Adaptive Coaching with High-Fidelity Physics for Virtual Reality Badminton Training

Category: Poster

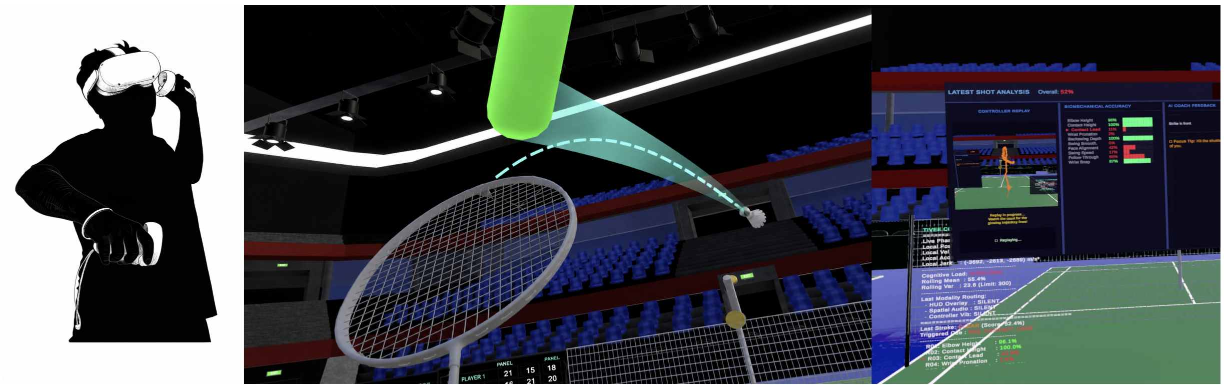

Figure 1: A user performing a badminton smash in *XR Badminton*. The system decomposes the stroke into seven biomechanical phases, provides real-time multimodal feedback, and adapts output modalities for accessibility.

Abstract

We present *XR Badminton*, a standalone VR coaching system for Meta Quest 2/3 that combines deterministic biomechanical analysis with generative Large Language Models (LLMs). The system segments overhead strokes into a seven-phase sequence using a Holt-Linear filtered state machine and evaluates ten expert-derived rules by comparing real-time kinematic trajectories against a procedurally generated biomechanical reference template. A hybrid neuro-symbolic feedback engine routes symbolic diagnoses to an LLM, which acts as a neural verbalizer to synthesize personalized spoken coaching, falling back locally to deterministic audio cues when offline. Dynamic challenge-point pacing adapts feedback modalities based on rolling performance volatility to prevent cognitive overload. Interactive ex-situ 3D volumetric replays let users visually compare their swings against the reference. We describe our prototype, including real-time standalone execution.

Index Terms: Virtual reality, sports training, badminton, biomechanical feedback, accessibility, coaching systems.

1 Introduction

Badminton is the fastest racket sport, with professional smash speeds exceeding 400 km/h [9]. Badminton moves require precise coordination of elbow elevation, contact height, wrist pronation, and follow-through timing, a chain of movements that novices find difficult to self-diagnose.Traditional coaching is constrained by instructor availability, subjective observation, and limited accessibility for users with hearing or visual impairments.

Virtual reality (VR) offers a compelling alternative: repeatable, instrumented practice with real-time feedback. Recent work in VR tennis demonstrated the feasibility of phase-based biomechanical coaching [6]. Applying this to badminton requires addressing unique shuttle aerodynamics and wrist-dominant power generation [9]. Furthermore, purely rule-based deterministic coaching engines often produce rigid, repetitive feedback that fails to contextualize compound errors, increasing cognitive load without proportional learning gains [8].

We propose *XR Badminton*, a neuro-symbolic system architecture that addresses these limitations. This paper presents:

1. A **real-time tracking pipeline** utilizing a Holt-Linear filtered state machine to segment complex overhead strokes into a seven-phase kinematic sequence, evaluating performance across ten expert-derived biomechanical rules against a procedurally generated reference template.

2. A **physics-informed training environment** featuring a procedural, aerodynamic shuttlecock simulation with nose-heavy physical balance and quadratic drag trajectory prediction.

3. A **hybrid neuro-symbolic feedback engine** combining a deterministic local rule engine for symbolic deficit classification with an LLM neural verbalizer to generate personalized verbal cues, supported by interactive 3D volumetric replays.

2 Related Work

VR-based sports training has progressed from early proof-of-concept simulators, such as the basketball free-throw trainer by Covaci et al. [2], to systems capable of real-time biomechanical coaching. VRTennis [6] advanced this line of work by introducing phase-based stroke segmentation with extensive coaching rules evaluated in a formal user study. Our work extends this paradigm to badminton, which presents distinct challenges in shuttle aerodynamics, overhead stroke geometry, and wrist-dominant power generation.

