Visual Analytics of Impacts of Behavior Adjustments on Scoring Rates in Table Tennis

Anonymous cvm submission

Paper ID 317

Abstract

Markov chain models have been adopted in table tennis analytics to examine the impacts of behavior adjustment of players on scoring rates. Analysis of these impacts is of great significance for understanding players' performance. However, understanding the reason behind the impacts estimated by these complex models is non-trivial due to the complex transition relations between the adjustments and the impacts. In this work, we worked with domain experts to define typical behaviors of players and characterize the problem of understanding the impacts of behavior adjustment estimated by a modified Markov chain model. We further designed and developed a visualization system to provide a multi-level analysis of the impacts. The system consists of four views, namely, a player view summarizing the impacts for various players, a behavior view presenting impacts of different behaviors of a player, a flow view tracking the accumulation of impacts throughout a rally, and a matrix view revealing the generation process of impacts at each tactic. With the system, not only existing hypotheses were tested for the first time, but also a set of new hypotheses were formulated. We demonstrated the effectiveness of the system with two case studies.

1. Introduction

Improvements in the behaviors of players in table tennis matches are important for coaches and analysts. Players' behaviors are commonly characterized by various stroke attributes (e.g., stroke technique). As fine-grained data on table tennis matches have been acquired, a set of statistical methods and mathematical models [12, 10, 26, 9, 7, 6, 16, 24] have been proposed to help analysts better understand players' behaviors. Among these methods, the finite Markov chain model [6] is well applied to performance analysis in table tennis [16, 24]. In this model, a table tennis rally is viewed as a stochastic process representing attribute values randomly changing over strokes. Making adjustment to probabilities of attribute values during this stochastic process simulates behavior adjustment in a table tennis rally. The adjustment will change the estimated scoring rates of the model and these changes are the impacts of behavior adjustment on scoring rates.

However, existing works that applied this model to analyzing table tennis data failed to explore when and how the estimated impacts of adjustments occur. During a table tennis rally, which comprises a sequence of strokes, adjustments to behaviors exert impacts on the scoring rate at many strokes. The total impact estimated by the model are the accumulation of impacts at certain strokes. In addition to the total impact of adjusting the behavior, the timing of the accumulation of impacts in a rally and the manner through which the behavior adjustment exerts impacts at certain strokes are also required by the experts for an in-depth analysis of the impact. However, tracking the step-wise impacts and examining their reasons quickly and effectively are beyond the capabilities of traditional analytical methods. This work hence adopted visualization techniques to allow a comprehensible multi-level analysis of these impacts and reasons.

The major challenge of this work is to propose a visual analytics pipeline to decompose and visualize the total impacts of behavior adjustments layer by layer. The behaviors of table tennis players are not defined clearly, and a set of indicators can be used to describe them. Defining the typical behaviors that include major technical features and consider tactical associations is difficult. Behavior adjustments of players also exert complex multistep impacts on the scoring rate many times throughout a rally. Identifying the impacts at each time and revealing the patterns of the multistep generation process of the impacts are hence challenging. We worked closely with the experts to address the challenge. First, we identified the typical two-stroke behaviors characterized by the stroke technique and placement. Second, we proposed to decompose a rally into a set of three-stroke tactics of the analyzed player and measured the impact of his/her behavior adjustment at each tactic. Furthermore, we divided each tactic into two phases to help analysts understand the multi-step generation process of the impacts. We finally developed a visual analytics system on the basis of the analysis pipeline to provide a multi-level presentation of the impacts of behavior adjustments of various players. The system consists of four views, namely, a player view, a behavior view, a flow view, and a matrix view. The player view provides an overview of the statistical information about impacts of adjustments for different players. The behavior view presents impacts of adjustments to various behaviors for a player. The flow view adopts flow charts to provide visual tracking of the impact at each tactic. The matrix view uses a pair of matrices to reveal the patterns of the generation process of the impacts. The contributions of this work are as follows:

- Identification of a set of typical two-stroke behaviors and a method of quantifying the impacts of behavior adjustments on the scoring rate at each tactic;
- A visualization system that supports the multi-level exploration of the impacts of behavior adjustments on scoring rates in table tennis;
- Case studies that illustrate the augmented capacity of the in-depth analysis of the impacts of behaviors and a set of insights obtained for an insightful guidance on adjustments to players' behaviors.

2. Related Work

This section presents existing analytical methods for table tennis, and an overview of visual techniques in sports visualizations.

2.1. Analytical methods for Table Tennis

The common approach to evaluating the performances (i.e., traits of techniques and tactics) of table tennis players is the use of performance indicators statistics. Existing studies analyzed shot characteristics [12], game structure indicators [10], skill evaluation indicators [26], and the integration of fundamental indices [9]. These works studied the statistics of game observations collected from the matches. However, these statistics ignore the dynamic interactions of two players and do not consider the game observations systematically [7]. Lames [6] proposed the Markov Chain for simulating matches and identifying behaviors with most impacts on the scoring rate. This model considers the context of observations and builds a structure on the game. On this basis, Pfeiffer et al. [16] developed four different state-transition-models to describe the impacts of behaviors in a table tennis match. Wenninger et al. [24] used a similar model to reveal the factors that influence the scoring rate. However, these models only consider attributes of one stroke and relations between two strokes in the model due to the limitation of model complexity. They ignore the combinational use of stroke technique and placement and the three-stroke tactical associations among strokes, thus characterizing players' behaviors inadequately. Moreover, interpreting the computational methods and estimated results was difficult. Experts could not understand the reasons behind the impacts.

Unlike these works, our work provided a method of examining when and how the estimated impacts of behavior adjustments occur. Powered by the modified Markov chain model and a set of concise visualizations, the system supports a multi-level presentation of the impacts of behavior adjustments.

2.2. Sports Visualization

We classified existing works on sports visualization into three parts on the basis of the design goals.

Seeking spatial and temporal patterns. Most sports visualization methods aim to detect spatial or temporal patterns in sports data. Relevant works aimed to detect patterns in spatial distributions of observations in sports. Counterpoints [2] presents the spatial patterns of players' statistics in basketball, such as disruption scores. Baseball4D [1] and SnapShot [17] present the spatial patterns of hit points in baseball and shot lengths in ice hockey, respectively. StatCast Dashboard [5] explores the spatial distributions of various kinds of data in baseball. Previous works have also explored the patterns of trajectories in sports, used for obtaining the attack patterns of teams [14], detecting anomalous events [3], and examining movements at different levels [19]. Temporal patterns in sports data were also worth exploring. BKViz [11] reveals the patterns of temporal observations, such as the play types and point outcomes, of a basketball team in a match and throughout a season, respectively. TenniVis [18] and iTTVis [25] present the temporal patterns of varying scores and rally lengths in tennis and table tennis, respectively. These works utilized the high-bandwidth channel provided by visualizations and delivered spatial-temporal patterns to sports experts.

