



HAL
open science

Study on the Theory and Practice of Data Visualization

Quan Wu, Xiaochen Li, Danqiong Wang, Weijie Jiao, Xue Han

► **To cite this version:**

Quan Wu, Xiaochen Li, Danqiong Wang, Weijie Jiao, Xue Han. Study on the Theory and Practice of Data Visualization. 10th International Conference on Computer and Computing Technologies in Agriculture (CCTA), Oct 2016, Dongying, China. pp.1-9, 10.1007/978-3-030-06155-5_1. hal-02179978

HAL Id: hal-02179978

<https://inria.hal.science/hal-02179978>

Submitted on 12 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Study on the theory and practice of data visualization

Quan Wu^(✉), Xiaochen Li, Danqiong Wang, Weijie Jiao, and Xue Han

Remote Sensing Application Centre, Chinese Academy of
Agricultural Engineering, Beijing, 100125, China
{477781686, 1351123026, 1226773217, 327352947,
346334593}@qq.com

Abstract. Data visualization is a science and technology research on data visual form, which is originated in the 50's of the twentieth Century. It went from scientific visualization, information visualization to data visualization. Based on computer graphics and pictures data visualization display and reveal the main information contained in data. Data visualization contains data acquisition, data analysis, data processing and data modeling. Based on psychology the motivation of data visualization is to expand the scope of visual perception which is easier to get information than other biological perceptions while the effect is better. Modern science and technology greatly extends the human sensory "arms". Visual perception has very long "arm" and often "robs" other perceptions' "business". This paper introduces the basic concept and main content of data visualization, reveals the essence of data visualization. Through the "Agricultural planning visualization" case, the technology process of data visualization applied in production has been established, which can be described as 4 steps operated by order. The first step is to determine and decompose a target. The second step is data acquisition and processing. The third step is models designing and expression. The last step is video production.

Keywords: Data visualization, Computer graphics, Image, 3D, Sense organ, Vision, Vision perception, Psychology, GIS, RS, Agricultural planning

1 Introduction

As a kind of technology, visualization applied to the scientific research and production practice leads to many research and practice fields[1][2][3][4][5][6][7]. Data mining is an example to use visualization technology in scientific research[1]. Data visualization is a science and technology research on data visual form, which is originated in the 50's of the twentieth Century. It went from scientific visualization to information visualization and then data visualization. In 1987, Bruce McCormack, American computer scientist, proposed the definition of scientific visualization which

* Quan Wu, Professor, Chinese Academy of Agricultural Engineering, No. 41 Maizidian Street Chaoyang District, Beijing, China.

is a digital form to create a visual image, help people understand the science and technology concept or result using computer graphics. Information visualization rising in 1990s was born in scientific visualization, which emphasizes to express the abstract information using computer graphics and images. The term of data visualization covering the fields of scientific visualization and information visualization was born in 2000s.

Based on computer graphics and images data visualization display and reveal the main information contained in data[2]. Data visualization contains data acquisition, data analysis, data processing and modeling. It is necessary to sufficiently analyze the research or production target. Data acquisition includes the scope of data collection and depth. Data analysis is to abstract, classify and synthesize data collected according to their characteristics. The basic steps of data processing and modeling are to select model types, produce data according to the corresponding model types and design models.

Based on psychology the motivation of data visualization is to expand the scope of visual perception because it is easier to get information than other biological perceptions while the effect from visual perception is better. Modern science and technology greatly extends the human sensory "arms". Visual perception has a very long "arm" and often "robs" other perceptions' "business". Graphics and images are more easily accepted and understood than sound, etc., thus they are more popular. Vivid images and rich colors will usually leave an unforgettable impression while data can easily be converted to graphics and images by computer, which can be demonstrated by the three-dimensional moving form.

Data visualization is not only applied to scientific research, but also widely used in production practice. Data visualization is mainly the process of research activities in scientific research while it is the result of productive labor in practice. The result of production can be called intellectual products which maybe is a report expected to visually expressed.

2 The psychology foundation of data visualization

The basic purpose of data visualization is to make the information contained in the data convey to the audience with the highest efficiency. The highest efficiency means that the information is fastly, comprehensively, impressively accepted by the audience. In the human sense organs, only the eyes have the capacity to act as the role.

2.1 The psychophysics of sensation

Sensation is reaction to the individual property of an objective thing which stimulates the brain. The sensory activities of human being rely on sensory organs to complete. The sensory organs are responsible for translating physical energy into nerve energy and producing the corresponding psychological events which can be described by psychological quantity[8]. The psychological quantity is directly proportional to

sensibility, which not only depends on the strength of the stimulus, but also depends on the sense organs. Different sense organs have different sensory thresholds. The sensory threshold is more low, the amount of stimulation of causing sensation, called physical energy, is more small and the sensibility is more high. The sensory thresholds of different sense organs to common things are listed in table 1[8].

