Cubic Gantt Chart as Visualization Tool for Learning Activity Data

Shohei NAKAMURA[a], Kosuke KANEKO[b], Yoshihiro OKADA[a, b], Chengjiu YIN[c], Hiroaki OGATA[b, c]

[a]*Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan*
[b]*ICER(Innovation Center for Educational Resources), Kyushu University Library, Japan*
[c] *Faculty of Arts and Science, Kyushu University Fukuoka, Japan*

Abstract: This paper proposes a new tool called Cubic Gantt Chart, 3D version of Gantt Chart for visualizing e-learning activity data such as how long and how often students look at each page of an e-learning material like PowerPoint slides. Using a dedicated tool called BookLooper can collect these activity data. The analysis of e-learning activity data is significant for instructors to understand students' learning states and the difficulty levels of e-learning materials because those can suggest more efficient instructional ways. So, in this paper, the authors propose Cubic Gantt Chart as a visualization tool for the analysis of students' e-learning activity data and justify its usefulness by showing visualization examples of actual data collected by BookLooper.

Keywords: Visualization, Activity Data, e-Learning, Instructional Design

1. Introduction

Recently, researches on cyber-physical systems have become more popular because various electric appliances have sensor equipment to gather some data through the Internet called the cyber world and by analyzing those data, it becomes possible to provide more efficient services to the physical world. As for e-learning based education environments, if we could understand students' e-learning activities by analyzing their activity data, we can employ more efficient instructional ways more suitable to those students. To do so, instructors need any tools that help them for collecting and analyzing students' e-learning activity data.

We employ BookLooper for collecting those data. It watches students to store their e-learning activity data such as how long and how often the students look at each page of an e-learning material like PowerPoint slides. We explain for details about BookLooper in section 4.1.

On the other hand, one of the most effective tools for analyzing those data is the visualization tool. It is important to apply the statistical methods to testify the fact of data user finds, but user needs to know what the data mean directly at first. That is more effective to notice the new facts, build up the hypothesis and discuss about data. Therefore, we propose a new visualization tool called Cubic Gantt Chart, 3D version of Gantt Chart in this paper. We also justify the usefulness of the proposed Cubic Gantt Chart by showing visualization examples of actual data collected by BookLooper.

Gantt Chart (2D version) is well known as a project management tool. It is one of the bar chart and proposed by Henry L. Gantt. Vertical line represents a kind of work and horizontal line represents time flow. Cubic Gantt Chart is the better tool to visualize the data which has multidimensional attributes. User select three from the attributes and each attribute is assigned to the each axis and he/she can look down the whole graph or look in any orthogonal direction, for example, X-Y plane, Y-Z plane and Z-X plane.

The remainder of this paper is organized as follows. First of all, Section 2 describes related works of educational data and visualization tools for them. Section 3 introduces basic functionalities of Cubic Gantt Chart we proposed. Then, Section 4 describes the usefulness of Cubic Gantt Chart by adopting actual data collected by BookLooper. In Section 5, we discuss what those data mean when we visualize them by Cubic Gantt Chart. Finally, we conclude the paper and describe our future work.

2. Related works

2.1 Learning data

The analysis of students' learning activities is more important for the effective education rather than the check of semester's end tests. According to Darrell et al [1], using online tools allow instructors to recognize students' activities, for instance, how long they read online electronic textbooks and how they understand the important parts of them. There is also the research about relationship between Self-Regulated Learning (SRL) and students' learning behavior by using learning data [2]. When students do e-learning with a smart phone, SCROLL is one of the effective e-learning systems using learning activity data [16].

2.2 Analyzing data

The way which decide the solution by analyzing data has been used and popular since a decade ago. However, this does not always lead a correct strategy. Picciano et al [3] stated "As data-driven decision making has entered learning analytics era of the big data, these new approaches are not silver bullets while they may be the part of solution. Higher education administrators should employ them by evaluating whether they can be used in their institutions and by recognizing what roles they can play."