Revealing patterns of relations and structures. Sports visualizations revealing the relations in sports data have also been proposed. For instance, several works aimed to reveal patterns in team rankings. A Table [15] reveals the evolving rankings of soccer teams. Many works aimed to demonstrate the tree structure in matches or tournaments. Tan et al. [22] visualized the tree structure of a tournament to provide an understandable representation of the process and allow nonlinear predictions, respectively. TennisViewer [4] provides clear tree structures of tennis match data. Lu et al. Other works proposed the goals of displaying the many-to-many relations in sports data. SoccerStories [14], iTTVis [25], and TacSimur [23] use matrix views to visualize the passing rates among players in soccer match, , the correlations among attribute values over strokes, and tactical patterns respectively. These works help users clarify and analyze the complex relationships in sports data.

Achieving a glyph-based and annotation aided analysis. Many works on sports visualization aimed to achieve an intuitive analysis by using glyphs and metaphors, or directly adding annotations on the video. MatchPad [8] and iTTVis [25] use a set of intuitive and clear glyphs that imitate the actions of players to help experts grasp the events in a rugby game "in a glance" and recognize stroke positions in table tennis tactics. TenniVis [18] adopts several glyphs to intuitively present the game observations, such as the score and serve information. Another set of works integrated visualizations into the original videos of sports. Parray et al. [13] introduced a video storyboard that visually depicted and annotated events in snooker videos. Furthermore, Director's Cut [20] and Bring it to the Pitch [21] use computer vision techniques to extract the trajectory data and add visualizations of the important distance and area to the soccer match videos. These works integrated real elements of sports into the analysis process of sports data, thus allowing an intuitive analysis.

Our work also aimed to reveal patterns of relations and structures. Specifically, we aimed to display the many-to-many transition relations in each phase and the connections between two phases in a tactic. The difficulty and focus of our work, however, were to propose a visual analytics pipeline to navigate the exploration on the generation process of impacts throughout a rally. Previous visualization works on analyzing many-to-many relations in sports data could not be directly applied to our problems.

3. Background and System Overview

This section presents the background, data description, requirement analysis, and system overview.

3.1. Background and data description

A behavior's main features are characterized by two attributes. Stroke technique and placement are two most important attributes in describing the general behaviors of players. Specifically, stroke technique describes the technical details of a player's stroke, and stroke placement records the drop point of the stroke. Stroke technique values indicate whether the player is attacking or controlling at this stroke, and stroke placement values indicate where to attack or control.

A behavior of a player concerns strokes of two players. A table tennis match is an antagonistic process where two players exhibit highly interactive strokes. During this process, players need to not only consider how to give a good stroke, but also think about how to influence the opponent's strokes for enhanced chances of attack. The final scoring rates are the result of a game between two players. Therefore, the behaviors of players in a match are described by the strokes of both players.

A tactic consists of two phases. Experts in table tennis regard three consecutive strokes as a tactic of the player who gives the first in three strokes. Specifically, the first and third strokes in the tactic are given by one player (player A), and the second stroke is given by the other player (player B). The detailed process of a tactic is as follows. When player A gives the first stroke, he considers giving a stroke kind that will influence player B, thus forcing the latter player to respond with a certain stroke desired by the former player. The first phase concerning the first two strokes describes the proactive behavior of player A. After the opponent gives the expected stroke as the second stroke, player A rapidly gives the planned third stroke to score. The second phase concerning the last two strokes describes how the behavior of A in the prior phase exerts impacts.

The table tennis data was manually collected from the video. Each match was collected as a CSV file with hundreds of rows, and each row recorded different stroke attributes. Among the attributes, stroke technique and placement are the two most important attributes in describing the behaviors of players. We used 306 table tennis matches among 21 top players (10 males and 11 females) from 2005 to 2012.

3.2. Requirement Analysis

The experts we collaborated with were senior analysts who worked for the national table tennis team; they were professional in the analysis of techniques and tactics of players in table tennis. These experts approached us and introduced the Markov chain model, which builds a structure on a game and precisely simulates the interactive process between two players. The experts had used this model to estimate the total impacts of behavior adjustments on the scoring rates of players. However, they could not explore how the impacts accumulate over strokes and why they occur at each stroke because of the lack of a clear analysis pipeline and powerful analytical techniques. These technical details are significant for the in-depth analysis of the impacts.

We hence worked with the experts closely to propose a method of quantifying the impacts at each tactic and to acquire a visual analytics system and thus support flexible navigation and pattern unfolding of the generation process of impacts. During this process, we proposed typical two-stroke behaviors, quantified the impacts of adjustments at each tactic, and developed prototypes to obtain a proper visual design for a comprehensible multi-level presentation of the impacts. We condensed the requirements as follows.

R1 ***What are players' impact distributions over different beahavior adjustments?*** Players have individual styles of playing, and the impacts of behaviors differ by player. The experts hoped to browse impact distributions over behavior adjustments of various players.

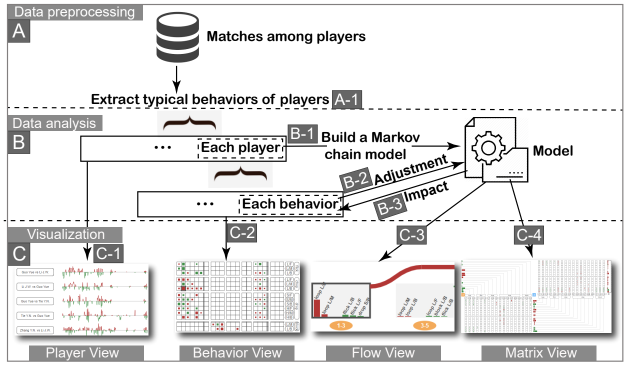

Figure 1. System overview. (A) Data processing component. (B) Data analysis componennt. For each player, this component builds a Markov chain model (B-1). The total impact of each behavior adjustment of the player (B-2) is estimated by the model (B-3). (C) Visualization component. This component visualizes the multilevel impacts received from the model.

R2 ***What are patterns on impacts of a player's different behavior adjustments?*** Adjustments to different behaviors exert varying impacts for a player. The experts needed to detect the patterns of different behavior adjustments and locate the interesting ones.

R3 ***How do impacts of a beahavior adjustment accumulate over strokes?*** A rally in table tennis consists of a sequence of strokes. Adjustments to the behavior of a player exert impacts on the scoring rate at certain strokes. Experts aimed to examine how the impacts accumulate throughout a rally.

R4 ***How do a behavior adjustment exerts impacts on the scoring rate at each tactic?*** Each score can be regarded as the result of a three-stroke tactic. The experts needed to examine the adjustments at the first two strokes of a tactic and identify changes to the probabilities of states at the last two strokes caused by the adjustments. Furthermore, they needed to identify the correlations between the adjustments and changes and how the changes in the second phase aggregate into variations in the scoring rate.

