P3VS: A Visual Analytics System for Structural Analysis of Pitching Sequence Trajectories

Ryosuke Tsujino
Graduate School of Business Administration and Information Science
Hokkaido Information University
Ebetsu, Japan
s2481107@s.do-johodai.ac.jp

Masahiko Itoh
Department of Information Media
Hokkaido Information University
Ebetsu, Japan
imash@do-johodai.ac.jp

***Abstract*—Pitch sequences in baseball reflect a pitcher's tactical intent, yet existing tools often overlook their structure by focusing only on frequency or location. This paper presents P3VS, an interactive visualization system that models each at-bat as an ordered trajectory of pitch zones, enriched with pitch type, velocity, and outcome. Using a modified Levenshtein Distance, the system clusters recurring sequence patterns and visualizes them as networks. Users can filter sequences by zone, pitcher, batter stance, and outcome to explore differences across contexts. P3VS supports real-time, outcome-specific analysis and reveals structural variations in pitch sequences that may influence game results.**

***Index Terms*—Baseball Analysis, Visualization, Interactive Web Application, User Interface.**

I. Introduction

Baseball strategy encompasses a wide range of elements, including base running, defensive positioning, and pitching. Among these, pitching plays a key role, as the sequence and selection of pitches directly influence the outcome of each at-bat and the overall flow of the game. For analysts, sequence-level data provides insights into pitcher intent and tactical decision-making. For viewers, understanding pitch sequences deepens engagement by building anticipation. This need has become more pronounced in recent years, as pitch types have diversified, new breaking balls such as sweepers have emerged, and high-velocity fastballs have become more common across the league [1], [2]. These developments emphasize the growing importance of analyzing pitch sequences rather than treating each pitch as an isolated event.

Although various tools and studies have been developed to support baseball analysis [3]–[5], most focus on pitch frequency or location independently and overlook the sequential structure of pitch combinations. While some studies acknowledge the importance of pitch order [6]–[8], they typically analyze only short-term or pairwise relationships, without modeling pitch sequences as structured patterns. Few systems provide interactive or visual tools for exploring these sequences. Our previous work [9] also focused solely on the order of pitch locations, without incorporating pitch type or enabling interactive exploration.

To address these limitations, we propose a system called P3VS (Pitcher Pitching Patterns Visualization System). It represents pitch sequences as trajectories, defined as ordered series of pitches annotated with pitch type, velocity, and at-bat outcome. Using clustering techniques, the system extracts recurring strategic patterns and visualizes them through an interactive interface. Users can filter sequences by zone, pitcher, and outcome, and explore them in real time. The system also highlights differences associated with specific at-bat results, such as strikeouts or hits.

The main contributions of this work are:

- A method for modeling pitch sequences as trajectories that integrate pitch type, velocity, and location to extract strategic patterns through clustering.
- An interactive system that enables real-time exploration of pitch sequences filtered by zone, pitcher, and outcome.
- A visualization method that enables users to compare pitch sequences across different outcomes, such as strikeouts or hits, to identify strategic differences.

II. Related Work

Research on pitching behavior has primarily focused on pitch type and location [6], [7]. Roegele [8] and Clemens [10] highlighted the strategic importance of pitch sequencing. Our prior system [9] clustered pitch sequences based solely on pitch location order and visualized them using network diagrams. Tools such as Baseball4D [4], Statcast Dashboard [5], VideoLens [11], Bloomberg Sports [12], Pitcher Plinko [13], and ESPN [14] provide visualizations of pitching trends and spatial data. However, these systems primarily focus on frequency-based or spatial analysis and generally do not capture the sequential nature of pitch sequences or support in-game reasoning. In contrast, our approach integrates pitch type, velocity, and outcome into a unified representation to support structural analysis and interactive exploration.

Other studies have modeled pitch selection as a decision-making process. Cox et al. [15] applied the Matching Law to explain pitch choices through reinforcement theory. Nakahara et al. [16] used stratified analysis with propensity scores to assess pitch effectiveness, and Hoffmann et al. [17] simulated at-bats to optimize decisions between pitchers and batters. However, these models treat pitches as independent events, overlooking the sequence-level structure underlying in-game strategies. Instead of treating pitches as isolated events, our

method analyzes the structure of entire pitch sequences to reveal how pitchers build their strategy across an at-bat.