Visualization of data is one of the ways to analyze them. User can notice the new fact by looking at visualization results of the data however it can be found by statistical methods. It can be said that analysis by looking at the data enables the user to understand what the data mean more directly, i.e., interactively and intuitively. Paul et al [4] can insist a certain hypothesis by visualizing the multivariate data using multiple connected view.

2.3 Visualizing methods

There are a lot of visualization tools so far [5]. For instance, Tree Map, Bullet Chart, Parallel Coordinate and so on. It is too many to explain them, respectively. Although the visualization of time oriented data is not easy task, various visualization methods have been proposed so far [6].

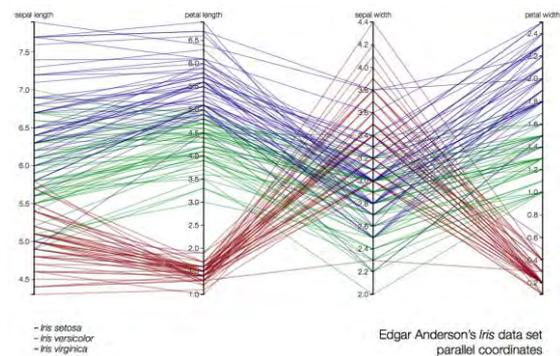Figure 1. Treemap (from D3.js example)

Figure 2. Parallel coordinates (from D3.js example)

Treemap is one of the popular visualization tools. It is proposed by Ben Shneiderman[13]. It displays hierarchical data as a set of nested rectangle. Each rectangle's area is decided by the weight value assigned to the corresponding node. The color can be used as another information of the node. Parallel coordinates is one of the methods to visualize multidimensional data. It is proposed by Alfred Inselberg et al [12]. Parallel coordinates has parallel multi-lines in 2-D plane, each of which represents an axis for each attributes of data. Making the line chart by connecting vertices on each axis, user can see relationships among all attributes of each data. Aigner et al [10] pointed out time, data and representation as the three main criteria for the visualization of time oriented data. Time means how time axis is represented, data and representation mean what kind of data and how those are visualized, i.e., dynamic graph or static graph. Time-tunnel is a visualization tool for time-series numerical data proposed by Akaishi, et al [7, 8, 9]. Time-tunnel is mainly composed from three components those are called Data-wing, Time-plane and Time-bar respectively. Time-bar is a long thin cylinder works as a

time-axis. Data-wing is like a sheet. One time-series numerical data is displayed as one chart on each Data-wing. Each Data-wing is connected to Time-bar and by rotating several Data-wings and putting them together, it is possible to compare the multiple time-series numerical data displayed on them. Time-plane is also like a sheet connected to Time-bar perpendicular to Data-wing. It is used for displaying a radar chart. Users can understand the relationship among multiple data which are extracted at any time.

3. Basic functionalities of Cubic Gantt Chart

This section introduces our proposed visualization tool called Cubic Gantt Chart. Cubic Gantt Chart is 3D version of Gantt Chart. Before explaining its details, we introduce the standard Gantt Chart (2D version).
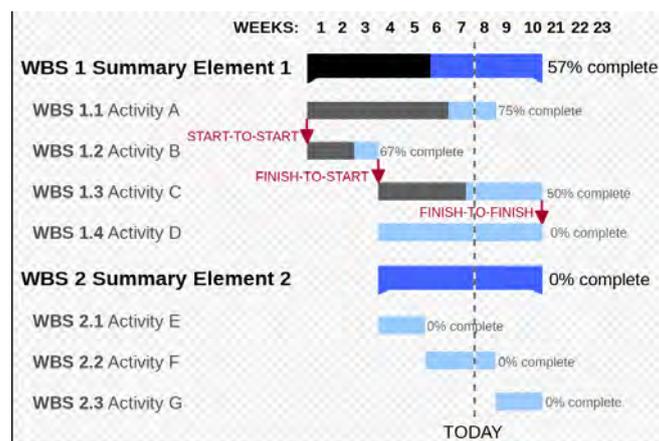
3.1 Gantt Chart (2D version)

Figure 3. Gantt chart (from Wikipedia)

Gantt Chart proposed by Henry L. Gantt in 1910s is a kind of bar chart. It is well known as a project management tool. Vertical line represents a kind of works and horizontal line represents time flow. Each job is illustrated by rectangles. Left and right side of each rectangle are the start and end time of the corresponding job. This graph allows users to understand the below at once.

