A TOOL FOR CHECKING GRAPHIC LOGIC IN SVG MAPS

Masatoshi ARIKAWA[*], **Fumio FUJIMORI**[*] and **Takashi MORITA**[**]

* Center for Spatial Information Science, The University of Tokyo

Email. {arikawa, fujimori}@csis.u-tokyo.ac.jp

** Department of Civil Engineering, Hosei University

Email. morita@k.hosei.ac.jp

ABSTRACT

SVG (Scalable Vector Graphics) is expected to be a new standard for graphic data on the Internet. In the near future, many maps on the Internet will be formatted as SVG. SVG-formatted maps are relatively easy to be understood by machines because of their internal design structures compared with image-formatted maps which have been mainly used as graphics on the Web. This paper introduces backgrounds and basic frameworks of our proposed tool *map checker* to help people produce better SVG maps using the knowledge of graphic representation and cartography.

1. INTRODUCTION

SVG (Scalable Vector Graphics), which has been developed in W3C (World Wide Web Consortium), is expected to be a new standard for XML-based graphics descriptions[1][2]. In the near future, many maps on the Internet will be formatted as SVG. SVG-formatted maps are relatively easily understood by machines because of their internal design structures, compared with image-formatted maps that have been mainly used as graphics on the Web. This paper introduces basic principles of our proposed tool called *map checker* for checking SVG maps using the knowledge of cartography. The map checker for SVG maps may become something similar to spelling and grammar checkers for text documents. It may help people create better maps to be published on the Internet, and learn cartography and graphic design through the process of improving maps.

Maps should communicate the logic of graphic relations represented in the maps with viewers. The graphic relations used in our research are based on visual variables proposed by Jacques Bertin [3][4]. Many maps on the Internet include inappropriate or wrong graphic relations for visual communication. These bad maps on the Internet are generated through current GIS (Geographic Information Systems), or created by people using some drawing software. Present GIS people have a tendency to think that the geographic databases are essential and maps are only temporal views generated from the databases. As a result, GIS people tend to forget the importance of visual communication, and they are producing poor geographic databases with less imagination, due to a lack of experience in creating effective visual communication. Our proposed map checker will be useful not only for ordinary on-line map creators but for GIS people to improve maps as views of geographic databases, and to learn how important the visual communication is for producing sophisticated geographic databases.

2. PROCESS OF MAP READING

2.1 Human's map reading

Reading maps means extracting information from maps. The information contained in the maps can be considered *relations of features* that are visualized as *graphic relations* in maps (Figure 1) according to the Jaques Bertin's theory, that is, the semiology of graphics. There are three well-known graphic relations: difference, order and proportionality. We give a definition of *good* maps by the meta-relation between the relations of features and the graphic relations. Good maps must be the ones that are easy for humans to extract the right graphic relations from maps by human's visual perception. Also, the levels of importance for features on maps should be visualized appropriately. In other words, important features should be visualized as distinct graphic representations, and less important features should be represented with weak visual effects. Human's map reading process can be interpreted as the transfer of relations in maps to relations in a human's brain. Human's map reading process is difficult to be defined in the framework of digital computing. One solid fact in the process is that the interfaces of transferring relations with human's visual perception are physical visual media such as papers and display screens. Human's map reading process is quite complex and excellent. It is almost impossible for computers to simulate all of human's map reading process.