In other sports domains, tactical visual analytics systems have been developed to explore evolving patterns and strategy. HoopInsight [18] compares basketball shooting strategies across players. In racket sports, systems such as iTTVis [19], Tac-Simur [20], and TacticFlow [21] visualize temporal tactical changes. Surveys on sports visualization [22]–[24] emphasize the importance of interactivity, structural awareness, and domain-specific design. However, few systems explicitly address the sequential nature of pitch combinations in baseball. Our system fills this gap by enabling users to explore pitch intent and flow through clustered trajectory patterns grounded in real game contexts.

III. Proposed System

A. System Requirements

Our system supports a broad range of users, including casual viewers, analysts, commentators, sportswriters and baseball researchers. Based on domain-specific tasks and prior experience, we define the following requirements for interactive analysis of pitching sequence trajectories.

R1. Zone-Based Filtering and Sequence Extraction: When investigating pitching patterns, users often start by selecting specific zones of interest. Yet, most existing tools do not support sequence filtering based on zone combinations. The system should allow the selection of one or more zones to extract relevant trajectories, facilitating intuitive, scenario-based exploration.

R2. Discovery of Strategic Pitch Patterns: Pitch sequences reflect strategies, but such patterns are hard to extract from raw data. The system should support clustering to reveal representative sequence structures employed in various contexts.

R3. Outcome-Aware Visualization: Understanding how sequences relate to outcomes (e.g., strikeouts, hits) is crucial, yet most systems do not support this linkage. The system must enable filtering and comparison by result category to assess the effectiveness of pitch sequence strategies.

R4. Interactive and Real-Time Exploration: Real-time exploration is important during live games and broadcasts, but static tools limit in-game analysis. The system should support interactive filtering, clustering, and visualization updates.

R5. Parameter Customization for Expert Analysis: To support expert-level and scenario-specific analysis, the system must allow users to specify key input parameters such as pitcher name, season, batter stance, number of clusters, and at-bat outcome categories. In addition, the system should enable detailed filtering and clustering across multiple dimensions including pitch type, speed, and location.

B. System Overview

Our previously developed system visualized pitch combinations by clustering sequences based on the order of pitching locations [9]. While this approach effectively summarized sequencing tendencies, it was limited to post-game analysis and lacked real-time and interactive capabilities. In this study, we redefine a pitch sequence as a trajectory composed of an ordered list of pitch zones with additional attributes such as pitch type, velocity, and outcome. We extract structural patterns from these trajectories using clustering. The updated version of P3VS introduces zone-based filtering, outcome-specific visualization, and live interactivity, enabling more flexible and user-driven exploration.

Figure 1 illustrates the system interface. Users first specify input parameters, including pitcher name, season, batter stance (all, left-handed, or right-handed), number of clusters, and at-bat outcome categories (Figure 1A) **(R5)**. Based on these inputs, the system retrieves the corresponding pitch data and constructs trajectory sequences for each at-bat (Figure 2 (i)). Users then select specific zones via the Zone Selection panel (Figure 1B) **(R1)**, which filters the data to include only trajectories that begin in the selected zones. The system calculates the distances between filtered pitch sequences (Figure 2 (ii)), performs clustering (Figure 2 (iii)), and visualizes the resulting pitch sequence patterns (Figure 2 (iv)) **(R2)**. This allows users to understand how pitchers structure their pitch sequences.

The system also supports outcome-based filtering, allowing users to compare clustered trajectory patterns across different at-bat results, such as strikeouts or outs (Figure 1C) **(R3)**. The system supports dynamic updates in response to user interaction, making it suitable for live gameplay analysis **(R4)**. P3VS is built using Dash[1] and Dash Cytoscape,[2] allowing browser-based deployment with responsive and interactive network visualizations. The following subsections describe each system component in detail.

C. Pitching Sequence Trajectory Data Extraction

We utilized pitch-by-pitch data from Baseball Savant [25] a comprehensive dataset containing approximately 100 variables describing pitch characteristics, game context, and at-bat outcomes. The data were provided in CSV format, and the key attributes used in our analysis are listed in Table I. Among them, the *zone* attribute specifies where the pitch crossed home plate, based on a 13-zone classification system comprising nine strike zones (*a* through *i*) and four ball zones (*j* through *m*) (Figure 3, left). To ensure consistency, we flipped the default catcher's viewpoint to the pitcher's perspective. The *events* attribute records the result of each at-bat, such as a strikeout or home run.