3.3. System Overview

The system is a web application with three parts, namely, the data preprocessing, data analysis, and visualization components. The data preprocessing component extracts the typical behaviors of players (Fig. 1A-1) from the CSV data tables that record matches among players. The data analysis component builds the Markov chain model (discussed in Section 4) for each player (Fig. 1B-1). For each behavior adjustment of the player (Fig. 1B-2), the model estimate the total impact on the scoring rate (Fig. 1B-3). The model also output the impact at each tactic (Fig. 1C-3) and the intermediate results that explain for the impacts (Fig. 1C-4). The visualization component uses vue.js to interactively visualize the multi-level impacts received from the model. This component consists of four views, namely, the player view, behavior view, flow view, and matrix view. The player view presents the impact distributions of different table tennis players (Fig. 1C-1). The behavior view presents the impacts of adjustments to various behaviors of a player (Fig. 1C-2). The flow view allows visual tracking of the accumulation of the impacts caused by the behavior adjustment over tactics (Fig. 1C-3). The matrix view, which comprises a pair of matrix views, presents the intermediate results during the generation process of impacts at each tactic (Fig. 1C-4). The four well-coordinated views provide a multi-level presentation of the impacts caused by adjustments to players' behaviors.

4. Model

For a table tennis match, building a structure to quantify the relations of observations on the game is significant [7]. Existing statistical methods for analyzing table tennis focus on certain statistics and cannot build an adequate structure on a game. Lames [6] hence proposed the Markov chain model to help structure a game and provide a uniform performance criteria. In the model, the analysis unit is the rally, which is regarded as a stochastic process over strokes of two players.

4.1. Lames' model

Previous works [16, 24] employing Lames' model used a certain attribute of a stroke, such as the technique of a stroke, as the transition state of the Markov chain. In this manner, the state varies over each stroke of the rally, namely, stroke 1, stroke 2, ..., stroke n. The state values are $value_{1,p1}$, $value_{1,p2}$, $\ldots$, $value_{n,p1}$, $value_{n,p2}$, $score_{p1}$, and $score_{p2}$. $value_{1,p1}, value_{1,p2}, \ldots, value_{n,p1}$, and $value_{n,p2}$ are all attribute values of the selected stroke attribute (e.g., stroke technique) of two players (P1 and P2). $score_{p1}$ and $score_{p2}$ represent the scoring state values of two players. These two values are the absorting state values; they will not transform into other values.

As shown in Fig. 2B, M is an *empirical transition matrix and V is the* empirical initial probability vector. The row headers and column headers of the matrix are the state values and each matrix entry presents the transition probability from the row header to the column header. Each element in vector V presents the initial probability of the corresponding state value. The transition matrix and initial probability vector are estimated from the analyzed matches. We calculated the ratio between times that state value A transforms into state value B and times that state value A transforms into all state values during collected matches. We used the ratio as the probability in the matrix entry with row header A and column header B. We calculated the ratio between

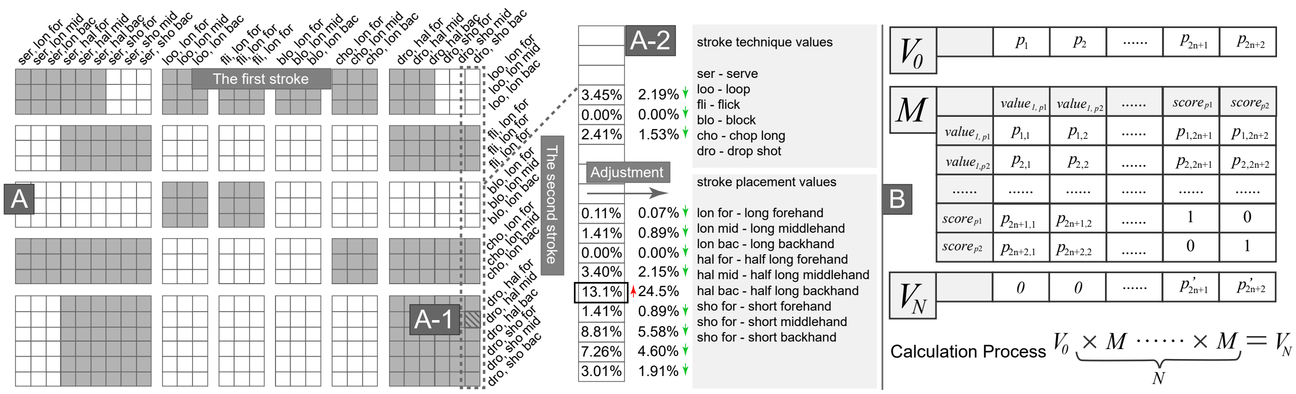

Figure 2. (A) We array 243 two-stroke behaviors as gray entries in a matrix to thoroughly understand their meaning. The column headers of the matrix are the combinations of stroke technique and placement values at the first stroke of the behavior, and row headers are those at the second stroke. (A-1) is a highlighted behavior. (A-2) gives an example of the adjustment to a behavior. (B) Illustration of the Markov chain model, where M is the transition matrix and V_0 and V_N are the state probability vectors.

times that state value A appears in the first stroke of a rally and times that all state values appear in the first stroke. We used the ratio as the probability in the initial probability vector V. During the computation, the initial probability vector is multiplied by the transition matrix repeatedly until the probabilities of all state values except absorbing state values approach zero (to five decimal points). The probabilities of $score_{p1}$ and $score_{p2}$ in the absorbing vector V_N are regarded as the final scoring rates of the two players. Making adjustments to behaviors of a player is achieved by tuning the empirical transition matrix M. Changes in final scoring rates quantify the impacts. The specifics of the adjustments are introduced in our modified model.

Existing works that applied the model to analyzing table tennis data had two limitations. First, these works only used an attribute of a stroke, such as technique of a stroke, to describe the state in the model. They ignored the joint use of stroke technique and placement and the tactical associations among strokes, thus characterizing players' behaviors inadequately. Second, existing works failed to explore when and how the impacts of adjustments occur in a rally.

4.2. Modified model

We held weekly discussions with the experts and decided to use the stroke technique and placement to describe the features of strokes and use two consecutive strokes to describe the behaviors of players. We decomposed a rally into a set of tactics, and divided each tactic into two phases to measure the stepwise impacts of behavior adjustments.

As discussed in the Section 3.1, stroke technique and placement are the two most representative features that describe the behaviors of players. Players' behaviors are described not only by their strokes, but also by the strokes next to their own strokes. We hence defined the behavior of a player as the stroke technique and placement of two consecutive strokes. Specially, the first of the two strokes is given by the player, and the second is given by the opponent. We used the behavior of the two players as the transition state of the Markov chain. In this manner, the state varies over two consecutive strokes of a rally, namely, stroke 1-2, stroke 2-3, ... stroke n-1-n. After iterative discussions with the experts and explorations on the collected data, we selected 243 typical two-stroke behaviors (Fig. 2A). In this case, the state values are $value_{1,p1}$, $value_{1,p2}$, $\ldots$, $value_{243,p1}$, $value_{243,p2}$, $score_{p1}$, and $score_{p2}$. The transition matrix M and state probability vector V are estimated in the manner similar to that of Lames' model.

