Integrating Retinal Variables into Graph Visualizing Multivariate Data to Increase Visual Features

Hong Thi Nguyen[1], Lieu Thi Le[2], Cam Thi Ngoc Huynh[3], Thuan Thi My Pham[4], Anh Thi Van Tran[5], Dang Van Pham[6,7], and Phuoc Vinh Tran[4(✉)]

[1] University of Information Technology, Vietnam National University – HCMC, Ho Chi Minh City, Vietnam
hongnguyen1611@gmail.com
[2] Thu Dau Mot University, Binhduong, Vietnam
lethilieu.qb@gmail.com
[3] Kiengiang Department of Education and Training, Kiengiang, Vietnam
camhuynhit@gmail.com
[4] Hochiminh City Open University, 35 Ho Hao Hon, Dist. 1, Hochiminh City, Vietnam
{thuanptm.178i,phuoc.tvinh}@ou.edu.vn, phuoc.gis@gmail.com
[5] Hochiminh College of Economics, Hochiminh City, Vietnam
anhttv@kthcm.edu.vn
[6] Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam
pvdang@ntt.edu.vn
[7] Graduate University of Science and Technology, Hanoi, Vietnam

Abstract. The efficiency of a graph visualizing multivariate data is not only subjectively evaluated by human visual perception but also objectively estimated by visual features of graph. For a designed graph, it is necessary to improve visual features to enable human to extract better information from data. Integrating retinal variables into graph is an approach to increasing visual features of graph. In this study, the constituents of graph are grouped into classes of marks by qualitative and quantitative characteristics. The retinal variables are studied and structured to integrate into the classes of marks. A process of five steps is proposed to increase visual features by integrating retinal variables into graph. The process is illustrated with two case studies, increasing visual features of bus space-time map with qualitative mark classes and increasing visual features of the graph representing the data of hand-foot-mouth epidemic in Binhduong with qualitative and quantitative mark classes.

Keywords: Visual features · Retinal variable · Visual perception · Visualization · Multivariate data

P. C. Vinh and A. Rakib (Eds.): ICCASA 2019/ICTCC 2019, LNICST 298, pp. 74–89, 2019.
https://doi.org/10.1007/978-3-030-34365-1_7

1 Introduction

Information exists in various forms of data as sound, languages, words, signs, glyphs, graphs, pictures, and so on. The data shaped image contain much information than others and people can easily perceive information implicit in it [1]. Visualization refers to an approach to converting data of various forms to graphs to enable human to extract better information. In a visualization system, human is a component of the system and plays an important role in understanding the insights of graph [1]. Accordingly, the efficiency of a structure of graph is evaluated by human visual perception [2]. The challenge in increasing the efficiency of graph is that the improvement of visual features of graph is objectively evaluated while human perception of graph is subjectively evaluated. The problem to be solved is how to improve visual features of graph consistently with human perception.

The main idea is to integrate retinal variables into graph to improve its visual features. This study approaches the structure of graph to grouping its constituents into classes of structured marks by qualitative and quantitative characteristics consistent with human perception. Retinal variables are studied and suitably structured to integrate into the mark classes. This paper proposes a process integrating retinal variables into graph to increase visual features of the graph and simultaneously respond human capacity of visual perception.

This paper is structured as follows. The next section interprets related works together with the concepts as data variable, visual representation, visual features, structural marks, planar marks, and retinal variables. The third section proposes the process integrating retinal variables into multidimensional graph to increase its visual features. The fourth section applies this process for two case studies, improving visual features of bus space-time map and multidimensional graph representing the happenings of hand-foot-mouth epidemic in Binhduong province, Vietnam. The fifth section resumes the results of the paper.

2 Related Works and Conceptual Framework

2.1 Visual Representation of Multivariate Data

A dataset is considered as a combination of several subsets each of which is called a data variable depicting a unique attribute, i.e. each variable is a set of values. Visual representation of a set of multivariate data is a mapping of the set onto a multidimensional graph, called structural graph, where each data variable is converted to a structural variable shaped an axis of the graph, each data value or a data tuple is converted to a structural mark shaped a point, a line, or a polygon of the graph. Visual display is a mapping of structural graph onto planar graph on planar screen, where each structural mark is converted to a planar mark on planar screen. The integration of retinal variables into variables and marks of a graph, structural or planar, by visualization techniques converts the graph to visual graph.

2.2 Visual Features

The efficiency of a visual graph is subjectively evaluated by the levels of human perception. A visual graph is considered to be efficient if the duration to perceive necessary information is short, shorter and shorter [3]. An efficient graph enables human to extract significant information and discover valuable laws. A visual graph enables human to perceive the significance of data at the levels of structure of the graph according to associative, selective, ordered, quantitative characteristics, and length, which are also considered as visual features [3–5]. Accordingly, the efficiency of a visual graph can be objectively estimated by the levels of visual features.

