VR System for Spatio-Temporal Visualization of Tweet Data

*

Kaya Okada
Grad. School of Humanities and Sciences
Ochanomizu University
Tokyo, Japan
kaya@itolab.is.ocha.ac.jp

Mitsuo Yoshida
Dept. of Computer Science and Engineering
Toyohashi University of Technology
Toyohashi, Japan
yoshida@cs.tut.ac.jp

Takayuki Itoh
Grad. School of Humanities and Sciences
Ochanomizu University
Tokyo, Japan
itot@is.ocha.ac.jp

Tobias Czauderna
Faculty of Information Technology
Monash University
Victoria, Australia
tobias.czauderna@monash.edu

Kingsley Stephens
Faculty of Information Technology
Monash University
Victoria, Australia
kingsley.stephens@monash.edu

***Abstract*—Social media analysis is helpful to understand the behavior of people. Human behavior in social media is related to time and location, which is often difficult to understand the characteristics appropriately and quickly. We chose to apply virtual reality technologies to visualize the spatio-temporal social media data. This makes us easier to develop interactive and intuitive user interfaces and explore the data as we want. This paper presents our visualization of tweets of microblogs with location information. Our system features a three-dimensional temporal visualization which consists of the two-dimensional map and a time axis. In particular, we aggregate the number of tweets of each coordinate and time step, calculate scores and display them as piled cubes. We highlight only specific cubes so that users can understand the overall tendency of datasets. We also developed user interfaces for operating these cubes and panels which indicate details of tweets.**

***Index Terms*—temporal visualization, virtual reality, social media, tweet data, immersive visualization**

I. Introduction

Social media analysis brings us important knowledge. Data on Twitter and Facebook indicates where, when and why events are held, or people gathered. Many analysis methods such as text mining or human network analysis have been applied to social media analysis. However, it often happens that these methods are not sufficiently effective to understand the complex data of human behavior involving various factors. Meanwhile, interactive visualization is also useful to understand complex social media data. Thus, we applied immersive analytics techniques using Virtual Reality (VR) technologies for social media exploration. VR is suitable for not only entertainment fields but also for visualization. It makes us easier to observe the spacial data intuitively and in detail.

This paper proposes a 3D spatio-temporal visualization with an interactive user interface (UI) for tweets of microblogs. We developed the visualization by assigning a map of TDL on an xz 2D plane and the time as the y-axis (vertical direction against map). The proposed technique aggregates the number of tweets in each time and blocks divided into the appropriate sizes. Then, we set colors and highlight only important portions of the 3D space which has many tweets by controlling the transparency. Users can immerse into this map and observe overview of the data while they look at the time change of the number of the tweets in each time periods with a small map which is the duplicated map of the large one. Moreover, we implemented intuitive operations supposed to use the VR device "HTC Vive" [1]. By using VR, users are supposed to fly around and explore the data by themselves so they will experience the environment of this area. It will help users to memorize the map and the data they found.

In this study, we introduce the example applied tweets with location information in Tokyo Disneyland (TDL). The visualization brings knowledge of congestion or remarkable events for users who rarely go to TDL and makes feel them as if they are in there.

II. Related Work

Itoh [2] classified major temporal data visualization techniques into four categories, including polyline charts, heatmaps, 3D representations and linked views. In this section, we introduce 3D-based techniques which can represent both spatial and temporal information and the visualization using VR.