Making adjustments. The behavior of two players is used as the transition state and is thus adjusted by tuning probabilities of the corresponding state values in the probability vectors at certain steps of the Markov process. Changes in scoring rates quantify the impacts. Assuming that users analyze player A in the rallies player A serves, we proposed to decompose a rally into a set of tactics of player A, namely, strokes 1-3, 3-5, ... The impacts of behavior adjustment were quantified at each tactic, which was further divided into two phases (as discussed in Section 3.1). Adjustments were made at the first phase of the tactics, such as at strokes 1-2 and 3-4, because this phase describes the behaviors of player A. The generation processes of impacts are examined at the second phase of the tactics, such as at strokes 2-3 and 4-5, because this phase describes the impacts of behavior adjustments of player A in the first phase.

As shown in Fig. 3B, adjustments ΔV_0 and ΔV_2 are applied to the probability vectors at strokes 1-2 and 3-4, respectively. The probability vectors at strokes 2-3 and 4-5 receive a boost of $\Delta V_0 M$ and $\Delta V_0 M^3 + \Delta V_2 M$ (Fig. 3A), respectively. The scoring rates after strokes 2-3 and 4-5 are subsequently influenced (Fig. 3C). Consequently, we can determine how the behavior adjustment of player A at the first phase of his tactics changes the probabilities of two-

		Calculation process	State probability vector	Adjustments	Scoring rate
Tactic 1-3	Stroke 1-2		V_0 A	$+\Delta V_0$ B	C
	Stroke 2-3	$(V_0+\Delta V_0)M =$	$V_1+\Delta V_0 M$	$\longrightarrow$	$0\ +\Delta S_{1\text{-}3}$
Tactic 3-5	Stroke 3-4	$(V_1+\Delta V_0 M)M =$	$V_2+\Delta V_0 M^2$	$+\Delta V_2$	
	Stroke 4-5	$(V_2+\Delta V_0 M^2+\Delta V_2)M =$	$V_3+\Delta V_0 M^3+\Delta V_2 M$	$\longrightarrow$	$\Delta S_{3\text{-}5}+\Delta S_{3\text{-}5}$
...		...	...		
		$V'_{N-1} M =$	V'_N		

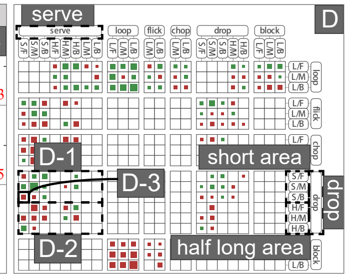

Figure 3. Illustration of the adjustments and impacts during computation process of the modified model. The items in red are caused by adjustments at the current tactic, and the items in blue are caused by adjustments at previous tactics. (A) presents the changed state probability vector at each step in the Markov process. (B) presents the adjustment made at each tactic. (C) presents the changed scoring rate at each step in the Markov process. (D) presents the behavior view of Ma Lin for the case study discussed in Section **??**.

stroke state values at the second phase, and finally changes his scoring rate after the tactics. Specifically, for the variations in probability vectors at stroke 4-5, $\Delta V_2 M$ is caused by the behavior adjustment at stroke 3-4, and $\Delta V_0 M^3$ is caused by adjustments at previous tactics. Our system distinguishes these two impacts. The specific values of the adjustments are calculated in the following methods.

Assuming that we need to adjust the state value drop to short backhand and drop to half middlehand (Fig. 2A-1). This adjustment represents that player A forces the opponent to use drop-half middlehand more to respond when he gives drop-short backhand. We firstly increase the probability of this state value and reduce probabilities of other state values in the same column in ratio to guarantee that the total probability of the column is constant (Fig. 2A-2). According to Pfeiffer et al. [16], the function for deflection is as follows (All probablities in the column are normalized through division by the total probability of the column.):

$$\delta P_x = C + B * 4 * P_x * (1 - P_x)$$

where P_x is the normalized probability of the state value; δP_x is the change of the probability; C is a constant that describes the deflection in the border probabilities; B is a constant that describes the maximum value of the relative magnitude of deflection; and 4 is a normalization factor that allows the constant B to be equal to the maximum value of deflection. In current work, the constant C = 0.05 and B = 0.25, which were determined on the basis of the previous work [16] and discussions with the experts. According to Pfeiffer et al. [16], the compensation function is

$$\delta P_{yi} = -(P_{yi}/(1 - P_x)) * \delta P_x$$

where P_{yi} is the normalized probability of other state values in the column and δP_{yi} is the change in the probability. Fig. 2A-2 shows an example for the adjustments.

5. Visual Design

This section discusses a set of design goals to follow, then describes the detailed designs of the system. **Blue** and **orange** stand for two players and **red** and **green** stand for the increase and decrease in the scoring rate and probability.

5.1. Design Goals

A visualization system is demanded to fulfill the requirements discussed in Section 3.2. However, existing visualization works on sports data cannot be directly applied to displaying the specific transfer process of impacts in two phases of each tactic. We summarized the design goals established with the experts as follows:

- G1 ***An overview of the impacts of adjustments to diverse behaviors for different players.*** The experts hoped to browse the impact distributions over different behavior adjustments of different players to identify features of each player and locate the interesting behavior adjustments (R1, R2). The system needed to provide an overview and flexible navigation of behaviors of different players.
- G2 ***Multi-level presentation of the impacts caused by adjustments to players' behaviors.*** The experts required a multilevel analysis of the impacts caused by behavior adjustments (R1, R2, R3, R4). In this manner, the experts can decompose the impacts layer by layer and obtain a better understanding. Therefore, the system should support the multi-level presentation of the impacts.
- G3 ***A timeline view that provides visual tracking of the accumulation of impacts.*** The experts aimed to track the accumulation of the impacts of a behavior adjustment over strokes (R3). A timeline view that supports visual tracking of the impacts was hence needed to be integrated into the system.
- G4 ***Pattern-unfolding and connection-preserving view of the generation process of the impact through two phases of a tactic.*** The experts hoped to examine the detailed process that the adjustment exerts the impact in a tactic (R4). Firstly, the system needed to enable pattern unfolding of the adjusted probabilities of

states in the first phase and the changed probabilities of states in the second phase. Furthermore, the system needed to reveal the connections between changes of two phases and how changes in the second phase aggregate into the variation of the scoring rate.

5.2. Player View

The player view (Fig.4A) presents players' impact distributions over different behavior adjustments (G1, G2).

This view contains a table that lists all top table tennis players. The line chart (Fig.4A-4) in each row in the table presents a player's impact distribution over different behavior adjustments during matches between the player and other players of a certain type. For instance, the line chart (Fig.4A-4) presents Ma Long's impact distribution over different behavior adjustments during matches between Ma and right shake-hand players (Fig.4A-3). The red parts of the line chart denote the behavior adjustments with positive impacts while the green parts denote the behavior adjustments with negative impacts. Two switch buttons (Figs.4A-1 and 4A-2) allow switching between serve rallies and receive rallies and between male players and female players.

5.3. Behavior view

After users select a player (e.g., player A) for analysis, the behavior view (Fig.4B) provides an overview of the impacts on the scoring rate by adjusting player A's different behaviors (G1, G2).