Associative Feature. The associative feature refers to human perception on the similarity of marks, i.e. the associative feature enables human to group marks sharing one characteristic into a group.

Selective Feature. The selective feature refers to human perception on the selection of marks in accordance with given characteristic. In other words, the selective feature enables human to perceive the distinction between mark groups of different attributes or titles.

Ordered Feature. The ordered feature of a graph refers to human perception on the order of marks or mark groups representing data variables, such as one is bigger or smaller than another, one is higher or lower than another, one is nearer or farther than another, one is left or right from another, one is front or back from another, one is above or below from another.

Quantitative Feature. The quantitative feature refers to human perception on marks representing scale values or the ratio of two marks representing two real values.

Length. The length feature refers to the number of planar marks of a planar variable. This feature relates to the size and resolution of displayed screen, where the possible number of planar marks on a planar variable has to be more than the number of values of the data variable which it represents.

2.3 Visual Variables

A planar graph representing a set of multivariate data on 2-dimensional displaying environment is constituted by two types of visual variables, position variables and retinal variables [3, 5].

2.3.1 Position Variable

The position variable is a set of marks of points, lines, polygon, bar, pie, etc. positioned on structural graph. For visualization of a set of multivariate data, a mapping is applied to convert the constituents of the structural graph which is designed to represent data variables to position variables and marks on planar screen. In other words, a graph on displayed environment is the set of planar marks representing data values or data tuples.

2.3.2 Basic Retinal Variable

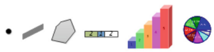
Basic retinal variables are composed of shape variable (S), size variable (Z), brightness variable (B), texture or symbol variable (L), color variable (C), and direction variable (D) [3, 5, 6] (Table 1).

Shape Variable, S. The S*hape* variable refers to geometrical planar marks. According to the concepts of geographic information science, the geometrical marks representing geographic objects comprises point, line, polygon or area, surface, and volume [7–10]. Point is indicated as a geometrical point positioned on planar screen. Line is traced as a curve or a polyline. Polygon or area refers to the area within a polygon of which vertices are on the same plane. Surface refers to a polygon of which vertices are not on the same plane. Volume refers to the space within a volume limited by planes. The *Shape* variable responds the associative and selective features. The *Shape* variable does not respond the ordered and quantitative features. The Shape variable can have numberless elements.

Size Variable, Z. The *Size* variable refers to the difference on size of marks. The *Size* variable responds associative, selective, and ordered features. The *Size* variable is hard perceived quantitative feature. The number of elements of *Size* variable is finite because of the resolution of human eyes and the size of planar screen.

Brightness Variable, B. The *Brightness* variable refers to human visual perception on the luminance of graphs. The *Brightness* variable is combined with the color variable for use. The *Brightness* variable responds associative, selective, and ordered features. The *Brightness* variable hard responds quantitative feature. The length of *Brightness* is finite because it depends on human's visual resolution.

Table 1. The table of basic retinal variables and their visual features.

Basic retinal variables	Elements of basic retinal variable (some illustrated examples)	Visual features				
		associa-tive	selec-tive	ordered	quanti-tative	Length (finite)
Shape *S*		✓	✓			
Size *Z*		✓	✓	✓	✓	✓
Brightness *B*		✓	✓	✓	✓	✓
Symbol *L*		✓	✓			
Color *C*		✓	✓			✓
Direction *D*		✓	✓	✓		✓

Symbol Variable, L. The *Symbol* variable refers to figures formed by points, line, curve, geometrical shape, star, cross, and so on which are arranged according to various types. The *Symbol* variable responds associative and selective features. The *Symbol* variable has not ordered feature. The *Symbol* variable does not limit the number of elements.

Color Variable, C. The *Color* refers to human visual perception by the discrimination of light waves reflecting from objects in real world. The *Color* variable refers to color spaces, where the usual space is RGB. In the RGB space, colors can be generated from three main colors, red, green, blue. Each generated color is defined by its coordinates on RGB space. The *Color* variable responds associative and selective features, but do not respond ordered and quantitative features. The number of colors which can be used depends on human color resolution.

Direction Variable, D. The *Direction* is generally indicated by the longitudes and latitudes of the Earth and presented on a plane. Each direction is indicated by the angle which is formed by the vector positioning object with the reference axis. The *Direction* variable responds associative, selective, and ordered features. The *Direction* variable does not respond quantitative feature. The number of directions is theoretically infinite but the demand of associative and selective features limits the number of directions.

