Visual Data Mining in Immersive Virtual Environment Based on 4K Stereo Images

Tetsuro Ogi, Yoshisuke Tateyama, and So Sato

Graduate School of System Design and Management, Keio University
4-1-1 Hiyoshi, Kohoku-ku, Yokohama 233-8526, Japan
ogi@sdm.keio.ac.jp, tateyama@sdm.keio.ac.jp,
sosato@z8.keio.jp

Abstract. In this study, super high-definition immersive visual data mining environment using 4K stereo projector was developed. In this system, data can be represented with high accuracy in the three-dimensional space using the super high-definition stereo images, and the user can recognize the relation among several kinds of data by integrating them in the immersive environment using the plug-in function. This system was applied to the seismic data visualization and the effectiveness of this system was evaluated.

Keywords: Visual Data Mining, Immersive Virtual Environment, 4K Stereo Image, Seismic Data Analysis.

1 Introduction

In recent years, a large amount of data has been recorded in various fields according to the advance of database technologies. However, since the discovery of valuable information from the huge data is difficult, it is desired to establish a methodology to utilize these data effectively. Though data mining is one of the methods to solve such a problem, it was difficult for the user to understand the process of mining data and evaluate the result of it, because processing data is performed only in the computer in the usual data mining method [1]. Therefore, this study focuses on the visual data mining in which the user can analyze data interactively with the computer by visualizing the process and the result of the data mining.

In the visual data mining, it is expected that an new information can be discovered and analyzed by combining the ability of data processing and high-speed calculation of the computer with the human's ability such as intuition, common sense and creativity [2][3]. Particularly, this study aims at improving the effect of visual data mining by enhancing the ability of expressing data and interaction function by using the advanced display technologies. In this paper, system architecture and visualization ability of the super high-definition immersive visual data mining environment that was developed in this study are discussed and it was applied to the seismic data analysis.

2 Concept of Immersive Visual Data Mining

Visual data mining supports the user to analyze a large amount of data and to find new information by integrating the abilities of computer and human. The computer

R. Shumaker (Ed.): Virtual and Mixed Reality, LNCS 5622, pp. 472–481, 2009.

has large memory and high speed calculation capability. On the other hand, human is excellent in intuition, commonsense and creativity based on the visual information. By combining the information processes performed by computer and human, a large amount of data could be analyzed effectively to discover new information.

In this case, visualization technology is used to transmit the information from the computer to the human and interactive interface environment is used to support the collaboration between human and computer. If the expression abilities of information visualization and the interaction function between human and computer were increased, the performance of visual data mining would be improved based on the improvement of information processing of human and computer.

As for the increase of the expression ability in the information visualization, super high-definition image would enables the transmission of the detailed information from the computer to the human, and three-dimensional stereo image would be used effectively to represent the relationship among several kinds of data in the three-dimensional space. And as for the interaction between human and computer, immersive virtual environment could be used to enable the user to operate the visualized data directly and to explore data space as a first person experience [4][5][6]. Figure 1 shows the concept of super high-definition immersive visual data mining that is proposed in this study. In this figure, it is shown that the performance of visual data mining can greatly be improved by increasing the expression ability and the interaction function between computer and human.

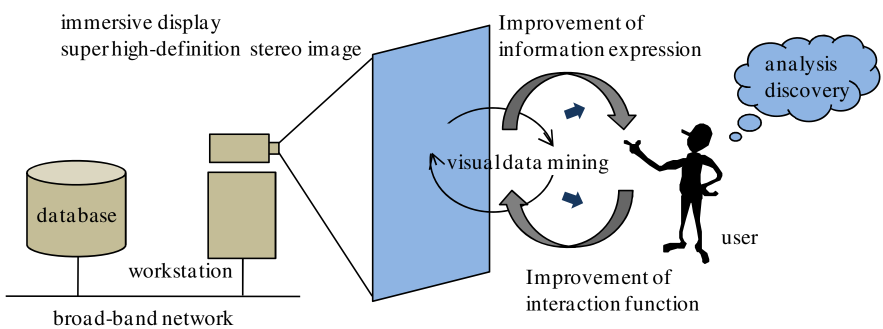

Fig. 1. Concept of super high-definition immersive visual data mining. Improvement of information expression and interaction function increase the performance of visual data mining.

3 Visual Data Mining Environment Using 4K Stereo Image

3.1 CDF System

In this study, CDF (Concurrent Design Facility) system that uses 4K stereo image was developed as a super high-definition immersive visual data mining environment. Figure 2 shows the appearance of the CDF system. CDF is designed for various purposes such as the collaborative design, presentation, and lecture in the university as well as for the visual data mining environment in which users can analyze data and find new knowledge while sharing the visualized data among users [7].

