rural/plane area* *Calm* palette/*urban/sea area* *Estonia/urban/sea area*

	Least	*UK/mountainous area* *UK/rural/plane area* *Soft* palette/*mountainous area* *Traditional/mountainous area* *Traditional* palette/*rural/plane area* *Tropical* palette/*mountainous area.*
Precise	*Most*	*Germany/urban/sea area* *Calm* palette/*mountainous area* *Calm* palette/*urban/sea area* *Estonia/urban/sea area.*
	Least	*UK/mountainous area* *UK/rural/plane area* *UK/urban/sea area* *Traditional* palette/*mountainous area* *Traditional* palette/*rural/plane area* *Tropical* palette/*mountainous area* *Tropical* palette/*rural/plane area*
Original	*Most*	*Earthy* palette/*mountainous area* *Earthy* palette/*urban/sea area* *Powerful* palette/*mountainous area* *Powerful* palette/*rural/plane area* *Powerful* palette/*urban/sea area* *Regal* palette/*urban/sea area* *Rich* palette/*urban/sea area* *Soft* palette/*urban/sea area* *Traditional* palette/*urban/sea area* *Tropical* palette/*urban/sea area* *Ukraine/urban/sea area*
	Least	*Germany/mountainous area* *Germany/rural/plane area* *Germany/urban/sea area* *Estonia/rural/plane area* *UK/rural/plane area*
Balanced	*Most*	*Estonia/urban/sea area*
	Least	*Faroe/mountainous area* *Faroe/rural/plane area* *UK/mountainous area* *UK/rural/plane area* *Traditional* palette/*mountainous area* *Traditional* palette/*rural/plane area*

Table 2. Map groups.

According to this classification some remarks can be given:

- many maps are qualified *very original*, i.e. those maps are designed differently from the usual mapping uses;
- few maps are very balanced. We suppose that some colors predominate among others;

- the color palette *calm* generates *very realistic* maps.

Furthermore, these map groups can be offered to mapmakers in Geoxygene: a GIS/map making software by IGN.

The translation of ordered qualitative variables in quantitative variables is possible because the study only considers one property at a time and can only be based on average and standard deviation. Nevertheless, it would have been more correct to design the test answers as a mark on a scale of 5 or more than as a qualitative answer among 5 ordered qualitative answers.

4. Conclusions and future perspectives

The survey answers have been organized into a database. The latter constitutes a source of information which is reusable for other work. In fact, additional variables such as age, sex, job, may be used to refine the result survey.

This work is the first step to analogy-assisted map making. The next step is in progress. It will consist in deducing global visual properties from elementary salient features. A semiotic study (Landragin 2006) according to type of landscape is being performed. It aims to highlight the map salient features, the elements that influence the choice, of people being surveyed, relative to properties. This analysis will answer the question: *how can we map out our data in order to get which <property> map*?

References

Bressoux P (2010) Modélisation statistique appliquée aux sciences sociales. Bruxelles : De Boeck (2ème éd.)

Christophe S, Ruas A (2009) A process to design creative legend on-demand, 24th International Cartographic Conference.

Dominguès C. (2008) Description de cartes géographiques. 8èmes journées francophones "Extraction et Gestion des Connaissances" - Workshop on "User modeling and Web interface customization", p. 59-68.

Dominguès C, Bucher B.(2006) Application d'aide à la conception de légende. Colloque International de Géomatique et d'Analyse Spatiale SAGEO'06. Strasbourg (France)

Jolivet L.(2009) Characterizing maps to improve on-demand cartography - the example of European topographic maps., 17th Conference on GIScience and Research in UK (GISRUK'09), poster session. Durham (UK)

Landragin F. (2006) Visual Perception, Language and Gesture: A Model for their Understanding in Multimodal Dialogue Systems. *Signal Processing* 86 (12), Elsevier, Amsterdam, The Netherlands, p.3578-3595.