J Vis
https://doi.org/10.1007/s12650-021-00772-0

REGULAR PAPER

Ji Lan · Jiachen Wang · Xinhuan Shu · Zheng Zhou · Hui Zhang · Yingcai Wu

RallyComparator: visual comparison of the multivariate and spatial stroke sequence in table tennis rally

Received: 8 July 2021 / Accepted: 18 July 2021

Abstract Table tennis matches consist of many stroke sequences where two players hit the ball interactively and consecutively until one fails to hit the ball. Players usually employ many complicated playing techniques at each stroke in highly antagonistic, variable, and flexible matches. In-depth comparative analyses of players' stroke sequences are necessary to obtain insights into the technical playing patterns of players. Experts commonly use spreadsheets to browse and compare strokes one by one, and this process is tedious and prone to errors. Statistical analyses are limited to well-defined patterns (e.g., value distribution and relation significance) and fail to present complex and peculiar patterns. We collaborated with experts to dig out soft patterns of stroke sequences and proposed a novel interactive visualization system to present and compare the patterns. The main visualization challenge is to display the multivariate stroke sequence and the spatial variation patterns. We designed a glyph-based pattern view to solve the challenge. These comprehensible visualizations and coordinated views in the system allow efficient comparative analysis of stroke sequence patterns and are highly commended by domain experts, who have identified several new and interesting patterns using the system. We demonstrated the effectiveness and usability of the visualization system through case studies with table tennis experts.

Keywords Sports visualization · Visual analytics · Glyph design

1 Introduction

Analyzing the stroke sequence in a table tennis rally is essential and challenging in sports science. In a table tennis match, a rally is the shortest scoring unit. A player will score at least after it. Meanwhile, it is the longest unit within which all strokes are consecutive. Analysis of table tennis data at the level of a rally involves both high-level scoring information and low-level sequential structures. Previous studies in table tennis have analyzed the statistics of a player's performance in a rally, such as shots within and rest time between rallies (Loh and Krasilshchikov 2015). However, these indicators fail to consider the dynamic interactions between players in a rally and are hence unable to measure the performance of players stably (Lames and McGarry 2007). Mathematical models, such as the Markov chain model (Pfeiffer et al. 2010;

J. Lan · J. Wang · Y. Wu (✉)
State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China
E-mail: ycwu@zju.edu.cn

X. Shu
Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

Z. Zhou · H. Zhang
Department of Physical Education, College of Education, Zhejiang University, Hangzhou, China

Published online: 06 September 2021

Wenninger and Lames 2016) have been used to simulate the sequential transitions of strokes in a rally. However, The scalability of the transition matrix in the model is limited. A table tennis stroke has five attributes and the Markov chain model can only simulate value variations of one attribute at a time. Therefore, analysis of the stroke sequences in rallies is a difficult task in the table tennis domain.

The analysis of the stroke sequence in a rally is similar to the event sequence analysis. Two difficulties in event sequence analysis are the volume and variety of data (Du et al. 2017). In the perspective of volume, a table tennis match contains around a hundred rallies and the analysis of multiple matches involves several hundred or thousand rallies. For variety, a rally contains a stroke sequence with variant length and a stroke can belong to multiple types according to its attributes. For these two difficulties, it is hard to browse the stroke sequences one by one or obtain a representative overview of the stroke sequences. Integrating a visualization system into the data mining process provides new possibilities to handle the scalability and heterogeneity problems. Two main kinds of visualization studies using event sequence mining have been proposed. Studies based on sequential pattern mining (SPM) (Liu et al. 2017) apply frequency-based sequence mining algorithms to generate a large number of patterns and employ a visualization interface to reduce the efforts of analysts in pruning the patterns to make them manageable. However, it is not easy to mine stroke sequences through SPM because patterns in stroke sequences are not based on simple and strict rules. The pruning process will filter out a large set of useful patterns. Studies based on sequence similarity and clustering (Chen et al. 2018) aggregate the sequences into an overview through clustering feature vectors of sequences and employ a visualization interface to help the analysts look at the original sequences to understand the aggregation and verify the results. This study employs the second method.

Besides the difficulties in previous studies of event sequence mining and visualization, two new difficulties are posed in this study. First, the visualization of a multivariate stroke sequence pattern is difficult. A stroke sequence needs to be analyzed simultaneously in terms of five stroke attributes, namely, strike technique kind, technique, position, spin and ball position. For instance, domain experts of table tennis can hardly make sense of a stroke sequence in terms of merely strike technique or position value. In previous event sequence analysis cases, such as the analysis of web clickstreams (Liu et al. 2017), the event is only divided according to values of one attribute, the event type. Moreover, the values of ball position and strike position are relevant to positions in the table tennis table. Therefore, designs for these two attributes are required to outstand the spatial variation patterns of strokes. Second, the comparison of multiple stroke sequence patterns is difficult. Rallies can be naturally classified into different groups based on the rally identity attributes, such as whether a player scores or serves in the rally. Table tennis analysts need to compare patterns in different kinds of rallies to point out why a player scores at a rally or how a player performs in a rally she/he serves.

We worked closely with table tennis domain experts to identify the important score and serve attributes of rallies for the classification of rallies and a set of stroke attributes to cluster stroke sequences in rallies. We also proposed glyph-based designs to visualize stroke sequence patterns and relevant layout and interaction designs for effective comparison and interpretation of the patterns. Our contributions are as follows:

- Characterization of the design goals for visual analytics of multivariate and spatial stroke sequences in table tennis;
- Glyph-based designs that visualize multivariate stroke sequence patterns with spatial variation in table tennis;
- A visualization system integrating the clustering algorithm and visualization designs that supports extraction and comparative analysis of patterns in stroke sequences;and
- New insights into the comparisons of rally stroke sequences under different classifications (using a real dataset of matches between a Japanese player and Chinese players).

2 Related work

This section reviews visualizations of sports data and comparison in visualization.

2.1 Racquet sports visualization

Sports visualization studies have investigated many sports domains, such as soccer (Wu et al. 2019; Xie et al. 2021), basketball (Chen et al. 2016), and racquet sports (Polk et al. 2014; Wu et al. 2018). Increasing interest has been shown in racquet sports visualization recently. Racquet sports are games where players use rackets to hit a ball. Commonly, the hits are given by two sides of players interactively. The innate characteristics of racquet sports induce a set of visualization studies. Racquet sports have a hierarchical game structure. For instance, in table tennis, a player needs to win many rallies to win a game and win many games to win a match. Displaying the hierarchical scoring data can help domain experts in racquet sports review and browse the scoring process and gain insights into which parts can be enhanced. Courtime (Polk et al. 2020) and TenniVis (Polk et al. 2014) visualized the hierarchical scoring data in tennis using glyphs and small multiples. Besides, iTTVis (Wu et al. 2018) visualized this data in table tennis using multiple tailored line charts. Tactics are also widely analyzed in racquet sports. Tac-Miner (Wang et al. 2021) proposed a glyph-based design to visualize multiple features of tactics. TacticFlow (Wu et al. 2022) aims to present the transformations among tactics in racket sports. The racquet sports are also featured with sequential multivariate hits. As the hits in racquet sports are given alternatively and consecutively by two sides of players until one side of players fails to hit the ball, a sequence of hit data is required to be analyzed. iTTVis (Wu et al. 2018) proposed a matrix flow design to visualize the hit correlations in hit sequences in table tennis. More general, Wu et al. (2020) proposed a set of glyphs for multivariate hit sequence visualization. Visualization studies have also explored other potentials aspects of racquet sports data. For instance, ShuttleSpace (Ye et al. 2021) and TIVEE (Chu et al. 2022) employ immersive visualization to consider the height of badminton data. VisCommentator (Chen et al. 2022) aims to augment sports videos with visualization to better communicate insights. Tac-Simur (Wang et al. 2020) explores the visual simulative analysis of hit performance. Existing studies clarify essential features in racquet sports data and provide baseline work for references. However, this study aims to visualize multivariate stroke sequence and needs to outstand the spatial variation pattern in the sequence. Previous designs in racquet sports visualizations fail to achieve this requirement.