Before extracting trajectories, the dataset is filtered based on user selections, including pitcher, season, batter stance (right-handed, left-handed, or all), and outcome category (e.g., strikeout, hit). The system supports multiple seasons and outcome types simultaneously (Figure 1.A). We then extract pitching sequence trajectories using the method proposed by Miyagi et al. [26], which detects recurring patterns in general sequential data. In this study, we apply their approach to sequences of pitch zones. For example, a sequence of pitches thrown to zones *m*, *m*, *e*, and *m* is represented as the trajectory

[1]https://dash.plotly.com/

[2]https://dash.plotly.com/cytoscape/

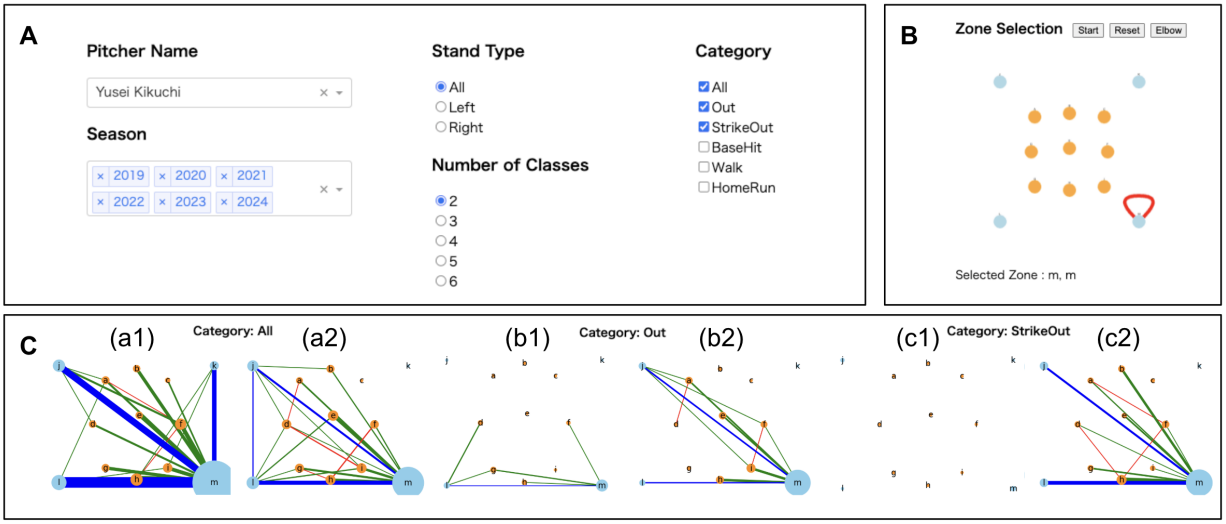

Fig. 1. User interface of P3VS. (A) Filtering options, (B) Zone selection panel, (C) Visualization of clustered trajectory patterns by at-bat outcome (All, Out and StrikeOut).

