Spatial Vision Analysis of Non Spatial Data

Rui Xin
Wuhan University /
School of Resource and
Environmental Sciences
129 LuoYu Road
Wuhan, China
xinrui@whu.edu.cn

Tinghua Ai
Wuhan University /
School of Resource and
Environmental Sciences
129 LuoYu Road
Wuhan, China
tinghuaai@whu.edu.cn

Xiongfeng Yan
Wuhan University /
School of Resource and
Environmental Sciences
129 LuoYu Road
Wuhan, China
xiongfeng.yan@whu.edu.cn

Abstract

With the help of a person's sense of space and familiar with the map, this study uses map-like visualisation to spatially represent and analysis the non spatial data. Map-like visualisation can display data and data structure explicitly, which is easy for the audience to understand and grasp. The addition of visual analysis methods can take the advantages of human visual perception to make the analysis process more intuitive and sensitive, and help discover hidden rules in the data. Experiments under real data show that the map-like visualisation in this paper can complete mining and analysis of data on the basis of correct expression of the data.
Keywords: map-like visualisation; visualization analysis; colour mixing analysis; map expression

1 Introduction

The visual mining can effectively make use of human visual perception talent to discover potential patterns in data. Faced with so many visualization technologies, it is urgent to choose one method to visualize data more effectively. Map is a good spatial data carrier. In the process of spatial data mining, a map is often used to display and analyse data. The map can also be used as data carrier in analysis and expression of non spatial data, which can be called map-like visualisation. The map-like visualisation transforms unfamiliar and abstract non spatial data into familiar and sensitive map elements by spatialization (Skupin, 2007).

Shneiderman (2006) divides information visualization data into 1-dimensional data, 2-dimensional data, 3-dimensional data, multi-dimensional data, temporal data, hierarchical data, network data and map-like visualisation researches are usually carried out around the latter four kinds of data (Mark, 2001; Auber, 2013; Gansner, 2010). In the field of theoretical research, Researchers have carried out studies on the basic theory of definition, theoretical framework, and so on (Fabrikant and Skupin, 2005; Skupin and Buttenfield, 1997). Many studies about map-like visualization have emerged on the basis of these theoretical guidance (Biuk-Aghai et al., 2015). At present, the analysis ability of map-like visualisation is weak. It is often used to display data structure such as hierarchical data or network data but lack of analysis of the data. This may lie in lacking of effective use of map methods, and rarely use relevant methods of visual analysis.

With the aid of spatialization, a map-like visualisation expresses multi dimensional data to a two-dimensional plane through the dimension reduction technique (Skupin, 2000). It determines the locations of the elements, but also limits the subsequent analysis that it needs visualization analysis technology to express and analyse the multiple attribute information in limited plane map space. Colour is a commonly used data rendering tool which can be used to express multidimensional attribute data (Silva, 2011). Colour mixing technology can use colour fusion characteristic to break the space limitation because a new colour can be generated by blending different colours at the same place (Gama, 2014; Gossett, 2005). Colour mixing follows certain rules, and there is a causal relationship between the mixing mode and the mixing results, which provides an intuitive logical basis for data analysis. Both its simple principle and readability provide cognitive convenience for the audience.

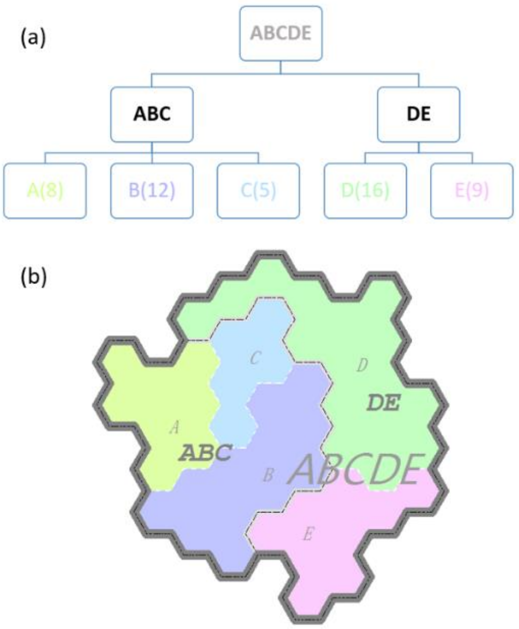
This paper expresses abstract data through map-like visualisation in which the colour analysis is introduced. In this way, we use the map carrier to display the structure of abstract data, and analyse the relationship between data through colour mixing technology. It's different from previous studies in which the colour is just used to render results. The colour in this study runs through the whole process in expression and analysis of abstract data. The colour features and the data features are bound to analyse data through the colour, so that the analysis process is easy to understand and the analysis results are more sensitive. Through the mixed colour analysis in map-like visualisation audience can not only grasp the data structure but also intuitively observe the data relationship, making it easy to excavate and discover hidden patterns in data.