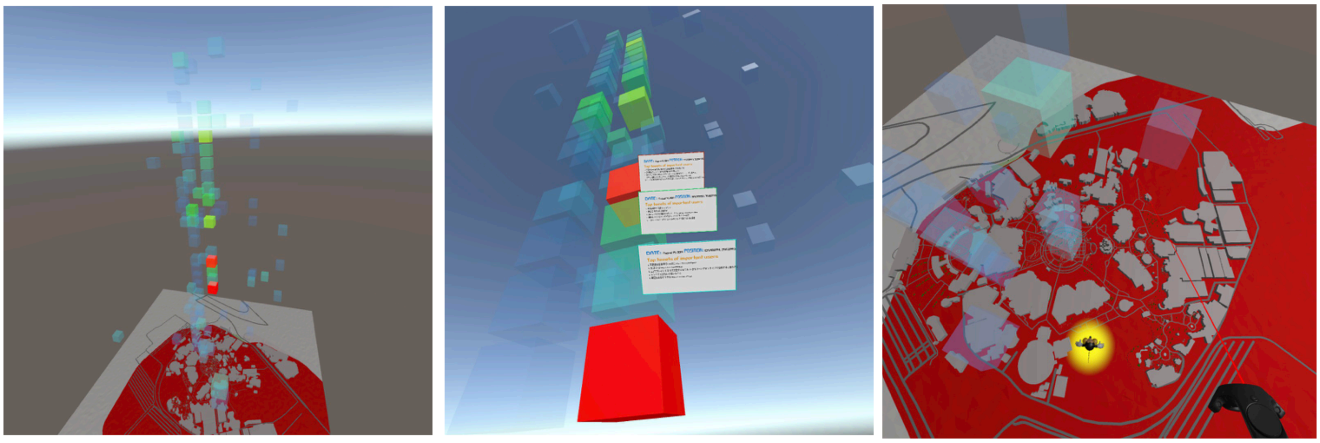

Fig. 1. WorldView. (Left) Temporal change of the number of tweets is shown as the overview in a VR space. (Center) Panels which display actual tweets included in each cube appeared when users select each cube. (Right) The character icon with yellow highlight indicates the position of a user to prevent missing his/her current position.

A. Spatio-Temporal visualization

Space-Time Cube [3] is a representative study on 3D temporal data visualization. It is often used for visualization of human behavior as the forms of spatio-temporal paths and geo-time. Fukada et al. [4] developed a method to represent walking routes and congestion areas by visualizing mobility of sightseeing behavior with GIS (Geographic Information System). They compared the proposed method with an existing representation which just draws spatio-temporal paths. Cuboid Matrix [5] arranged dynamic network information in a 3D space consisting of a plane and a time axis. Users can observe overview and detail of spatio-temporal information. However, the technique often caused insufficient readability of crowded regions, and therefore required operations for breakdown displays into 2D spaces.

B. Visualization of tweets data using VR

Guttentag [6] demonstrated that VR has potentials for tourism and marketing. This paper also claims that VR models allow planners to observe an environment from an unlimited number of perspectives instead of just applying a bird's-eye view. As a result, travelers can make appropriate decisions based on the information displayed graphically and had practical expectations. Moreover, the experiential nature of VR makes it an optimal tool for providing rich data to tourists. This study concluded that VR has a potential to revolutionize the promotion and selling of tourism.

Immersive Analytics [7] is a recent framework for supporting the analytics of real data. Virtual reality environment such as a large-sized touch panel, Oculus Rift [8], Cave2 [9] and tracking devices like Kinect make users immerse into the data. Specialists and Analysts have natural access to large complex data. ImAxes [10] is a typical example of immersive analytics. Users can generate visualization displays freely by using Vive controller. A scatterplot is generated when we select and combine any axes in a VR space, and a PCP link appears after we put together multiple scatterplots. This method makes user experiences immersing in the data and users operate the visualization display and search information by themselves.

Moran [11] visualized tweets in a VR space. They indicate characteristics of tweets as object attributes. Users can observe details of the tweets. They focus on characteristic individual tweets contrary to our method focuses on characteristic tweets of certain time and location. Also, their method does not show a temporal variation of the number of tweets at the same time.

Based on the above, we chose to visualize the temporal change of the tweets in a VR space because users themselves can go through and operate the data. This study visualized the tweets around Tokyo Disneyland (TDL) as an example the experiential nature provides the best affection. Moreover, the operation of an unlimited number of perspectives using VR solves the problem of insufficient data comprehensibility in 3D spaces.

III. Proposed Technique

This section presents the data structure and processing flow of the proposed visualization. It consists of the following two displays.

WorldView: 3D spatio-temporal view for the overview of a certain period of time (one-month data in this study)

MiniMap: a small map for operation and viewing the time change of each time zone (one-day data in this study)

A. Tweet data

Tweet objects have various attributes such as coordinates (latitude/longitude), created time, user id, hashtags, and text. We gathered these data via Twitter API and saved in JSON format. We applied approximately 16,000 tweets with location information around TDL in August 2014.

B. VR environment

We developed the visualization on Unity3D [12] game engine. Unity Assets have rich supports including the SDK (Software Development Kit) for the VR devices which makes

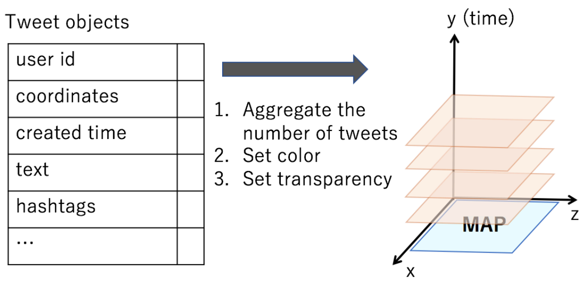

Fig. 2. Processing flow of cubes generation in WorldView.