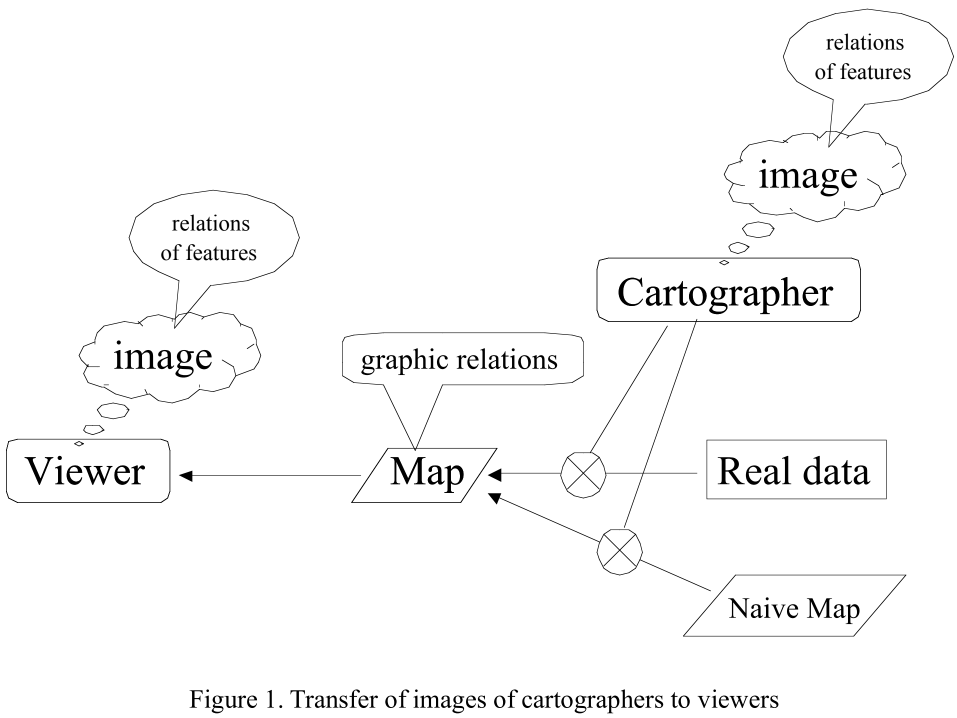

Figure 1. Transfer of images of cartographers to viewers
through graphic relations on maps.

2.2 Computer's map reading

As we described, it seems to be impossible for computers to read or understand maps. Computer's map reading can be defined as extracting graphic relations from maps. While humans can extract graphic relations as well as importance levels of them with no difficulties, computers can neither extract all graphic relations nor define importance levels of them. The first step of computer's map reading process is converting pixel data to graphic data. The second step is converting graphic data to geographic features and relations. The first step has been studied in the research field of computer vision for the last several decades. Many efforts have been paid to revealing and making the same mechanisms of human's visual perception to understand the real world through pixel data. For computers, it is much easier to understand the pixel data representing *maps* than to understand the data taken from *the real world*. But, it is still difficult to achieve the same levels for reading maps as human's visual perception. Thus, the first step of computer's map reading process has been very hard to realize.

There is a good tendency that SVG (Scalable Vector Graphics) is going to be popular on the Internet. SVG is a description language for vector graphics based on XML (eXtensible Markup Language). Data of SVG are basically textual data, which are descriptions of graphic features, graphic attributes of them and tree structures for grouping the graphic features. Assuming that the use of SVG for representing maps becomes popular, we can avoid a hard obstacle, that is, the first step of computer's map reading process, which was not overcome for decades. Thus, we can start from the second step, that is, converting graphic data to geographic data. The second step is similar to the process of human's understanding maps. In SVG, it is easy for computers to extract graphic features from SVG maps, because primal elements on them are not pixels, but graphic elements. Also, it is easy to extract attribute values of graphic features because the attribute values are described explicitly as XML data of SVG maps. The attribute values of graphic features correspond to visual values in the Bertin's theory. It may be possible for computers to extract graphic relations from maps using visual values. A shortcoming of SVG is that there is almost no description concerning geographic features unless humans explicitly describe geographic identifiers or attributes for graphic features of SVG map data. The computer's map reading process can be interpreted as extracting geographic features from a set of graphic features, and then extracting geographic relations from a set of graphic relations. Thus, the process of computer's map reading can be modeled in the similar process of human's map reading. This suggests great potentials for innovating new intelligent tools to support the process of human's map reading and *creating* a new process in various ways, because computers will be able to simulate part or all of the process.

2.3 Human's process of reading maps generated from GIS

Generally speaking, maps generated form GIS are not sophisticated for people to read. The geographic databases are considered the most important and core components in GIS. The databases are designed based on feature models [5][6]. From this point of view, GIS is suitable for computers to understand the databases themselves rather than maps generated form the databases. Thus, the current GIS focuses on the process of computer's understanding databases, ignoring the process of human's visual perception of maps and imagination.