- What kind of activities there are
- When each activities starts and ends
- How long each activities are
- What activities overlap to other activities
- How far project goes
- When project starts and ends

3.2 Cubic Gantt Chart (3D version of Gantt Chart)

This visualization tool uses a 3D space so that it can treat three attributes, one additional attribute besides two attributes of the standard Gantt Chart, each of those is assigned to each X, Y and Z axes of the 3D space, separately. By looking at the tool in the direction toward any face, i.e., X-Y plane, Y- Z plane or Z-X plane, the user can understand relationships between the corresponding two attributes of the data.
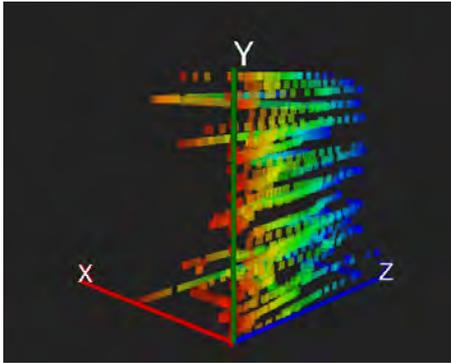
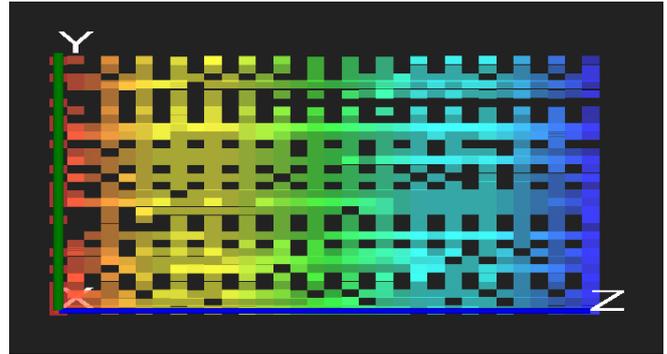Figure 4. Perspective view of Cubic Gantt Chart

Figure 5. Orthogonal view of Cubic Gantt Chart

When a user looks at this graph, he/she can select the perspective or orthography camera (See Fig. 4 and Fig. 5). A user selects the perspective camera if he/she wants to look down at the whole graph and user selects orthography camera if he/she wants to look only two attributes relationship such as X-Y plane, Y-Z plane or Z-X plain.

Each voxel located in this visualization tool has semi-transparent. Therefore, users can understand how voxels are overlapped by checking those color depth and it represents the information same as that of heat map. A user can easily find out where he/she should pay attention to at first.

The length between the two sides of each voxel represents the value of that attribute if the attribute is continuous variable. If the attribute has discrete values, the length is always one and the position of each voxel treats as the same meaning as that of a standard Gantt Chart. The standard Gantt Chart represents time flow in horizontal axis and a kind of works in vertical axis. If the horizontal length of a rectangle is one that means the task takes one minute. The first location at vertical axis represents a first task, the second location represents a second task and so on.

The color of each voxel is reddish if it is located in at begin point and bluish if it is located end point in Z axis. We explain the details what colors represent in Section 4.

Cubic Gantt Chart offers not only the functionalities mentioned above but also the functionality similar to the standard Gantt Chart. By selecting any plane, i.e., any two attributes of three attributes at a certain value of the remaining attribute, the proposed tool visualizes the selected data similarly to the standard Gantt Chart. Users can understand the features of the selected data quickly like the standard Gantt Chart.