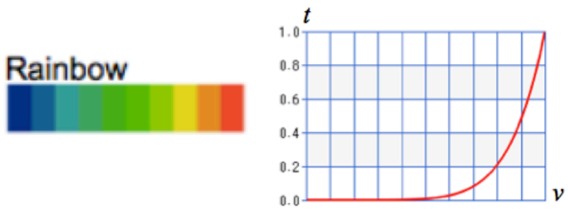

Fig. 3. Colormap and the transfer function for setting transparency.

us easier to develop the complicated applications. We used HTC Vive Virtual Reality Headset.

C. WorldView

Fig. 1 shows the large map which users immerse into. Users can fly around the objects of visualization and observe outline of the data by operating TrackPad of HTC Vive. WorldView gives us the remarkable time and area especially. Thus users can seek the data with a focus on this part. At this moment, an icon is set on the map just under the users not to lose the position of themselves. The icon informs users the positions of cubes and the users themselves. We used UtyMap [13] to reproduce the map of TDL and obtained the map data from Open Street Map (OSM). Peculiar interactions of VR make us easier to understand the data. We explain technical components in WorldView and provided operations below.

D. Cubes

This technique regards the frequency of tweets as temporal data and represents as a set of cubes. Our visualization consumes tweets data written in a JSON file as tweet objects. The processing flow of cubes generation is shown in Fig. 2. In the following equations, n_{ij} denotes the frequency of tweets at the i-th date in the j-th block, and v_{ij} denotes its score.

1) Aggregation: For the comprehensive understanding of tweets data and people behavior, the technique divides the map into appropriate sizes of blocks and aggregates the number of tweets in each block and time step (one day in this case). It then normalizes the number of tweets included certain blocks and time step to a range[0,1].

$$v_{ij} = \frac{n_{ij} - n_{min}}{n_{max} - n_{min}} \tag{1}$$

2) Setting of Colormaps: Next, the technique prepares colormaps automatically to represent values of cubes. Borland et al. [14] describes the correspondence between data types and colormaps. We applied a common rainbow colormap based on the hue. Hue h of the HSV color space is generated as follows.

$$h_{ij} = \frac{160}{240}(1.0 - v_{ij}) \tag{2}$$

3) Setting of Transparency: Then, we define the transfer function for setting transparency. Cluttering is a common problem of 3D visualization techniques depending on viewpoint setting. The technique represents important portions which have projecting values opaquely, and other portions transparent, to prevent the cluttering and improve the comprehensibility. Thus, the important parts are only highlighted. We define the transfer function as an exponential function of the values of cubes. The equation is as follows and the graph is shown in Fig. 3.

$$t_{ij} = v_{ij}^{a} \tag{3}$$

E. Panels

Fig. 4 shows the panels, which appear when users select a cube they want to see the details with the pointer. Panels show details about the tweets corresponding to the selected cubes, including date, coordinates (latitude/longitude), and texts of representative tweets. This technique selects the tweets which include important words and written by important users. Currently, we specify important users based on the following criteria:

- Number of followers.
- Degree of activity. We define the activity as the number of tweets of the user dividing by the number of days since their account was created at.

Moreover, we apply the tf-idf method to the tweets to extract important texts and remove general-meaning tweets like "I'm at Tokyo Disneyland." These panels make easier for users to find when, where and why people were gathered or events have occurred. Also, panels and the operation of them make the switchover between overview and detail more intuitive.

F. Operation procedure in WorldView

Following is the list of controller operations in WorldView.

TrackPad of the left controller
Moving and flying around the map and cubes.

Trigger button of the right controller
Pointing for the selection of cubes and panels. Cubes will be highlighted when users select them.

Grip button of the right controller
Grabbing the panels.

Snapping the right controller
Destroying the panels. Snap upward rapidly.

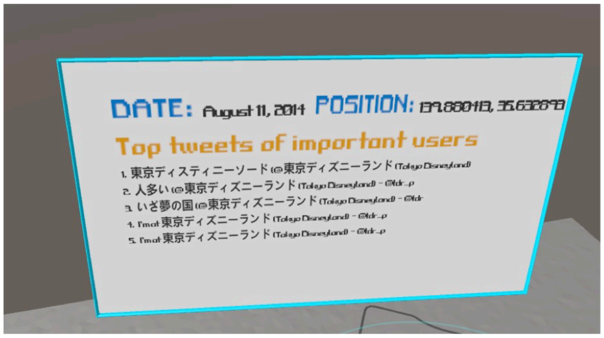

Fig. 4. The panel displays the detailed information of each cube. It contains the date, coordinate and text of representative tweets and gives us the knowledge about the reason why people made many tweets in a certain area and time.