Graphic representation has a double function in the Bertin's theory: as an *artificial memory* and as a *tool for discovery*, by the huge power of visual perception. Most of GIS people take it for granted that maps, which are composed of graphic features, are derived from features stored in geographic databases. We should remember that at first there were the real world and humans before both computers and papers were invented (Figure 2). Features in human's brain can be created by human's visual perception of information including features of the real world and graphic representation such as maps. Human makes schemas of both maps and geographic databases with the power of imagination using graphic representation. In other words, the schemas of geographic databases are designed after we create mental features in our brains and graphic features on our maps. The current GIS does not support the capability of human's discovery using graphic representation.

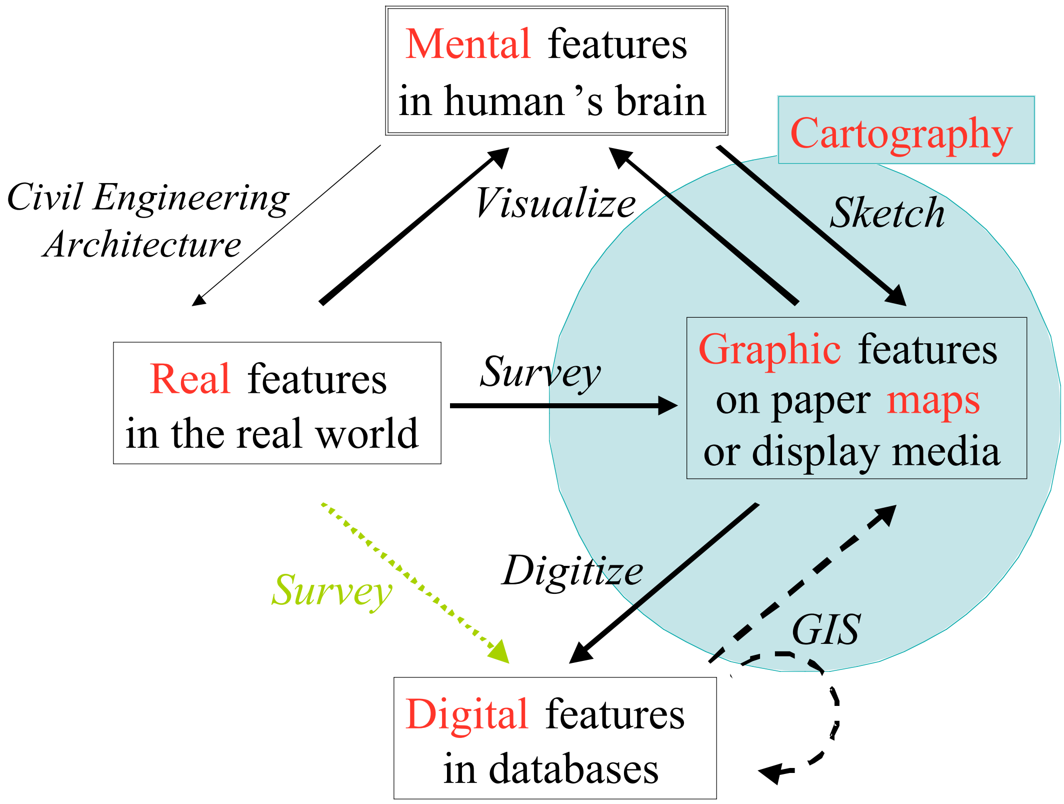

Figure 2. Various features for information transfer and creation.

3. DEFINITION OF MAP CHECKERS

There are useful tools for checking digital documents such as spell and grammar checkers. These tools are generally bundled with word processors, and help humans produce digital documents. It is expected that similar tools for checking digital maps will be created. We call such tools *map checkers* for supporting production of digital maps as well as paper maps. However, it seems to be impossible to create ideal map checkers that work well just like humans checking maps. We should therefore start from developing naive map checkers, which are far from the high levels of capability provided by experts of cartography, but can assist humans to produce maps at lower levels. Then, we continue to improve the map checkers. Everyone, even a cartographer, makes simple mistakes when he or she produces maps. For example, simple mistakes may happen in choosing fonts, colors, textures, positions and directions of layout, and so on. Most of these mistakes can be interpreted as problems of setting visual values according to the Bertin's theory. Various levels of mistakes happen in producing maps. Some mistakes are trivial ones; other mistakes are ones at high levels such as in abstracting multiple geographic data and producing artistic representations for maps. Mistakes in choosing visual values on maps make it difficult for viewers to read the maps. Serious mistakes in visual values may cause errors, which communicate wrong information to viewers.