A key question in any coaching system is what constitutes a "correct" stroke. Zhao and Li [9] provided a biomechanical foundation by quantifying kinematic differences between professional and amateur smashes, while BadminSense [1] established stroke-quality dimensions from inertial data across 848 strokes. Building on these insights, we employ a player-centric spatial coordinate alignment to evaluate a user's stroke against a procedurally generated reference template, enabling precise evaluation of kinematic snapshots.

Figure 2: Five-layer system architecture. Sensing data flows upward through Motion Analysis, the Neuro-Symbolic Coach, and the Challenge & Access layer, with an Analytics bus recording telemetry asynchronously at 72 Hz.

Equally important is how feedback is delivered. Wang et al. [8] demonstrated that restrained, well-timed corrections, including deliberate silence windows, outperform continuous verbal instruction in guided sports communication. Motion-to-language systems like ViSTAR [7] have utilized classifiers to prioritize errors before LLM verbalization in AR. Unlike server-side AR systems, XR Badminton targets fully standalone VR execution under strict frame budgets (13.8 ms) with a local deterministic audio fallback. Furthermore, inclusive design frameworks like BRIDGE [5] emphasize embodiment-aware representations in adaptive coaching; our architecture directly incorporates this ethos via first-class modality substitution routing (HUD text, spatial audio, and haptics) to accommodate diverse learner needs. We combine this with interactive 3D volumetric replay, which leverages self-observation, one of the most effective mechanisms in motor learning, to let users visually compare their execution against the professional reference.

3 Proposed System Architecture

The proposed *XR Badminton* framework is organized as a five-layer pipeline (Fig. 2).

Layer 1: Sensing. The system reads six-degree-of-freedom (6-DoF) pose data from the head-mounted display (HMD) and the dominant-hand controller (racket) via the Meta Quest runtime at 72 Hz.

Layer 2: Motion Analysis. Raw controller poses are transformed into a *player-centric coordinate frame*: the HMD forward vector is projected onto the horizontal XZ plane, creating a local frame aligned with the player's horizontal facing direction (yaw), thereby decoupling stroke kinematics from locomotion and jumping. Positions $\mathbf{p}_t$ and velocities $\mathbf{v}_t$ are smoothed at each frame t (time step $\Delta t \approx 0.0138$ s) using a Holt-linear (double exponential) filter [4] with smoothing factors $\alpha = 0.25$ and trend factors $\beta = 0.10$:

$$\hat{\mathbf{p}}_t = \alpha \mathbf{p}_t + (1-\alpha)(\hat{\mathbf{p}}_{t-1} + \mathbf{b}_{t-1}\Delta t) \tag{1}$$

$$\mathbf{b}_t = \beta \frac{\hat{\mathbf{p}}_t - \hat{\mathbf{p}}_{t-1}}{\Delta t} + (1-\beta)\mathbf{b}_{t-1} \tag{2}$$

where $\hat{\mathbf{p}}_t$ is the filtered position and $\mathbf{b}_t$ is the trend vector. Accelerations $\mathbf{a}_t$ and jerks $\mathbf{j}_t$ are computed using second-order accurate three-point causal backward differences to eliminate phase lag and look-ahead:

$$\mathbf{a}_t = \frac{3\hat{\mathbf{v}}_t - 4\hat{\mathbf{v}}_{t-1} + \hat{\mathbf{v}}_{t-2}}{2\Delta t} \tag{3}$$

$$\mathbf{j}_t = \frac{3\mathbf{a}_t - 4\mathbf{a}_{t-1} + \mathbf{a}_{t-2}}{2\Delta t} \tag{4}$$

A seven-phase state machine (Ready → Preparation → Backswing → Acceleration → Contact → Follow-Through → Recovery) segments each stroke using expert-driven kinematic thresholds.

Layer 3: Neuro-Symbolic Feedback & Volumetric Replay. To overcome the limitations of rigid decision trees and ensure offline robustness, Layer 3 employs a hybrid neuro-symbolic architecture:

1. *Symbolic Diagnosis & LLM Synthesis (Primary Mode)*: The local C engine performs high-precision, deterministic calculations (Table 1) comparing the user's spatial trajectories against a procedurally generated biomechanical reference template. This ideal reference trajectory is dynamically constructed using kinematic formulas that model a perfect, high-reaching forward smash arc, scaled and aligned to the player's anatomical contact coordinate. The system evaluates the scores $S_i \in [0, 100]$ for all active rules i. If a score falls below its respective flaw threshold T_i, a weighted deficit is calculated as:

$$W_i = (100 - S_i) \cdot w_i \tag{5}$$

   where w_i is the structural priority weight of the rule. The hybrid engine then selects the rule with the maximum weighted deficit to dictate the coaching correction cue:

$$i^* = \arg \max_{i: S_i < T_i} W_i \tag{6}$$

   This selected rule ID and score are passed as a structured *symbolic diagnosis* to the LLM (Gemini 2.5 Flash), which acts as a neural verbalizer to synthesize a personalized, context-aware coaching comment (max 25 words).