3 The Integration of Retinal Variables into Graph

Visual representation of data refers to the conversion of various modes of data to image shaped graph of which significance is easily perceived by human. The level of human perception of graph depends on its visual features. The perception can be improved by integrating reasonably retinal variables into graph. The integration of retinal variables into a graph is a conversion of the multidimensional graph *G'* to the graph *G* of higher visual features. Technically, retinal variables can be integrated into structural or planar marks of structural or planar graph, respectively. The following is the process integrating retinal variables into graph.

3.1 Step 1. Grouping Marks of Graph into Mark Classes

A multidimensional graph G' displayed on planar screen is considered as set of planar marks $g_n | n = 1, 2, \ldots$. The planar marks are grouped into mark classes by qualitative or quantitative characteristics to meet the demand of problems extracting information or analyzing data. Each qualitative class is a set of marks of the same nominal characteristic or the same attribute, each quantitative class is a set of marks referring to values belonging to a defined segment.

Mathematically,

$$G' = \{g_n | n = 1, 2, \ldots\}$$
$$(G_m \cap G_{m'} = \emptyset) \vee (G_m \cap G_{m'} \neq \emptyset) | \forall (m \neq m')$$
$$(G_m \subset G' | \forall m) \wedge (G' = \bigcup_m G_m)$$

where $G_m | m = 1, 2, \ldots$ are mark classes which can be combined according to various modes G^k to restore multidimensional graph G':

$$G^k | k = 1, 2, \ldots \equiv G' = \{G_m | \bigcap_m G_m = \emptyset | m = 1, 2, \ldots\}$$

where modes $G^k | k = 1, 2, \ldots$ are restored from mark classes G_m with dimensional integrality [11].

3.1.1 Grouping Marks by Qualitative Characteristics

The constitution of mark classes by qualitative characteristics refers to the significance of constituents of graph and human perception. Mark classes are constituted by applying Gestalt principles [11–13], where the marks of the same nominal characteristic or the same attribute such as bus route, bus station, ground trajectory, the number of patients, humidity, rainfall, temperature, and so on are grouped into a mark class by qualitative characteristic, briefly called qualitative mark class. The mark classes may intersect one another or not, i.e. a mark may simultaneously belong to one or more classes, e.g. a bus station may be used for one or more bus routes, a ground trajectory belongs to both ground trajectory class and bus route class.

3.1.2 Grouping Marks by Quantitative Characteristics

The constitution of mark classes by quantitative characteristics, briefly called quantitative mark class, refers to values of data variables. The range of values of a variable can be divided into segments. The marks representing the values of a segment are grouped into a class, called a scale, where each scale is a subset of a class grouped by qualitative characteristics. The following are the ways to group the marks of a class into scales.

Number Table. Number table which refers to a data table of a data variable comprises two columns of which elements associate with one another on each line as a data tuple. As an example, wind speed is recorded in a table of two columns, one for wind speed in *m/s* and one for wind force scale in natural numbers. The wind speed table may be represented on a 2-dimensional orthogonal coordinates, where one coordinate indicates wind speed in *m/s*, another indicates wind force scale in natural numbers (see Fig. 1). In visualization, marks representing wind speed in real values are grouped into segments each of which is a mark class, briefly called scale.

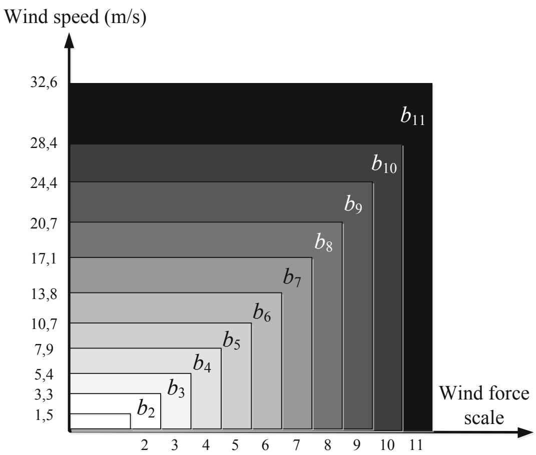


Fig. 1. Wind speed of real numbers mapped onto wind force scale of natural numbers.

Transfer Function. Mathematically, the relation between real values and scales is defined by a transfer function which is linear or nonlinear (see Fig. 2). Marks of a scale are grouped into a class, briefly called scale.

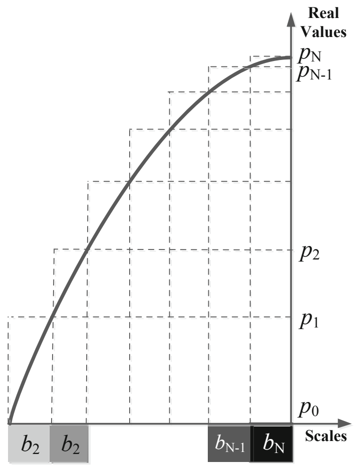


Fig. 2. The transfer function representing the relation between real values and scales of a data variable.