The process of human's improving contents on maps is divided into two sub processes. The first sub process is to find inappropriate visual values from maps through human perception. The other sub process is to change visual values and layout of graphic features on maps by hands. Map creators should check if maps are easy to read or not, and if maps may possibly convey wrong information or not. In order to check these problems in maps, we usually check the maps directly, but it will be more useful if we have new tools for visualizing visual values. This enables us to check maps indirectly. The followings are major checkpoints when creating maps.

(I) atomic elements and logic structures
(II) layout or positioning
(III) ambiguity
(IV) metrics and topology
(V) class order or importance
(VI) visual complexity and its limits
(VII) contrast
(VIII) legibility and its limits
(IX) interactivity and its limits
(X) consistency
(XI) continuity and direction

The new tools can provide visualization of visual values in maps, or graphic representation of graphic representation. The tools are useful for checking both of the above two sub processes of improving maps; finding bad points and changing visual values. These tools are like analyzers for experts such as stethoscopes for doctors and circuit testers for electronic engineers. Maps formatted as SVG can be structured as trees by the definition of XML. Using the tree structures of SVG, we can change visual values for units of groups. Graphic elements such as points, lines and polygons, do not necessarily correspond to features, which are meaningful units of objects such as roads and buildings.

The functions of grouping make a set of graphic elements into a graphic feature, and make a set of graphic features into a composite feature on a map. In other words, groups can be interpreted as individual features as well as layers for classes of features. However, the static tree structures of SVG maps are generally not enough to cover a wide variety of needs and perspectives. There may be no complete solution to satisfy both the variety of needs and perspectives. Even if there is no complete solution for deriving tree structures for conceptual models, we can derive naive but better tree structures of geographic features from tree structures of graphic features, which are SVG maps themselves, using rules of visual values and usage for graphics. These functions of map checkers can be helpful hints for map creators in producing maps effectively and efficiently. Figure 3 illustrates a flow diagram for humans to improve a SVG map with a map checker. The map checker selects graphic values from XML data of SVG maps, and visualizes the graphic values as visual meta-representation of maps.

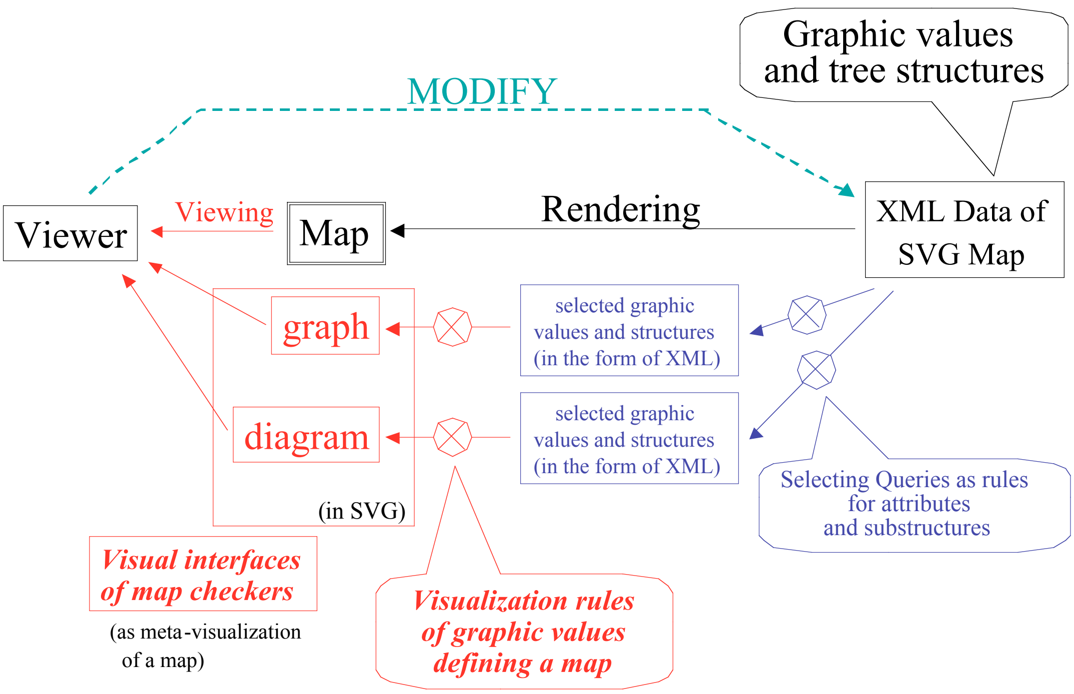

Figure 3. Flow of improving a map using a map checker.

