



Characterizing Serves in Table Tennis

Aymeric Erades, Thomas Papon, Romain Vuillemot

► To cite this version:

Aymeric Erades, Thomas Papon, Romain Vuillemot. Characterizing Serves in Table Tennis. Machine Learning and Data Mining for Sports Analytics, Sep 2024, Vilnius (Lituanie), Lithuania. pp.11. <hal-04691690>

HAL Id: hal-04691690

<https://hal.science/hal-04691690v1>

Submitted on 27 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Characterizing Serves in Table Tennis

Aymeric Eradès^{1,2}, Thomas Papon^{1,2}, and Romain Vuillemot^{1,2}

¹ École Centrale de Lyon, France

² LIRIS CNRS UMR 5205

Abstract. In table tennis, serves play a crucial role as they are the first and only shot during which players have full control of the game. In this paper, we explore serve techniques and tactics using a dataset of 9 games and a total of 510 serves collected semi-automatically for 5 players. We first provide a descriptive analysis and group the data into clusters to identify what we refer to as a serve repertoire, which are serve categories specific to players. We then identify which serve is used during the game and the tactic players employ during competitions. In particular, we provide a better understanding of the notion of variation, a concept often used in table tennis, and we show the differences in serves that are used in key moments of a game (e. g., during score domination, decisive points).

Keywords: Sports Analytics · Table Tennis · Serves

1 Introduction

Serves are the very first action in racket sports, in particular in table tennis. According to Larry Hodges [7] they are *the most strategic and tactical part of your game*. In table tennis, serves consist of a ball bouncing on the two sides of the table, following a set of specific rules. There are several types of serves, each with unique techniques and tactical purposes. Examples of techniques are numerous. The *topspin serve*, which generates forward spin to make the ball bounce higher and faster. The *backspin serve*, which produces reverse spin, causing the ball to slow down and stay low. The *side-spin serve*, which gives a horizontal spin to make the ball curve sideways. And the *no-spin serve* which aims to surprise the opponent by having minimal to no spin. Serves can also be grouped by ball placements with *short* and *long serves*, and *pivot serves* between the forehand and backhand area. Further nuances can be added by adjusting the speed, angle, ball toss of the serve, offering players a wide palette of tactical options to initiate play effectively. Some serves also have colorful names such as the Ghost Serve (heavy backspin), Tomahawk (the racket is swung like a tomahawk), Pendulum (the server swings the paddle in a pendulum-like motion generating a lot of side-spin), or Lollipop (slow shot with little spin). In general, there is no official or formal classification in particular related to data that can be collected for this sport. Regarding the tactical aspect of serves, the goal is to find efficient serves according to the opponent and the score of the game, with variations to make the opponent's prediction and anticipation difficult.

In this paper, we aim to characterize serves in a deeper way. Indeed, while both technique and ball placement can characterize them, they are limited for advanced tactical analysis. The first limitation is that the current classification of techniques and ball placement does not capture the intent of the player to anticipate the next shot. The second is that this classification is too coarse, so similar serves may be grouped in different categories, and different serves can be grouped in the same category. Besides making it impossible to accurately classify serves, the key tactical notion of variations cannot be fully explored. If accurate categorization is achieved, then many longstanding open questions in table tennis can be answered: What is the total number of serves?³ What are the signature serves and the ones used for critical points of the game? What are the main profiles of servers?

To achieve this, we first collected a detailed dataset of serves that includes both the players' positions and techniques, ball placement in an accurate way, and the opponent's returning stroke. This provided a way to group serves by second ball bounce placements to reveal clusters on the opponent's side. To refine such clusters, we split them by the first ball bounce and the technique of the server as a style prior [15]. This led us to a specific number of clusters we referred to as the *serve repertoire* from which the server can pick. We then conducted an analysis of the use of this repertoire to identify when they are used first, their variations, and their use in key moments. We released our code and data as an open-source project to foster more research in this area.

2 Background and Related Work

Serves follow specific rules beyond the two bounces. Each player serves two consecutive points, alternating throughout the game, with adjustments during the final game at deuce. The server must place the ball on their open hand, throw it vertically at least 16 cm, and strike it so it bounces on their side before crossing the net to the opponent's side. Players have to hit the ball behind the edge of their side of the table. The serve must be visible to the opponent, and the free hand must not obstruct the view. Let serves, which touch the net but land correctly, are replayed, while fault serves result in a point for the opponent.