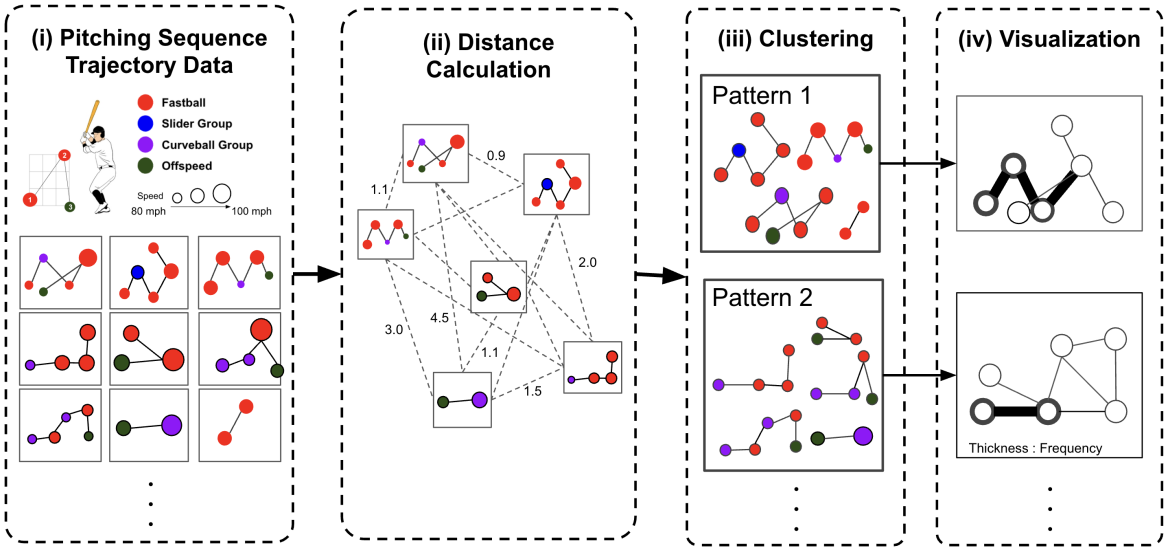

Fig. 2. Clustering process from pitching sequence trajectories to pattern visualization.

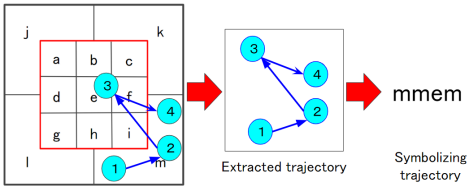

Fig. 3. Extraction of pitch trajectories from pitch-by-pitch data.

mmem (Figure 3). We excluded all single-pitch sequences and analyzed only those consisting of two or more pitches.

D. Zone Selection

We developed a zone selection feature that allows users to filter pitching sequence trajectory data based on an ordered sequence of selected zones (Figure 1.B). Selecting a zone filters the trajectory data described in Section III-C. The interface, built with Dash Cytoscape, arranges 13 nodes (*a* to *m*) based on the zone table in Figure 3 (left). This zone layout is based on [25]. Zones *a* to *i* (strike zone) are shown in orange, and *j* to *m* (out-of-zone) are shown in light blue.

When a node is selected, the system extracts trajectories that begin with the selected zone sequence. For example, if the user selects zone *m* first and zone *m* second, only trajectories beginning with *mm* are extracted, and the corresponding edge *m*-*m* is displayed (Figure 1.B).

The Zone Selection panel also includes three buttons: Start runs the analysis; Reset resets the interface; and Elbow displays the elbow method results for selecting the optimal number of clusters.

TABLE I
SELECTED ATTRIBUTES FROM THE BASEBALL SAVANT DATASET.

Attribute	Description
game_date	Date of the game
batter	MLB player ID of the batter
pitcher	MLB player ID of the pitcher
zone	Zone where the pitch crossed home plate
stand	Batter's stance (R = Right-handed, L = Left-handed)
inning	Inning number
inning_topbot	Half-inning indicator (Top or Bottom)
effective_speed	Pitch speed at the plate (mph), adjusted for extension
pitch_name	Pitch type (e.g., Four-Seam Fastball, Slider)
pfx_x	Horizontal movement (inches)
pfx_z	Vertical movement (inches)
game_pk	Unique game identifier
at_bat_number	At-bat number within the game
pitch_number	Pitch number within the at-bat
events	At-bat result (e.g., strikeout, single)

E. Trajectory Clustering

To quantify similarity between pitching sequence trajectories, we adopted a modified Levenshtein Distance framework. Pairwise distances were calculated based on location, pitch type and speed. The Levenshtein Distance measures the minimum number of insertions, deletions, or substitutions to convert one sequence S_1 into another S_2, where smaller values indicate higher similarity. For example, the distance between *mmem* and *mmmfj* is 3, with two substitutions and one insertion (Figure 4). Substitution costs were assigned based on the degree of difference between elements, while insertion and deletion costs were kept uniform. These costs were weighted and normalized to ensure comparability across different factors.

Zone Distance, Pitch Type and Speed. Understanding the spatial relationship between pitch zones is crucial for analyzing how pitchers structure their sequences. Transitions between zones, such as moving from outside to inside, often carry strategic intent. However, previous systems [9] assigned uniform costs to all transitions, ignoring physical distance and overlooking subtle spatial patterns. To address this, we define zone distance as the Euclidean distance between zone coordinates. This spatial metric captures relative placement differences and enables more fine-grained analysis than uniform-cost approaches.