4. A PROTOTYPE OF A NAIVE MAP CHECKER

We have been developing a prototype system of a map checker based on the ideas described in the previous sections. As we described, there are various levels of map checkers. The first prototype system of our map checker has only simple functions to visualize graphic values of SVG maps as visual graphs. Figure 4 shows an example of the interface of a naive map checker, which includes a SVG map, a histogram, and controllers for selecting layers and graphic values. When a viewer selects layers or graphic values of the SVG map, he can immediately check the distribution of the graphic values depending on the change of selected layers and graphic values. If a viewer finds problems in the histogram as a visualization of the SVG map, the viewer changes graphic relations to improve the SVG map. Figure 5 and 6 show two histograms for the ratio of colors in a map. Figure 5 shows the distribution of the number of elements in XML data of a SVG map for the range of colors. Figure 6 shows the distribution of area for the range of colors. Figure 5 visualizes the internal structure of the XML data, but does not tell the distribution of the ratio of colors in the map. The number of elements in the XML data is meaningless for visual perception and design, but may be meaningful for some other purposes. On the other hand, Figure 6 shows the distribution of area for colors, which is meaningful and natural for visual perception and design. These examples indicate an essential difference between the structures of XML data and structures of graphic representations. Our map checker provides viewers with functions and interfaces to change histograms to be suitable for checking SVG maps by choosing both graphic values and methods for visualization.