2. *Deterministic Rule Evaluation (Offline Fallback)*: If the API request fails or times out (guarding the connection with a 2.0 s timeout window), the system falls back locally to the symbolic engine's output, playing prerecorded audio cues (e.g., `R02.mp3`).

3. *Volumetric Replay*: Following a stroke, users can trigger a 3D volumetric replay. The system instantiates a ghost racket avatar of the user's stroke, projected 2 m in front of the player (ex-situ) to prevent camera clipping, overlaid with the professional reference trajectory.

Layer 4: Challenge-Point Pacing & Accessibility. A modality-substitution engine maps output feedback based on user preferences. Guided by the Challenge Point Framework [3], when high performance volatility is detected (indicating the player is struggling to stabilize their motor execution), the pacing gate dynamically enforces consecutive flawed swing requirements and suppresses verbal speech feedback to prevent cognitive overload.

Layer 5: Analytics. An asynchronous telemetry bus writes frame-by-frame CSV logs at 72 Hz and per-stroke summary JSON, capturing phase scores, coaching-rule adherence, and trajectory data for offline analysis.

Table 1: Active Coaching Rules and Mathematical Mappings

ID	Coaching Rule	w	Phase	Mapping Shape
R01	Elbow Height at Backswing	1.8	Backswing	Sigmoid
R02	Contact Height/Extension	2.0	Contact	Linear (Calibrated)
R03	Contact Lead Position	1.7	Contact	Gaussian
R04	Wrist Pronation Speed	1.2	Accel.	Sigmoid
R05	Backswing Depth	1.5	Backswing	Linear
R06	Swing Smoothness	1.1	Accel.	Decel. Ratio (Linear)
R07	Racket Face Alignment	1.9	Contact	$\cos^2$ Profile
R08	Swing Speed at Contact	1.4	Contact	Sigmoid
R09	Follow-Through Path	1.3	Follow-Thr.	Dual Linear
R10	Wrist Snap / Flexion	1.4	Contact	Sigmoid

4 Prototype and Ongoing Implementation

We have developed a prototype of the proposed *XR Badminton* framework in Unity 2022 LTS using the Universal Render Pipeline (URP), targeting Meta Quest 2/3 standalone headsets. Currently, the core layers are in varying stages of implementation:

Sensing & Kinematics (Layers 1 & 2). Sensing and kinematics are fully implemented. We retrieve HMD and controller tracking poses at 72 Hz via the Meta Quest SDK. Racket-tip velocity is computed as $\mathbf{v}_{\text{point}} = \mathbf{v}_{\text{linear}} + \boldsymbol{\omega} \times \mathbf{r}$ (combining linear controller velocity with angular velocity at the hit offset) with a kinematic-tracking fallback for grabbed rackets. Player-centric coordinate normalization and Holt-linear filtering are active. The Holt-linear filter ($\alpha = 0.25$, $\beta = 0.10$) introduces a group delay of ≈ 1.5 frames (≈ 20 ms at 72 Hz), which lies safely below the human visual lag threshold (~ 50 ms) and is compensated in ex-situ volumetric rendering. To support standalone VR execution (HMD and a single dominant controller) without full-body tracking or inverse kinematics (IK) overhead, joint-level rules are operationalized via direct kinematic proxies: elbow elevation is approximated using racket-height relative to HMD at the deepest backswing point ($R01$), and wrist snap/pronation are measured via controller angular velocity pitch/roll components at contact ($R04, R10$). The 7-phase state machine successfully segments strokes. The local C# processing overhead is under 4.2 ms, fitting comfortably within the standalone 13.8 ms frame budget.