3.3 Web based visualization tool

For implementing the proposed visualization tool, we selected JavaScript as a programming language although there are many programming environments. Its reason is that JavaScript contents can be used in any web browser on any hardware platform without installing the dedicated software. Another reason is that there are a lot of libraries such as Three.js can be used when making applications like visualization tools.

4. Visualization Examples about eBooks Access Logs

This section presents visualization examples of Cubic Gantt Chart about students' e-learning activity data actually collected using BookLooper. Cubic Gant Chart is most suitable for the visualization of students' e-learning activity data because it can visualize the transition pattern among pages of PowerPoint slides and the stay time of each page about each student, and it also makes possible to atomically understand the same information among all students. In the following subsections, first of all, we introduce the detail of BookLooper data. After that, we shows the visualization examples of Cubic Gantt Chart.

4.1 Detail of BookLooper data

BookLooper developed by KYOCERA Communication Systems Co., Ltd is one of the cloud services in which users can read electronic textbooks stored into the service through the Internet. Users can access this service on any platform, i.e., any device and any OS. This service gathers users' learning activity data of reading electronic textbooks, e.g., how long each user reads an electronic textbook, from and to which pages of the textbook the user traverse. The attributes of such data are date, time, user name, material name, activity information represented as a sequence of reading page indices. As other functionalities of Booklooper, users can add a marker to important places and can also write memo.

Our university has promoted BYOD (Bring Your Own Device) and asked students bring their own device for taking a class since 2013 and also asked to use an electronic textbook instead of the paper textbook through BookLooper. As a result, a huge number of learning activity data have been stored in a database.

In this paper, we employ learning activity data of information technology class collected by BookLooper from October 2014 to January 2015.

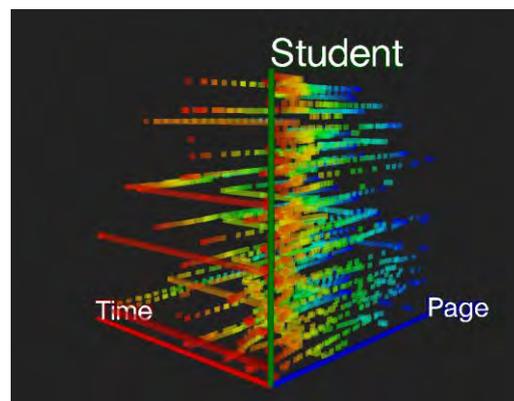## 4.2 Visualization Examples

Figure 6. Perspective view of Cubic Gantt Chart.

In the followings, we introduce several visualization examples of the proposed Cubic Gantt Chart for students' learning activity data actually BookLooper gathered. First of all, Fig. 6 is the graph when the user select the perspective camera to look down in this graph. There are three long red voxels, which mean three students open the textbook at first page and those students did not look at other pages.
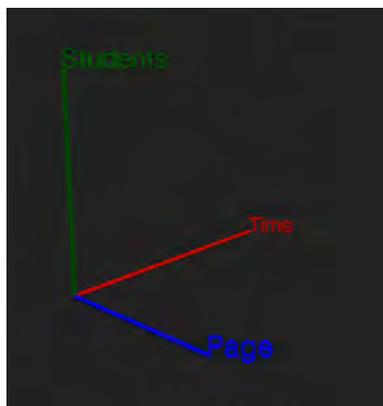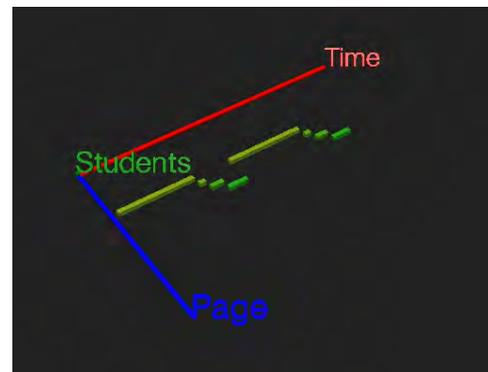

Figure 7. The axes in Cubic Gantt Chart.