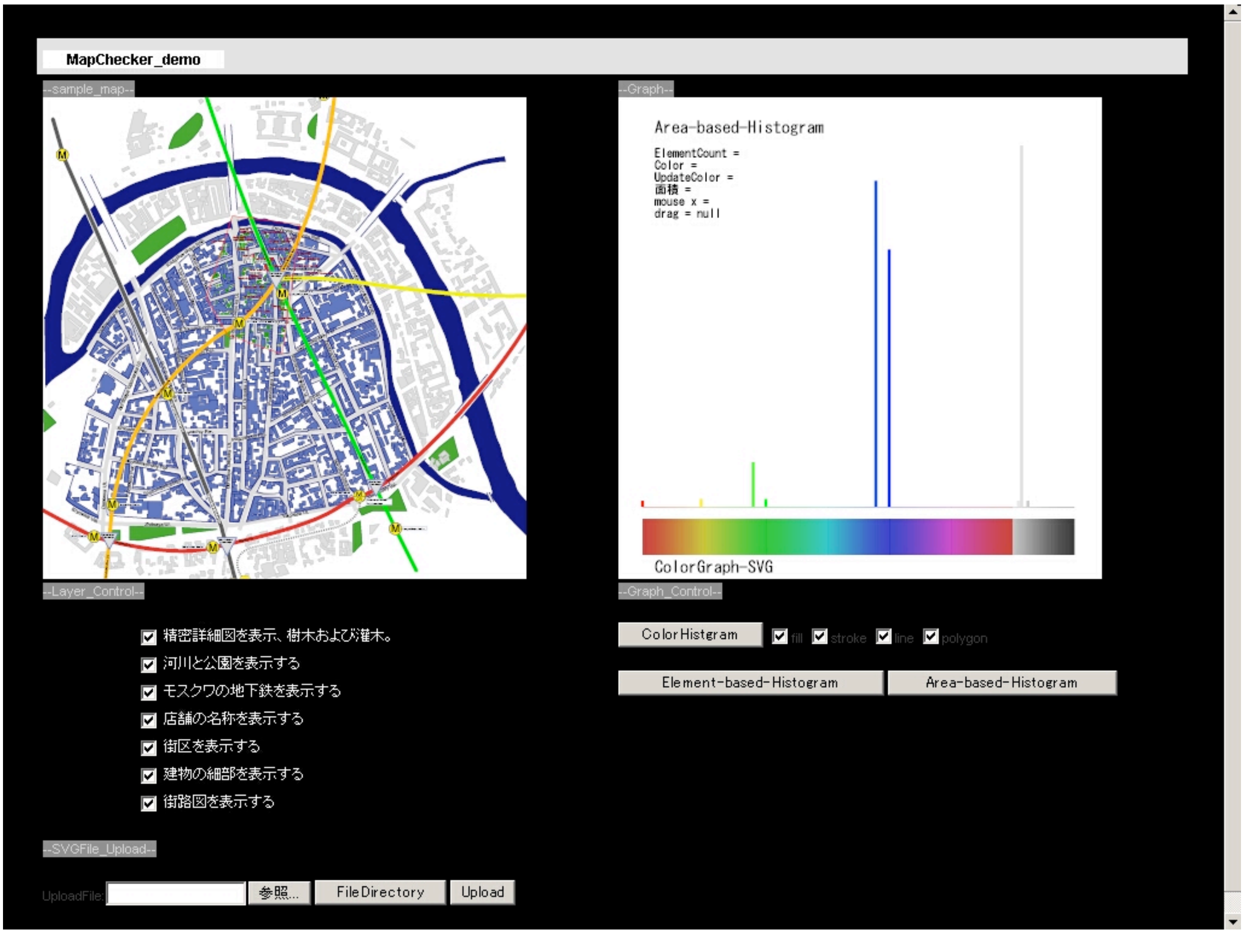

Figure 4. A visual interface of a naive map checker including a SVG map, selectors of its layers, a histogram in a map checker, and selectors of graphic classes.

(A SVG map of Adobe System Inc. [7] is used as a test data for the map checker.)

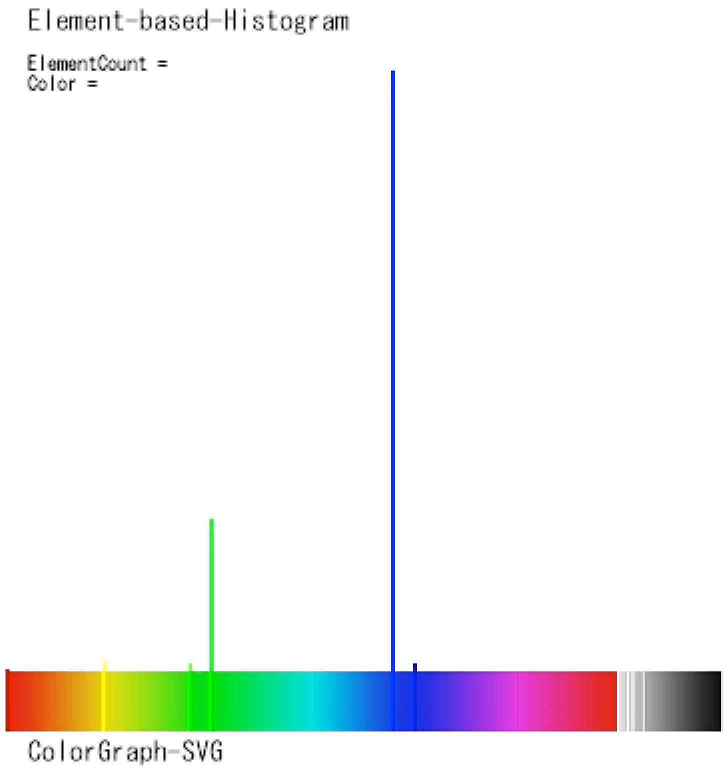

Figure 5. A histogram of the number of graphic elements for the range of colors in a SVG map.