Local Maximum or Minimum. When marks representing local maxima or minima of a variable need to be dominated to evaluate the relation among variables, these marks and their neighbors are grouped into a class, briefly also called scale.

3.2 Step 2. Constituting Generated Retinal Variables

Generated retinal variables integrating into mark classes are constituted from basic retinal variables based on the number of mark classes, the characteristics of each class, and human capacity of visual perception. Each generated retinal variable is a subset of a basic retinal variable or a product of basic retinal variables. The number of mark classes, the characteristics of mark classes, the requests of visual features of classes are referred to form generated retinal variables.

3.2.1 Generated Retinal Variable for Qualitative Mark Classes

The integration of generated retinal variables into qualitative mark classes aims at dominating their qualitative characteristics. The generated retinal variables are designed as subsets of basic retinal variables. Each element of a generated retinal variable, called visual mark, which is used to integrate into a qualitative mark class has to be suitable for the characteristics and significance of the class, as well as the user's demands in extracting information. For example, the symbol of line is suitable for trajectories, the symbol of polygon for administrative units, the symbol of bar for the number of patients, the color is utilized to group or select the constituents of a qualitative mark class. The number of visual marks of a generated retinal variable is equal to the number of qualitative mark classes which they integrate into.

The problem to be solved to increase visual features of structural graphs is that the number of visual marks of a basic retinal variable can not respond all qualitative mark classes or that human visional resolution limits the number of visual marks of a generated retinal variable. Indeed, it is impossible to use all colors to increase visual features of graph because human is difficult to perceive the difference of two quite similar colors, human can not perceive the difference of two line marks of two different attributes.

Generated retinal variables are sets of visual marks extracted from basic retinal variables or from the product of two or more basic retinal variables. Mathematically, if let $R' = \{S, Z, B, L, C, D\}$ be the set of basic retinal variables, generated retinal variables are the subsets of the sets S, Z, B, L, C, D or the subsets of product sets SZ, ZB, BL, LC, CB, DS, and so on. The set of generated retinal variables is $R = \{R^S, R^Z, R^B, R^L, R^C, R^D, R^{SZ}, R^{ZB}, R^{BL}, R^{LC}, R^{CB}, \ldots\}$, where $R^S \subset S$, $R^Z \subset Z$, $R^B \subset B$, $R^L \subset L$, $R^C \subset C$, $R^D \subset D$, $R^{SZ} \subset SZ$, $R^{ZB} \subset ZB$, $R^{BL} \subset BL$, $R^{LC} \subset LC$, $R^{CB} \subset CB$, and so on.

The number of visual marks of a product set is equal to the product of the number of marks of individual sets. As an example, if the basic retinal variables $S = \{s_1, s_2, \ldots, s_M\}$ contains M visual marks, the basic retinal variables $Z = \{z_1, z_2, \ldots, z_N\}$ contains N visual marks, then the product set SZ has $M.N$ visual marks $SZ = \{s_1z_1, s_1z_2, \ldots, s_1z_N, s_2z_1, s_2z_2, \ldots, s_2z_N, \ldots, s_Mz_1, s_Mz_2, \ldots, s_Mz_N\}$. The increase of visual marks by the product of basic retinal variables results in the increase the number of mark classes which can be integrated visual marks.

3.2.2 Generated Retinal Variables for Scale Mark Classes

The integration of a generated retinal variable into mark classes of scales aims at dominating the quantitative characteristics of scale marks. In visualization, the values and scales of a data variable are simultaneously represented on graph, the marks representing real values are integrated by visual marks, the marks representing scales are also integrated by visual marks of generated retinal variable. Accordingly, each scale mark is integrated by two visual marks, one for characteristic of data variable and one for scale of data variable. In this case, the set of marks representing a data variable is visualized by two generated retinal variables, one retinal variable associates marks of the same characteristic or attribute of the data variable and another selects different scales of the data variable.

3.3 Step 3: Processing the Intersections of Mark Classes

Many marks of graph belong to two or more classes, such as some stations in bus space-time map simultaneously belong to two or more different classes of routes [14]. Mathematically, a mark belonging to many classes only shares one or few attributes or characteristics which it represents. Hence, a mark shared by different mark classes must represent common and individual attributes. In other words, a shared mark must satisfy all significances of the mark classes which it belongs to. For example, a station shared by several route represents a common attribute being station, and individual attributes being the names of bus routes to which it belongs as the bus station of routes 15. For displaying, a shared mark is structured as a combination of visual marks of the classes which it belongs to.