Figure 8. The length of a voxel toward x-axis represents reading time of a certain student for a certain page.

We used reading time of each page, student ID and page index of PowerPoint Slides those are assigned to X (red), Y (green) and Z-axes (blue), respectively as shown in Fig. 7. As shown in Fig. 8, each rectangular solid called a voxel means one recorded data consisting of page index, student ID and

reading time of the student for the page. The length of each voxel toward x-axis represents reading time of the student for the page. For example, the length of voxel toward x-axis has one if the student look at the page for one minute. Therefore, users can understand how long and which page in this material students read. The each voxel of the material represents start time and end time like the standard Gantt Chart. So, users can find out when and which page in this material students read.

Each voxel has a color value corresponding to its attribute value. For instance, early pages of the material have a reddish color and ending pages have a bluish color. Since it also has a semi-transparency value, a deep color means a higher density area of voxels. Therefore, as shown Fig. 9, by checking the color depth, it is possible to easily understand which pages are frequently read, who spend much time to read the page.
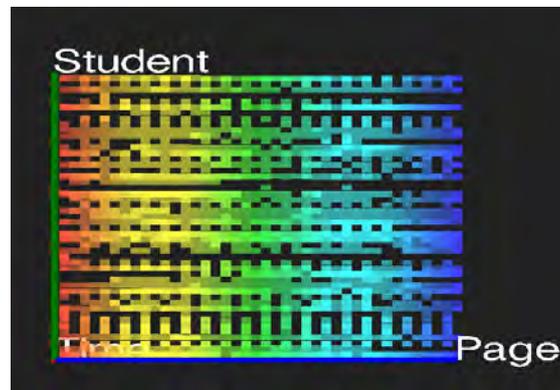

Figure 9. Orthogonal view toward the plane of student ID and Page indices.

Fig. 9 indicates relationships between students and page indices of the material. Voxels in some lines has a deep color which mean those students read the page many times or by spending much time. You can also understand some students read pages skipping hear and there, and some students did not read any pages at all because no voxels are in those students' line. There are not only students who learn hardly but also students who do not learn at all. And also, you can understand which pages students read frequently by checking the color depth. You can also understand that some pages are difficult for almost students to understand in this material if those pages need to be looked at for long time. For example, the areas of yellowish and sky bluish have a deep color so that it can be said that those pages are difficult for students because they looked at many times or by spending much time.
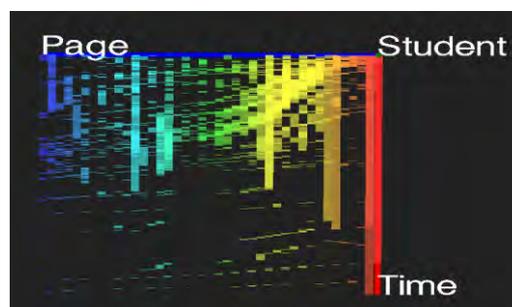

Figure 10. Orthogonal view toward the plane of reading time and Page indices.

Fig. 10 indicates relationships between page indices of the material and reading times for them. Firstly, it can be seen that some students looked at the ending pages for long time. By checking reading times, all students take the same actions at early time because of those color depth, and there are some students who gradually came to spend time back to early page and some students who came to spend time toward ending pages. There are also some students who have been looking at a first page. They may listen to teacher hardly or they may be asleep after opening their textbook in this lecture. Students reading the ending page at early time may do homework assigned in this lecture. The areas of sky bluish voxels and yellowish voxels have a deep color at middle time of this lecture. So, it can be said that at that time, almost students read the pages of sky bluish or yellowish colors many times with spending much time.
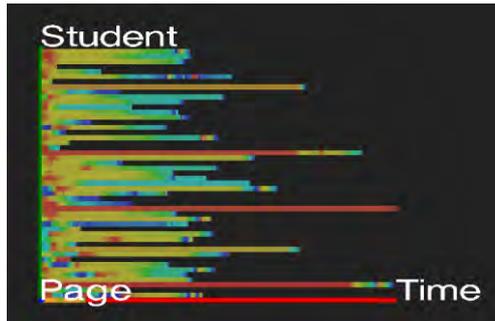
Figure 11. Orthogonal view toward the plane of student ID and reading time.