This view contains a matrix similar to that in Fig.2A. Each entry of the matrix presents a behavior of player A. The column headers denote the combinations of stroke technique and placement values at the first stroke (given by the opponents of player A) of the behavior and row headers denote those at the second stroke (given by the opponents of player A). The area of the rectangle in an entry encodes the impact of adjusting the corresponding behavior; adjusting a behavior is that letting the opponents of player A use the specific second stroke of the behavior more frequently after player A uses the specific first stroke. When users hover on the entry, the size of impact will be displayed (Fig.4B-3). This matrix hides the rows and columns with small values to lay stress on the key behavior adjustments.

5.4. Flow View

After users select the adjustment to a behavior, the flow view (Fig.4C) supports visual tracking of the accumulation of impacts on the scoring rate at each tactic (G2, G3).

This view contains a flow chart that represents the variation in the scoring rate of player A after adjusting a behavior at each of his tactics (Fig.3C). The flow chart is adopted because it intuitively illustrates the process that the impacts at all tactics accumulate gradually and finally aggregate into the total impact. The width of the main flow (Fig.4C-4) encodes the accumulated changed scoring rate and the width of branches (Fig.4C-1) encodes the variation at each tactic. When users hover on each branch, the original scoring rate (above) and changed scoring rate (below) will be displayed in a detail view (Fig.4C-3). The changed scoring rate is further divided into two parts. A pie chart is employed to present ratios of two parts. One part is caused by the adjustment at current tactic and the other is caused by adjustments at previous tactics.

The bar chart (Fig.4C-5) of each tactic shows six kinds of strokes with certain attributes values. The scoring probabilities of the first three of strokes are increased most at the last stroke of the tactic; the increments in them contribute most to positive impacts on the scoring rate at the tactic. The scoring probabilities of the final three of strokes are decreased most at the last stroke of the tactic; the decrease in them contribute most to negative impacts on the scoring rate at the tactic. The height of the red bar encodes the increase in scoring probability, while that of green bar encodes the decrease.

5.5. Matrix View

After users select a tactic, the matrix view (Fig.4D) supports the pattern unfolding of the impacts' generation process through the two phases of the selected tactic (G2, G4). *Justification:* We employ flow charts to present the detailed generation process within tactics for intuitiveness initially. However, the probabilities of many state values are influenced by the adjustments, and multiple impacts are needed to be displayed, which results in visual clutters appear in the flow charts. We hence used a clutter-free matrix view to help browse changed probabilities caused by the adjustments rapidly and clearly. We further added leader lines, separated columns of the matrices, and laid the plus signs between the columns to provide a comprehensible manner of presenting the transformation during the tactic. Probabilities of the first stroke are transformed into vectors of sub-probabilities, which are further aggregated into probabilities of the second stroke and then are transformed again. *Description:* The matrix view uses a pair of matrices to represent the changed probabilities of the state values at the two phases of a tactic, thus explaining the varying scores at the end of this tactic. The first matrix (Fig.4D-6) that concerns the first two strokes of the tactic displays how adjustments are made in the first phase. The second matrix (Fig.4D-2) that concerns the final two strokes displays how the adjustment changes the probabilities of the state values at the second phase and further changes the scoring rate. The three columns of *stroke bars* (Figs.4D-1, 4D-8, and 4D-3) from left to right represent the changed probabilities of employing various combinations of technique and placement values at the three strokes in a tactic. These probabilities are

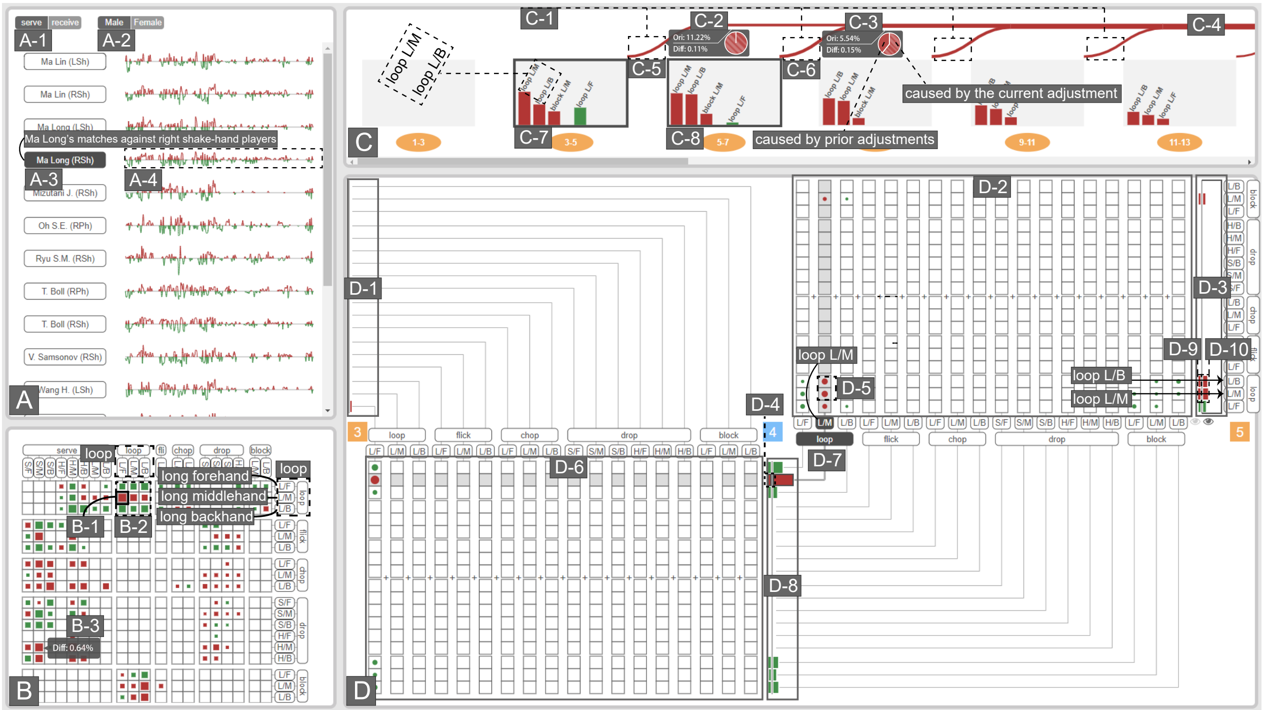

Figure 4. System Interface. The interface comprises four well-coordinated views: the player view (A), the behavior view (B), the flow view (C), and the matrix view (including a pair of coordinated matrices (D-6) and (D-2)).

derived by the addition of the probabilities of the columns in the matrices. The detailed encodings are as follows.