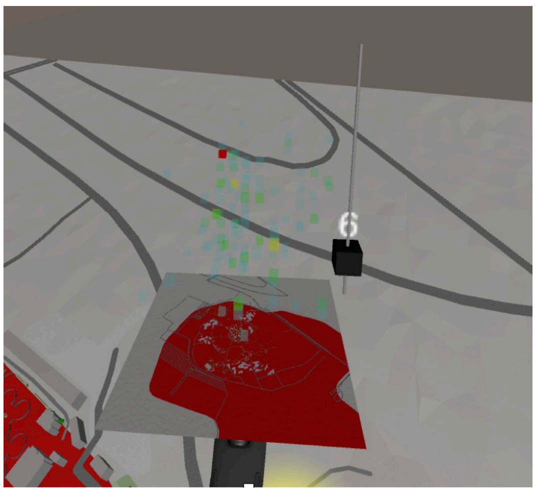

Fig. 5. MiniMap views the time change of the number of the tweets in a day. Cubes in MiniMap indicate the aggregation of one-hour data different from those in WorldView. Users can observe the time change more specifically.

G. *MiniMap*

Fig. 5 shows MiniMap, a small duplication of "WorldView" attached to the left controller. It appears by turning over the controller. This map also provides the overview, but this provides a time change of the data in a shorter time range differently from WorldView. Our current implementation of MiniMap represents a particular day, where cubes of MiniMap depict the aggregation per hour. Users can select the date to observe details with the slider attached to MiniMap. Users are required to press the trigger and the grip button at the same time while using the slider. They can also warp to the positions of the corresponding cubes of WorldView after the users point the cubes of MiniMap by pressing the trigger button on right controller and releasing. This operation saves time to move to the distant cubes and explore the tweets included in these cubes.

IV. Example

This section introduces an example which applied the dataset described in Section 3.1. In this example, we applied 16,000 tweets those latitude and longitude were inside the rectangular area surrounding TDL. Our implementation adopted a picture of Mickey Mouse for the icon of WorldView. We developed effects like fireworks of TDL for destroying the panels in WorldView.

At first, we realized that opaque cubes concentrated around the middle of this month, from the overview shown in the upper-left figure of Fig. 6). TDL was crowded with people in this period because it was during the summer vacation week in Japan, and therefore tweets also increased in this period. Then, we moved the viewpoint to the position of cubes, as shown in the upper-right figure of Fig. 6. We found that many people tweeted in front of Cinderella castle located at the center of TDL from this figure.

The rest figures in Fig. 6 show the real tweets in front of the Cinderella castle. In the lower-left figure, several tweets which are not related to Disneyland are picked up. On the other hand, tweets related to Disneyland are displayed in the lower-right figure. However, the corresponding cube does not have many tweets. It is predictable that many tourism had appeal tweets like "I'm at Tokyo Disneyland" that they were at TDL in this position. It is likely to be increased such noisy information while the number of tweets is increased, so extracting meaningful tweets is important. This is an essential future issue, and we still have room for improvement.

A. *Conclusion*

This paper proposed a spatio-temporal visualization technique in a VR space consisting of "WorldView" and "MiniMap." We aggregated the number of tweets in each block of coordinates and time step and represented the scores as cubes applied color and transparency. At the same time, we combined the UI for displaying the detail of tweets included in cubes with this visualization display. This enables users to grasp the characteristics of human behavior in one month and observe the critical times and regions. Moreover, they can get the detailed information in the particular times and regions corresponding to the remarkable cubes. Users can also observe the time change of the number of tweets in each period. Simple association of these outline view and details is realized by VR. Users will be familiar with the environment while exploring and experiencing the data by themselves.

We have several future issues. At first, we would like to improve text extraction methods so that we can retrieve more meaningful tweets. We would like to apply various method in addition to the tf-idf method. We also have ideas to develop new functions. For example, we would like to trace the tweets of an individual to find how the regular customers behave. Also, we would like to automatically arrange the panels.

In addition, evaluations relating to reality and operability should be conducted. For example, whether users feel as immersed in TDL and how long they take to master the manipulation. The result of the evaluations would help the further development of more effective visualization methods and UI.