2 Method

2.1 GosperMap

This study takes hierarchical data as the research object, at the same time, the Gosper map is chosen as the spatial expression vector of abstract data. The Gosper map in this study is drawn on the experience of Auber (2013). In the hexagonal base map, assign the corresponding hexagonal units to the leaf nodes under the guidance of Gosper curve. Based on the inclusion relation between nodes, the map build is completed by generating the parent node region from merging child node regions bottom-up. Under the constraint of cartographic rules, the content and structure of abstract data are expressed on the map. As shown in Figure 1, each node of the hierarchical data is mapped to the corresponding area on the map, and there is a positive proportion relationship between the area size and the attribute value of the corresponding elements. Following the linear guidance sequence, the adjacent relations of nodes in hierarchical tree are represented as topological adjacency relations among map domains.

Figure 1: Gosper map corresponding to the hierarchical tree



The map uses hexagon as the basic composition unit which can achieve seamless splicing in space area. Compared with other regular polygons such as triangle, square, the adjacency form of shared edge help it avoid the regional discontinuity caused by the point connection. Through the Gosper map, the audience can obtain a concrete understanding of content and structure in the abstract data. However, the content expressed in Gosper map still belongs to shallow layer of information. Each region represents the data item composed of multidimensional attributes while the potential relationships between the attributes are difficult to analyse and express in the figure. The introduction of additive colour mixing technique will be helpful to solve this problem.

2.2 Additive colour mixing principle

Additive colour mixing, that is, the mixing of lights to produce another colour. Light mixing acts on the human eye, which produces a new sense of light colour due to the different colours stimulate the retina simultaneously or in a very short period of time. Red, green and blue lights are called additive primary colours because by the combination of these three primaries in different proportion, various colour lights will be produced. For example, by mixing the primary colours with each other equally, we can obtain green, magenta and yellow. The following basic laws are obeyed when colour mixing occurs: the colour energy after mixing is the sum of each light that participates in the mixing process; when the two primary colours are not mixed in equal quantity, the result colour will tend to the one in larger quantity. By using the above rules, we can deduce energy and proportion relation of the primary lights through result colour. This can provide a logical basis for subsequent colour mixing analysis.

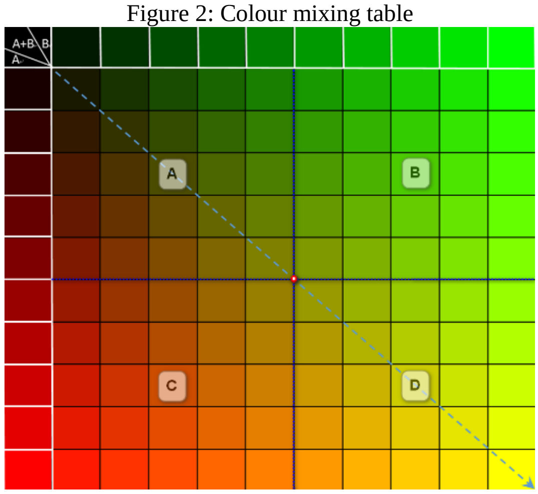
2.3 Additive mixed colour analysis

In operation, we choose RGB color space whose three color channels correspond to the three additive primary colors. At the same time, its simplicity also brings great convenience to our implementation. Each parameter value in RGB corresponds to the energy of the primary colour (range 0 to 255), the greater the value, the brighter the colour. This study uses equal proportion method to map attribute value to integer range 0-255. Therefore, the relative size of attribute value is represented by magnitude of colour energy. As shown in (1), V represents the attribute value needs to be converted, V_{max} and V_{min} represent the maximum and minimum value of corresponding attribute respectively while C is the colour value after rounding.

$$C = \lfloor \frac{(V - V_{min})}{V_{max} - V_{min}} \times 255 \rfloor \quad (1)$$

For the two attributes A and B involved in the analysis, colour mixing process is as follows: the converted values C_A and C_B are assigned to two primary colour domains, such as R domain and G domain and set the value of B domain to 0. The colour represented by RGB parameter (C_A, C_B, 0) is the mixed colour result. The analysis can be carried out when colour mixing result is assigned to corresponding region. The results of colour mixing can be explained b y the colour table in Fig. 2 in which the colour in mixing area is generated by addition of corresponding red and green lights in the header area. The yellow light around the diagonal area in figure 2 represents the relative strength of the two attributes is equal. Along the diagonal arrow direction, the colour intensity becomes larger and larger, representing the increasing value of the two attributes involved in colour blending. Extending from the center in diagonal to the lower left corner or the upper right corner, the colour mixing results are gradually dominated by one light, which represents the relative intensity of imbalance, one colour gradually occupies the peak.