2.2 Event sequence mining and visualization

A broad of studies have explored visualization techniques to reduce the efforts of analysts in processing the event sequence data. As the volume and variety grow very large, machine learning and data mining algorithms are proposed to aggregate the data into a manageable size. Most of the studies are either based on SPM (Liu et al. 2017) or sequence similarity and clustering (Chen et al. 2018). The studies based on SPM mined hard patterns (Gotz 2016) and commonly generated a large number of patterns. Filtering and pruning the patterns will still be tedious and unavoidably lose useful patterns. Visualization systems are integrated into this process to reduce the efforts of analysts and improve accuracy. The studies based on sequence similarities mined soft patterns (Gotz 2016). However, the generated patterns are commonly not easy to interpret directly. Analysts need to examine the original sequences to verify and interpret the aggregated patterns, which is tedious and error-prone. Visualization techniques are integrated to help analysts interpret and verify the results efficiently.

This study also mined the soft patterns based on sequence similarity and clustering methods. However, the presentation of stroke sequence is more difficult than common event sequence. A set of studies (Du et al. 2016; Guo et al. 2020; Wongsuphasawat et al. 2011) visualized the event sequence with one attribute (commonly the event type). In this study, strokes need to be analyzed in terms of more than one attribute. Another set of studies have proposed techniques, such as glyph (Wu et al. 2020) and exploration rules (Cappers and van Wijk 2017) to enhance visual analytics of multivariate event sequences. However, spatial variation patterns in table tennis stroke sequences cannot be displayed by these techniques. Nevertheless, these designs inspire us to design a new glyph for visualizing multivariate event sequences with spatial variation information.

2.3 Comparison in visualization

Comparison is a common task in many visual analytics systems (Guo et al. 2018; Li et al. 2020; Glueck et al. 2018). Kehrer and Hauser (2013) indicated that comparison is a typical task in the visual analysis of scientific data. Gleicher (2018) reported that visualization could often help analysts conduct comparison tasks.

The three approaches for visualization comparisons are juxtaposition, superposition, and explicit encoding (Gleicher et al. 2011). The juxtaposition layout is commonly used and many recent visualization studies have juxtaposed multiple ensembles, such as threads (Guo et al. 2018), radial charts (Chen et al. 2021; Filipov et al. 2021), glpyh (Weng et al. 2021; Wu et al. 2020), regional patterns (Jin et al. 2021) and graph (Jin et al. 2021) for comparison. Moreover, Nebula (Chen et al. 2021) proposed a grammar for creating coordination for juxtaposed views. According to Gleicher et al. (2011), the juxtaposition layout can reduce viewers' burden. Instead of using only one comparison approach, Alexander and Gleicher (2016) employed direct encodings to represent the distance from one document to others and juxtaposed these encodings of multiple documents for a side-by-side comparison. He et al. (2020) employed direct encodings to represent the overall dissimilarity of two information groups and juxtaposed the distributions in two groups for a side-by-side comparison. A survey (Gleicher 2018) demonstrates that interactions, such as brushing and linking and detail on demand, can also augment the comparison between juxtaposed views. For instance, PhenoLines (Glueck et al. 2018) coordinates corresponding parts of multiple radial charts and highlights all charts when a chart is hovered over. EventThread (Guo et al. 2018) allows for a multi-scale exploration of each thread and only compares high-level features side by side, thus reducing the visual clutter of juxtaposed threads.

Previous explorations of visual comparison inspired our design. In this study, we aimed to compare stroke sequence patterns. We employed a juxtaposition layout for a side-by-side comparison of stroke sequence patterns. Besides, we calculated the overall similarities among patterns and encoded them explicitly.

3 Background

This section introduces domain data and requirements.

3.1 Domain knowledge and data description

Table tennis is a sport in which two players hit a lightweight ball back and forth across a table [1]. The action wherein a player uses a table tennis bat to hit the ball once is a *stroke*. The time during which the ball is continuously in play is a *rally*. A rally contains a sequence of strokes with variable length. If a player wins a rally, she/he will obtain a rally score. The player who obtains 11 rally scores first will win a *game*. A game contains a set of rallies and can be divided into start, middle, key, and control *phases* according to rally scores. The player who wins a game will obtain a game score. The player who obtains three or four game scores (according to different rules) first will win a *match*. We worked with table tennis experts to use the rally as the analysis unit. The data are composed of hundreds of rallies and attributes of rallies are identified as follows.

Identity attributes: A rally has its inherent attributes in a table tennis match based on rules. These inherent attributes generate natural classifications of rallies for comparisons and analysis.

– *Scoring player:* The player who wins the rally;
– *Serving player:* The player who serves in the rally;
– *Game ID:* The ID of the game the rally belongs to;
– *Phase ID:* The ID of the phase the rally belongs to.

Stroke attributes: A rally also contains a sequence of strokes with variable length. A stroke has five attributes describing the technical features. The sequential patterns of these attributes are insightful for characterizing and understanding the technical features of rallies.

– *Strike technique:* Fourteen strike technique values at the top of Fig. 2e denote the techniques a player uses in a stroke. Further details can be found on a Wikipedia page [2].
– *Strike state:* Four state values at the bottom of Fig. 2e denote the states of a player in a stroke.
– *Ball position:* Nine ball position values (Fig. 2b and c) define all possible drop points of a stroke. They can be grouped into three parts, namely, forehand, middle, and backhand by horizontal variation (Fig. 2b), or into long, half-long, and short by vertical variation (Fig. 2c).

– *Strike position:* Four strike position values (Fig. 2a) illustrate all possible positions where a player gives a stroke.
– *Strike spin:* Six strike spin values (Fig. 2d) describe all possible types of the spin of the ball.

The table tennis data were manually extracted from an original match video and organized as a data table. Each row of the table presents the attributes of a stroke in the match.

3.2 Requirement analysis

We collaborated with sports analysts (hereinafter referred to as experts) A and B from the national table tennis team with a focus on analyzing the techniques and tactics of table tennis players. Specifically, expert A had worked for the national table tennis team. Expert B was a postdoctoral researcher majoring in sports science. He was also a table tennis athlete and a senior analyst of table tennis data. Both experts attended the entire process of identifying the requirements.

The experts aimed to explore the stroke sequences in rallies based on the log data for insights. Following the methodology proposed by Sedlmair et al. (2012), we worked closely with experts for a year. In this process, we cooperated with the experts through interviews, formal meetings, and online discussions. In particular, we observed that they analyzed the stroke sequence using Excel functions. The experts were often overwhelmed by the multifarious patterns and tedious operations. We introduced that many visualization systems and tools had been proposed to help analysts identify patterns and stories from data (Wang et al. 2020; Chen et al. 2015; Mei et al. 2019; Zhao et al. 2020; Shi et al. 2020; Tang et al. 2020; Shu et al. 2021). The experts were then convinced that visualization techniques could help them analyze the stroke sequences. We worked with experts and synthesized the analysis requirements from their analysis process as follows:

R1 *During which rallies does a player score or lose more?* Rallies are classified into different groups according to their inherent identity attributes according to the match rules. Experts are familiar with these identity attributes and they need to find out which during which group of rallies a player score or lose more.

R2 *What are the typical patterns of rallies with specific identity attributes?* A group of rallies with specific identity attributes contain variable-length stroke sequences and each stroke has five attributes indicating technical features. Experts need to identify the sequential trends of attributes in a group of rallies to describe their technical features.

R3 *What are the similarities and differences among the patterns of rallies with different identity attributes?* Experts hope to compare patterns in rallies with different identity attributes to analyze their similarities and differences.

R4 *What are the features of original rallies underlying a pattern?* Experts need to check the original rallies to understand and verify the patterns obtained by the data mining algorithms.

4 Sequence clustering

A *stroke* is donated as K. A *stroke attribute* (introduced in Sect. 3.1) is donated as A. A stroke K has five stroke attributes A_d. A *stroke sequence* is an ordered list of strokes K_i, where i denotes the index of the stroke in the sequence: $S = \{K_1, K_2, \ldots, K_n\}$. In table tennis, a rally contains a stroke sequence.