As shown in Figure 5, we grouped pitch types into Fastball, Offspeed, Curveball Group, and Slider Group based on the classification from Baseball Savant [25]. We computed the average horizontal and vertical movement (pfx_x, pfx_z) for each pitch type per pitcher, and used the Euclidean distance between the corresponding centroids when types differed. This approach reflects physical and strategic differences between pitch types, improving over uniform-cost methods. For example, Fastball (F) and Curveball Group (C) are positioned far apart, indicating distinct characteristics, while Fastball (F) and Offspeed (O) are close together, suggesting similar movement.

We categorized pitch speed into Slow, Middle, and High based on each pitcher's individual speed distribution. Substitution costs varied according to these categories, allowing for finer comparisons across pitchers.

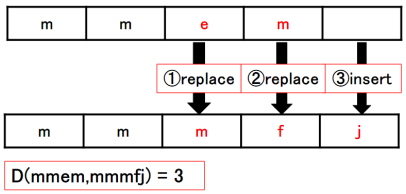

Fig. 4. Example calculation of the Levenshtein Distance between *mmem* and *mmfj* ($D = 3$).

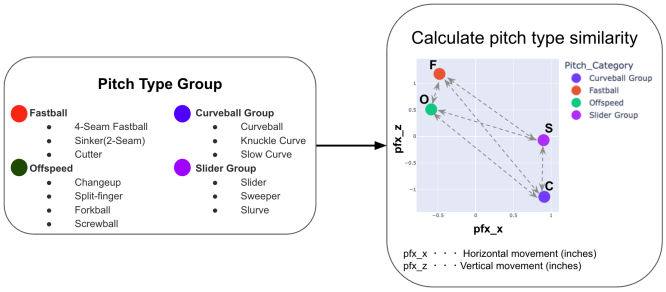

Fig. 5. Pitch Type Grouping and Similarity: Pitch types are classified into four functional groups. Inter-type similarity is assessed based on average pitch movement in horizontal and vertical dimensions.

Integrated Distance and Clustering. We applied K-medoids clustering to group pitch sequences using an integrated distance matrix that combines location, pitch type, and speed. Unlike K-means, K-medoids supports heterogeneous data and selects actual sequences as cluster centers. This approach groups structurally similar sequences and extracts pitch sequence patterns.

F. Visualization of Pitching Sequence Trajectory Data

We visualized the clustering results using network diagrams, where each network represents a trajectory cluster composed of ordered pitch zone sequences from individual at-bats. For example, the trajectory *mmem* (Figure 3) produces the nodes and edges listed in Table II. We used an undirected graph model, treating bidirectional edges such as *e-m* and *m-e* as a single edge (*e-m*).

TABLE II
NODES, EDGES, AND FREQUENCIES FOR TRAJECTORY *mmem*.

Nodes	*e*, *m*
Edges	*m-m*, *m-e*, *e-m*
Node frequencies	*e*: 1, *m*: 3
Edge frequencies	*e-m*: 2, *m-m*: 1

We generated the visualizations using NetworkX,[3] drawing one network per cluster (Figure 1C). The 13 pitch zones (*a* to *m*) were initially placed based on the strike zone layout in Figure 1B. Node size was scaled based on the frequency within each cluster (100× frequency), and edge thickness and color also reflected cluster-specific frequencies: red for strike-to-strike, green for strike-to-ball, and blue for ball-to-ball tran-

[3]https://networkx.org

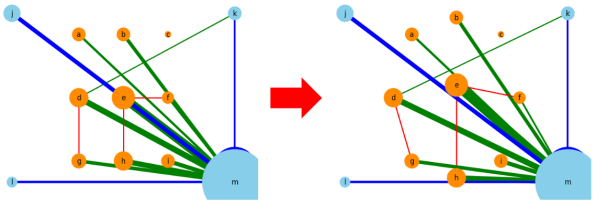

Fig. 6. Left: Original layout. Right: Adjusted layout.

sitions. To reduce clutter, only edges with a frequency of two or more were displayed. Self-loops, which indicate consecutive pitches to the same zone, were omitted at rendering time to improve visual clarity.