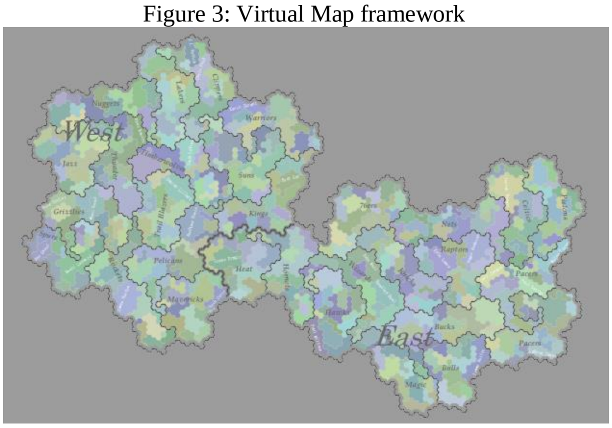
Figure 2: Colour mixing table



3 Experiments and discussions

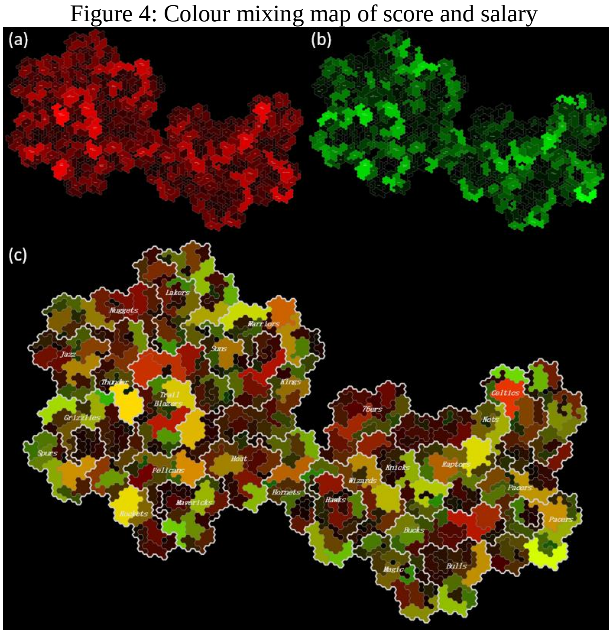
Based on the above theory, this study selects NBA data to carry out the experiment, including player scores, rebounds, assists and other technical data. The hierarchical tree is constructed according to the hierarchical relationship among players, teams and divisions. After that, the hierarchical tree is transformed into a virtual map. We allocate the area according to the score of players, the area size of map region is proportional to the score. In order to take into account the relative orientation between the Eastern Division and the West Division, the map is rotated and adjusted. Finally, the result is obtained as shown in Figure 3.

Figure 3: Virtual Map framework



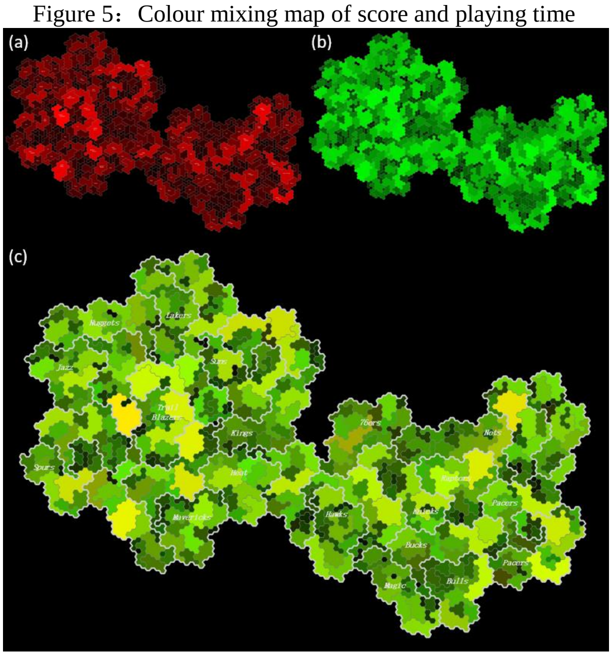
Calculate the colour value corresponding to the player's average score by the formula (1) and assign it to parameter R. The other two parameters G and B are set to 0. The result is shown in Figure 4(a). The bright red regions represent the players with high scoring ability, their average scores are generally higher, while the black areas show that the players have lower average scores. Similarly, the colour value calculated by the player's annual salary is given to G, and the other two parameters are assigned 0. The result is shown in Figure 4(b). As the contract salary belongs to pre competition rather than post game evaluation, the player performance may be out of balance with the salary level. As figure 4(c) shows, we perform colour blending process for the average score and the annual salary after which the result is overlaid with the team boundary. The bright yellow parts are high score and high salary shared areas, these players have achieved high scores matched with their high salary. The score of players in the regions of darker yellow tone is still commensurate with salary, but the positions of them are relatively low in sequences of their respective attribute values. The players with strong scoring ability but low salary are in red regions. On the contrary, green areas represent players with higher pay but without corresponding score. Through simple recognition of the colour, various types of players are visually reflected on the map.