The attributes of stroke have different importance in determining the stroke similarity. We hence worked with experts to assign different weights to attributes in the calculation of stroke similarity. The formulas for the similarity of two strokes are as follows.

$$Similarity_1(K_i, K_j) = \begin{cases} \sum_d f(A_{i,d}, A_{j,d}) \cdot aw_d & K_i, K_j \neq null \\ -p_k & K_i \ or \ K_j = null \end{cases}$$

$$f(A_{i,d}, A_{j,d}) = \begin{cases} r & A_{i,d} = A_{j,d} \\ -p_a & A_{i,d} \neq A_{j,d} \end{cases}$$

where p_k is the punishment for similarity if one of the two compared strokes is null. aw_d is the weight assigned to attribute A_d. r is the reward for similarity if two attribute values are equal. p_a is the punishment for similarity if two attribute values are unequal. Based on this similarity, the formulas for the similarity of two stroke sequences are as follows.

$$Similarity_2(S_a, S_b) = \sum_i Similarity_1(K_{a,i}, K_{b,i}) \cdot kw_i$$

$$kw_i = \begin{cases} weight_1 & i \leq 4 \\ weight_2 & 4 < i \leq 6 \\ weight_3 & i > 6 \end{cases}$$

where kw_i is the weight assigned to stroke K_i. According to the experts, the first four strokes are more important than other strokes. Therefore, we assigned different stroke weights according to the sequence number.

Stroke sequence group: A group of stroke sequences are derived by classifying all rallies according to the identity attributes. For instance, all rallies can be classified into two groups, namely, rallies player A scores and player B scores, by the *scoring player*. Stroke sequences of rallies in each group are denoted as a stroke sequence group: $G = \{S_1, S_2, \ldots, S_m\}$.

Stroke sequence cluster: A cluster of stroke sequences are derived by clustering all stroke sequences in a stroke sequence group: $C = \{S_1, S_2, \ldots, S_l\} \subseteq G$.

Based on the definition of sequence similarity, we calculated the similarity matrix of all sequences in a stroke sequence group G and employed spectrum clustering (von Luxburg 2007) to cluster the sequences. In each derived stroke sequence cluster C, we chose the sequence with the highest average similarities to other sequences in the cluster as the representative pattern. In this manner, the pattern is also interpretable.

5 Visualization system design

5.1 Design goals

As discussed in the related work section, existing visualization techniques are unsuitable for the comparative analyses of multivariate and spatial stroke sequences. We summarized the design goals based on the requirements.

G1 *A table overview of analyzed identity attributes values to allow easy navigation.* Experts need to classify rallies into different rally groups according to the analyzed identity attributes and compare the rally groups (R1). When experts need to analyze multiple identity attributes simultaneously, multi-level classifications are imposed on rallies. The overview should allow easy navigation of the derived rally groups. We decided to employ a table-format overview, which the experts are familiar with, to achieve this goal.

G2 *Integration of data mining algorithms into the system and interactive adjustments to extract patterns of stroke sequences in rallies efficiently and precisely.* Experts need to extract patterns from a set of rallies to reduce the volume and obtain an aggregate overview (R2). Data mining algorithms are effective methods to process the data into a manageable size. Therefore, we integrated a clustering algorithm into the visualization system and provided interactive adjustments of the parameters.

G3 *Glyph-based flows that can display the multiple attributes and spatial variations of strokes in sequence patterns.* Experts need to detect the sequential trends of multiple stroke attributes quickly from the extracted patterns (R2). As discussed in Sect. 2.2, glyph-based designs have been employed to visualize multivariate event sequences. A glyph offers a means of visual fusion of multiple visual channels (Fuchs et al. 2017). We aimed to design a glyph-based design to present a multivariate stroke sequence and visually emphasize the spatial variation of strokes. Besides, the glyph design should be relevant to encoded attributes to be more comprehensible for expects.

G4 *Juxtaposition of patterns and explicit encodings of pattern similarities for efficient comparison.* Experts need to compare patterns in rallies with different identities efficiently (R3). As discussed in Sect. 2.3, juxtaposition, superposition, and explicit encoding (Gleicher et al. 2011) have been proposed for visual comparison. We deserted the superposition of patterns as a large amount of information is encoded in a pattern. We chose to juxtapose patterns for side-by-side comparison. We also needed to encode the similarities of patterns explicitly to enhance comparisons.

G5 *Detail of demand for interpretation and verification of patterns.* Experts hope to look at original rallies and figure out their relations with the aggregate patterns (R4). We designed a detail view that presents similarities of the pattern and original rallies at each stroke step and concrete attribute values for interpretation and verification of the derived patterns.

5.2 System overview

The system consists of three views, namely, a table view, a pattern view, and a detail view. The table view (Fig. 1a) provides an overview of the analyzed identity attributes (G1). The pattern view (Fig. 1b) contains glyph-based flows (Fig. 1b2) to visualized patterns (G3), which are extracted through the clustering algorithm (G2), in different stroke sequence groups. All the flows are juxtaposed and aligned based on the sequence number axis and allow the explicit presentation of the pattern similarities (G4). The detail view (Fig. 1c) is displayed when a pattern is clicked to help users interpret and verify the aggregated stroke sequences (G5). The colors purple and orange stand for two players in the system.

5.3 Table view

We provided a table-format overview (Fig. 1a) of analyzed identity attributes (G1). A table is an arrangement of data in rows and columns and is widely used in data analysis. Commonly, univariate, bivariate, and three-variate analyses are conducted in a table. Our table view also supports them as follows:

The table view supports bivariate analysis. When a user analyzes two identity attributes of rallies. She/He can drag the two identity attributes to the column and row panels separately (Fig. 1a1). Then the column and row titles of the table will become values of these two attributes and the rallies are classified accordingly. For instance, the user drags "Serving player" (values including "PlayerA Serves" and "PlayerB Serves") and "Scoring player" (values including "PlayerA Scores" and "PlayerB Scores") into the column and row panels separately (Fig. 1a1) and the values appear in the titles (Fig. 1a). The rallies are hence classified according to the combinations of the values. The table view also supports univariate and three-variate analyses. When a user analyzes one identity attribute, she/he only drags one attribute into the column panel. When a user analyzes three identity attributes of rallies. She/He can drag the two identity attributes to the column and one to the row panel. The column titles will be hierarchical (Fig. 5b).

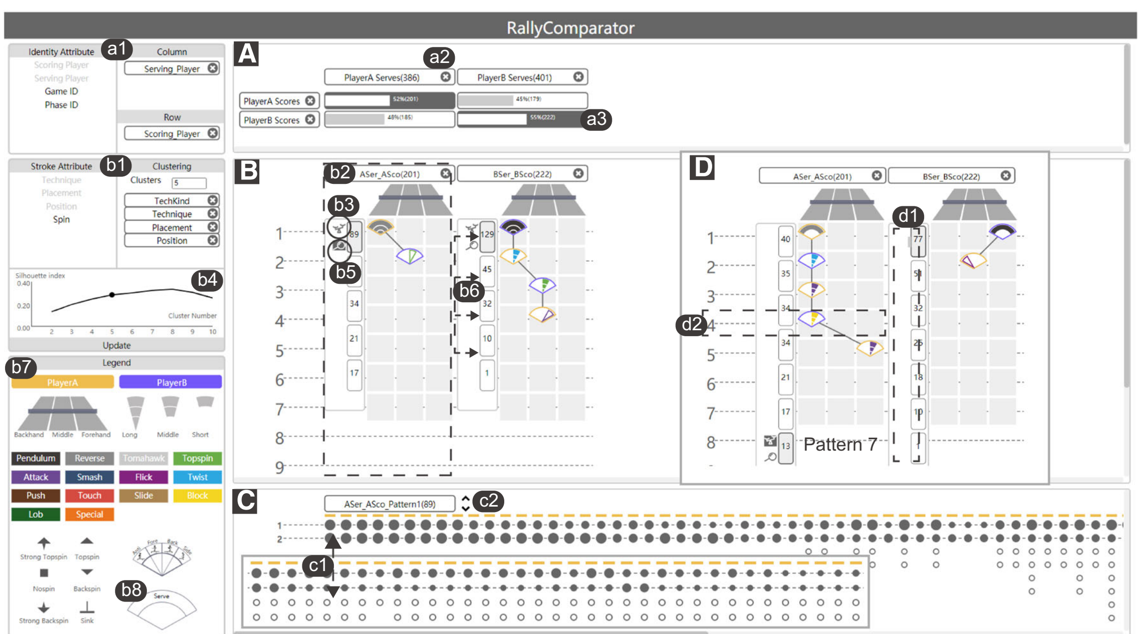

Fig. 1 The system interface comprises a table view **a** (including a control panel (a1) and a table), a pattern view **b** (including a control panel (b1), a legend panel (b7), and multiple flow views (b2)), and a detail view **c** (including the display of similarities (c1) and sequences (Fig. 4b))