Table 1. The sensory thresholds of several common things

Sense organs	Sense objects	Background conditions	Thresholds	Units
Eye	Candlelight	Clear night sky	48.3	km
Ear	Needle sound	Quiet environment	6	meter
Tongue	Sweet taste	one teaspoon of sugar in water	9	liter
Face skin	touch	Bee wings	1	centimeter

From table 1, it is evident that the sensory threshold of vision is the lowest in several senses. In other words, vision is the easiest to cause sensation.

2.2 The special characteristics of vision

Vision is obviously different from other senses. The eyes are the visual organ which not only has the ability to feel the light, but also can distinguish details and colors. Light wave has intensity, wavelength, and purity of three attributes, which decide the brightness, hue and saturation of visual sense of human being. Visual sense also has the properties of being aware of space, time and motion. In contrast, another important sense is hearing which only has pitch, loudness and timbre of three attributes. It is due to a variety of visual characteristics, which determines vision has extraordinary ability to access information.

2.3 Visual perception and consciousness

Perception is human brain's reflection to the whole objective things which act on sense organs. Perception is generated from senses, which integrate and interpret sensory information. Human perception has many attributes in which constancy is particularly important[8]. The constancy includes size, shape and orientation constancies provided mainly by visual perception. Space perception is the basic condition for human being to understand the world[9]. The perception of size, shape, depth and orientation constitute the space perception. These perceptions mainly depend on visual perception. At the same time, time and motion perception for human being to know themselves and the world is also essential. The two perceptions also rely on visual perception. Consciousness is the function of human brain and physical reaction. Perception is the premise of consciousness. Visual perception is the priority among perceptions of human being. On the other hand, language is not only the carrier of consciousness, but also a tool for creates consciousness. The existence and development of language mainly depends on visual perception. Therefore, from a certain angle, visual perception determines consciousness.

3 The basic content of data visualization

Computer graphics is the basis for the development of data visualization. The spatial and temporal attributes of data are the most difficult tasks in research and practice for data visualization. The expression mode is the main and important content of data visualization.

3.1 Data visualization based on computers

The development of computer science and technology has opened a broad prospect for data visualization. From the two-dimensional histogram to the 3D animation, data visualization comes a long way. This is a process from simple to complex, from the abstract to the image. Based on the theory of computer science and technology, the technology development of computer software and hardware makes the process possible. The development of computer hardware technology expands the scope of computer vision, in which human eyes can browse through digital-to-analog converter. Undoubtedly, computer graphics plays an important role in it. Meanwhile, computer graphics is also the design foundation of computer software related with graphics or images. It is necessary for data visualization to select and use these softwares.

3.2 The spatial and temporal attributes of data

The world consists of matter and consciousness. Space and time are the basic forms in which matter exists. Because the consciousness is the reflection on matter, the consciousness must have the brands of space and time through the matter. Data is an abstract concept, which itself does not have the spatial and time attributes. But data always exists as the attributes describing things. There is no data which has no owner. In other words, Data visualization is concerned with the owner of data rather than the data itself. It is possible to realize data visualization through the description to spatial or temporal attributes of data. In fact, the spatial or temporal attributes of data is often a breakthrough in data visualization.

3.3 The expression for data visualization

The expression of data visualization not only depends on the data type, but also depends on the level of computer technology. Different data has different visual characteristics. The progress of computer technology is making the visual scene from 2D to 4D[6]. GDP is usually used to describe the overall economic status of a government in the form of 2D curve. The geographical information data can well demonstrate the process of a city expansion. This data can be displayed with RS images in 2D, 3D or 4D based on a computer and related software. Similarly, a planning for the future of a region can also be shown by this way, but it is necessary to design models and add them into the RS images.

Although the 2D or 3D images can make people understand a scene at a glance, the technology of camera roaming or animation can make this process more vivid[10]. The two techniques meet different needs. It is relatively easy to realize the camera roaming with professional software. The animation usually needs to design models of buildings, objects, persons and their environment. The 4D images are mainly displayed in the form of animation. Naturally, the technology of voice synthesis is used in adopting animation technology. The technology of virtual reality is a branch or an extreme form of data visualization, in which human tactile function is utilized[10][11].