Figure 4: Colour mixing map of score and salary



We switch the analysis perspective to team. Through an overall view of the map we can find that, the players with high salary and high score are less in number and scattered in the Cavaliers, the Warriors, the Thunder, the Rockets and other teams. These players are so-called superstar players. As the core players of their teams, they led the teams in bright yellow areas into the playoffs and even the finals in recent years. Teams in dark zones which represent low pay and low scoring such as the 76ers, the Nets and the Lakers have been at the bottom of the League rankings in recent years.

Figure 5：Colour mixing map of score and playing time



The colour value corresponding to the average playing time is assigned to the G parameter. The map of playing time is also blended with the score map to analyse the two attributes. As shown in Figure 5 (c), there is no widespread distribution of yellow areas which represents equal scores and playing time from the overall perspective of the map. On the contrary, the map shows a green hue on the whole. The bright yellow tone represents the corresponding players who has a higher scoring efficiency. They are usually core players of their teams and get more time to play rely on their outstanding scoring ability. From the point of view of team, the other members in the same team with these core players often show green hue in their regions. This may be attributed to that the team score often rely on core players and the offensive organizations are carried out around them while the other members are mostly responsible for assistance and cooperation so the chances of scoring are affected in the same time.

4 Conclusions

This study uses map-like visualisation to achieve spatialization to express and analyze non spatial data. Compared with other visualizations, it incorporates the features of the map. Combined with the map operation, the abstract data can be observed in the form of map browsing. The potential of the application of maps is great. Based on the classical methods and theories in the field of maps, it would provide a new perspective for us to observe and solve problems. The inclusion relation of the region in the map corresponds to the data hierarchy. The attributes in the data and their relationships can be displayed by visual symbols. They all transform the abstract contents and relations into visual symbols and use maps which are familiar in our daily life as data expression vectors. However, the other potential of the map remains to be further excavated. In this study, by introducing additive colour mixing technology, the relationships between data items are mined. Combined with certain data we can explore the principle reasons behind these relationships. The technology and method of current application are still simple and the law we find are relatively shallow, which is also the focus and motive of the future research.

References

Auber, D., Huet, C., Lambert, A., Renoust, B., Sallaberry, A., & Saulnier, A. (2013). Gospermap: Using a gosper curve for laying out hierarchical data. *IEEE Transactions on Visualization and Computer Graphics*, 19(11), 1820-1832.

Biuk-Aghai, R. P., Yang, M., Pang, C. I., Ao, W. H., Fong, S., & Si, Y. W. (2015). A map-like visualisation method based on liquid modelling. *Journal of Visual Languages & Computing*, 31(PA), 87-103.

Gama, S., & Gonçalves, D. (2014). Studying Color Blending Perception for Data Visualization. *Eurovis 2014 - Eurographics Conference on Visualization*.

Gansner, E. R., Hu, Y., & Kobourov, S. G. (2010). Visualizing graphs and clusters as maps. *IEEE Computer Graphics & Applications*, 30(6), 54-66.

Gossett, N., & Chen, B. (2005). Paint Inspired Color Mixing and Compositing for Visualization. *Information Visualization*, 2004. *INFOVIS 2004*. *IEEE Symposium on* (pp.113-118). IEEE.

Mark, D. M., Skupin, A., & Smith, B. (2001). Features, Objects, and other Things: Ontological Distinctions in the Geographic Domain. *Spatial Information Theory*. Springer Berlin Heidelberg.

Shneiderman, B. (2003). The eyes have it: a task by data type taxonomy for information visualizations. *Craft of Information Visualization*, 364-371.

Silva, S., Santos, B. S., & Madeira, J. (2011). Using color in visualization: a survey. *Computers & Graphics*, 35(2), 320-333.

Skupin, A. (2000). From metaphor to method: cartographic perspectives on information visualization. *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on* (pp.91-97). IEEE.

Skupin, A., & Fabrikant, S. I. (2007). Spatialization. *The handbook of geographical information science*, 61-79.