TennisProfile: Reliability-Aware Groundstroke Profiling from Monocular 3D Human Recovery

Jinkun Peng[0009−0000−1168−395X]

Independent Researcher, Shenzhen, China
pengjk0@gmail.com

Abstract. Monocular 3D human recovery enables low-cost tennis motion analysis, yet stroke labels provide limited information about how individual strokes differ from expert technique. To address this limitation, we present TennisProfile, a reference-based protocol for forehand and backhand groundstrokes in controlled public RGB video. HMR2-recovered 3D body motion is aligned by three reviewed events and summarized into 11 rotation, timing, and wrist-motion features. Rather than reducing a stroke to a skill label, each profile reports the three largest standardized deviations from a stroke-specific expert reference, together with feature-specific interpretation levels. On THETIS, held-out expert profiles remained closer to the reference than beginner profiles (0.49 versus 0.76 SD mean absolute deviation; 0.0% versus 9.7% outside-range rate). Matched SportsPose checks identify a reliability hierarchy, with lower cross-source discrepancy for rotation-range features than peak-speed features (15.4% versus 43.2%). These checks support feature-level tennis motion review while assigning evidence-based interpretation levels to recovered cues.

Keywords: Monocular 3D human recovery · Tennis motion analysis · Expert-reference profiling · Reliability-aware motion profiling · Sports biomechanics

1 Introduction

Computer-vision motion capture is making low-cost human motion analysis feasible beyond laboratory motion-capture settings. Tennis is a compelling target: it is widely practiced, visually structured, and increasingly studied with machine-learning tools for performance analysis [26]. Existing tennis video analysis often reports stroke categories, action labels, or skill groups. These outputs are useful for indexing but provide limited information about how an individual stroke departs from expert movement patterns. We present TennisProfile, a reference-based protocol that converts monocular 3D recovery into feature-level groundstroke profiles and makes the interpretation level of each recovered movement variable explicit. Intended for clip-level review by coaches and players, the output prioritizes variables and stroke phases for inspection against the source video.

Tennis groundstrokes are event-structured rather than unstructured clips. Forehands and backhands combine clear contact events with fast whole-body coordination, and prior biomechanics studies describe them through body rotation, segmental timing, and upper-limb motion around ball contact [9, 11, 25]; stance, trunk rotation, and hip-mediated rotation further shape their mechanics [1, 3]. While marker-based motion capture can quantify these variables precisely under controlled conditions, its equipment, calibration, and operator demands severely limit repeated, low-cost deployment outside laboratory settings [20, 21].

In this setting, monocular human mesh recovery supplies the visual backbone. Models such as HMR2 estimate 3D body joints from RGB video [12, 17, 18, 22]. TennisProfile turns these trajectories into named movement variables and makes reliability part of the output: stability checks trigger review markers, and matched cross-source checks define feature-specific interpretation levels.

The design combines reference-population movement indices [5, 27] with tennis kinematics and contact-centered timing [21]. Reviewed HMR2 trajectories are converted to named variables and standardized against a THETIS expert reference. Each stroke reports its three largest deviations with reliability markers and video-review visualizations.

We evaluate held-out expert consistency, beginner–expert group differences, and profile stability on THETIS [13], then assess feature consistency on matched SportsPose RGB videos. The evaluation yields two main findings: held-out experts lie closer to the reference than beginner subjects (0.49 versus 0.76 SD mean absolute deviation; 0.0% versus 9.7% outside-range rate), and the rotation-range feature group receives the strongest cross-source support (15.4% versus 43.2% relative discrepancy compared with peak-speed features). Speed, fastest-motion timing, and wrist motion add complementary cues for reviewed stroke profiles.

Our contributions are threefold: a public-video protocol that compares recovered strokes with an expert reference; a compact output with top-three deviations, a five-phase map, and a rotation-range score; and a reliability study showing held-out experts closer than beginners and stronger matched-check support for rotation-range features.

2 Related Work

TennisProfile sits at the intersection of monocular pose recovery, markerless biomechanical evaluation, reference-based profiling, action quality assessment, and tennis groundstroke biomechanics. The missing link among these areas is a public-video protocol that represents each groundstroke as feature-level deviations from a stroke-specific expert reference while making the evidence level of each recovered feature explicit.

Monocular pose recovery. Human mesh recovery estimates articulated 3D bodies from image or video evidence [17, 18], commonly using the SMPL model [22]. 4DHumans/HMR2 reconstructs and tracks people from monocular video with a transformer-based model [12]. BioPose maps virtual mesh markers to a biomechanical skeleton through neural inverse kinematics [19]. In tennis, CalTennis finds joint-angle recovery more reliable than metric depth and foot-contact estimates across synchronized views [8], motivating feature-level reliability assessment.

Markerless biomechanical evaluation. Sport studies increasingly use deep learning pose estimation [2]. Markerless biomechanics systems translate pose estimates into movement variables: OpenCap estimates kinematics and dynamics from synchronized smartphone video [29], while OpenCapBench evaluates estimators through downstream biomechanical consequences [14]. Tennis-specific work also validates mechanical-work measurement during strokes [10]. These studies establish downstream variables as an appropriate evaluation target.