3.4 Step 4: Integrating Generated Retinal Variables into Mark Classes

The integration of generated retinal variables into mark classes converts graph G' to visual graph G. Mathematically, the integration of generated retinal variables into mark classes is a mapping of G' onto G by the Cartesian product of G' and the set of generated retinal variables R,

$$
\begin{aligned}
G = R \times G' &= \{R^S, R^Z, \ldots\} \times \{G^1, G^2, \ldots\} \\
G = R \times G' &= \{R^S \times G^1, R^Z \times G^2, \ldots\}
\end{aligned}
$$

The products can be generally defined as follows.

$$
\begin{aligned}
R^S \times G^1 &= \{s_1, s_2, \ldots\} \times \{G^1_1, G^1_2, \ldots\} = \{s_1 G^1_1, s_2 G^1_2, \ldots\} \\
R^Z \times G^2 &= \{z_1, z_2, \ldots\} \times \{G^2_1, G^2_2, \ldots\} = \{z_1 G^2_1, z_2 G^2_2, \ldots\} \\
R^B \times G^3 &= \{b_1, b_2, \ldots\} \times \{G^3_1, G^3_2, \ldots\} = \{b_1 G^3_1, b_2 G^3_2, \ldots\} \\
R^L \times G^4 &= \{l_1, l_2, \ldots\} \times \{G^4_1, G^4_2, \ldots\} = \{l_1 G^4_1, l_2 G^4_2, \ldots\} \\
R^C \times G^5 &= \{c_1, c_2, \ldots\} \times \{G^5_1, G^5_2, \ldots\} = \{c_1 G^5_1, c_2 G^5_2, \ldots\} \\
R^D \times G^6 &= \{d_1, d_2, \ldots\} \times \{G^6_1, G^6_2, \ldots\} = \{d_1 G^6_1, d_2 G^6_2, \ldots\} \\
&\ldots
\end{aligned}
$$

where $G^k_m | k = 1, 2, \ldots$ is a mode combining mark classes G_m.

$$G = R \times G' = (R^S \times G^1) \cup (R^Z \times G^2) \cup (R^B \times G^3) \cup \ldots$$

The classes of marks representing data of real values and scales may be integrated by a generated retinal variable which is a product of two basic retinal variables. For example,

$$R^{BC} \times G^7 = \{b_1c_1, b_2c_1, \ldots\} \times \{G_1^7, G_2^7, \ldots\} = \{b_1c_1G_1^7, b_2c_1G_2^7, \ldots\}$$

where:

$$R^B = \{b_1, b_2, \ldots\} \subset B$$
$$R^C = \{c_1, c_2, \ldots\} \subset C$$

$G^7 = \{G_1^7, G_2^7, \ldots\}$ is a mode of G', where $G_m^7 | m = 1, 2, \ldots$ is the mark class G_m of the combining mode G^7.

The product $R^{BC} \times G_m = \{b_1c_1, b_2c_1, \ldots\} \times G_m$ is considered that the color c_1 is integrated into all marks of the mark class G_m meanwhile scales of G_m are integrated by different brightness $b_1, b_2, \ldots$.

Generally, the integration of generated retinal variables into mark classes can be visually designed with a 2-dimensional matrix, where a dimension indicates mark classes, another indicates generated retinal variables and visual marks. In addition, the integration can also be designed with mathematical expressions. The design with mathematical expression is suitable for graphs visualizing data variable of quantitative characteristics as scales.

3.5 Step 5: Displaying Visual Graph

The graph G visualized from G' by integrating retinal variables is displayed on screen with various modes to enable user to extract information and discover knowledge.

4 Case Studies

4.1 Case Study 1. Visualizing Space-Time Map of Bus

Bus space-time map is considered as a graph representing bus network of a city on space-time cube [14, 15]. The density of marks on the map results in the decrease of visual features and user is difficult to extract information. The efficiency of the map is improved by integrating retinal variables into its marks in accordance with the following process.

Step 1. Grouping Marks of Bus Space-Time Map. The marks of bus space-time map is divided into the mark classes.

– *The class of stations.* Bus stations of all routes are grouped into the mark class of station.

– *The class of space-time points*. Space-time points are marks of points associating bus stations with time points, where each space-time point associates with a station and a time point when a bus calls at the station. All space-time points are grouped into a mark class of space-time points.
– *The class of ground trajectories*. Ground trajectories are polylines connecting bus stations of the same route. Ground trajectories of all routes are grouped into a mark class of ground trajectories.
– *The class of space-time trajectories*. Space-time trajectories are polylines connecting space-time points of the same trip. Space-time trajectories of all routes are grouped into a mark class of space-time trajectories.
– *The class of route*. Each route are constituted by a ground trajectory, stations combining with the ground trajectory, space-time points, and space-time trajectories of the route.