Shuttlecock & Trajectory Physics. The shuttlecock's physics, including quadratic drag ($\mathbf{F}_d = -\frac{1}{2}\rho C_d A|\mathbf{v}|^2\hat{\mathbf{v}}$, terminal velocity ≈ 6 m/s, and cork-first Slerp alignment), is fully implemented. Predicted trajectories are rendered via Euler integration (50 Hz, 250 steps) with color-coded paths (pink for smashes > 15 m/s, gold for drives > 8 m/s, and cyan for drops < 8 m/s) and are currently being optimized to ensure Quest frame rates remain stable at 90 fps. The virtual shuttlecock launch thresholds (e.g., smashes > 15 m/s) are deliberately scaled down compared to empirical professional smashes (> 70 m/s) to: (1) prevent optical tracking loss of controllers during ultra-high-speed swings on Quest cameras; and (2) provide reaction times suitable for recreational players on standard Quest frame budgets.

Feedback & Accessibility (Layers 3 & 4). The 10 coaching rules are implemented, utilizing spatial differences against the procedurally generated biomechanical reference template. For vertical extension (Rule R02), the target height is dynamically calibrated based on the player's maximum observed reach in the HMD-relative local frame:

$$y_{\text{target}} = \text{clamp}(y_{\text{max\_obs}}, 0.70, 1.10)\,\text{m} \tag{7}$$

and the score is computed by mapping the contact height piecewise-linearly from $y_{\text{target}} - 0.40$ m (score 0) to y_{target} (score 100). For swing smoothness (Rule R06), the score is evaluated by calculating the cumulative deceleration drops $d_t < -4.0\,\text{m/s}^2$ during the forward phase relative to the peak forward acceleration a_{peak}, and mapping this ratio linearly to a score:

$$S_{\text{R06}} = \max\left(0, 100\left(1 - \frac{\sum |d_t|}{0.35 \cdot a_{\text{peak}}}\right)\right) \tag{8}$$

The priority-weighted deficit selection and rolling score variance calculations for challenge-point pacing adaptation are active. Modality substitution routes the LLM's verbal cues to HUD text or haptic vibration profiles.

LLM Coaching & Volumetric Replay. The LLM coaching client is integrated via an asynchronous, main-thread C# Web Request to the Gemini 2.5 Flash API with a native `CancellationToken` timeout guard and a 1.5 s cooldown to prevent rate-limit exhaustion. The client passes the symbolic diagnosis via structured payloads and uses Gemini's `responseSchema` to enforce strict key mapping for Unity's `JsonUtility`. Spoken audio is generated using a local JNI bridge to the native Android TTS engine with explicit reference disposal to prevent memory leaks on the Quest. Under testing, the online API request exhibits a mean round-trip latency of 1.64 s ($SD = 0.22$ s), with the local 2.0 s timeout guard reliably executing fallback audio cues in 100% of simulated connection dropouts. The volumetric replay renders ex-situ comparative overlays with depth-cued trails.

5 Evaluation and Pilot Results

To validate the system architecture and coaching efficacy, we conducted a pilot usability study ($N = 10$) followed by the design of a randomized controlled trial (RCT) protocol.

Pilot Study Usability Results. A usability evaluation was conducted with a cohort of $N = 10$ participants (8 male, 2 female; age $M = 19.8$ years, $SD = 1.48$). The cohort represented casual (60%), intermediate (20%), and competitive (20%) players.

The system achieved a mean System Usability Scale (SUS) score of 73.50 ($SD = 12.97$), which lies above the standard usability benchmark of 68.0, indicating good system usability. For the NASA-TLX (measured on a 1–10 scale), the mean subscale ratings were: Mental Demand 3.00 ($SD = 1.70$), Physical Demand 4.60 ($SD = 2.55$), Temporal Demand 3.50 ($SD = 1.90$), Effort 5.30 ($SD = 2.54$), and Frustration 2.70 ($SD = 2.00$). The mean self-reported Performance score (where 1 is perfect and 10 is failure) was 6.00 ($SD = 2.40$). The overall mean Raw Task Load Index (R-TLX) was 4.18 ($SD = 1.14$), indicating a highly manageable cognitive workload.

Qualitative feedback was positive, highlighting the high immersion of the virtual court and the utility of the ex-situ 3D volumetric replay. Two primary design refinements were implemented in response to pilot issues: (1) we resolved an accidental grip-trigger racket drop conflict by adjusting the trigger-button activation thresholds; and (2) we optimized

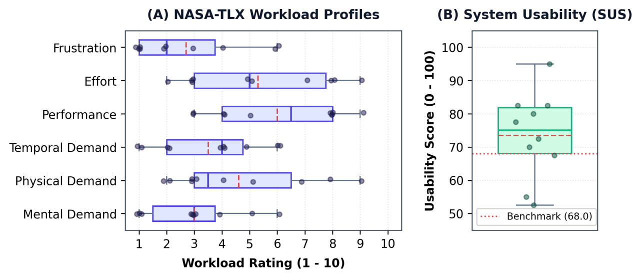

Figure 3: Pilot study results ($N = 10$): (A) NASA-TLX workload dimensions (1-10 scale; dashed red lines indicate mean values, black dots represent raw participant points); and (B) System Usability Scale (SUS) distribution relative to the standard usability benchmark of 68.0 (grade 'B').

the LLM cooldown gate to prevent repetitive feedback on successive swings.