Each cell in the table view is a rally group. The number of rallies is displayed directly. In a table, analyzers commonly analyze the attribute in the row in terms of the column. Therefore, we encoded the percentage of rallies in each cell versus rallies in the column as a bar. Interactions are summarized as follows:

5.4 Pattern view

After a user clicks on a cell in the table view (Fig. 1a3), a flow view (Fig. 1b2) that presents the corresponding stroke sequence group *G* (introduced in Sect. 4) will appear in the pattern view (Fig. 1b). The number of stroke sequences is displayed in the title of the flow view. The back-end algorithm clusters the sequences in the group into different stroke sequence clusters. Both the stroke attributes and cluster number for clustering can be adjusted interactively (G2, Fig. 1b1). A line chart is given in the control panel (Fig. 1b4) to help users quickly browse the Silhouette index (Rousseeuw 1987) of different cluster numbers and select an appropriate one. The Silhouette index is a widely used internal clustering validation measure that indicates the compactness within each cluster and the separation between clusters. The line chart helps users evaluate the quality of the clustering results when different cluster numbers are used.

Each cluster is represented by a stroke sequence with the highest similarities to other sequences in the cluster. The reason why we use a real stroke sequence instead of a cluster center or sequence summary is that it is easier for experts in table tennis to interpret a real sequence example. The flow view (Fig. 1b2) encodes a representative stroke sequence as a glyph-based flow, which satisfies multivariate display and sequential trends detection. Representative stroke sequences of multiple clusters in a stroke sequence group can be quickly examined by clicking corresponding labels (G3, Fig. 1b6). Besides, multiple flow views are juxtaposed and aligned based in the pattern view for efficient comparison. The pattern view can display at most six flow views, which is enough for the analysis requirement because experts in the table tennis domain usually compare two or four groups of stroke sequences. For instance, two groups of stroke sequences were compared in the case studies (Fig. 1b). Comparing more than four groups requires a heavy cognitive burden as each group still has many clusters of stroke sequences. After a user clicks on a focused pattern in a flow

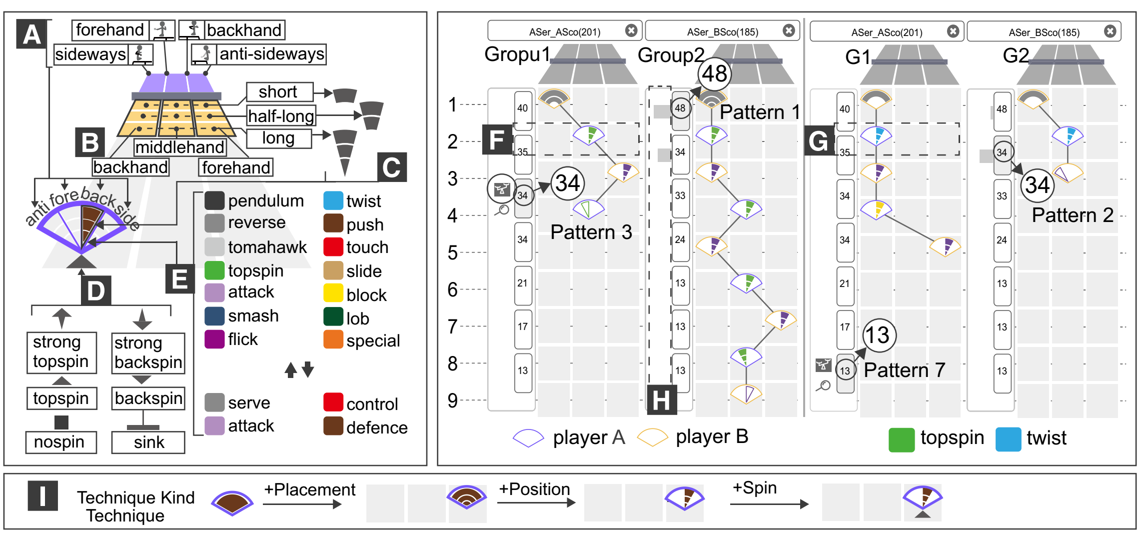

Fig. 2 (**a**)–(**e**) and (**i**) illustrate encodings of five technical attributes in the fan glyph. (**a**) Four strike position values, which relate to different positions where the player hits the ball, are encoded as four orientations in the fan glyph. (**b**) The horizontal ball position values are encoded as three horizontal positions of the fan glyph. (**c**) The vertical ball position values are encoded as the number of sectors in the fan glyph. (**d**) The spin values are encoded as different arrow-shaped geometries. (**e**) The 14 technique values or six state values are alternatively encoded as color hues. Notably, only a few technique values appear frequently. Experts commonly analyze at most five technique values simultaneously. (**i**) The fan glyph originates from a simple fan geometry encoding the "technique" and other attributes are loadable into the initial glyph (**f**)–(**h**) illustrate similar patterns in group one and two are detected in flows one and two separately in the case study discussed in Sect. 6.1.2. The left part shows pattern 3 in group one and its most similar pattern (pattern 1) in group two. The right part shows pattern 7 in group one and its most similar pattern (pattern 2) in group two. Player B attacked with the technique "topspin" (**f**) or "twist" (**g**) after player A served in these patterns. Supporting sequences in group one (34 and 13) are less than those in group two (48 and 34). (**h**) shows the bars indicating similarities of patterns in group two to pattern 3 in group one

view, the pattern's overall similarities to patterns in other flow views will be explicitly displayed (G4, Fig. 2h). We will introduce the designs in detail as follows.

5.4.1 Glyph design

Spatial variation of ball and player positions. Fig. 2b shows the horizontal ball position values. According to experts, the horizontal difference in the ball position is essential. We thus used an outstanding channel, the position (Munzner 2014) (Fig. 2b), to represent the difference. As this attribute indicates the horizontal variation of the ball positions in the table tennis table, we placed a table tennis table above the flow and used the horizontal side of the table as the x-axis of the flow. The horizontal position of a glyph thus encodes the horizontal position value of the encoded stroke. This design refers to the meanings of attributes and is comprehensible for experts.

As shown in Fig. 2c, the ball positions in the table can be divided into three groups in the vertical direction to indicate their different distances from the middle of the table. We hence used the number of the sectors to encode this vertical distance (Fig. 2c). The longer the distance the ball position is from the middle of the table, the longer the sectors. In this way, the experts in the table tennis domain can browse how long the ball positions are in different strokes intuitively.

Figure 2a shows four distinct strike position values related to four positions where the player hits the ball. The ball thus comes from different directions of the different positions. We employed a fan to represent a stroke and used four orientations within each fan to represent four strike position values (Fig. 2a). Legg et al. (2017) suggested that orientation is a well-separable visual channel in the glyph. The experts commented that the direction in each glyph shows the variation of spatial patterns over the stroke sequence very well. A serve stroke does not have the strike position; therefore, we used the glyph without an orientation to represent it (Fig. 1b8).