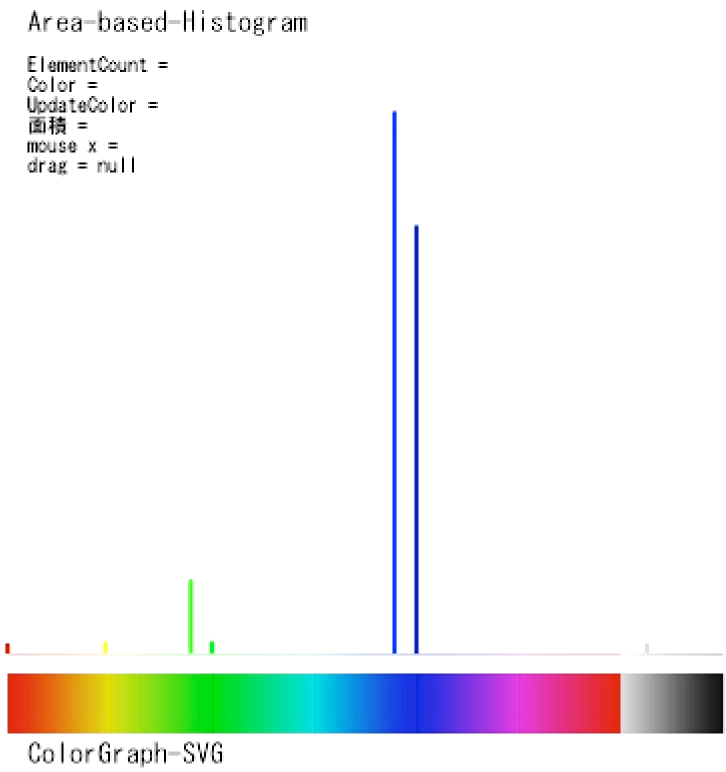

Figure 6. A histogram of area of graphic elements for the range of colors in a SVG map.

5. SUB SYSTEMS OF MAP CHECKERS

We have been developing a map checker to solve simple problems that frequently occur in creating maps at first, and then we are going to enhance its functions gradually. To realize a huge system of the map checker, we divide it into several primal sub systems as follows. These sub systems can be considered software component modules to interact with one another, and units of tools and interfaces for people.

(1) Map Analyzer:

Map analyzer extracts properties of maps from the DOM (Document Object Model) of the corresponding XML documents for SVG maps, and provides graphic values and relations of the properties as visual graphs, etc. It allows people to distinguish appropriate visual relations from inappropriate ones in maps.

(2) Map Synthesizer:

Map synthesizer provides graphical user interfaces, which allow people to change properties in graphics of maps. It is a good environment for people to make better graphic relations in maps.

(3) Map Articulator:

The DOM structure of SVG maps is not necessarily matched with meaningful structures of the maps. Map articulator extracts meaningful clusters from the maps, such as a combination of geographic features and labels.

(4) Map Corrector:

Map corrector finds inappropriate relations from SVG maps and gives candidates for the replacement of property values representing the inappropriate relations. To achieve this, there must be databases of patterns of both typical inappropriate relations and typical modifications for them.

(5) Map Translator:

Map translator extracts spatial relations from SVG maps, and generates some sentences of natural languages from the spatial relations. This function is useful for navigation maps, but not for survey maps. Also, if maps can be translated into sentences or words, these results are applied for text information retrieval. This function enables us to retrieve maps on the Internet by keywords through search engines.

6. CONCLUSION

We focused on the emergence of SVG and new tools for assisting human's map creation. Assuming that the popularity of SVG-formatted maps on the Internet continues to grow, we have discussed a future visual environment based on *map checkers* proposed in this paper. The map checkers are expected to decrease non-creative work in human's map creating process, and to increase the chances of activating the power of human's imagination. The power of human's imagination using graphic representations has been ignored in the field of GIS. The emergence of map checkers will provide GIS people with more chances to know how much the knowledge of cartography is important when we use maps on computers. This paper has presented the basic principles of map checkers. The essential aspects of map checkers include very difficult problems to realize on computers, but it is important to find simple solvable problems to realize useful and simple functions as part of map checkers. We will deal with real problems and solve practical problems among them in realizing the map checkers while developing an advanced system of the map checkers.

REFERENCES

[1] Scalable Vector Graphics (SVG), W3C, <http://www.w3.org/Graphics/SVG/>

[2] eXtensible Markup Language (XML), W3C, <http://www.w3c.org/xml/>

[3] J. Bertin, "La graphique et le traitement graphique de l'information", Flammarion (1977) (Japanese edition by T. Morita: Heibonsha, Tokyo (1982))

[4] J. Bertin and W. Berg, "Semiology of Graphics: Diagrams, Networks, Maps", Univ of Wisconsin Press (1983)

[5] ISO/TC 211, Geographic information/Geomatics, <http://www.isotc211.org/>

[6] Open GIS Consortium, <http://www.opengis.org/>

[7] Adobe SVG, <http://www.adobe.com/svg/>