4 The practice of data visualization

The practice is the cognitive purpose. The theory of data visualization like any other theory must be applied to production practice. This is not only the requirement of practice, but also it is the requirements of cognition, because only in practice the theory will be tested and improved.

4.1 The principle of the practice of data visualization

The theory of data visualization is produced in production and practice, which should be applied to the production and practice in order to play its role. This is the second leap of cognition. That means that the operational method or production flow should be established. Moreover, because the data visualization faces various fields, this method or process should be macro and strategic, which can be further refined in different production fields.

4.2 The production process of data visualization

The macro and strategic method can be described with the production flow chart below. Data visualization method includes target decomposition and object classification, data acquisition, data analysis, data processing and modeling, etc.. Data acquisition is the first step in data visualization. The data type determines its visual form. Data analysis is a profound understanding to the collected data. Finally, according to the data types and characteristics, the model construction can be completed by data processing in accordance with the rules of visualization. Video production is the last step, in which pictures, animation, video, voice and subtitles will be synthesized.

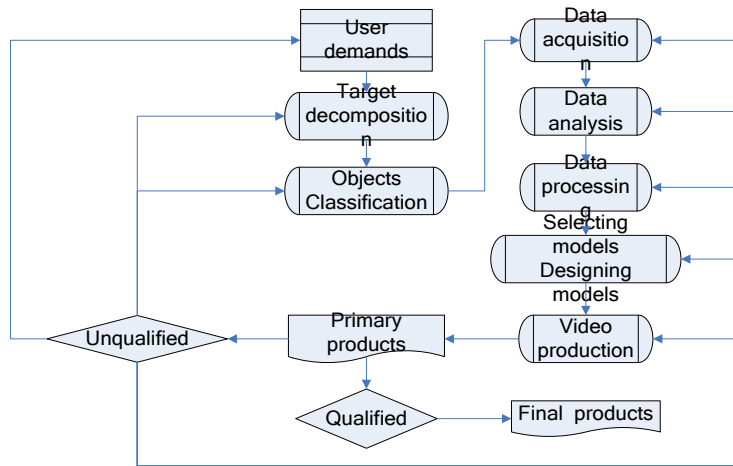


Fig. 1. The production flow chart for data visualization

4.2.1 Target decomposition and objects classification

Data visualization will be facing many problems in production. All the problems can be summarized as a general goal. This is also a process of understand the production task. It is necessary to decompose the target because the task needs to decompose. The purpose of the decomposing task is to discover the work objects. The bigger the goal is, the more the task is, the more the objects are. Depending on different work objects the methods of data visualization may be different. Therefore, classifying the objects is an important part in data visualization.

4.2.2 Data acquisition

Data is the carrier of information which is different from data itself. The fundamental purpose of data visualization is information transmission with effective forms. Therefore, data acquisition is to collect the data containing main information. That means making a choice is necessary. A kind of information has many data forms. It is a wise choice for selecting the data which is easy to be converted into figure or image.

4.2.3 Data analysis

As mentioned above, there are judgment and discrimination in the process of collecting data, which is preliminary. Data analysis is to judge the authenticity, evaluate applicability and necessity to collected data. This process is usually time-consuming. Sometimes, tests have to be done. In the process of data analysis, it is possible to find valuable or new data. There are two main ways to obtain these data, which is to collect it again and another is to produce new data from existing data.

4.2.4 Data processing and designing models

The first step of data processing is to determine the model types which will be used according to the data types, and then the second step is to process data and make it meet the requirements of the models. The content of data processing varies with the

types of data, which include the normalization of attribute data, spatial registration of geographic data and resampling, data calculation and so on. The designing models is a process of selecting the appropriate and professional software to design models with the processed data. In order to achieve a production target, the multiple models may be needed to design.

4.2.5 Video production

The last is producing the product of data visualization. The main content is to synthesize a variety of designed models and other materials into a video in video production software. Obviously, the other materials not only include subtitles, narration, music, etc., but also include other photos, videos, etc. Of course, the video needs the user to accept it. It is general to modify the primary product according to user's requirements. The final product should be satisfied with the user.

4.3 A case--- Agricultural planning visualization

The general goal of the visualization of Agricultural planning is to transform the paper-based traditional report of Agricultural planning into a video file. Although there are many types of Agricultural planning, but it is its basic characteristics to analyze the present situation and design the future targets to a planned area. Therefore, the visualization task can be divided into two categories. One is the description to the current situation. The other is the future planning.

The status data can be summarized as two kinds of data of Agricultural natural conditions and the social economy. The former, as a class of work objects, can be expressed by modeling with professional GIS software[9]. With office software, the latter, as another class of work objects, can be displayed by histograms, pie charts, curve graphs and other forms. As an example, using GIS software modeling, the picture below is a 3D display of terrain and landform of parts of Alukerqin county in Inner Mongolia, in China.