In table tennis, most analyses are based on *rallies*, which can be seen as multivariate sequences and are the focus of most analytical methods. In *TacticFlow* [16], a mining method is proposed to mine frequent patterns and detect how these patterns change over time using MDL (Minimum Description Length). In [5], sequence mining is achieved using the SPADE algorithm to determine the most frequent tactics and the WRAcc measure to keep the most relevant ones. In [17], they reveal the hit correlations and hit sequences. In [14], a Markov chain model is used to characterize and simulate the sequence process in table tennis. In these works, the serve is only considered as a particular technique attribute.

³ Some players claim they have around 100 serve variations <https://www.canalplus.com/sport/extrait-interieur-sport-la-nouvelle-dynastie-les-freres-lebrun/h/24369553.50001> (in French)

They eventually serve as a descriptor of the type of hit in [8], with explicit use of the pendulum, reverse, and tomahawk serves. However, as far as we know, there does not exist a detailed analysis of serves, particularly due to the categorization of the ball placement on the 3×3 grid [5] [2] [?] or in a $3 \times 3 \times$ grid in 3D [4], which does not capture placement subtleties. Additionally, none of them focus on the exact ball position and first rebound analysis.

Other similar racket sports (e. g., tennis, badminton) share similarities as two players alternate hitting the ball, with one player serving and one player winning the point. Still there is little work focusing on serves analysis. [19] introduces a serve prediction method in table tennis. They extensively review the current performance of prediction models. However, they only rely on shot-by-shot data and do not leverage the potential of tracking data [11], particularly what Hawk-Eye-like systems could provide [10]. [4] analyze rallies in badminton solely based on trajectories reconstructed from the shuttle’s kinematic features. Both works are difficult to translate to table tennis, where there is a first bounce and heavy ball spin that is difficult to detect.

3 Data Collection and Exploratory Data Analysis

Our first step was to collect a detailed and representative dataset of serves. We used TV broadcast videos available on the ITTF Channel on YouTube to collect data from professional games. We employed a combination of automated (e. g., OpenPose [3]) and semi-automated computer vision techniques to extract players’ positions, ball hit locations, and event characterizations. We picked this approach as there is currently no fully automated and accurate table tennis game reconstruction despite recent advances in this area [13]. We collected full rallies, but in this paper, we focus only on the analysis of the first stroke (serve) and the return technique used (not its placement). We also collected context data such as the outcome of the point (winning/losing) and contextual elements such as scores and players’ names. We collected a total of 9 games and up to 510 serves for 5 players. The dataset has been released publicly and is available online⁴. We removed data from left-handed players as they provided spatial inconsistencies, so we only investigated games against right-handed players. We augmented the collected dataset using several post-processing steps to calculate the score and various metrics [2] for the section related to tactical analysis. We also reconstructed 3D ball trajectories [1] to gather details on the server’s technique. We stored both the collected data and the augmented data in a database that can be quickly queried and joined with other metadata.

We then conducted an exploratory data analysis [12] of serves to grasp the distribution of serves and identify the attributes and parameters that could separate different types of serves. This step also helped to diagnose any data inaccuracies and inconsistencies. Fig. 1 shows that standard ways to explore ball placement (e. g., by position on the table) are not sufficient to distinguish