Fig. 2. CDF (Concurrent Design Facility) system. This system consists of 4K stereo projector and two LCD monitors.

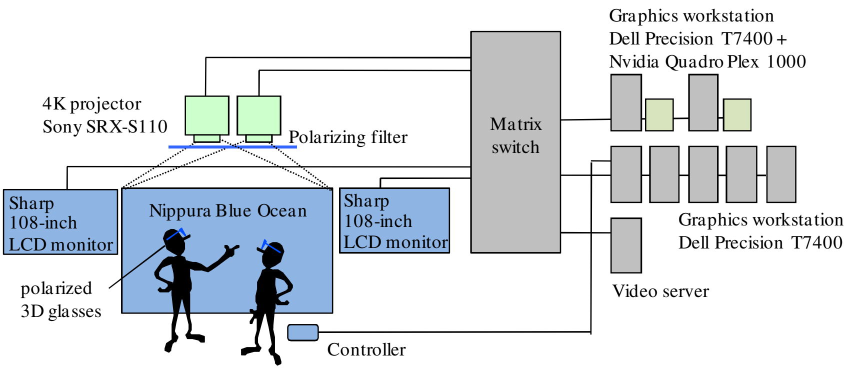

Fig. 3. System configuration of CDF system. 4K projectors and LCD monitors are connected to the graphics workstations through matrix switch.

The CDF consists of a 180-inch center screen for the 4K stereo projectors and two LCD 108-inch high-definition monitors placed at both sides. The center screen can be used not only to project 4K resolution stereo image but also to display four high-definition divided images as a multi screen monitor. Therefore, the whole of CDF system can display up to six high-definition images and it can be used as a data visualization and analysis environment in various application fields.

In order to generate images, eight graphics workstations (two workstations for 4K stereo image and six workstations for high-definition images) are used, and output image of the arbitrary graphics workstation can be sent to the arbitrary display area freely by switching the connections in the matrix switch. Figure 3 shows the system configuration of the CDF system.

3.2 4K Stereo Image

The main component of the CDF is 4K stereo projector system. 4K means super high resolution image of 4,096 x 2,160 pixels, and its image quality would be more than four times of the usual high-definition image [8]. This system uses stacked two 4K projectors (Sony, SRX-S110), and the images corresponding to the right eye and left eye positions of the user are projected. The image output from the projector is rear projected onto the screen through the polarizing filter, and the user can see the passive stereo image based on the binocular parallax by wearing the polarized 3D glasses.

The brightness of each 4K projector is 10,000 ANSI lumens, and an acrylic rear projection screen (Nippura, Blue Ocean Screen) is used to represent a high contrast image. As a result, the brightness of the images displayed on the center screen and the LCD monitors placed at both sides becomes almost same levels.

As for the computer to process data and generate 4K stereo image, two high-end workstations (Dell Precision T7400, 2xQuad Core Xeon 3.2GHz) with the graphics engine that has a genlock function (NVIDIA Quadro Plex 1000 Model IV) are used. In this system, interface devices such as an USB game controller and a wireless touch panel are used. The game controller is used to change the visualization method or move the view point, and the touch panel is used to switch the connections between the input images and the output devices. This system configuration of the CDF system realizes the super high-definition immersive visual data mining environment using the 4K resolution stereo image.

4 Data Representation Ability of CDF

In the data visualization using the 4K image, information can be represented with high accuracy based on the high resolution image compared with the visualization using the usual XGA or SXGA resolution image. Particularly, when the 4K stereo image is used, it is expected that the three-dimensional image with high spatial resolution can be displayed. In this study, the spatial resolution with which the visualized data can be represented in the three-dimensional space was experimentally measured.

In the experiment, the subjects sat at a position 3m away from the screen, and two parallel perpendicular lines with random intervals of 0.5mm, 1mm, 2.5mm, 5mm, or 10mm were displayed in the direction from the front to left. The subjects were asked to move the parallel lines in the depth direction using the controller, and to stop them at the limit position where they can recognize the displayed image as two lines and this position was recorded. Figure 4 shows the appearance of this experiment and Figure 5 shows the result of the experiment for five subjects with visual acuity of more than 1.0. In this graph, the contour lines are drawn by connecting the average values of the perceived depth positions for each interval of the parallel lines.