Fig. 11 indicates relationships between students and their reading times. From this viewpoint, you can compare students learning activities with each other by checking those color changes. For instance, you can understand learning activities are the same if those voxels' color changes and total lengths are the same.

In Figure 11, you can see three red long voxels those are corresponding to the first pages. This means that three students spent long time to read the first pages. Additionally, you can also see that yellow voxels occupy the most areas. So, some students spent moderate time. On the other hand, some students spend long time to understand this page. This means the page of the material must be important part in this lecture. If students can easily understand the material, the changing color becomes red to blue. If students read back frequently, the changing color becomes reddish on the way. You can also see long sky bluish voxel. This means students spent long time to read the ending pages.

The total length of voxels in each student is different. This means that students stopped the BookLooper application at different time or some students stopped on their learning. Anyway, this lecture ended earlier than the regular time because almost students stopped BookLooper at the half of the maximum total length.
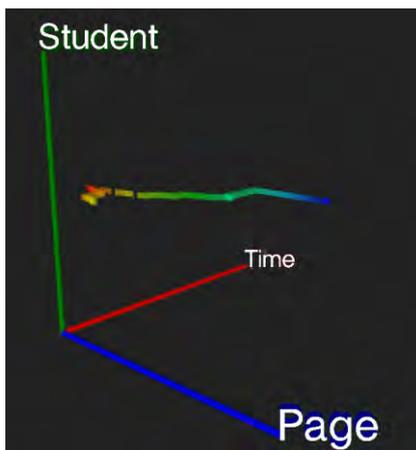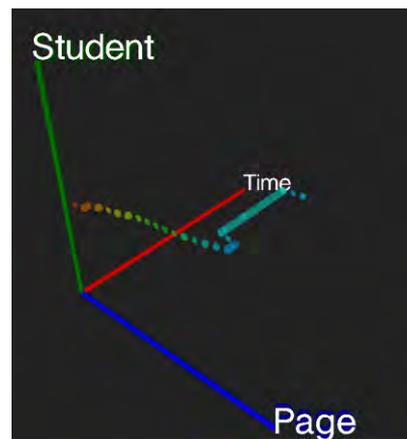

Figure 12. One student activity data

Figure 13. Another student activity data

It is possible to look at not all students' activity data as shown in Fig. 11, but also one of the students' activity data as shown in Fig. 12 and Fig. 13. In Fig. 11, there are several students who did not spend long time at the page represented as yellow voxels, but did spend long time at the page represented as sky blue voxels. Maybe, users want to check more detail of such students' activity independently. In this case, you can choose one of the students and can check details of the student' activity data by visualizing the data separately as shown in Fig. 12 and Fig. 13.
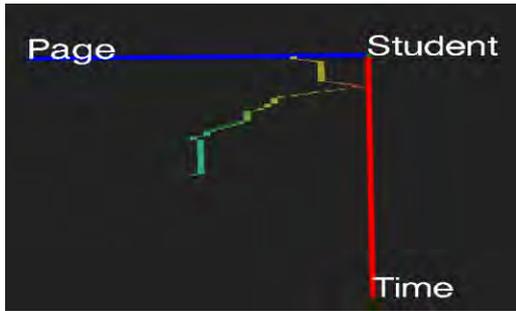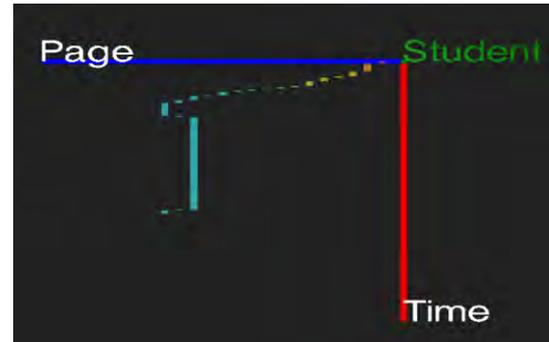
Figure 14. Orthogonal view of Fig. 12.