⁴ <https://github.com/centraleyon/table-tennis-services/>

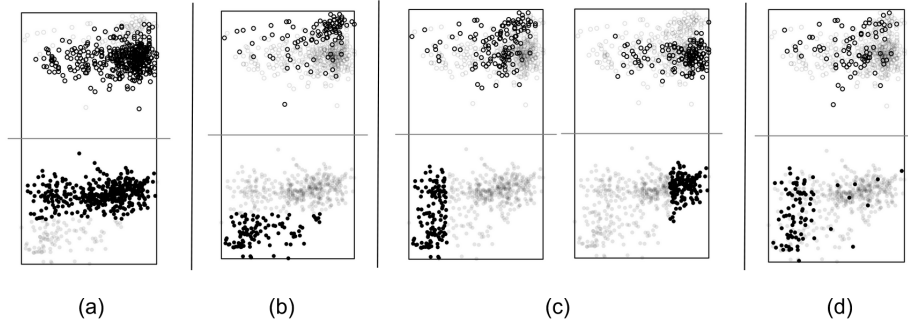


Fig. 1: Spatio-temporal representation of serves we collected, normalized, and filtered. The server is at the top; the dots on the upper part are the first bounce, the ones below are the second bounce, and the receiver is at the bottom. Gray dots represent all the serves, and the black ones: (a) short serves, (b) long serves, (c) side serves, and (d) pivot serves.

between types of serves. However, it revealed some interesting patterns, such as regions that are physically impossible to reach due to ball rebound physics (e. g., close to both sides of the net). Also, many serves are not attempted because their trajectory would lead to a fault if both bounces occur on the same side. Some possible bouncing regions are intentionally avoided for tactical purposes, such as those close to the receiver’s forehand, which is the most efficient stroke. Another tactical pattern that was revealed was outliers, which are uncommon and may surprise the opponent. These outliers can be easily observed as their ball placement and trajectories are distinctive.

4 Serves Categorization

We detail in this section our method to group serves in a way to reveal players repertoire by grouping similar serves as a combination of second-bounce ball placement and technique.

Our goal is to group balls that land on the other side of the table into meaningful clusters. Similarly to [15] relied upon K-means clustering to automatically find a way to partition 2D space based as the clusters we are looking for are roughly circular, centered around an aiming placement. Fig. 2 (b) shows a clear separation with centroid representing the center of the target point and the spread around it (likely due to inaccuracies). Other methods could be used such as HDBSCAN [9] used in Badminton [18] or in Tennis [19]. But those methods are trajectory based and we consider that the current method we used was not sufficiently accurate to reconstruct them. We used K-means combined with the *elbow method* (an heuristic aiming at maximizing similarity within clusters while minimizing similarity between clusters) to determine the number of clusters. Despite the simplicity of this method, we visually assessed the resulting clusters and found them relevant in both the number, shape and centroid of the groups,

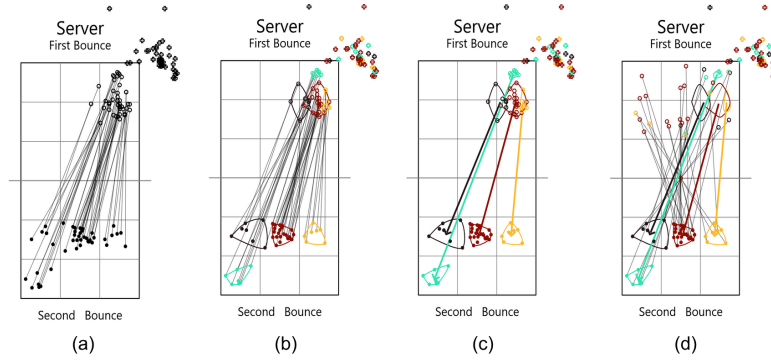


Fig. 2: Steps to create serves clusters (a) a single game is picked for a player, (b) K-means clusters are created based on the second-bounce, (c) the clusters are connected to their corresponding first-bounces, (d) we show the corresponding returns by the other players with lines.

to globally separate serves. We also found those clusters relevant to group serves from different games with the same players.

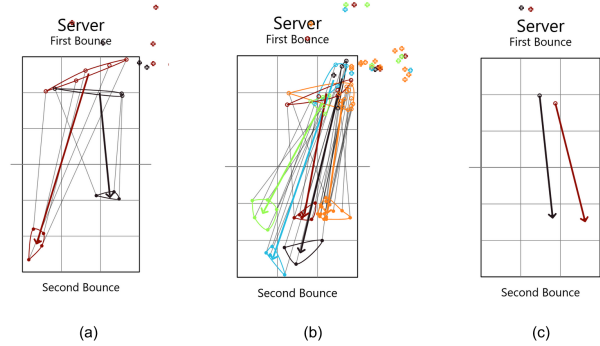


Fig. 3: Each clusters were then separated using serving technique (a) forehand serve with right side-spin, (b) forehand serve with left side-spin (c) backhand serve with left side-spin.