From this graph, we can understand that the data can be represented at the resolution of 10mm intervals at the distance of about 10m from the subjects, and it can be represented with high accuracy at the resolution of 0.5mm intervals at the distance of about 50cm. Though these values are a little bad compared with the theoretical values calculated from the interval of one pixel (0.97mm) on the screen, we can consider it is caused by the stereo vision image. In the inclined direction, the distance in which the subjects recognize the parallel lines becomes a little far, since the pixel interval on the screen looks narrow when it is seen from the slant direction. For example, the

Fig. 4. Appearance of the experiment on measuring the spatial resolution of data representation in CDF system

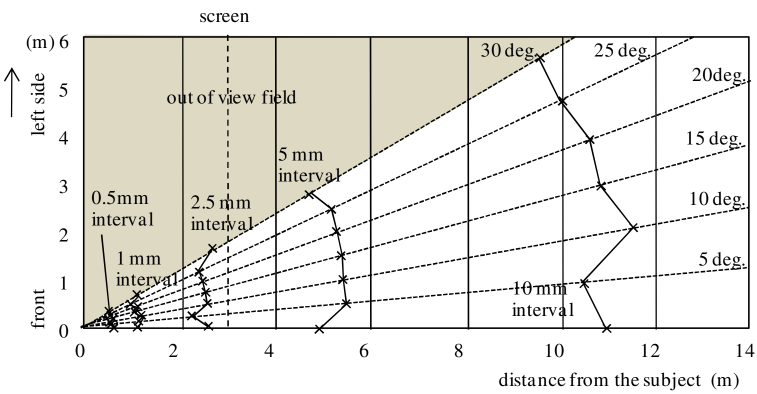

Fig. 5. Result of experiment on data representation ability of CDF system. Contour lines show the limit position where the subjects can recognize the parallel lines.

distance of the recognition of the parallel lines with 2.5mm interval in 30 degrees direction was about 1.2 times as long as the front direction. From the result, it is understood that the 4K stereo display can represent the detailed information with high accuracy, though the spatial resolution of the displayed image depends on the distance and direction from the user. Namely, this means that the visual data mining using the super high-definition stereo image has an ability to transmit a large amount of information from the computer to the user with high accuracy.

5 Application to Seismic Data Analysis

In this study, the super high-definition visual data mining environment was applied to the visualization and analysis of the earthquake data.

Fig. 6. Visualization of hypocenter data using 4K stereo projection system. User can recognize both each hypocenter data and distribution of whole data.

5.1 Distribution of Hypocenter Data

In Japan, since a lot of earthquakes occur every year, it is very important to analyze the feature of the hypocenter distribution for the earthquake prediction and the damage measure [9]. However, it was difficult to observe both the individual location and the overall distribution of the hypocenter data and to understand the feature of three-dimensional distribution of the earthquake by visualizing data using the conventional two-dimensional display. Therefore, the visualization system for the hypocenter data was constructed by using the super high-definition immersive visual data mining environment.

Figure 6 shows that the hypocenter data recorded in the past three years in Earthquake Research Institute of the University of Tokyo is visualized by using the 4K projection system. In the super high-definition display, the user can recognize the individual hypocenter while seeing the spatial distribution of the hypocenter in the whole of Japan, since the data can be represented with high resolution in the three-dimensional space with wide field of view. In this system, when the 1:10,000 scale map is displayed, the user can recognize the hypocenter positions located 5m away between each other at 50cm away in front of him. By using this system, the user could understand intuitively that the depths of hypocenters in west Japan is relatively shallow and hypocenters occurred in Tokai and Kanto area are distributed along a plate.

5.2 Relationship between Hypocenter and Other Data

Next, this system was applied to analyze the relationship between data by integrating different kinds of visualization data. In this application, map data, hypocenter data, terrain data, basement depth data and plate structure data were integrated in the immersive visualization environment. Though these data are stored in different database, they have location information. Therefore, they can be related with each other using the location information as a reference key. Table 1 to Table 4 show the database structure used in this system. Though the terrain data originally consist of altitude information corresponding to the latitude and longitude on the lattice, they are stored as data sets for the divided areas in the database.

In this system, the user can accesses these databases from the virtual environment and visualize the retrieved data by specifying the necessary condition. Thus, this system enables the user to understand the relationship among the hypocenter, terrain, basement depth, and plate structure data intuitively, and to analyze the feature of the earthquake phenomenon, by overlapping these data in the three-dimensional visualization environment.

Table 1. Database structure of hypocenter data

Latitude	Longitude	Depth	Magnitude	Date
36.3005	139.9837	40.68	0.8	2003-01-01
36.0927	138.7390	153.97	1.7	2003-01-01
36.2901	139.6655	121.42	1.3	2003-01-01
…..	…..	…..	…..	…..

Table 2. Database structure of basement depth data

Latitude	Longitude	Depth
36. 505	138.510	3.057
36.505	138.515	2.801
36.505	138.520	2.661
…..	…..	…..

Table 3. Database structure of plate structure data

Latitude	Longitude	Depth
36.510	138.460	151.827
36.510	138.465	151.103
36.510	138.470	151.371
…..	…..	…..