Bars for strokes. Three columns of bars (Figs.4D-1, 4D-8, and 4D-3) present the changed probabilities of employing various combinations of technique and placement values at the first, second, and third strokes. The height of each bar encodes the changed probability of each combination. If a probability increases or decreases due to the adjustments, then the height of a red or green bar is used to encode the enhanced or reduced probability. The probability of each combination is divided into two parts. The first part is changed scoring probability, i.e., the change in the probability that the player gives this stroke with the stroke attribute values and scores directly. This part is encoded by bars on the left (The left part of the first stroke is not shown because it is not considered in the current tactic). The second part is the change in probability that the player gives this stroke with the stroke attribute values and does not score. This part is encoded by bars on the right. When users hover on each part of a bar, the original and changed probability will be displayed in the detail view (Fig.6C). A pie chart presents two parts of the changed probability. One part is caused by the adjustment at current tactic and the other is caused by adjustments at previous tactics.

Matrices for behaviors. A pair of matrices presents the changed probabilities of all state values in the two phases of a tactic. The adjustments (Fig.3B) are presented by the first matrix (Fig.4D-6) concerning the first two strokes in the tactic. The impacts on the second phase of the tactic (Fig.3A) are presented by the second matrix (Fig.4D-2). The final impacts on the scoring rate (Fig.3C) are presented on the left parts of bars for the third stroke (Fig.4D-3). Similar to the state values in Fig.2A, the two-stroke state values are arranged as a matrix with column and row headers as stroke technique and placement at the prior and next strokes of the two strokes, respectively. The headers are linked with bars representing aggregate changed probabilities. We separate the columns to express that the entries in a column are the column header's child nodes that represent changed sub-probabilities. All columns are aggregated into the bars that represent changed probabilities of various attribute values at the next stroke. The changed probabilities branch into sub-probabilities or merge into aggregate probabilities in a manner similar to flows, thus enhancing the understanding of changed probabilities at each phase and connections between two phases in a tactic. When the serve tactic (i.e., the tactic consisting of strokes 1-3) is presented in the matrix view, the columns of the first matrix are altered because the first stroke in serve tactic must be serve stroke.

In each matrix entry, the area of the circle encodes the changed probability transforming from the column header to the row header. If the probability increases or decreases due to the adjustments, then the area of a red or a green circle is used to encode the enhanced or reduced probability, respectively. A switch button is placed at the bottom of the bars for the third stroke. The right part of the button (corre-

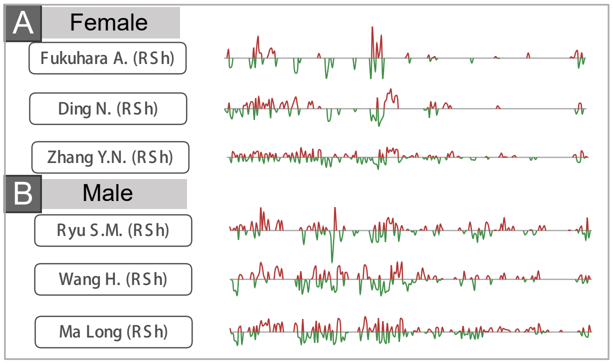

Figure 5. The impact distributions of six players (three females (A) and three males (B)) in the player view.

sponding to the right part of the bars) is enabled by default. Circles in entries of the second matrix represent the probabilities from the column headers to the row headers. When users switch to the left part (corresponding to the right part of the bars) of the button, the circles in the entries of the second matrix are transformed into rectangles whose area represent the scoring probabilities from the column headers to the row headers. When users hover on an entry, a detail view (Fig.6B) similar to that of the bar will be displayed.

6. Case Studies

We invited two experts we worked with through the design process (hereafter called A and B) to use the system. Expert A worked for the national table tennis team and was an authority on the analysis of techniques and tactics of players in table tennis. Expert B was a senior PhD student majoring in sports science. He was both a table tennis athlete and an experienced analyst on table tennis data. We gave a tutorial of how to explore the impacts on the system. The experts used the system for two weeks, during which we answered their questions at any time. Afterward, we worked with these experts to select two insightful patterns and summarized their exploration processes as two case studies. We presented parts of insights obtained by the experts and collected the feedback. The case studies were conducted on the Google Chrome on a PC (equipped with Intel Xeon E3, 32GB of memory, and a 1920*1080 display).

6.1. Impact distributions of different players

In this case study, the expert browsed the player view to examine distributions of impacts over behavior adjustments of different players. Specifically, the expert picked out line charts of three male players and three female players with distinct features. Fig.5A illustrates impact distribution over different behavior adjustments of three female players, namely, Fukuhara, Ding, and Zhang; the impacts were estimated by matches between three players and female right shake-hand players. The expert found that the variance of impacts caused by adjusting different behaviors of Zhang was smallest, while those of Ding and Fukuhara were larger. The expert commented that Zhang maintained a high performance because almost every behavior of her could not be dramatically improved. And the performances of Ding and Fukuhara were relatively low. Several behaviors of these two players could be further improved. Fig.5B illustrates impact distribution over different behavior adjustments of three male players, namely, Ryu, Wang, and Ma; the impacts were estimated by matches between three players and male right shake-hand players. The expert found that the variance of impacts caused by adjusting different behaviors of Ma and Wang were relatively small, while that of Ryu was large. The expert commented that Ma and Wang had a relatively high performance because only small improvements could be made on the behaviors of them. And the performance of Ryu was relatively low. A few behaviors of Ryu could be dramatically improved.

This case study demonstrated that the expert could browse the impact distributions over behaviors adjustments of different players and obtain the overall evaluation players' behaviors with the player view.

6.2. Good performance in receiving loop to middlehand

This case study were about matches between Ma Long and right shake-hand players. The expert found interesting behavior adjustments of Ma Long in these matches in the behavior view (Fig.4B); the behavior adjustments were that letting the opponents employ strokes with loop more frequently after Ma's strokes with loop.

The expert browsed the entries (Fig.4B-2) which represent impacts caused by the adjustments that increasing probabilities of the opponents' strokes with loop after Ma Long's strokes with loop. Then he found that the scoring rate was always increased due to the adjustments that increasing probabilities of the opponents' strokes with loop to long middlehand area (the colors of rectangles in most entries are red), while always decreased due to the adjustments that increasing probabilities of the opponents' strokes with loop to long backhand area and long forehand area (the colors of rectangles in most entries are green). The expert commented that when a player let the opponents use more strokes with loop instead of block, he would commonly decrease the probabilities that he employed strokes with loop at the following stroke and scored, thus reducing his scoring rate. It was hence against common knowledge that the scoring rate was increased due to the adjustments that increasing probabilities of the opponents' strokes with loop to long middle area after Ma Long's strokes with loop. The expert hoped to find out the reasons. He clicked on the entry with the biggest rectangle of the three (Fig.4B-1).