We refined the grouping based on the technique used when serves as two different techniques could lead to the same second-bounce cluster but with a different incoming trajectory (Fig. 3). Also, as the clusters were sometimes too large, and differences were caused by the effect in the serve. Fig. 2 (c) shows such deeper separation by including the serve technique. Finally we found a total of N clusters which will be identified later as $C_1, C_2, \dots, C_N$, ranked by number of serves in each. Each cluster C_i is represented by an area, a density and a distance to the first-bounce cluster (to indicate if this is a short or a long serves).

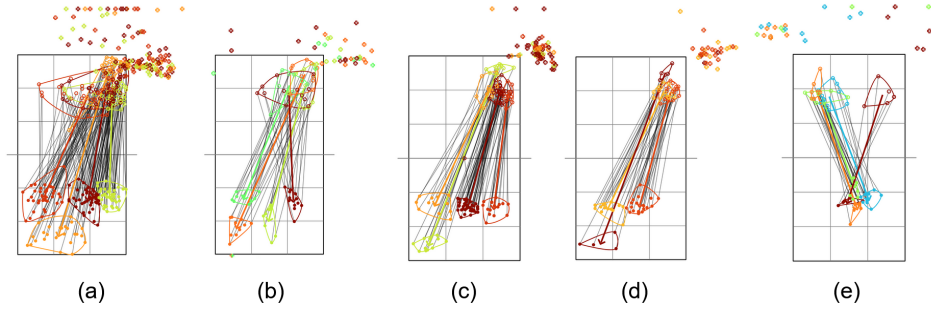


Fig. 4: Server’s repertoire for top players: a small first bouncing area ((d) **Player4**, (c) **Player3**), which is explained by a static position; spreading the first bounce across the width ((e) **Player5**); the same serve but from different initial positions ((a) **Player2**, (b) **Player1**).

We displayed on Fig. 4 the results from the application of the methodology we previously used by picking a single game for our collection of players.

5 Servers Tactics

In this section we define as tactics the use of efficient serves, and our goal is to reveal which serves are concerned and when they are used.

5.1 Serves Repertoire

Table 1 provides a shorter, more descriptive summary of all the serves clusters we identified for top ranked players. Those are only from male, right-handed players to obtain some comparable results. This enabled to characterize the general signature of players showing either that they limit to same serves in general or have a high level of diversity.

Player	$C_1(area, density, dist)$	C_2	C_3	C_4
Player1	(13%, 0.0196, 154)	(8%, 0.0139, 229)	(6%, 0.0160, 174)	(4%, 0.0192, 195)
Player2	(15%, 0.0127, 144)	(8%, 0.0085, 237)	(7%, 0.0079, 156)	(1%, 0.0437, 146)
Player3	(18%, 0.0056, 259)	(18%, 0.0017, 160)	(9%, 0.0034, 166)	(14%, 0.0018, 155)
Player4	(36%, 0.0023, 175)	(34%, 0.0036, 159)	(36%, 0.0007, 246)	
Player5	(3%, 0.0289, 145)	(2%, 0.0383, 160)	(1%, 0.0319, 131)	(1%, 0.0241, 218)

Table 1: Details of the 4 most frequent serves from each of the top-player repertoire we picked. The area % is calculated compared to one half of the table. The distance (*dist*) is in cm (length of a Table is 274cm and diagonal 313cm).

5.2 Serves Similarity

We define a way to group serves based on their similarities. An example of low similarity would be **Player3** who does a long serve C_1 and then a short serve C_3 . This aims to capture the level of difference between clusters (we assume inter-cluster serve variations are due to how serves were executed, with a certain level of imperfection). To formalize this intuition, we extend the *shot diversity* metric introduced in [2]. The metric used below follows the following order of priority: serves with different lateralities will be more dissimilar than serves with different side-spins, which in turn will be more dissimilar than serves that bounce a second time in different clusters. We will also consider the distance between different clusters, starting from the centroids of these clusters. Finally, to highlight the variations induced by long serves, we will introduce a higher coefficient for distances along the length of the table than for the width. Here is the exact definition of the distance between two serves used:

$$D(S_1, S_2) = d_{lat}(S_1, S_2) + d_{sidespin}(S_1, S_2) + d_{cluster}(S_1, S_2) \quad (1)$$

with

- $d_{lat}(S_1, S_2) = 0.3$ if both serves have different lateralities (say, one is a forehand while the other one is a backhand), and $d_{lat}(S_1, S_2) = 0$ otherwise.
- $d_{sidespin}(S_1, S_2) = 0.22$ if both serves have different sidespins (say, one is a Left side serve while the other one is a Right side serve), and $d_{sidespin}(S_1, S_2) = 0$ otherwise.
- $d_{cluster}(S_1, S_2) = \|C_2 - C_1\|_1$ if both serves have their second bounce inside a different cluster, and $d_{cluster}(S_1, S_2) = 0$ otherwise. We define C_1 and C_2 the centroids of the first serve's cluster and the second serve's cluster respectively.

Then again, we must define the L1-norm used here: let (x_1, y_1) and (x_2, y_2) be the centroids of two different clusters, C_1 and C_2 . Therefore,

$$\|C_2 - C_1\|_1 = \frac{|x_2 - x_1|}{76} \times 0.1 + \frac{|y_2 - y_1|}{137} \times 0.38$$

If 76 and 137 are the dimensions of half a table tennis table in centimeters, meant to homogenize the x and y distances, the coefficients 0.1 and 0.38 are chosen arbitrarily to allow the distance D between two serves to fall within $[0,1]$, and to highlight the variation introduced by a long serve within a series of serves. Such a definition then results in typical distances of 0.12 between a cluster of serves on the opponent's forehand side and another on the backhand side, and 0.16 between a short serve and a long serve, both in the middle of the table.

Finally, the term *diversity* is introduced to help quantify the diversity of serves used by a player over the course of a set. It is defined as:

$$diversity = \sum_{i=1}^{n-1} \left(\frac{D(S_i, S_{i+1})^2}{n-1} \right) \times 100$$

with S_i being the i^{th} serve of a predetermined player in a set, and n being the number of times he served this set. Below is the *diversity* results found for **Player2** and **Player3**, depending on the set :

Table 2: Diversity across sets for each player

Player / Set	Set 1	Set 2	Set 3	Set 4	Set 5
Player 3	0.77	1.09	2.2	0.95	0.45
Player 2	10.20	14.47	3.03	4.96	4.39

5.3 Serving Tactics

We are now interested in understanding the tactics that motivate the change of serves. We will explore such tactics based on diversity and variations of serves, e. g., to understand which players prefer to create surprise or the ones that stick to efficient serves.

Fig. 5 illustrates the distances and variation metrics of serves during a single game between **Player2** and **Player3**. It clearly shows that **Player2** tends to use a more diverse range of serves compared to **Player3**. Due to his extensive repertoire, **Player2** explores various serves while occasionally sticking to effective ones before introducing new tactics. Table 2 supports this observation, showing that **Player2** won the second set by employing a wide variety of serves, surprising his opponent with new tactics rather than relying on a single effective serve. In terms of the score’s influence, a notable trend is that dominating players often adopt a conservative approach, preferring to stick with reliable serves, which aligns with the tactic outlined by [7]: “*Keep using what works*”. On the other hand, **Player3** utilizes fewer serves, resulting in less variation overall, particularly in crucial points. Further investigation is necessary to refine the identified clusters and potentially uncover subtle variations.

6 Discussion and Perspectives

This paper is a preliminary attempt to characterize table tennis serves using detailed data on bouncing positions from both sides of the table. However, it has several limitations primarily related to the level of details and volume of the data used. Regarding the volume, we aim to collect more serves to build a representative repertoire, as the games we analyzed likely constitute only a subset of all serves used, particularly the less frequently observed ones. The reconstruction of detailed ball trajectories remains an ongoing challenge in computer vision [13]. To address this, we plan to advance research in serve detection by contributing our annotated dataset to computer vision benchmarks [6]. Nonetheless, numerous influential factors, such as players’ racket surfaces, deceptive body movements, and external conditions, are challenges to accurately predict ball trajectories.