Since the original layout led to overlapping elements (e.g., edge *e*-*m* overlapping with *j*-*m*), we manually adjusted node positions (e.g., *b*, *d*, *e*, *f*, and *h*) for readability (Figure 6).

P3VS also supports clustering and visualization of pitching sequence trajectories by at-bat result. The `events` attribute was categorized into five outcome groups (Table III), and the same visualization method was applied. Users can filter by outcome using checkboxes in the interface (Figure 1A).

Figure 1 shows an example where outcomes "Out" and "StrikeOut" were selected, and trajectories beginning with *mm* were visualized separately for each outcome, enabling direct comparison. In this figure, (a1) and (a2) show "All" outcomes, (b1) and (b2) "Out," and (c1) and (c2) "StrikeOut."

TABLE III
AT-BAT RESULT CATEGORIES BASED ON THE `events` ATTRIBUTE.

Category	Included results
StrikeOut	strikeout
Out	field_out, double_play, force_out, etc.
BaseHit	single, double, triple
Walk	walk, hit_by_pitch
HomeRun	home_run

IV. Case Studies

This section presents a case study using P3VS. We analyzed 23 pitchers: the top 20 by pitch volume from 2018 to 2024, plus Shohei Ohtani, Yu Darvish, and Yusei Kikuchi[4]. As an initial example, we examined all pitches by Kikuchi during this period (Figure 7) to illustrate the overall structure and functionality of P3VS. The visualization shows frequent use of outer zones such as *j* and *m*, providing a basis for selecting a specific zone (e.g., *j*) to explore subsequent pitch sequences.

Figure 8 shows pitch sequences starting from zone *j*, clustered by outcome. In (a1) (Out), transitions spread across zones such as *d*, *f*, and *h*, suggesting a flexible approach. In (b3) (StrikeOut), strong links from *j* to *m*, and then to *e*, indicate a focused strategy aimed at inducing swings and misses. These differences suggest that pitch sequence structure varies depending on the intended outcome.

Figure 1 shows sequences starting with two pitches to zone *m* ("mm"), clustered by outcome. Both categories share the edge *m*–*m*, but the subsequent sequences diverge. "Out" sequences tend to transition to zones outside the strike zone, such as *j* or *l* (b2), whereas "StrikeOut" sequences frequently involve the *l*–*m* transition and disperse across multiple zones within the strike zone (c2), as indicated by the dense and varied green edges. These differences highlight how sequences with identical openings can vary depending on the intended or resulting outcome.

Figure 9 presents "HomeRun" pitch sequences clustered into two groups. The left cluster displays numerous green edges that start in the out-of-zone regions *j* and *l* and flow into the strike zone. The right cluster, by contrast, concentrates pitches in the central zones *d* and *e*. Both clusters therefore reveal a common tendency for trajectories to converge on hittable locations near the heart of the zone.

These examples illustrate how P3VS allows users to interactively explore pitch sequences by specifying starting zones and filtering by outcomes. This functionality supports the discovery of strategic patterns and adaptive behaviors that may not be apparent through traditional aggregate analysis.

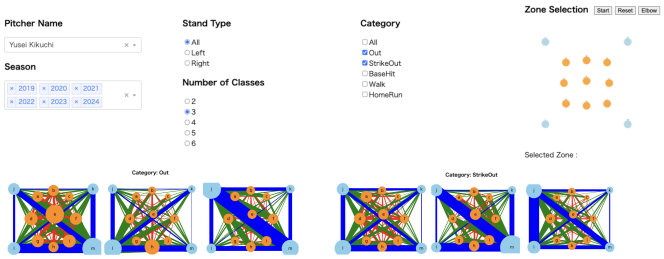

Fig. 7. Exploration of Yusei Kikuchi's pitch sequences (2019–2024), filtered by outcome categories "Out" and "StrikeOut."

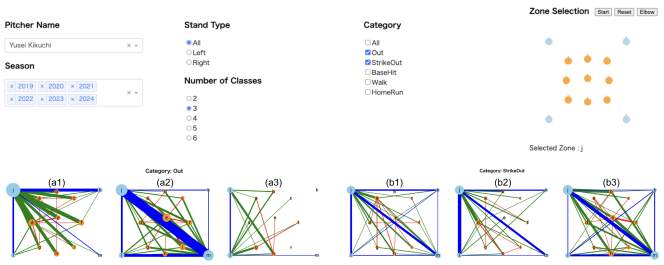

Fig. 8. Pitch sequences starting with zone "j" for Yusei Kikuchi, clustered into three groups by outcome ("Out" and "StrikeOut").