Then the accumulation process of this impact was displayed on the flow view (Fig.4C). The impacts accumulated

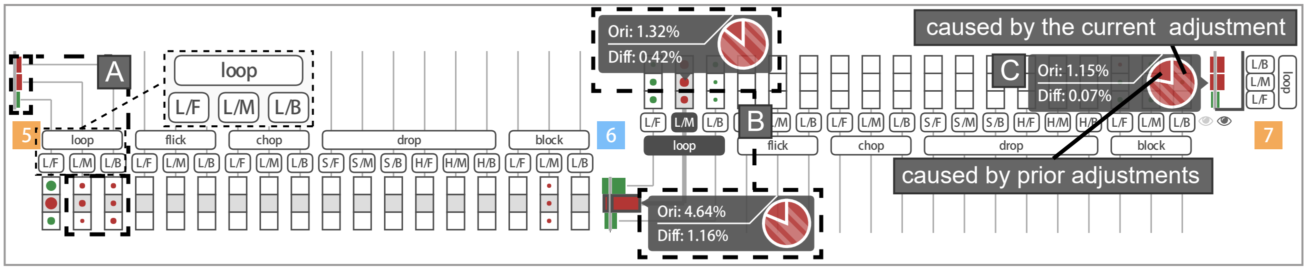

Figure 6. Illustration of the matrix view; the matrix view explains for that adjustments in previous tactics caused increments in the scoring rate at the tactic presented in Fig.4C-8.

at each tactic evenly (Fig.4C-1). The expert browsed the bar chart of the tactic consisting of stroke 3-5 (Fig.4C-5) because it was the first time the increments in the scoring rate took place. Scoring probabilities of strokes with loop to long backhand area (loop L/B) and loop to long middlehand area (loop L/M) were increased most at the last stroke the tactic; the increments in them contributed most to the total increments in the scoring rate at this tactic (Corresponding two bars are red and highest).

The expert clicked on this tactic (Fig.4C-5) to examine the generation of these increments in the matrix view (Fig.4D). Probabilities of strokes with loop L/M and loop L/B were increased, which was caused by the adjustments that increasing probabilities of the opponents' strokes with loop L/M (Fig.4D-7). The expert also found that the scoring probability at the second stroke was low (Fig.4D-4). This meant that it was difficult for the opponents to score when they employed strokes with loop L/M after Ma's strokes with loop, thus giving more chances for Ma Long to further used strokes with loop (Fig.4D-5) and scored (Fig.4D-9). The expert commented that Ma Long could switch between forehand and backhand easily. When the opponents gave the ball with loop to long middlehand, Ma Long could easily choose forehand or backhand to return, and thus increased the chances of scoring.

The expert browsed the flow view again and found that the increments in scoring rate at the tactic consisting of strokes 5-7 (Fig.4C-5) was larger than that of the tactic consisting of strokes 3-5 (Fig.4C-6, according to the width of the two branches). He further hovered on the branches of these two tactics and found in the detailed views that the increments at the later tactic consisted of two parts (Fig.4C-3). One part was caused by adjustments during the current tactic, while the other part was caused by adjustments in prior tactics. He further clicked on this tactic (Fig.4C-8) to examine how adjustments in prior tactics caused the increments in the scoring rate at the current tactic in the matrix view (Fig.6). He quickly found that the probabilities of strokes with loop L/M and loop L/B were increased at the first stroke of the tactic and probabilities of strokes with loop were increased at the second stroke (Fig.6A). The experts commented that the probabilities of these two kinds of strokes were increased at the last stroke of the prior tactic. A part of the increments joined in the scoring probabilities (Fig.4D-9) while another part of the increments were transferred into current tactic (Fig.4D-10). He hovered on several increased probabilities in this tactic and found in the detailed views that many of them were positively influenced by the adjustments at the prior tactic (Fig.6B). Finally, the increase in scoring probability at the last stroke of the tactic became larger due to the prior adjustments (Fig.6C).

This case study demonstrated that the assistance offered by detailed views in the flow view and matrix view in the multi-level analysis of impacts. The detailed views allowed the expert to distinguish between impacts caused by adjustments during the current tactic and prior tactics, thus helping the expert to track the chain effects of behavior adjustment.

6.3. Insights Obtained through the System

Experts obtained a set of insights from the table tennis data using our system. We selected several insights as follows. On the basis of these insights, not only existing hypotheses were tested for the first time, but also new hypotheses were formulated.

- ***Drop shot after serve.*** *The scoring rate is decreased due to the adjustments that letting the opponents employ strokes with drop shot to short area more frequently after serve strokes (the first kind of adjustments), while increased due to adjustments that letting the opponents employ strokes with drop shot to half long area more frequently after serve strokes (the second kind of adjustments).*
  - For most players, the main reason why the first kind of adjustments decrease the scoring rate is that the adjustments decrease the chances that the player attacks with loop and scores at the following stroke. (*verify an existing hypothesis*)
- ***Flick after serve.*** *The scoring rate is sometimes increased and sometimes decreased due to the adjustments that letting the opponents employ more strokes with flick after serve strokes*

- For most players, the reason why the adjustments sometimes increase the scoring rate is that the adjustments increase the probabilities that the player attacks with loop and scores in the following stroke. (*verify an existing hypothesis*)
- For a part of players, such as Guo Yue, the adjustments can always increase the scoring rate (during the matches between Guo and right shakehand players). The reason is that the adjustments can greatly increase that scoring probabilities of strokes with loop by Guo. (*propose an new hypothesis*)

• ***Loop after loop.*** *The scoring rate is decreased due to the adjustments that letting the opponents employ more strokes with loop after strokes with loop.*

- For most players, the reason why the adjustments decrease the scoring rate is that the adjustments decrease the chances that the player attacks with loop and scores at the following stroke. (*verify an existing hypothesis*)
- For a part of players, such as Ma Long, it will increase the scoring rate to let the opponents employ more loop to long middlehand after strokes with loop. It is because that Ma Long can switch between forehand and backhand easily and is hence good at receiving loop to long middlehand (Case study 3). (*propose an existing hypothesis*)

6.4. Domain Expert Feedback

We collected the feedback from the two experts. Overall, they thought the system was concise and practical for performance analysis of table tennis player. Two characteristics of the system are thought highly by the experts. First, the system, based on the analysis pipeline, revealed the detailed generation process of impacts of behavior adjustment. The experts were impressed by the multi-level presentation provided by the system as this was first time they had been able to examine the detailed process and reasons for the impacts. The expert A commented: "Firstly, this system takes more key features of the behaviors into consideration when calculating the impacts of the behavior adjustment; the key features are combinational use of stroke technique and stroke placement and the three-stroke tactical associations between strokes. Second, the system supports identification of the impacts on the scoring rate and their reasons at each tactic for the first time. I think it is very advanced and practical." Second, the tailored matrix view illustrated the generation process of the impacts comprehensibly. The experts appreciated the matrix view which enhanced the understanding of the transfer process of the changed probabilities. The expert B commented: "The matrix view comprehensibly presents how the changed probabilities branch, merge, and finally lead to the changed scoring rate. I think it is comprehensible and effective."

7. Discussion and Conclusion

Lessons Learned. We learned the lesson from this work that visualization practitioners should possess certain domain knowledge. Although visual applications are derived with the help of domain experts, visualization researchers should also learn domain knowledge more thoroughly. Designing visualizations for the data is not the same as automatic data processing. According to our discussion about the most insightful aspect of the data and the best way to visually present it, the practical significance of the data must be appreciated. Before this work, we had been working closely with the experts for one and a half years. During this time, we gradually determined the practical significance of most stroke attribute values and understood the tactical behaviors and technical details of table tennis players more thoroughly. On the basis of this prior knowledge, we were able to think from the perspective of experts when deriving the visual designs.