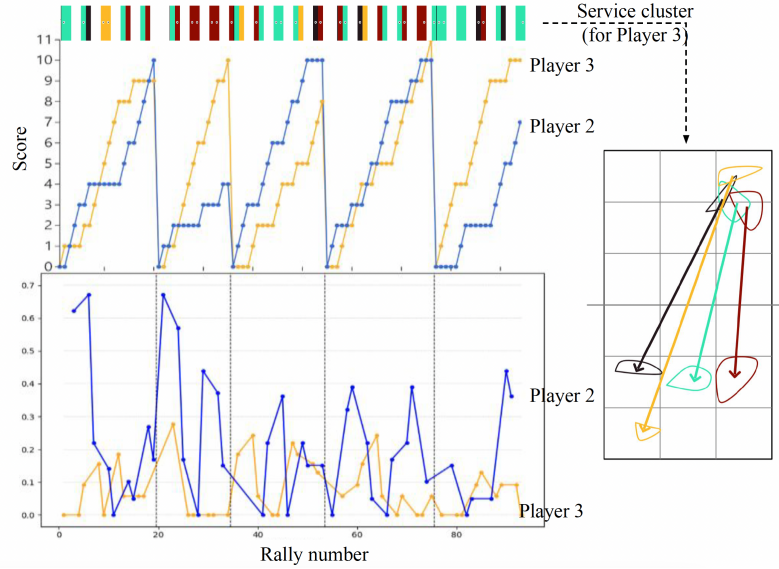


Fig. 5: From top to bottom: serves used by **Player3** against **Player2** (colors represent the point’s winner); score evolution; and distances between 2 consecutive serves (colors represent the server).

We also conducted quantitative evaluations of both the resulting clusters and the tactics employed during games. Our future plans include interviewing players from our dataset to gather feedback on the clusters and influential tactical factors (e. g., preparation, fatigue, risk-taking). Looking ahead, we intend to expand our dataset to include doubles matches, which offer more diverse combinations of servers and returners. Additionally, we aim to explore serves used by left-handed players, as well as how players adapt based on their handedness (left or right) or grip style (shake-hand or pen-hold). Further investigation into larger datasets could address questions such as: Do players have invariant routines or recurring serve sequences based on game context? Are there differences in serve tactics between top-ranked players and lower-ranked players? Lastly, do players possess secret serves reserved for critical moments in a match?

References

1. Calandre, J., Peteri, R., Mascarilla, L., Tremblais, B.: Extraction and analysis of 3D kinematic parameters of Table Tennis ball from a single camera. In: 2020 25th International Conference on Pattern Recognition (ICPR). pp. 9468–9475. IEEE, Milan, Italy (Jan 2021). <https://doi.org/10.1109/ICPR48806.2021.9412391>, <https://ieeexplore.ieee.org/document/9412391/>
2. Calmet, G., Erades, A., Vuillemot, R.: Exploring Table Tennis Analytics: Domination, Expected Score and Shot Diversity. In: Link, S. (ed.) Machine Learning and Data Mining for Sports Analytics. Communications in Computer and Information Science, Turin, Italy (Sep 2023), <https://hal.science/hal-04240982>