Table 4. Database structure of terrain data

North edge in latitude	West edge in latitude	South edge in latitude	East edge in latitude	Filename
36.6051	139.8046	36.5895	139.8202	N36.6051_E139.8046.dat
36.6051	139.8202	36.5895	139.8358	N36.6051_E139.8202.dat
36.6051	139.8358	36.5895	139.8514	N36.6051_E139.8358.dat
…..	…..	…..	…..	…..

As for the mechanism of integrating the visualized data in the virtual space, the plug-in function of the OpenCABIN library was used [10]. In this method, each data is visualized by different application programs and they are integrated in the three-dimensional space in the runtime. Figure 7 and Figure 8 show that several visualization data are integrated in the same space. In these examples, visualization programs for hypocenter data, terrain data, basement depth data and plate structure data are plugged-in to represent the relation among the data. The hypocenter data is visualized using the sphere, and the size of it indicates the magnitude of the earthquake. The terrain data is created by mapping the texture image captured from the satellite onto the shape model. And the basement depth and the plate structure data are represented using the colors that indicate the depth values.

Fig. 7. Integrating visualization data in CDF. The relation between terrain data and hypocenter data are shown.

Fig. 8. Integrating visualization data in CDF. User can understand the relation among hypocenter data, basement depth data and plate structure data.

When the several visualization programs are plugged-in, the toggle buttons that show the conditions of each data are displayed. The user can switch the visible condition of each data by using the toggle button in the virtual space. For example, the user can change the visualization data from the combination of hypocenter data and basement depth data to the combination of hypocenter data and plate structure data, while running the application programs in the virtual space. By using this method, the user can intuitively understand the feature of the distribution of hypocenter and the relation with other data. For example, user could see whether the attention data of the earthquake occurred on the plate or in the plate structure. Thus, this system could be effectively used to represent the relationship among several data in the three-dimensional space and to analyze the earthquake phenomenon.

6 Conclusions

In this study, the super high-definition immersive visual data mining environment that uses the 4K stereo projector was constructed. In this system, the effect of visual data mining based on the collaborative process between the user and the computer would be greatly improved, since the 4K stereo image increases the amount of information transmitted from the computer to the user and the interactive interface enables the user to explore the data space. The developed prototype system was applied to the seismic data analysis, and several kinds of visualization data such as map, hypocenter, terrain model, depth of basement and plate structure were integrated. Then, the effectiveness and possibility of the intuitive analysis in the immersive visual data mining environment were confirmed through the interactive operation to the visualized data. Future research will include developing more effective visual data mining method through the collaboration with the earthquake experts and applying this technology to other application fields.

Acknowledgements

This research was supported by G-COE (Center of Education and Research of Symbiotic, Safe and Secure System Design) program at Keio University. And we thank Dr. Takashi Furumura of the University of Tokyo, Dr. Shoji Itoh of Riken, and Satoshi Oonuki of University of Tsukuba for their support.

References

1. Kantardzic, M.: Data Mining: Concepts, Models, Methods, and Algorithms. Wiley-IEEE Press (2002)
2. Wong, P.C.: Visual Data Mining. Computer Graphics and Applications 19(5), 20–21 (1999)
3. Keim, D.A.: Information Visualization and Visual Data Mining. IEEE Transactions on Visualization and Computer Graphics 7(1), 100–107 (2002)
4. Ammoura, A., Zaïane, O.R., Ji, Y.: Immersed visual data mining: Walking the walk. In: Read, B. (ed.) BNCOD 2001. LNCS, vol. 2097, pp. 202–218. Springer, Heidelberg (2001)
5. Wegman, E.J., Symanzik, J.: Immersive Projection Technology for Visual Data Mining. Journal of Computational and Graphical Statistics 11, 163–188 (2002)
6. Ferey, N., Gros, P.E., Herisson, J., Gherbi, R.: Visual Data Mining of Genomic Databases by Immersive Graph-Based Exploration. In: Proc. of 3rd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia, pp. 143–146 (2005)
7. Ogi, T., Daigo, H., Sato, S., Tateyama, Y., Nishida, Y.: Super High Definition Stereo Image Using 4K Projection System. In: ICAT 2008, Proceedings of 18th International Conference on Artificial Reality and Telexistence, pp. 365–366 (2008)
8. Sony Press Release: Sony Unveils New "4K" Digital Cinema Projector (June 2004), `http://news.sel.sony.com/pressrelease/4864`

9. Furumura, T., Kennett, B.L.N.: Subduction Zone Guided Waves and the Heterogeneity Structure of the Subducted Plate: Intensity anomalies in northern Japan. Journal of Geophysical Research 110, B10302.1–B10302.27 (2005)
10. Tateyama, Y., Oonuki, S., Ogi, T.: K-Cave Demonstration: Seismic Information Visualization System Using the OpenCABIN Library. In: ICAT 2008, Proceedings of 18th International Conference on Artificial Reality and Telexistence, pp. 363–364 (2008)