Table 2. The table of the integration of generated retinal variables into mark classes.

Generated retinal variables	Visual Marks	Mark classes of				
		stations	space-time points	ground trajectories	space-time trajectories	routes
Shapes ↓	Point	●	●			
	Line			╱	╱	
Symbols ↓	Square	□				
	Dash				⋰	
	Arrow			↗		
Colors ↓	Red	■ 1	●	↗	⋰	Route 1
	Green	■ 1	●	↗	⋰	Route 2
	Blue	■ 1	●	↗	⋰	Route 3

Step 2. Constituting Generated Retinal Variables. The shapes of point and line are used to indicate marks of bus space-time map on displayed plane. The symbols of dash, arrow, triangle, square, and filled circle are designed to integrate into marks of lines or points. The color variable is designed to integrate into marks of bus routes.

Step 3. Processing the Intersections of Mark Classes. In a bus space-time map, some bus stations are shared by different routes. The marks indicating these stations need to be processed to connect simultaneously to different routes. In other words, each common station must be integrated the visual marks of all routes which it belongs to.

Step 4. Integrating Generated Retinal Variables into Mark Classes. The shape of point is assigned to class of stations and class of space-time points. The shape of line is assigned to the class of ground trajectories and the class of space-time trajectories to represent route trajectories and trip trajectories, respectively. The symbol of square is integrated into the mark class of stations to represent bus stations. The symbol of dash is integrated into the mark class of trip trajectories to represent trip trajectories. The symbol of arrow is integrated into the mark class of route trajectories to represent route trajectories and the direction of bus movement on the route. All constituents of a route are integrated by a color to differentiate a route from others. The integration generated retinal variables into mark classes is designed on a table (see Table 2).

Step 5. Displaying Visual Map. After being integrated retinal variables, the bus space-time map is improved its efficiency because of the increase of visual features (see Fig. 3).

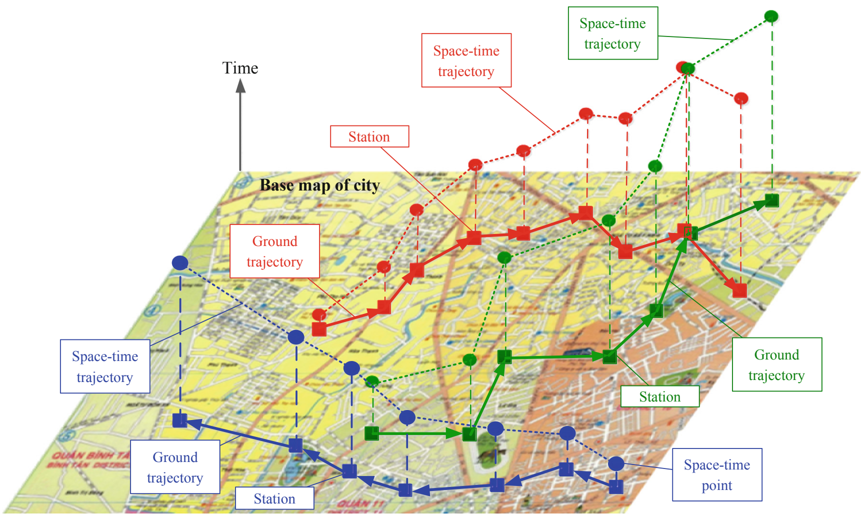


Fig. 3. An illustration of the integration of retinal variables into bus space-time map

4.2 Case Study 2. Representing the Data of Hand-Foot-Mouth Epidemic

The happening of hand-foot-mouth epidemic in Binhduong province during 2012–2014 is represented as a multidimensional graph on a multidimensional cube [16, 17]. The graph represents four data variables, the number of patients, rainfall, humidity, and temperature which share the reference variable of time. The graph can be improved the efficiency by integrating retinal variables into graph to increase visual features.

Step 1. Grouping Marks of Graph. The marks representing the data of happening of hand-foot-mouth epidemic are divided into 4 mark classes corresponding to 4 data variables, the number of patients (P), rainfall (R), humidity (H), temperature (T). It is necessary to consider the relation among data variables because they together refer to time variable. Each of P, R, H, T variables is divided into 3 mark classes as subset of P, R, H, T, respectively, called scales, corresponding to 3 groups of values of data. Maximal scale involves marks representing the values of local maxima and their neighbors, minimal scale involves marks representing the values of local minima and their neighbors, and medial scale involves remainder.

Step 2. Constituting Generated Retinal Variables. The mark classes representing 4 data variables share the visual mark of bar shape to indicate the value of data. Each qualitative mark class as the number of patients, rainfall, humidity, temperature is integrated by a color (a visual mark of color variable). Each quantitative mark class as scales is additionally integrated by a visual mark of brightness variable.