3. Cao, Z., Hidalgo, G., Simon, T., Wei, S.E., Sheikh, Y.: OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **43**(1), 172–186 (Jan 2021). <https://doi.org/10.1109/TPAMI.2019.2929257>, <https://doi.org/10.1109/TPAMI.2019.2929257>
4. Chu, X., Xie, X., Ye, S., Lu, H., Xiao, H., Yuan, Z., Chen, Z., Zhang, H., Wu, Y.: TIVEE: Visual exploration and explanation of badminton tactics in immersive visualizations. *IEEE Transactions on Visualization and Computer Graphics* **28**(1), 118–128 (2021), publisher: IEEE
5. Duluard, P., Li, X., Plantevit, M., Robardet, C., Vuillemot, R.: Discovering and Visualizing Tactics in a Table Tennis Game Based on Subgroup Discovery. In: *Machine Learning and Data Mining for Sports Analytics*, vol. 1783, pp. 101–112. Springer Nature Switzerland, Cham (2023)
6. Erades, A., Martin, P.E., Vuillemot, R., Mansencal, B., Peteri, R., Morlier, J., Duffner, S., Benois-Pineau, J.: SportsVideo: A Multimedia Dataset for Event and Position Detection in Table Tennis and Swimming. *MediaEval Workshop 2023* (2023)
7. Hodges, L.: Table Tennis Tactics for Thinkers. CreateSpace Independent Publishing Platform, Leipzig (Feb 2013)
8. Lan, J., Wang, J., Shu, X., Zhou, Z., Zhang, H., Wu, Y.: RallyComparator: visual comparison of the multivariate and spatial stroke sequence in table tennis rally. *Journal of Visualization* **25**(1), 143–158 (Feb 2022). <https://doi.org/10.1007/s12650-021-00772-0>, <https://doi.org/10.1007/s12650-021-00772-0>
9. McInnes, L., Healy, J.: Accelerated hierarchical density based clustering. In: *2017 IEEE international conference on data mining workshops (ICDMW)*. pp. 33–42. IEEE (2017)
10. Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F., Benguigui, N.: The serve impact in tennis: First large-scale study of big Hawk-Eye data. *Statistical Analysis and Data Mining: The ASA Data Science Journal* **9**(5), 310–325 (2016). <https://doi.org/10.1002/sam.11316>, <https://onlinelibrary.wiley.com/doi/abs/10.1002/sam.11316>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/sam.11316>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/sam.11316>
11. Perin, C., Vuillemot, R., Stolper, C.D., Stasko, J.T., Wood, J., Carpendale, S.: State of the Art of Sports Data Visualization. *Computer Graphics Forum (EuroVis’18)* **37**(3), 663–686 (2018). <https://doi.org/10.1111/cgf.13447>, <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13447>
12. Tukey, J.W.: *Exploratory data analysis*, vol. 2. Reading, Mass. (1977)
13. Voeikov, R., Falaleev, N., Baikulov, R.: TNet: Real-time temporal and spatial video analysis of table tennis. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. pp. 884–885 (2020)
14. Wang, J., Wu, J., Cao, A., Zhou, Z., Zhang, H., Wu, Y.: TacMiner: Visual Tactic Mining for Multiple Table Tennis Matches. *IEEE Transactions on Visualization and Computer Graphics* **27**(6), 2770–2782 (Jun 2021). <https://doi.org/10.1109/TVCG.2021.3074576>, <https://ieeexplore.ieee.org/document/9411869>, conference Name: IEEE Transactions on Visualization and Computer Graphics
15. Wei, X., Lucey, P., Morgan, S., Carr, P., Reid, M., Sridharan, S.: Predicting serves in tennis using style priors. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 2207–2215 (2015)

16. Wu, J., Dongyu Liu, Guo, Z., Xu, Q., Wu, Y.: TacticFlow: Visual Analytics of Ever-Changing Tactics in Racket Sports. *IEEE Transactions on Visualization and Computer Graphics* **28**, 835–845 (2022). <https://doi.org/10.1109/TVCG.2021.3114832>, <https://ieeexplore.ieee.org/document/9552436>
17. Wu, Y., Lan, J., Shu, X., Ji, C., Zhao, K., Wang, J., Zhang, H.: iTTVis: Interactive Visualization of Table Tennis Data. *IEEE Transactions on Visualization and Computer Graphics* **24**(1), 709–718 (Jan 2018). <https://doi.org/10.1109/TVCG.2017.2744218>, <https://ieeexplore.ieee.org/document/8017600>, conference Name: IEEE Transactions on Visualization and Computer Graphics
18. Ye, S., Chen, Z., Chu, X., Wang, Y., Fu, S., Shen, L., Zhou, K., Wu, Y.: ShuttleSpace: Exploring and Analyzing Movement Trajectory in Immersive Visualization. *IEEE Transactions on Visualization and Computer Graphics* **27**(2), 860–869 (Feb 2021). <https://doi.org/10.1109/TVCG.2020.3030392>, <https://ieeexplore.ieee.org/document/9222313/>
19. Zhu, Y., Naikar, R.: Predicting Tennis Serve Directions with Machine Learning. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) *Machine Learning and Data Mining for Sports Analytics*. pp. 89–100. Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-27527-2_7