Techniques and spins. Figure 2e shows the strike technique and state values. According to the experts in the table tennis domain, analysts do not need to analyze the two attributes simultaneously because the strike state and technique have great overlap in meaning. We thus employed an outstanding channel, the color hue Munzner (2014) of the sector, to alternatively encode one of them. As there are fourteen kinds of strike techniques, we carefully selected fourteen color hues to encode them. However, fourteen colors are still too many for users to remember. Therefore, these colors are mainly used to distinguish different technique values. When users want to identify which technique value a color encodes, they can either refer to the legend on the left of the system or hover on the glyph to examine the information.

The spin attribute is closely related to the upward and downward directions. We used arrow-shaped geometries to encode spin values (Fig. 2d).

These encodings of the fan glyph are good metaphors for the corresponding stroke attributes and they are well coordinated. A question is that whether so many encodings in a glyph will require a heavy cognitive burden. Based on our observations, experts in the table tennis domain are very familiar with the table tennis data and often analyzed the multiple attributes of strokes simultaneously in videos and Excel tables. Therefore, all they need to do is mapping the attributes they are familiar with to the visual encodings. Although they still need some time to learn and remember the mapping, they are willing to pay the time to learn the encoding because the visualization tool will help them greatly in analyzing the stroke sequences. According to the experts, their analysis of stroke attributes strictly follows the pipeline illustrated in Fig. 2i. Therefore, the fan glyph originates from a simple fan geometry encoding the "technique" and other attributes are loadable into the initial glyph.

5.4.2 Glyph-based flow view

A glyph represents a stroke and a glyph-based flow represents a stroke sequence pattern. Multiple patterns can be detected in a sequence group. They can be examined one by one by clicking the labels in the flow view. The number of sequences that support the pattern is displayed on the label. The glyph-based flow view facilitates analysts' abilities to browse the multiple attributes of strokes and detect the trend in a sequence.

Multiple flow views can be placed side by side for comparison. As shown in Fig. 1b, the flow views are aligned according to the sequence number axis. Notably, the sequence number in stroke sequence is meaningful and important, which is slightly different from common even sequence, e.g., web clickstreams (Liu et al. 2017). Therefore, we assigned stroke weights according to different sequence numbers in the clustering process (introduced in Sect. 4). We also employed a sequence number axis and auxiliary lines in the pattern view to notify the sequence numbers of strokes in patterns and provide direct links among corresponding strokes. When a user focuses on a pattern, she/he can further click two icons (Fig. 1b3 and b5) beside the label of this pattern. The first icon (Fig. 1b3) is for displaying direct encodings of similarities among this pattern and patterns in other groups. For instance, when a user clicks the first icon of pattern 7 in group one, a bar appears in each pattern in group two (Fig. 2h). The height of the bar encodings the similarity with the focused pattern. In this case, no patterns in group two are similar to pattern 7 and no high bars can be detected. These encodings and interactions help the user to quickly find out whether there are similar patterns of the focused pattern in other groups and what they are. A second icon is for displaying detailed sequences of the focused pattern. After clicking the icon, a detail view will appear. Interactions are summarized as follows.

– *Adjust parameters for clustering:* Users can add or delete stroke attributes for clustering by dragging and clicking and adjust the cluster number in the control panel (Fig. 1b1).
– *Switch displayed patterns in a flow view:* Users can examine patterns in a sequence group quickly by switching labels (Fig. 1b6) in a flow view.
– *Click the comparing icon for direct encodings of overall similarities:* Users can click the comparing icon beside the label of a focused pattern (Fig. 1b3) to display direct similarities of this pattern and patterns in other sequence groups. The similarities will be encoded by the height of bars beside labels of other patterns.
– *Click the detail icon to examine the sequences of a pattern:* Users can click the detail icon beside the label of a focused pattern (Fig. 1b5) to examine sequences of this pattern for interpretation and verification.

5.4.3 Design alternatives

Two alternative glyphs, namely, the text and table glyphs were proposed. The text glyph (Fig. 3a) directly lists the names of attribute values. We used this glyph as an alternative because experts used to employ this kind of encoding to browse and compare strokes in Excel tables. We added a color label at the bottom of each attribute value to denote the attribute it belongs to for enhancing the glyph. The text glyph is abstract but familiar to the experts. It presents the stroke clearly and accurately. However, multiple text glyphs are not appropriate for users to browse at the same time. According to the experts, it is not efficient for users to browse the text glyphs than fan glyphs.

The table glyph (Fig. 3b) refers to the tactic view of iTTVis (Wu et al. 2018). We modified this glyph to encode the same information as the fan glyph. A half table tennis table is divided into nine grids and a filled grid presents horizontal and vertical positions of the ball position. The color of the filled grid encodes the strike technique or strike state. The figure in the top left of the glyph imitates the state of players and represents the strike position. The same set of geometries in the bottom left of the glyph encode the strike spin. The table glyph is intuitive and metaphoric. However, the glyph is not unified. That is, the components in the glyph are separate and not well-organized. Therefore, this glyph is not so easy to be perceived as a whole stroke compared to the fan glyph (Fig. 3c). The table glyph is also not suited for displaying the vertical variation of ball position in when a sequence of table glyphs are arranged vertically. The vertical spatial variation of the ball position is difficult to detect. The experts prefer the fan glyph to the table glyph.

5.5 Detail view

The detail view is displayed when the detail icon of a focused pattern is clicked for interpretation and verification of this pattern (G5). This view presents all sequences that support the pattern. The users can

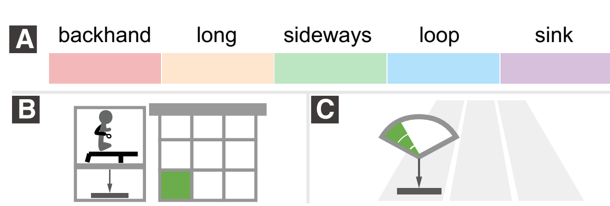

Fig. 3 Three alternative glyphs: **a** text glyph, **b** table glyph, and **c** fan glyph

examine these sequences and their similarities and differences compared to the pattern for a better understanding of the aggregation.

Display of Similarities. The detail view displays stroke-specific similarities of sequences and the pattern. As shown in Fig. 1c, each column is a sequence and the area of each circle in the column encodes the similarity of a stroke in the sequence and the corresponding one (with the same sequence number) in the pattern. If the the sequence is longer than the pattern, the extra strokes will be encoded as empty circles to indicate that these strokes are not similar to those in the representative sequence. The sequences are ordered according to the overall similarities to the pattern from left to right. The display of similarities enables the user to browse and detect the distributions and trends of similarities of sequences and the pattern.

Display of Sequences. A more detailed display of sequences is required to further interpret the distributed similarities (Fig. 4b). When the user clicks the expand button (Fig. 1c2), the detail view is expanded and detailed attributes of sequences are displayed for browsing and understanding. All sequences are encoded in the same way with the pattern view and ordered according to overall similarities to the pattern (Fig. 4b). The display of sequences enables the user to browse and detect the distributions and trends of attributes of sequences to understand and verify the aggregated pattern.

6 Evaluation

This section presents the evaluation part of this study. We employed two case studies with the experts to demonstrate the effectiveness and usability of the system. The study was conducted on Google Chrome on a PC (equipped with Intel Xeon E3, 32GB of memory, and a 1920*1080 display).

6.1 Case studies

We invited expert A and B (introduced in Sect. 3.2) to use the system and detect patterns. The data used for analysis are on seven recent matches between the Japanese player Itou Micheng (player A) and Chinese players (player B). The Chinese players included Wang Manyu, Chen Xingtong, Liu Shiwen, Ding Ning, and Sun Yingsha. Experts aimed to find the differences in the strokes of Ito Micheng and the Chinese players and explore the reasons for the differences. We offered a tutorial on how to use the system for expert B and answered his questions. Experts were then asked to use the system and explore the data on his own. We were nearby to provide help all the time. After that, experts conducted an in-depth comparative analysis of the strokes of players A and B and derived insightful patterns. The results were summarized as the case studies. We then conducted a semi-structured interview with the two experts and collected their feedback on the system.

6.1.1 Pattern examination in a Stroke sequence group

In this case study, the experts selected sequence groups based on the patterns in the table overview (G1), adjusted the cluster numbers interactively (G2), compared patterns in the pattern view and found a unique pattern in group one (G3, G4), and verified and interpreted the pattern in the detail view (G5).