$R^S = \{bar\} = \{barP, barR, barH, barT\}$, where *barP* is the class of marks shaped bars representing the number of patients; *barR* is the class of marks shaped bars representing rainfall; *barH* is the class of marks shaped bars representing humidity; *barT* is the class of marks shaped bars representing temperature.

$R^C = \{c_1, c_2, c_3, c_4\} = \{red, green, blue, orange\}$ is the generated color variable.
$R^B = \{b_1, b_2, b_3\} = \{dark, medium, light\}$ is the generated brightness variable.
$R^{BC} = \{b_1, b_2, b_3\} \times \{c_1, c_2, c_3, c_4\}$ is the product of two retinal variables.

$$R^{BC} = \{\{b_1, b_2, b_3\}c_1, \{b_1, b_2, b_3\}c_2, \{b_1, b_2, b_3\}c_3, \{b_1, b_2, b_3\}c_4\}$$

$$R^{BC} = \{\{b_1c_1, b_2c_1, b_3c_1\}, \{b_1c_2, b_2c_2, b_3c_2\}, \{b_1c_3, b_2c_3, b_3c_3\}, \{b_1c_4, b_2c_4, b_3c_4\}\}$$

Step 3. Processing the Intersections of Mark Classes. The intersection of a qualitative mark class with scales is processed by the product of brightness variable and color variable, where color variable is integrated into qualitative mark classes and brightness variable is integrated into scales.

Integrating the product of generated retinal variables into mark classes to represent simultaneously qualitative mark classes and scales:

$$R^{BCS} = R^{BC} \times R^S$$

$$R^{BCS} = R^{BC} \times \{barP, barR, barH, barT\}$$

Step 4. Integrating Generated Retinal Variables into Mark Classes. Mark classes representing the number of patient composed of qualitative marks and scales are integrated by the generated retinal variable:

$R^{BCP} = \{dark.red.barP, medium.red.barP, light.red.barP\}$ composed of 3 visual marks corresponding to 3 scales of the variable of the number of patients.

Mark classes representing rainfall composed of qualitative marks and scales are integrated by the generated retinal variable:

$R^{BCR} = \{dark.green.barR, medium.green.barR, light.green.barR\}$ composed of 3 visual marks corresponding to 3 scales of the variable of rainfall.

Mark classes representing humidity composed of qualitative marks and scales are integrated by the generated retinal variable:

$R^{BCH} = \{dark.blue.barH, medium.blue.barH, light.blue.barH\}$ composed of 3 visual marks corresponding to 3 scales of the variable of humidity.

Mark classes representing temperature composed of qualitative marks and scales are integrated by the generated retinal variable:

$R^{BCT} = \{dark.orange.barT, medium.orange.barT, light.orange.barT\}$ composed of 3 visual marks corresponding to 3 scales of the variable of temperature.

Step 5. Displaying Visual Graph. The dark colors representing the maximum marks of the number of patients, rainfall, humidity emphasizes the relation among variables. The marks of dark colors of visual graph representing the data of hand-foot-mouth epidemic in Binhduong are perceived the correlation among the variables of the number of patients, rainfall, humidity. The variable of temperature is not perceived any correlation. If this correlation is confirmed by bigger dataset, the result may be used to predict the happening of hand-foot-mouth epidemic when viewing visual graph representing data in real time (Fig. 4).

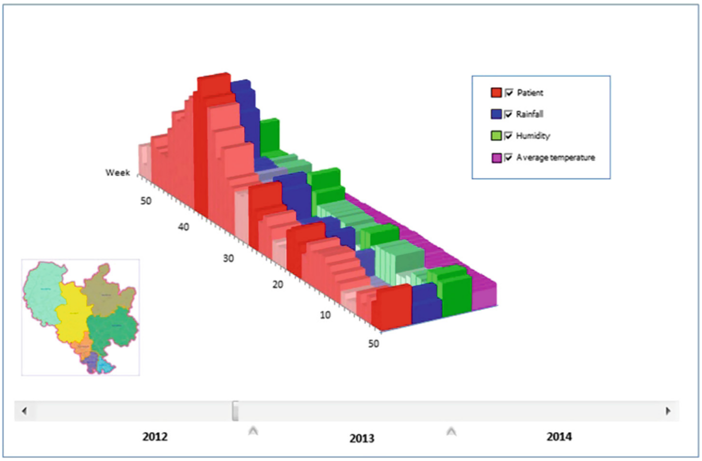


Fig. 4. The visual graph representing the dataset of hand-foot-mouth epidemic in Binhduong province, Vietnam, during 2012–2014, shows the relation among the number of patients, rainfall, and humidity, meanwhile temperature does not affect the number of patients [16, 17].