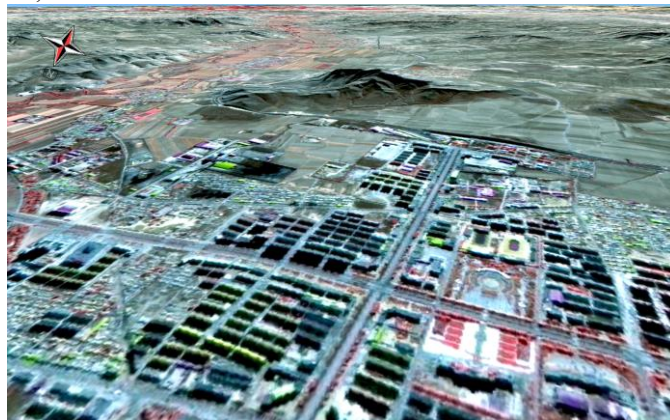


Fig. 2. The terrain and landform of parts of Alukerqin county in Inner Mongolia (Image: GF1-PMS, B432; DSM, patial resolution: 15meters; Author: Li Xiaochen)

The data for future planning is mainly explained by various models. These models include scenes, Agricultural facilities, crops, animals, and so on. It is often needed to design these models, rather than using photos, with professional software. The picture below is a scene model designed by a student of Mars Time Inc.



Fig. 3. A scene model (Author: Zhu Yihan, a student of Mars Time Inc)

5 Conclusions

Data visualization is a science and technology research on data visual form, which is related with scientific visualization and information visualization. Based on computer graphics and images data visualization displays and reveals the main information contained in data. The basic purpose of data visualization is to make the information contained in the data convey to the audience with the highest efficiency. Data visualization has a solid psychological foundation, which expands the scope of visual perception. The development of computer science and technology has opened a broad prospect for data visualization. Three-dimensional or four-dimensional animation is the advanced form of data visualization. The technology of virtual reality is a branch or an extreme form of data visualization. The method for data visualization includes target decomposition and object classification, data acquisition, data analysis, data processing and modeling, etc..

Acknowledgement

This paper is supported by Innovation Team of Crop Monitoring by RS (CMIT), authorized by Chinese Academy of Agricultural Engineering (CAAE), in 2015.

References

1. Cortez, P. & M.J. Embrechts.: Using sensitivity analysis and visualization techniques to open black box data mining models[J]. *Information Sciences*. 2013, 225:1-17
2. Belilovsky, E. & K. Gkirtzou & M. Misyrlis. etc.: Predictive sparse modeling of fMRI data for improved classification, regression and visualization using the k-support norm[J]. *Computerized Medical Imaging and Graphics*. 2015, 46:40-46
3. Cheng, T. & J. Teizer.: Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications[J]. *Automation in Construction*. 2013, 34:3-15
4. Shao, Y.L. & Y.S. Liu & C.G. Li.: Intermediate model based efficient and integrated multidisciplinary simulation data visualization for simulation information reuse[J]. *Advances in Engineering Software*. 2015, 90:138-151
5. Marcus, D.S. & M.P. Harms & A.Z. Snyder. etc.: Human connectome project informatics: quality control, database services, and data visualization[J]. *NeuroImage*. 2013, 80:202-219
6. Shimizu, F. & M. Uehara & M. Oatari. etc.: Three-dimensional visualization of the human face using DICOM data and its application to facial contouring surgery using free anterolateral thigh flap transfer[J]. *Journal of Plastic, Reconstructive & Aesthetic Surgery*. 2016, 69:e1-e4
7. Kim, J.H. & V. Lyer & S.B. Joshi. etc.: Improved data visualization techniques for analyzing macromolecule structural changes[J]. *Protein Science*. 2012, 21:1540-1553
8. Meng, Z.L.: *General psychology*[M]. Beijing: Peking University Press, 2014, 83-153 (in Chinese)
9. Yang, J.F. & Q. Zhang.: Research on spatial cognition based on geographic information visualization[J]. *Geomatics & Spatial Information Technology*. 2013, 36(7) :12-14(Chinese)
10. Zong, C.Y.: The research of virtual reality technology in the animation and roaming building models[J]. *China Science and Technology Information*. 2015, 2:92-93(Chinese)
11. Hao, S.: The prospect of virtual reality technology and its application[J]. *Coal Technology*. 2013, 32(5):160-162(Chinese)