The experts compared sequences in rallies that player A (group one) and player B (group two) served and scored. He selected these two groups because the table view shows that both players score more in their

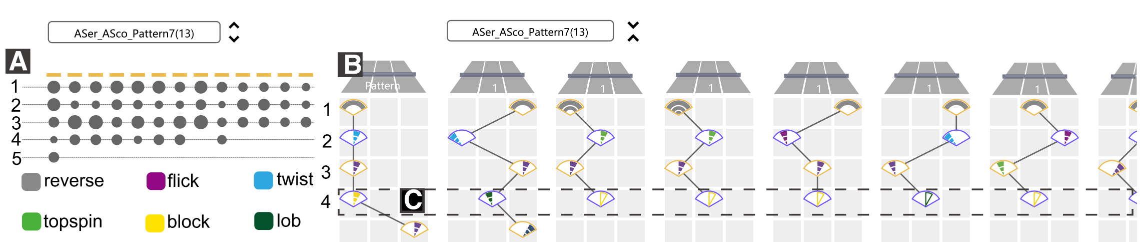

Fig. 4 The detail view of pattern 7 in the case study discussed in Sect. 6.1.1. **a** shows similarities of all supporting sequences to pattern 7. Similarities in the first and third rows are larger than those in the second and fourth rows. **b** shows pattern 7 and detailed supporting sequences side by side. The second and fourth strokes vary in colors indicating different technique values, while the first and third strokes are almost identical in colors. **c** All fourth strokes are defense strokes according to their colors

serve rallies (The height of bars in these two selected cells are both higher than those in the other cells in the columns in Fig. 1a).

The experts firstly clustered the sequences in each group into five clusters by default. Then he examined the detail view of the pattern with the most support of group one (Fig. 1b5) and found that half of the rallies in this cluster are very different from the other half (Fig. 1c1). The lengths of rallies are different (the numbers of circles are different) and the similarities of strokes are low (the areas of circles are small). He hence added up the cluster number to seven to split the largest cluster.

He then obtained seven appropriate patterns and examined each pattern in group one. He found that pattern 7 in group one is not similar to any patterns in group two (He clicked the comparing icon besides pattern 7 and all bars in group two are very low in Fig. 1d1). He browsed the patterns in group two and commented that pattern 7 in group one is special for containing a defense stroke by player B at the fourth stroke (Fig. 1d2). This pattern that player B defended at the fourth stroke will only appear in the rallies that player A served and scored (group one).

He examined the pattern in detail for interpretation and verification. He looked at the detail view and quickly found that the length range of sequences in this cluster is between three and five (according to the numbers of circles in all columns in Fig. 4a). The first and third stroke in sequences is more similar to the pattern sequence compared to the second and fourth stroke (the areas of the circles in the second and fourth rows are smaller than those in the first and third rows). In order to understand the distribution, the experts expanded the detail view to browse all sequences (Fig. 4b). He found that the first and third strokes in all sequences were mostly strokes with the techniques “Reverse Serve” and “Flick” (according to the colors of strokes). The second strokes varied among techniques “Twist,” “Topspin,” and “Flick,” and were all attack strokes. The fourth strokes (Fig. 4c) varied between techniques “Block” and “Lob”, and were all defense strokes. This information helped the experts understand the similarity distribution detected just now and convinced him of the specialty of the clustered pattern that the fourth stroke is a defense stroke.

The pattern and associated detail view inspired the experts to pay attention to the defense at the fourth stroke for its effects on scoring.

6.1.2 Pattern comparison among Stroke sequence groups

In this case study, the experts located the sequence groups he wanted to explore based on his experiences in the table overview (G1), and identified interesting pairs of patterns for insights through comparing clustered patterns of two groups in the pattern view (G2, G3, G4).

Experts selected two stroke sequence groups, namely, sequences in rallies player A scores (group one) and loses (group two) when she serves (Fig. 5a), for comparison based on his interest. The experts also specified the three stroke attributes and set the cluster number as seven.

In the pattern view, the experts quickly examined the seven patterns in group one and compared them with patterns in group two. He found insights in two kinds of patterns (Figs. 2 and 6). The first kind of patterns are those where player B attacked after player A served (patterns 3 and 7 in group one in Fig. 2). As shown in pattern 3 in group one, player B attacked at the second stroke with the technique “Topspin” (according to the color of the stroke in Fig. 2f) and the rally did not end. After the experts clicked the comparing icon of this pattern, he found the most similar pattern, pattern 1 in group two. He examined

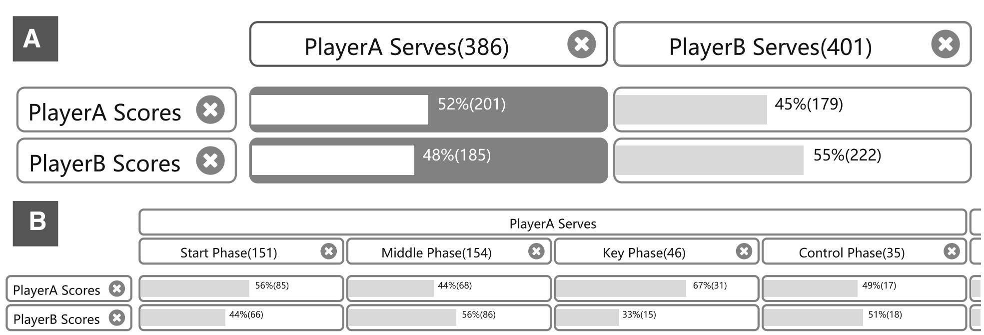

Fig. 5 **a** The expert selected two cells, namely, player A serves&scores and player A serves but player B scores, in the table view in the case study discussed in Sect. 6.1.2. **b** When a user drags two and one identity attributes to the column and row panels separately, the column titles will be hierarchical

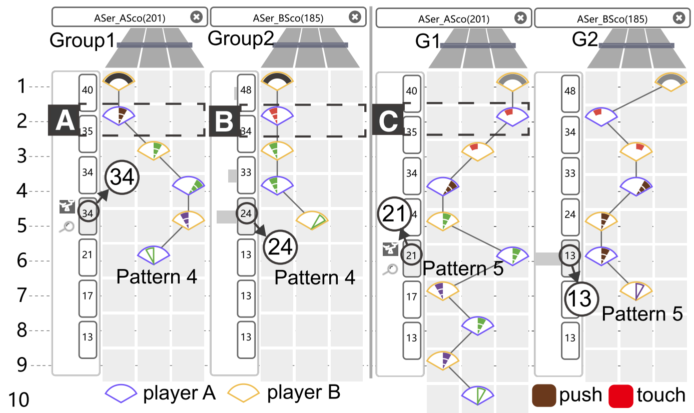

Fig. 6 Similar patterns in groups one and two are detected in flows one and two separately in the case study discussed in Sect. 6.1.2. The left part shows pattern 4 in group one and its most similar pattern (pattern 4) in group two. The right part shows pattern 5 in group one and its most similar pattern (pattern 5) in group two. Player B controlled with technique "push" (**a**) or "touch short" (**b** and **c**) after player A served in these patterns. Supporting sequences in group one (34 and 24) are more than those in group two (21 and 13)

pattern 1 and confirmed the similarity. The number of sequences that support pattern 1 in group two is larger than pattern 3 in group one. Also, the experts examined pattern 7 in group one. In the pattern, player B also attacked at the second stroke but with another technique "Twist" (Fig. 2g). The most similar pattern in group two (pattern 2) was also examined by the experts and the similarity was confirmed. The number of sequences that support a similar pattern in group two is also larger than that in group one. Therefore, the experts commented that the pattern that player B attacked right after player A served appeared in both rallies that player A scored (group one) and lost (group two). However, the pattern appeared more frequently in rallies that player A lost. Player B hence had more chances to score when he attacked at the second stroke.