User Evaluation Protocol. The planned main study will employ a randomized controlled trial (RCT) between-subjects design ($N \geq 40$, based on a G*Power analysis for a medium-to-large effect size at $\alpha = 0.05$ and power $\beta = 0.80$) with participants stratified by prior experience. To isolate the effect of LLM personalization, both groups will receive the ex-situ 3D volumetric replay. The experimental group will receive the prioritized, LLM-guided feedback. The active control group will receive prioritized feedback using standard C# template audio clips. The protocol comprises: pre-test (10 shots) → system-guided training (20 mins) → post-test (10 shots) → 24-hour retention test (10 shots, no feedback) → questionnaires.

We will test three primary hypotheses:

- **H1**: Participants training with LLM-guided feedback will exhibit significantly higher post-training motivation (Intrinsic Motivation Inventory) and self-efficacy than those in the active control group.

- **H2**: The experimental group will exhibit a significantly greater pre-to-post reduction in trajectory deviation from the biomechanical reference template compared to the control group, and this improvement will be preserved during the 24-hour retention test.

- **H3**: Despite sharing volumetric overlays, the experimental group will report significantly lower Frustration and Mental Demand subscale ratings on the NASA-TLX compared to the control group due to LLM prioritizing and adaptive pacing.

6 Conclusion and Future Work

We presented the architecture of *XR Badminton*, a VR coaching framework that integrates phase-based biomechanical analysis, neuro-symbolic feedback, and accessibility-aware coaching. The framework defines a five-layer pipeline for stroke segmentation, biomechanical assessment, and adaptive feedback delivery under real-time standalone VR constraints. Our current prototype demonstrates the feasibility of the proposed architecture through real-time sensing, kinematic analysis, and physics-based shuttlecock simulation.

Future work will focus on evaluating long-term motor learning through longitudinal user studies, validating the framework with badminton players of varying skill levels, and extending the neuro-symbolic coaching engine with personalized biomechanical models that adapt to individual movement patterns. We also plan to investigate multimodal sensing using wearable IMUs and pressure-sensing insoles for full-body motion analysis, as well as exploring rehabilitation and inclusive coaching applications for users with diverse accessibility needs.

References

[1] T. Chen, K. Chen, X. Liu, P. Ke, and Z. Sun. BadminSense: Enabling fine-grained badminton stroke evaluation on a single smartwatch. In *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI)*, pp. 1–15. ACM, 2026. arXiv:2603.21825. 1

[2] A. Covaci, C.-C. Postelnicu, A. N. Panfir, and D. Talaba. A virtual reality simulator for basketball free-throw skills development. *Virtual Reality*, 22(3):183–200, 2018. doi: 10.1007/s10055-017-0320-5 1

[3] M. A. Guadagnoli and T. D. Lee. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36(2):212–224, 2004. doi: 10.3200/JMBR.36.2.212-224 2

[4] C. C. Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1):5–10, 2004. doi: 10.1016/j.ijforecast.2003.09.015 2

[5] H.-j. Park, J.-m. Shin, and G. Lee. BRIDGE: Empowering inclusive play and sports training for visually impaired users through embodiment-aware visualizations. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI)*, pp. 895–909. ACM, 2023. 2

[6] A. Sebernegg and P. Kán. VRTennis: Forehand training in virtual reality with rule-based motion analysis and multimodal feedback. In *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1085–1095. IEEE, 2025. doi: 10.1109/ISMAR67309.2025.00115 1

[7] C.-H. Wang, J. Yang, C. Liu, and S. Harrison. ViSTAR: Motion-to-language video synthesis and translation for sport activity coaching in augmented reality. In *Proc. ACM Symposium on User Interface Software and Technology (UIST)*, pp. 412–425. ACM, 2024. 2

[8] Q. Wang, S. Sabanović, and R. Gomez. Speech is silver, silence is golden: Analyzing micro-communication strategies between visually impaired runners and their guides. In *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI)*, pp. 1–18. ACM, 2024. doi: 10.1145/3613904.3642388 1, 2

[9] X. Zhao and S. Li. A biomechanical analysis of lower limb movement on the backcourt forehand clear stroke among badminton players of different levels. *Applied Bionics and Biomechanics*, 2019:7048345, 2019. doi: 10.1155/2019/7048345 1