Figure 15. Orthogonal view of Fig. 13.

Fig. 12 and 13 are perspective views. You can select orthogonal views for them as shown in Figure 14 and 15. From these views, you can clearly understand the students' learning activity transitions. As for the student of Fig. 14, maybe you wonder he/she started the middle page and went back to the beginning page. As for the student of Fig. 15, you can understand he/she looked at the pages in the normal order one by one, but he/she went back to the a few previous page around the page of sky bluish voxels. Anyway, it can be said that there is one common characteristic of the two students which they looked at the page of sky bluish voxels for much time just after looking at the next page. From this visualization result, we can understand that the pages of sky bluish voxels are more important than others in this material.

5. Discussion

Using our proposed visualization tool called Cubic Gantt Chart for students' e-learning activity data, mainly it becomes possible to understand the following three points.

- How long and which pages students look at
- How students transit among pages
- Similar transit patterns of pages among students

From these points, with the use of Cubic Gantt Chart, users can grasp all students' learning behaviors. For example, from Fig. 11, users can understand that some students spent much time at the pages represented as yellow voxels although most students could understand easily. A teacher should spend more time for such students when teaching the corresponding contents in a lecture. Then, the students will come to understand the contents more deeply and rapidly. As another example, although there is not the case in this paper, if it could be found that many students go back to the previous pages to conform the contents related to the current page, a teacher should spend more time for teaching the contents enough at the first time. Furthermore, if it could be found that there are several patterns of students' transitions among pages, a teacher can divide students into several groups according to the patterns and then will be able to teach adaptively to each group.

From Fig. 10, we can understand there are sigmoid patterns at the beginning part of the pages by focusing on the deep color voxels. Some classes have sigmoid patterns although most classes do not have. We wonder there are some factors that cause sigmoid patterns. In this study, although we have not used any statistic methods for analyzing those data, as our feature work, we may find such factors by adapting statistic methods.

Unfortunately, we have not yet received end semester test results of the students. If we have such test results, by investigating relationships between learning patterns of students and their obtained grade, it becomes possible to find out the learning patterns of high achievers and the learning patterns of low achievers, e.g., what pages high achievers spent much time and how different between the high achievers' and low achievers' patterns. As we stated in Sec. 4.2, as shown in Fig. 11, some students spend moderate time to look at the pages represented as yellow voxels and some students spend much time to understand those pages. This result is very important as a remarkable difference among students in this lecture if we can compare the students' learning pattern about the rest of lectures and find out how this result affects their grades. Furthermore, if there are any characteristics of the students of higher grade, by analyzing difference between learning patterns of those students, a teacher can find more

efficient instruction methods. In general, it is better to get higher grade by the study spending little time rather than spending much time. If low achievers can check and understand learning patterns of high achievers, they can improve their learning manner so as to get higher grade efficiently. In this way, the visualization of learning patterns of students is very significant for the both of teachers and students.

In this visualization tool, users can select any three attributes, e.g., in the case of this paper, X, Y and Z axes are time, students and indices of page. However, the data collected by BookLooper has more attributes such as device types, user access date and so on. We think that by applying any statistic methods to the data and by visualizing those results using Cubic Gantt Chart, it will become possible to analyze the data more and more efficiently.