The second kind of patterns are those player B controlled after player A served (player 4 and 5 in group one). As shown in pattern 4 in group one, player B controlled at the second stroke with the technique "Push" (Fig. 6a) and the rally went on. After he clicked the comparing icon of this pattern, he found the most similar pattern, pattern 4 in group two. He examined the similar pattern and confirmed the similarity. He explained that although the control technique "Touch" of the second stroke in group two (Fig. 6b) is different from that in group one, it was not a significant difference. The number of sequences that support this pattern in group one is larger than that in group two. Also, the experts examined pattern 5 in group one. In the pattern, player B also attacked at the second stroke with the control technique "Touch" and hit the ball to the short area (Fig. 6c). The very similar pattern in group two (pattern 5) also presents this. The number of sequences that support a similar pattern in group one is also larger than that in group two. The experts commented that the pattern that player B controlled right after player A served appeared in both rallies that player A scored (group one) and lost (group two). However, the pattern appeared more frequently in rallies that player A scored. Therefore, player B would have fewer chances to score after he chose to control at the second stroke.

These two patterns inspired the experts to attach importance to the two technique kinds, namely, attack and control, of the second stroke because they have influences on scoring.

6.2 Expert interview

Experts A and B have used the system to explore and compare patterns in stroke sequences and obtain insights into the typical patterns of sequences. The exploration processes are summarized as case studies as illustrated above. After that, we let experts answer open-ended questions about the effectiveness and usability of the system (10 minutes for each expert). Their feedback and suggestions are summarized as follows. Theoretically, the experts thought the system filled the gap of bottom-up pattern identification in stroke sequence. Experts in the table tennis domain have a top-down understanding of the typical patterns of stroke sequences based on their experiences. However, different experts will have different top-down understandings and it is hard to communicate and verify the typical patterns of stroke sequences in their minds. The system provides an objective and bottom-up way to extract patterns from stroke sequences. The empirical results can inspire experts to reflect on their understandings of typical patterns and provide evidence for hypotheses and conclusions.

Practically, experts appreciated the useful functions, effective designs, and commendable usability of the system. Expert A commented,"The system is very useful and thoughtful for advanced comparisons of players' strokes." Experts commented,"The way the flow view presents strokes matches well with the way we players look at the ball. It feels like the strokes flow out of the table. This view commendably brings reality into data analysis." He also commented,"The interactions of the system are direct and simple and the transitions between two views are smooth and understandable."

The experts also suggested the integration of more domain rules which we will discuss in the first limitation in the next section.

7 Discussion and conclusion

This section discusses the limitations of this study, lessons learned in the design process of the system, and the generalizability of the design and system.

7.1 Connections to visualization community

The visualization problem in this study can be regarded as an event sequence visualization. The first difficulty is that the event in the sequence is multivariate. At most four attributes of strokes are required to be examined simultaneously. We employed a glyph to display the multiple attributes of strokes. The second difficulty is to outstand the spatial variation, e.g., the spatial variation of the ball and player's positions. We employed the position and direction visual channels of the glyph to encode the spatial position. This study provides a good example of the problem of visualizing multivariate and spatial event sequences. The design can be referred to when visualization practitioners need to visualize spatial or multivariate event sequences in other domains. Moreover, our pipeline of comparative analysis of stroke sequences can be applied to other sports with similar sequential data and multiple attributes, such as tennis and badminton.

Another point is to discuss is what a visualization pattern is and how it is related to a domain. A visualization pattern is a combination of two or more related data components (Andrienko et al. 2021). Many previous studies have presented different visualization patterns (Zhao et al. 2019; Wang et al. 2020). In this study, the patterns are composed of multiple attributes of multiple strokes, which is based on table tennis domain knowledge. This indicates that the pattern can be nearly arbitrary in terms of visualization designs and the constraint is the meanings of the pattern based on domain knowledge.

7.2 Limitations

Integration of more domain rules. The system integrates a set of domain rules, such as the attribute importance and sequence number differences, in the clustering and display. However, the expert suggested that many other rules, such as "the stroke that the vertical position varies from short or half-long to long is important" and "the variation of horizontal position becomes more important after the first four strokes" are not considered. In fact, many untidy and intangible rules have been generated based on the rich experiences and long-term study in the table tennis domain. More efforts are needed to summarize and integrate this knowledge into the exploration of stroke sequences.

7.3 Conclusion

This study investigated the problem of comparative analyses of stroke sequences in table tennis. We worked closely with experts to summarize domain requirements about sequential patterns. We also developed a visualization system integrating a clustering algorithm for extracting patterns, a glyph-based flow view for the multivariate display of patterns, and a set of interactions for efficient comparison. The two case studies demonstrate the effectiveness and usability of the system. In the future, we plan to (1) refine the glyphs and use them as visual symbols of table tennis stroke attributes; (2) integrate more domain rules into the sequence clustering.

Acknowledgements The work was supported by Zhejiang Provincial Natural Science Foundation (LR18F020001). This project was also funded by the Chinese Table Tennis Association.

References

Alexander EC, Gleicher M (2016) Task-driven comparison of topic models. IEEE Trans Vis Comput Graphi 22(1):320–329
Andrienko N, Andrienko G, Miksch S, Schumann H, Wrobel S (2021) A theoretical model for pattern discovery in visual analytics. Vis Inf 5(1):23–42
Cappers BC, van Wijk JJ (2017) Exploring multivariate event sequences using rules, aggregations, and selections. IEEE Trans Vis Comput Graph 24(1):532–541
Chen H, Zhang S, Chen W, Mei H, Zhang J, Mercer A, Liang R, Qu H (2015) Uncertainty-aware multidimensional ensemble data visualization and exploration. IEEE Trans Vis Comput Graph 21(9):1072–1086
Chen W, Lao T, Xia J, Huang X, Zhu B, Hu W, Guan H (2016) GameFlow: Narrative visualization of NBA basketball games. IEEE Trans Multimedia 18(11):2247–2256
Chen Y, Xu P, Ren L (2018) Sequence synopsis: optimize visual summary of temporal event data. IEEE Trans Vis Comput Graph 24(1):45–55
Chen K, Wang Y, Yu M, Shen HW, Yu X, Shan G (2021) ConfVisExplorer: a literature-based visual analysis system for conference comparison. J Vis 24(2):381–395
Chen R, Shu X, Chen J, Weng D, Tang J, Fu S, Wu Y (2021) Nebula: A coordinating grammar of graphics. IEEE Trans Vis Comput Graph
Chen Z, Ye S, Chu X, Xia H, Zhang H, Qu H, Wu Y (2022) Augmenting sports videos with viscommentator. To appear in IEEE Trans Vis Comput Graph 28(1)
Chu X, Xie X, Ye S, Lu H, Xiao H, Yuan Z, Chen Z, Zhang H, Wu Y (2022) TIVEE: Visual exploration and explanation of badminton tactics in immersive visualizations. To appear IEEE Trans Vis Comput Graph 28(1)
Du F, Shneiderman B, Plaisant C, Malik S, Perer A (2017) Coping with volume and variety in temporal event sequences: Strategies for sharpening analytic focus. IEEE Trans Vis Comput Graph 23(6):1636–1649
Du F, Plaisant C, Spring N, Shneiderman B (2016) EventAction: Visual analytics for temporal event sequence recommendation. In: Proceedings of IEEE Conference on Visual Analytics Science and Technology, pp. 61–70
Filipov V, Schetinger V, Raminger K, Soursos N, Zapke S, Miksch S (2021) Gone full circle: a radial approach to visualize event-based networks in digital humanities. Vis Inf 5(1):45–60
Fuchs J, Isenberg P, Bezerianos A, Keim DA (2017) A systematic review of experimental studies on data glyphs. IEEE Trans Vis Comput Graph 23(7):1863–1879
Gleicher M (2018) Considerations for visualizing comparison. IEEE Trans Vis Comput Graph 24(1):413–423
Gleicher M, Albers D, Walker R, Jusufi I, Hansen CD, Roberts JC (2011) Visual comparison for information visualization. Inf Vis 10(4):289–309
Glueck M, Naeini MP, Doshi-Velez F, Chevalier F, Khan A, Wigdor D, Brudno M (2018) PhenoLines: phenotype comparison visualizations for disease subtyping via topic models. IEEE Trans Vis Comput Graph 24(1):371–381
Gotz D (2016) Soft patterns: Moving beyond explicit sequential patterns during visual analysis of longitudinal event datasets. In: Proceedings of the IEEE VIS Workshop on Temporal & Sequential Event Analysis
Guo S, Xu K, Zhao R, Gotz D, Zha H, Cao N (2018) EventThread: visual summarization and stage analysis of event sequence data. IEEE Trans Vis Comput Graph 24(1):56–65
Guo R, Fujiwara T, Li Y, Lima KM, Sen S, Tran NK, Ma KL (2020) Comparative visual analytics for assessing medical records with sequence embedding. Vis Inf 4(2):72–85
He W, Wang J, Guo H, Shen HW, Peterka T (2020) CECAV-DNN: collective ensemble comparison and visualization using deep neural networks. Vis Inf 4(2):109–121
Jin Z, Cao N, Shi Y, Wu W, Wu Y (2021) EcoLens: visual analysis of ecological regions in urban contexts using traffic data. J Vis 24(2):349–364
Jin Z, Chen N, Shi Y, Qian W, Xu M, Cao N (2021) TrammelGraph: visual graph abstraction for comparison. J Vis 24(2):365–379
Kehrer J, Hauser H (2013) Visualization and visual analysis of multifaceted scientific data: a survey. IEEE Trans Vis Comput Graph 19(3):495–513
Lames M, McGarry T (2007) On the search for reliable performance indicators in game sports. Int J Performance Anal Sport 7(1):62–79
Legg PA, Maguire E, Walton SJ, Chen M (2017) Glyph visualization: a fail-safe design scheme based on quasi-hamming distances. IEEE Comput Graph Appl 37(2):31–41
Li Y, Fujiwara T, Choi YK, Kim KK, Ma KL (2020) A visual analytics system for multi-model comparison on clinical data predictions. Vis Inf 4(2):122–131
Liu Z, Wang Y, Dontcheva M, Hoffman M, Walker S, Wilson A (2017) Patterns and sequences: interactive exploration of clickstreams to understand common visitor paths. IEEE Trans Vis Comput Graph 23(1):321–330
Loh TC, Krasilshchikov O (2015) Competition performance variables differences in elite and u-21 international men singles table tennis players. J Phys Edu Sport 15(4):829
Mei H, Chen W, Wei Y, Hu Y, Zhou S, Lin B, Zhao Y, Xia J (2019) Rsatree: Distribution-aware data representation of large-scale tabular datasets for flexible visual query. IEEE Trans Vis Comput Graph 26(1):1161–1171
Munzner T (2014) Visualization analysis and design. A.K Peters visualization series. A K Peters, Natick
Pfeiffer M, Zhang H, Hohmann A (2010) A markov chain model of elite table tennis competition. Int J Sports Sci Coach 5(2):205–222
Polk T, Yang J, Hu Y, Zhao Y (2014) TenniVis: visualization for tennis match analysis. IEEE Trans Vis Comput Graph 20(12):2339–2348