V. Discussion

P3VS, in its current form, has several limitations related to exploration support, visual representation, user interface design, and data accessibility. This section outlines these issues and proposes directions for improvement.

While users can freely explore pitch sequences by selecting starting zones and outcome categories, the system lacks interpretive guidance. Key differences or recurring patterns are not automatically highlighted, and users cannot adjust visual parameters such as node size or edge thickness. Adding

[4]Currently with the Los Angeles Angels

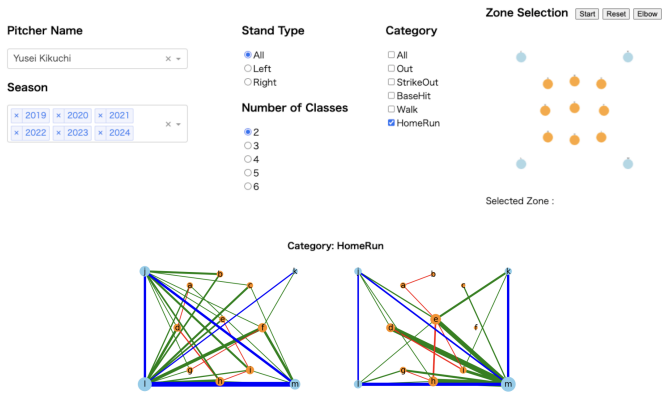

Fig. 9. Pitch sequence trajectory clusters for Yusei Kikuchi filtered by outcome category "HomeRun."

adjustable controls and basic pattern-emphasis features would improve clarity and adaptability.

Real-time usage remains challenging, not due to system latency, but because determining pitch location and categorizing it in real time is inherently difficult. In live settings, reliable input data is often unavailable right after each pitch, which limits the system's applicability for in-game tactical support.

The current system lacks directional edges and does not visualize pitch attributes such as type, velocity, or movement. Exploration is also limited to fixed positions (e.g., first or second pitch), making it difficult to analyze mid-at-bat developments or decisive finishing pitches. These issues could be mitigated by adding directionality, encoding pitch attributes in edge styles, and enabling user-defined starting points.

The interface currently supports pitcher-based filtering not batter selection, limiting analysis of pitcher–batter dynamics. Adding batter-level filters would support matchup evaluations.

VI. Conclusion

In this paper, we presented P3VS, a system for visualizing pitch sequence tendencies using data from Baseball Savant. For future work, we plan to conduct user evaluations to assess its effectiveness and practical utility, particularly in helping users better understand individual pitchers' strategic tendencies. While the current system provides structural insights into pitch sequences, it does not yet incorporate edge direction, making it difficult to represent the sequential flow of pitch decisions. Introducing directed graphs may improve interpretability by more clearly capturing the progression within each sequence.

Acknowledgment

This research was supported by the Hokkaido IT Creator Discovery and Development Project (Shinsetsu Program), organized by Shinsetsu, a general incorporated association. We sincerely appreciate their support.

References

[1] R. O'Connell, "Why Some MLB Pitchers Are Abandoning the Fastball," *FiveThirtyEight*, 2022. "Accessed on 2025-05-15". Available: https://fivethirtyeight.com/features/why-some-mlb-pitchers-are-abandoning-the-fastball/

[2] M. Petriello, "The Hot New Pitch That's Sweeping Across the Majors," *MLB.com*, 2023. "Accessed on 2025-05-15". Available: https://www.mlb.com/news/sweeper-slider-latest-pitching-trend-explained

[3] B. Lindbergh and T. Sawchik, *The MVP Machine: How Baseball's New Nonconformists Are Using Data to Build Better Players*. New York, NY, USA: Basic Books, 2019.

[4] M. Chen, A. Bezerianos, and P. Isenberg, "Baseball4D: A tool for baseball game reconstruction visualization," in *Proc. VAST2014*, pp.23–32, 2014.