6. Conclusion and future work

In this paper, we proposed new visualization tool called Cubic Gantt Chart, 3D version of Gantt Chart. Using a 3D space in Cubic Gantt Chart, we can employ one more additional attribute besides two attributes of the standard Gantt Chart. Also by adopting color and color depth information, the proposed visualization tool becomes useful for analyzing not only relationships between any pair of two attribute values but also relationships among the all three attribute values. In this paper, we clarified the usefulness of Cubic Gantt Chart for students' learning activity data actually collected by BookLooper. By using this visualization tool, users can understand how long and which pages of a learning material students look at, how students transit among pages and whether there are any similarities among students learning activity patterns. Although in this paper, we showed the usefulness of the proposed visualization tool especially for students' e-learning activity data collected by BookLooper, this tool is also available for visualizing other types of data those include more than three attributes, e.g., for visualizing network traffic data.

Finally, as future works, we will try to analyze more and more learning pattern of students by visualizing such data for making possible to suggest more efficient teaching strategies by improving the functionalities of the tool.

Acknowledgements

References

[1] West, Darrell M. "Big data for education: Data mining, data analytics, and web dashboards." *Governance Studies at Brookings(2012)*: 1-10.
[2] Yamada, Masanori, et al. "Preliminary Research on Self-regulated Learning and Learning Logs in a Ubiquitous Learning Environment."
[3] Picciano, Anthony G. "Ther Evolution of Big Data and Learning Analytics in American Higher Education." Journal of Asynchrous Learning Networks 16.3 (2013) : 9-20
[4] Klemm, Paul, et al. "Interactive visual analysis of image-centric cohort study data." *Visualization and Computer Graphics, IEEE Transactions on* 20.12 (2014): 1673-1682.
[5] Herman, Ivan, Guy Melanc, n, and M. Scott Marshall. "Graph visualization and navigation in information visualization: A survey." *Visualization and Computer Graphics, IEEE Transactions on* 6.1 (2000): 24-43.
[6] Aigner, Wolfgang, et al. "Visual methods for analyzing time-oriented data." *Visualization and Computer Graphics, IEEE Transactions on* 14.1 (2008): 47-60.
[7] Akaishi, M. and Okada, Y. : Time-tunnel: Visual Analysis Tool for Time-series Numerical Data and Its Aspects as Multimedia Presentation Tool, Proc. of 8th International Conference on Information Visualization (IV04), IEEE CS Press, pp. 456-461, London UK, July 2004.
[8] Notsu, H., Okada, Y., Akaishi, M., Niijima, K. : Time-tunnel: Visual Analysis Tool for Time-series Numerical Data and Its Extension toward Parallel Coordinates, Computer Graphics, Imaging and Visualization as Proc. of CGIV 05, IEEE CS Press, pp. 167-172, Beijing, China, July, 2005.

[9] Okajima, S., Okada, Y.: Interactive Analysis of Multidimensional Data on the Web by Using Time-tunnel, Proc. of 5th International Conference on Web Information Systems and Technologies (WEBIST 2009), INSTICC Press, pp. 415-418, March 23-26, 2009.

[10] Etemadpour, Ronak, et al. "Perception-based evaluation of projection methods for multidimensional data visualization." *Visualization and Computer Graphics, IEEE Transactions on* 21.1 (2015): 81-94.

[11] Yang, Jing, et al. "Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets." *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*. IEEE, 2003.

[12] A. Inselberg and B. Bimsdale, "Parallel Coordinates: A Tool for visualizing Multidimensional Geometry." Proc. IEEE Visualization 1990, IEEE CS Press, pp 361-378, 1990

[13] Shneiderman, B. (1992). Tree Visualization With Treemaps: A 2-D Space-Filling Approach. ACM Transactions on Graphics, 11(1), pp.92-99.

[14] James M. Wilson. "Gantt charts: A centenary appreciation" European Journal of Operational Research 149 (2003) 430-437

[15] Aigner, Wolfgang, et al. "Visualizing time-oriented data—a systematic view." *Computers & Graphics* 31.3 (2007): 401-409.

[16] Ogata, H., Li, M., Bin, H., Uosaki, N., El-Bishouty, M. M., & Yano, Y. (2010). Ubiquitous Learning Log: What if we can log our ubiquitous learning? *Proceedings of the 18th International Conference on Computers in Education* (pp. 360-367).