Polk T, Jäckle D, Häußler J, Yang J (2020) CourtTime: generating actionable insights into tennis matches using visual analytics. IEEE Trans Vis Comput Graph 26(1):397–406
Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65
Sedlmair M, Meyer MD, Munzner T (2012) Design study methodology: reflections from the trenches and the stacks. IEEE Trans Vis Comput Graph 18(12):2431–2440
Shi D, Xu X, Sun F, Shi Y, Cao N (2020) Calliope: automatic visual data story generation from a spreadsheet. IEEE Trans Vis Comput Graph 27(2):453–463
Shu X, Wu J, Wu X, Liang H, Cui W, Wu Y, Qu H (2021) DancingWords: exploring animated word clouds to tell stories. J Vis 24(1):85–100
Table tennis. https://en.wikipedia.org/wiki/Table_tennis (2018)
Tang T, Tang J, Hong J, Yu L, Ren P, Wu Y (2020) Design guidelines for augmenting short-form videos using animated data visualizations. J Vis 23(4):707–720
Types of strokes. https://en.wikipedia.org/wiki/Table_tennis#Types_of_strokes (2018)
von Luxburg U (2007) A tutorial on spectral clustering. Statist Comput 17(4):395–416
Wang J, Zhao K, Deng D, Cao A, Xie X, Zhou Z, Zhang H, Wu Y (2020) Tac-Simur: tactic-based simulative visual analytics of table tennis. IEEE Trans Vis Comput Graph 26(1):407–417
Wang J, Wu J, Cao A, Zhou Z, Zhang H, Wu Y (2021) Tac-Miner: visual tactic mining for multiple table tennis matches. IEEE Trans Vis Comput Graph 27(6):2770–2782
Wang X, Bryan CJ, Li Y, Pan R, Liu Y, Chen W, Ma KL (2020) Umbra: A visual analysis approach for defense construction against inference attacks on sensitive information. IEEE Trans Vis Comput Graph
Wang X, Chen W, Xia J, Chen Z, Xu D, Wu X, Xu M, Schreck T (2020) ConceptExplorer: Visual analysis of concept drifts in multi-source time-series data. In: Proceedings of IEEE Conference on Visual Analytics Science and Technology, pp. 1–11
Weng D, Zheng C, Deng Z, Ma M, Bao J, Zheng Y, Xu M, Wu Y (2021) Towards better bus networks: a visual analytics approach. IEEE Trans Vis Comput Graph 27(2):817–827
Wenninger S, Lames M (2016) Performance analysis in table tennis-stochastic simulation by numerical derivation. Int J Comput Sci Sport 15(1):22–36
Wongsuphasawat K, Guerra Gómez JA, Plaisant C, Wang TD, Taieb-Maimon M, Shneiderman B (2011) LifeFlow: Visualizing an overview of event sequences. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp. 1747–1756
Wu Y, Lan J, Shu X, Ji C, Zhao K, Wang J, Zhang H (2018) iTTVis: interactive visualization of table tennis data. IEEE Trans Vis Comput Graph 24(1):709–718
Wu Y, Xie X, Wang J, Deng D, Liang H, Zhang H, Cheng S, Chen W (2019) ForVizor: visualizing spatio-temporal team formations in soccer. IEEE Trans Vis Comput Graph 25(1):65–75
Wu Y, Weng D, Deng Z, Bao J, Xu M, Wang Z, Zheng Y, Ding Z, Chen W (2020) Towards better detection and analysis of massive spatiotemporal co-occurrence patterns. IEEE Trans Intell Transp Syst 22(6):3387–3402
Wu J, Guo Z, Wang Z, Xu Q, Wu Y (2020) Visual analytics of multivariate event sequence data in racquet sports. In: Proceedings of IEEE Conference on Visual Analytics Science and Technology, pp. 36–47
Wu J, Liu D, Guo Z, Xu Q, Wu Y (2022) TacticFlow: Visual analytics of ever-changing tactics in racket sports. To appear in IEEE Trans Vis Comput Graph 28(1)
Xie X, Wang J, Liang H, Deng D, Cheng S, Zhang H, Chen W, Wu Y (2021) PassVizor: toward better understanding of the dynamics of soccer passes. IEEE Trans Vis Comput Graph 27(2):1322–1331
Ye S, Chen Z, Chu X, Wang Y, Fu S, Shen L, Zhou K, Wu Y (2021) ShuttleSpace: exploring and analyzing movement trajectory in immersive visualization. IEEE Trans Vis Comput Graph 27(2):860–869
Zhao Y, Luo X, Lin X, Wang H, Kui X, Zhou F, Wang J, Chen Y, Chen W (2019) Visual analytics for electromagnetic situation awareness in radio monitoring and management. IEEE Trans Vis Comput Graph 26(1):590–600
Zhao Y, Jiang H, Qin Y, Xie H, Wu Y, Liu S, Zhou Z, Xia J, Zhou F et al (2020) Preserving minority structures in graph sampling. IEEE Trans Vis Comput Graph 27(2):1698–1708