[5] M. Lage, B. Harrison, and A. Endert, "Statcast dashboard: Exploration of spatiotemporal baseball data," in *Proc. VIS2026 (Posters)*, 2016.

[6] J. R. Bock, "Analysis of pitching patterns with regard to the pitch type and zone to be thrown next," *Sabermetric Research*, 2015.

[7] J. Z. Clay, "Strategic pitch location: The role of two-pitch sequences in pitching success," *J. Sports Analytics*, vol. 9, no. 2, pp. 87–101, 2023.

[8] D. Roegele, "Pitch sequencing and its impact on outcomes," presented at the SABR Analytics Conference, 2014.

[9] R. Tsujino and M. Itoh, "Visualization system for pitching sequence patterns based on pitching locations' order," in *Proc. PacificVis2024 (Posters)*, 2024.

[10] B. Clemens, "Pitch Type Sequence Similarity Ratio: Understanding the Role Pitch Sequencing Plays in the MLB," Uram Analytics Blog, 2023. "Accessed on 2025-05-15". https://www.uramanalytics.com/post/pitch-type-sequence-similarity-ratio-understanding-the-role-pitch-sequencing-plays-in-the-mlb

[11] J. Matejka, T. Grossman, and G. Fitzmaurice, "Video Lens: Rapid Playback and Exploration of Large Video Collections and Associated Metadata," in *Proc. UIST2014*, pp. 537–546, 2014.

[12] Y. Moon and R. Brath, "Bloomberg Sports: Visualization for pitch analysis," *CG&A*, vol. 33, no. 1, pp. 12–17, 2013.

[13] MLB Advanced Media, LP., "Pitcher Plinko," "Accessed on 2025-05-15". Available: https://baseballsavant.mlb.com/visuals/pitch-plinko

[14] ESPN Enterprises, Inc., "MLB on ESPN," "Accessed on 2025-05-15". Available: https://www.espn.com/mlb/

[15] D. J. Cox, J. J. Sosine, and J. Dallery, "Application of the Matching Law to Pitch Selection in Professional Baseball," *J. Appl. Sport Psychol.*, vol. 29, no. 2, pp. 202–217, 2017.

[16] H. Nakahara, T. Nagami, and Y. Yamashita, "Pitching strategy evaluation via stratified analysis using propensity score," *Front. Sports Act. Living*, vol. 4, pp. 1–12, 2022.

[17] G. Hoffmann, J. Bronn, and H. Roark, "Baseball Decision-Making: Optimizing At-bat Simulations," *SMU Data Science Review*, vol. 5, no. 2, 2022.

[18] Y. Fu and J. Stasko, "HoopInsight: Analyzing and comparing basketball shooting performance through visualization," in *Proc. VIS2023*, Oct. 2023.

[19] Y. Wu, Z. Liu, S. Liu, and K. Zhou, "iTTVis: Interactive Visualization of Table Tennis Data," in *Proc. PacificVis 2018 (Posters)*, 2018.

[20] X. Wang, Y. Wu, and H. Qu, "Tac-Simur: Tactic-based Simulative Visual Analytics of Table Tennis," in *Proc. PacificVis2019*, pp. 112–116, 2019.

[21] Y. Wu, Z. Liu, H. Qu, and S. Liu, "TacticFlow: Visual Analytics of Ever-Changing Tactics in Racket Sports," in *Proc. VIS2021*, 2021.

[22] C. Perin et al., "State of the Art of Sports Data Visualization," *Comput. Graph. Forum*, vol. 37, no. 3, pp. 675–699, 2018.

[23] O. C. Obi et al., "Data Science in Sports Analytics: A Review of Performance Optimization and Fan Engagement," *World J. Adv. Res. Rev.*, 2024.

[24] L. Shi, Y. Wu, and H. Qu, "Visual Analytics for Sports Data: A Survey," *TVCG*, vol. 30, no. 1, pp. 1–20, 2023.

[25] MLB Advanced Media, LP. *Baseball Savant*. Available at: https://baseballsavant.mlb.com/. Accessed 2025-5-15.

[26] Y. Miyagi, M. Onishi, C. Watanabe, T. Itoh, and M. Takatsuka, "Classification and visualization for symbolic people flow data," *J. Vis. Lang. Comput.*, vol. 43, pp. 91–102, 2017.