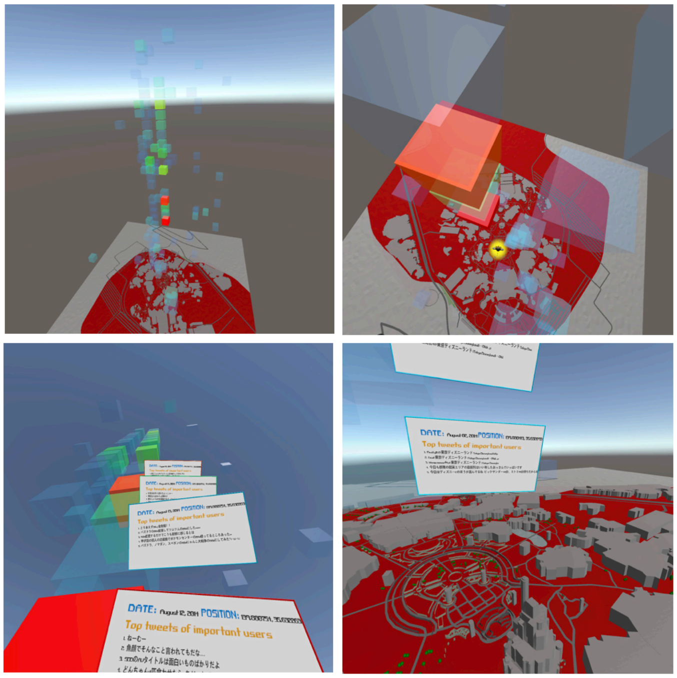

Fig. 6. Example. Top left figure shows people tweet a lot during Japanese summer vacation. Top right figure is the map looked down from the position of cubes. It indicates user is in front of cinderella castle. Panels of the both figure below pick up actual tweets.

Acknowledgment

This work has been partially supported by Japan Society of the Promotion of Science under Grant-in-Aid for Scientific Research.

References

[1] Htc vive. https://www.vive.com/jp/.

[2] Takayuki Itoh. Visual synergetics (5): Mapping techniques ...especially towards systemization of time-varying data visualization techniques. *The 17th Annual Conference of The Japan Society for Computational Engineering and Science*, Vol. 17, p. 4, 2012. (in Japanese).

[3] Benjamin Bach, Pierre Dragicevic, Daniel Archambault, Christophe Hurter, and Sheelagh Carpendale. A Review of Temporal Data Visualizations Based on Space-Time Cube Operations. In *Eurographics Conference on Visualization*, 2014.

[4] Hidemi Fukada, Yusuke Okuno, Shou Ohtsu, and Yuichi Hashimoto. Proposal of Technique for 3D Visualization of Behavioral Data on Scenic Walk using Geographic Information System. *Society for Tourism Informatics*, Vol. 8, No. 1, pp. 51–66, 2013. (in Japanese).

[5] Teseo Schneider, Yuriy Tymchuk, Ronie Salgado, and Alexandre Bergel. Cuboidmatrix: Exploring dynamic structural connections in software components using space-time cube. In *IEEE Working Conference on Software Visualization (VISSOFT)*, pp. 116–125, 2016.

[6] Daniel A Guttentag. Virtual reality: Applications and implications for tourism. *Tourism Management*, Vol. 31, No. 5, pp. 637–651, 2010.

[7] Tom Chandler, Maxime Cordeil, Tobias Czauderna, Tim Dwyer, Jaroslaw Glowacki, Cagatay Goncu, Matthias Klapperstueck, Karsten Klein, Kim Marriott, Falk Schreiber, et al. Immersive analytics. In *IEEE International Symposium on Big Data Visual Analytics (BDVA)*, 2015.

[8] Oculus rift. https://www.oculus.com/.

[9] Alessandro Febretti, Arthur Nishimoto, Terrance Thigpen, Jonas Talandis, Lance Long, JD Pirtle, Tom Peterka, Alan Verlo, Maxine Brown, Dana Plepys, et al. Cave2: a hybrid reality environment for immersive simulation and information analysis. *The Engineering Reality of Virtual Reality 2013*, Vol. 8649, p. 864903, 2013.

[10] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H Thomas, and Kim Marriott. ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 71–83, 2017.

[11] Andrew Moran, Vijay Gadepally, Matthew Hubbell, and Jeremy Kepner. Improving big data visual analytics with interactive virtual reality. In *IEEE High Performance Extreme Computing Conference (HPEC)*, pp. 1–6, 2015.

[12] Unity3d. http://unity3d.com/.

[13] utymap. https://github.com/reinterpretcat/utymap.

[14] David Borland and Russell M Taylor Ii. Rainbow color map (still) considered harmful. *IEEE computer graphics and applications*, Vol. 27, No. 2, 2007.