Limitations. Three limitations are listed as follows. First, the adjustments made in the system can be increased to be more flexible. In the system, when we increase the probabilities of a behavior, other relevant behaviors are evenly reduced to maintain the constant total probability. However, as several behaviors are more relevant while others are not, it will be better to reduce th behaviors individually. Second, the system cannot support explorations on adjustments to multiple behaviors. Considerable efforts can be further exerted to identify patterns of impacts caused by adjusting combinations of multiple behaviors. Third, the system can be further enhanced to distinguish the impacts of behavior adjustment under different conditions, such as different scores.

General Applicability. Game sports, such as squash, tennis, badminton, volleyball, and baseball, are most appropriately regarded as dynamical interaction processes between two parts [7]. Therefore, the Markov chain model can also simulate structures of these sports given properly identified state values. In this case, our system can be extended to these sports.

Conclusion. This work investigated the problem of visual analytics of impacts of behavior adjustments on scoring rates in table tennis. We worked with experts closely to define the typical two-stroke behaviors and proposed the method of decomposing a rally into tactics of a player to measure the impacts of adjustments at each stroke. We further developed a system that concentrates on the multi-level explorations on impacts of behavior adjustments. We conducted two case studies to demonstrate the effectiveness and usability of our system. The main implications of this work are as follows. First, this study investigated the problems of visually tracking and explaining the impacts of adjusting behaviors over strokes in table tennis and proposes a method of measuring the step-wise impacts of adjustments. Second, our system enables the experts to gain insights into the timings and reasons for impacts of behavior adjustments. A set of insights were detected from matches of top table tennis players.

References

[1] C. Dietrich, D. Koop, H. T. Vo, and C. T. Silva. Baseball4D: A tool for baseball game reconstruction &visualization. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, pages 23–32, 2014. 2

[2] A. Franks, A. Miller, L. Bornn, and K. Goldsberry. Counterpoints: Advanced defensive metrics for NBA basketball. In *Proceedings of MIT Sloan Sports Analytics Conference*, 2015. 2

[3] H. Janetzko, D. Sacha, M. Stein, T. Schreck, D. A. Keim, and O. Deussen. Feature-driven visual analytics of soccer data. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, pages 13–22, 2014. 2

[4] L. Jin and D. C. Banks. Visualizing a tennis match. In *Proceedings of IEEE Symposium on Information Visualization*, pages 108–114, 1996. 2

[5] M. Lage, J. P. Ono, D. Cervone, J. Chiang, C. Dietrich, and C. T. Silva. StatCast Dashboard: Exploration of spatiotemporal baseball data. *IEEE Computer Graphics and Applications*, 36(5):28–37, 2016. 2

[6] M. Lames. *Leistungsdiagnostik durch Computersimulation: ein Beitrag zur Theorie der Sportspiele am Beispiel Tennis*. PhD thesis, University of Mainz, Germany, 1991. 1, 2, 4

[7] M. Lames and T. McGarry. On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport*, 7(1):62–79, 2007. 1, 2, 4, 11

[8] P. A. Legg, D. H. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths, and M. Chen. MatchPad: Interactive glyph-based visualization for real-time sports performance analysis. *Computer Graphics Forum*, 31(3pt4):1255–1264, 2012. 3

[9] R. Leser and A. Baca. 30 practice oriented match analyses in table tennis as a coaching aid. *Science and Racket Sports IV*, page 214, 2008. 1, 2

[10] T. C. Loh and O. Krasilshchikov. Competition performance variables differences in elite and u-21 international men singles table tennis players. *Journal of Physical Education and Sport*, 15(4):829, 2015. 1, 2

[11] A. G. Losada, R. Therón, and A. Benito. BKViz: A basketball visual analysis tool. *IEEE Computer Graphics and Applications*, 36(6):58–68, 2016. 2

[12] I. Malagoli Lanzoni, R. Di Michele, and F. Merni. A notational analysis of shot characteristics in top-level table tennis players. *European journal of sport science*, 14(4):309–317, 2014. 1, 2

[13] M. L. Parry, P. A. Legg, D. H. S. Chung, I. W. Griffiths, and M. Chen. Hierarchical event selection for video storyboards with a case study on snooker video visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1747–1756, 2011. 3

[14] C. Perin, R. Vuillemot, and J. D. Fekete. SoccerStories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013. 2

[15] C. Perin, R. Vuillemot, and J.-D. Fekete. A Table!: Improving temporal navigation in soccer ranking tables. In *Proceedings of ACM Conference on Human Factors in Computing Systems*, pages 887–896, 2014. 2

[16] M. Pfeiffer, H. Zhang, and A. Hohmann. A markov chain model of elite table tennis competition. *International Journal of Sports Science & Coaching*, 5(2):205–222, 2010. 1, 2, 4, 6

[17] H. Pileggi, C. D. Stolper, J. M. Boyle, and J. T. Stasko. SnapShot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2819–2828, 2012. 2

[18] T. Polk, J. Yang, Y. Hu, and Y. Zhao. TenniVis: Visualization for tennis match analysis. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2339–2348, 2014. 2, 3

[19] D. Sacha, F. Al-amoody, M. Stein, T. Schreck, D. A. Keim, G. L. Andrienko, and H. Janetzko. Dynamic visual abstraction of soccer movement. *Computer Graphics Forum*, 36(3):305–315, 2017. 2

[20] M. Stein, H. Janetzko, T. Breitkreutz, D. Seebacher, T. Schreck, M. Grossniklaus, I. D. Couzin, and D. A. Keim. Director's cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, 2016. 3

[21] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim. Bring it to the Pitch: Combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):13–22, 2018. 3

[22] D. S. Tan, G. Smith, B. Lee, and G. G. Robertson. AdaptiviTree: Adaptive tree visualization for tournament-style brackets. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1113–1120, 2007. 2

[23] J. Wang, K. Zhao, D. Deng, A. Cao, X. Xie, Z. Zhou, H. Zhang, and Y. Wu. Tac-simur: Tactic-based simulative visual analytics of table tennis. *IEEE Trans. Vis. Comput. Graph.*, 26(1):407–417, 2020. 2

[24] S. Wenninger and M. Lames. Performance analysis in table tennis-stochastic simulation by numerical derivation. *International Journal of Computer Science in Sport*, 15(1):22–36, 2016. 1, 2, 4

[25] Y. Wu, J. Lan, X. Shu, C. Ji, K. Zhao, J. Wang, and H. Zhang. iTTVis: Interactive visualization of table tennis data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):709–718, 2018. 2, 3

[26] H. Zhang, W. Liu, J.-j. Hu, and R.-z. Liu. Evaluation of elite table tennis players' technique effectiveness. *Journal of sports sciences*, 31(14):1526–1534, 2013. 1, 2