5 Conclusion

The efficiency of a graph representing visually a dataset depends on human perception. The visual features are utilized to estimate objectively the efficiency of a visual graph according to the level of human perception of the significance of graph. The visual features of a graph are improved when retinal variables are suitably integrated into marks of graph. Graphs of good visual features enable user to extract valuable information and/or discover new knowledge by viewing and thinking.

The paper proposed a process of integrating retinal variables into graph representing multivariate data to increase visual features of the graph. Marks of graph are grouped into mark classes by qualitative and quantitative characteristics, where qualitative mark classes refer to nominal characteristics and quantitative mark classes refer to the values of data variables. Generated retinal variables used to integrate into graphs are formed from basic retinal variables. In that, each generated retinal variable is a subset of a basic retinal variable or a product of basic retinal variables. The process is illustrated with the visualization of bus space-time map and the multidimensional graph representing the data of happening of hand-foot-mouth epidemic in Binhduong province, Vietnam, during 2012–2014.

Acknowledgment. The paper is sponsored by Hochiminh City Open University, Vietnam.

References

1. Few, S.: Tapping the power of visual perception. Perceptual Edge, pp. 1–8, 4 September 2004
2. Dastani, M.: The role of visual perception in data visualization. J. Vis. Lang. Comput. **13**, 601–622 (2002)
3. Green, M.: Toward a Perceptual Science of Multidimensional Data Visualization: Bertin and Beyond. ERGO/GERO Human Factors Science. Citeseer (1998)
4. Few, S.: Data visualization for human perception. In: Soegaard, M., Dam, R.F. (eds.) The Encyclopedia of Human-Computer Interaction, 2nd edn. Aarhus, Denmark (2014)
5. Bertin, J.: General theory, from semiology of graphics. In: Dodge, M., Kitchin, R., Perkins, C. (eds.) The Map Reader. Theories of Mapping Practice and Cartographic Representation, pp. 8–16. Wiley (2011)
6. Bertin, J.: Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin (1983)
7. Card, S.: Information visualization. In: Chi, Ed. (ed.) Information Visualization, PARC 2007, pp. 510–543 (2007)
8. Jensen, J.R., Jensen, R.R.: Introductory Geographic Information Systems. Pearson Education (2013)
9. Raper, J.: Multidimensional Geographic Information Science. Taylor & Francis, London (2000)
10. Carpendale, M.S.T.: Considering Visual Variables as a Basis for Information Visualisation. Department of Computer science, University of Calgary, Calgary, AB, Canada (2003)
11. Wagemans, J., et al.: A century of gestalt psychology in visual perception II. Conceptual and theoretical foundations. Psychol. Bull. **138**, 1218–1252 (2012)

12. Nguyen, H.T., Pham, T.M.T., Nguyen, T.A.T., Tran, A.V.T., Tran, P.V., Pham, D.V.: Two-stage approach to classifying multidimensional cubes for visualization of multivariate data. In: Cong Vinh, P., Alagar, V. (eds.) ICCASA/ICTCC-2018. LNICST, vol. 266, pp. 70–80. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-06152-4_7
13. Alexandre, D.S., Tavares, J.M.R.S.: Introduction of human perception in visualization. Int. J. Imaging Robotics™, **4**, 60–70 (2010)
14. Nguyen, H.T., Duong, C.K.T., Bui, T.T., Tran, P.V.: Visualization of spatio-temporal data of bus trips. Presented at the IEEE 2012 International Conference on Control, Automation and Information Science, ICCAIS 2012, Hochiminh City, Vietnam (2012)
15. Nguyen, H.T., Ngo, D.N.T., Bui, T.T., Huynh, C.N.T., Tran, P.V.: Visualizing space-time map for bus. In: 6th International Conference on Context-Aware Systems and Applications, and Nature of Computation and Communication, ICCASA 2017, Tam Ky, Vietnam, pp. 38–47 (2017)
16. Nguyen, H.T., Tran, A.V.T., Nguyen, T.A.T., Vo, L.T., Tran, P.V.: Multivariate cube integrated retinal variable to visually represent multivariable data. In: EAI Endorsed Transactions on Context-aware Systems and Applications, vol. 4, pp. 1–8 (2017)
17. Nguyen, H.T., Tran, A.V.T., Nguyen, T.A.T., Vo, L.T., Tran, P.V.: Multivariate cube for representing multivariable data in visual analytics. In: Cong Vinh, P., Tuan Anh, L., Loan, N.T.T., Vongdoiwang Siricharoen, W. (eds.) ICCASA 2016. LNICST, vol. 193, pp. 91–100. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56357-2_10