Reference-based profiling and action quality assessment. Clinical gait analysis compares individual kinematics with a normative population. The Gait Deviation Index summarizes reference distance [27], while the Gait Profile Score preserves component-level deviations [5]. Action quality assessment instead predicts scores or parses phases [4, 23, 28], and skilled-activity resources add expert commentary [15]. TennisProfile brings feature-specific reference comparison to short, non-cyclic groundstrokes.

Tennis biomechanics. Tennis biomechanics supplies the movement variables used in the profile. Forehand stance has been studied through upper-extremity kinetics [3], hip-generated trunk rotation is central to one- and two-handed backhands [1], and skill-level forehand studies identify distinct rotation timing, kinematic, and muscle-activation patterns [21, 24]. Wrist kinematics also differ between expert and novice backhands [6]. Reviews synthesize these mechanical factors for forehands and backhands [11, 20, 25]. Together, these findings guide the rotation, timing, and wrist-motion feature groups reported here.

3 Data, Annotation, and Pose-Recovery Protocol

Terminology. Throughout the paper, THETIS subjects appear in three cohort roles. The expert reference set $\mathcal{E}_{ref}$ contains $N_{ref} = 18$ expert-labeled subjects and defines stroke-specific means and standard deviations. A held-out expert set, $\mathcal{E}_{hold}$, contains $N_{hold} = 3$ expert-labeled subjects excluded from reference construction. The beginner set $\mathcal{B}$ contains $N_{beg} = 15$ subjects with the dataset-provided beginner label. These dataset-provided proficiency annotations define the expert-reference, held-out expert, and beginner cohorts used for kinematic profiling.

3.1 Dataset and Split

The controlled public-video setting comes from THETIS [13], which contains 990 groundstroke RGB clips across 55 subjects and multiple forehand and backhand action classes. The selected 640×480 AVI clips have nominal frame rates from

10 to 19 fps, typically 18 fps; clip-specific time bases are retained for filtering and frame-to-millisecond conversion (Section 4.1).

The roster was fixed before final feature recomputation. Eligibility required a reviewed flat forehand and one-handed backhand for the same subject and successful HMR2 feature extraction; derived feature values and group-separation results were not used for selection. Manual-review coverage yielded 36 retained subjects. Of the remaining 19, two were removed after review (a duplicate visual identity and a non-independent stroke pair), while 17 were not sampled for review; the supplementary audit accounts for all 55 subjects.

Reference construction, same-label hold-out checking, and beginner comparison use disjoint subject roles. The roster contains 36 reference strokes from $\mathcal{E}_{ref}$, 6 held-out expert strokes from $\mathcal{E}_{hold}$, and 30 beginner strokes from $\mathcal{B}$. The indoor gym contains all 15 beginners, 14 reference subjects, and one held-out expert; the smaller indoor room contains four reference subjects and two held-out experts. Each setting uses a fixed monocular camera placement, and all retained players were visually verified as right-handed. The release records the split and audit, event table, derived features, and analysis scripts. The held-out set measures same-group reference consistency for unseen THETIS subjects.

3.2 Stroke Events

All reference comparisons use the same functional movement window. One study annotator completed two video-guided passes to define forward-swing start, ball contact, and swing end. Forward-swing start is the first frame in which the racket arm and torso begin forward motion toward the ball after preparation. Ball contact is the frame that visually best matches racket-ball contact. Swing end is the first frame after contact where the main forward swing has visibly completed and the body enters follow-through recovery. The final event table records the second pass, and the stability analysis treats these anchors as explicit inputs to the profile.

3.3 Pose Recovery

Before any stroke variable is computed, each clip is converted into a consistent frame-indexed body representation. HMR2 supplies this representation using the official default checkpoint [12]. Full-frame inference retains the largest detected person as the foreground target in each frame. The resulting SMPL/3D joints [22] are stored with trial ID, frame index, joint name, 3D coordinates, and confidence; the feature set uses pelvis, hip, shoulder, elbow, wrist, and hand joints. Figure 1 visualizes the recovered pose.

3.4 External Pose Source

The external check separates stroke-level profiling from source-consistency testing. THETIS provides reviewed stroke types and expert-reference cohorts; SportsPose probes feature extraction under a different skeleton and coordinate convention. SportsPose contributes 20 outdoor tennis sequences from four subjects,

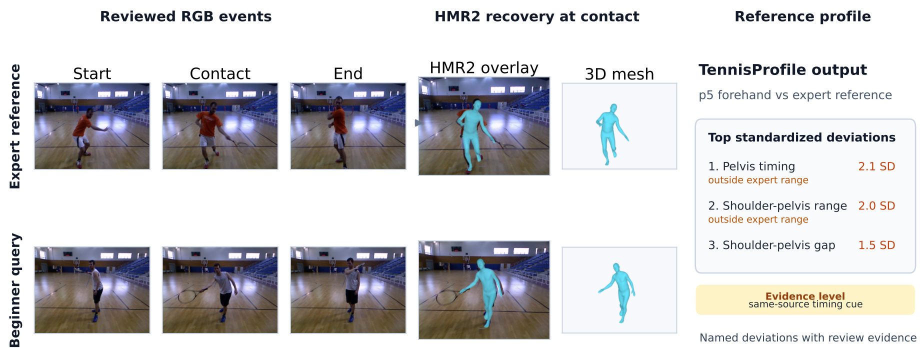

Fig. 1: TennisProfile overview. The figure contrasts an expert-reference forehand with a beginner-query forehand. Reviewed stroke events define the analysis window, HMR2 recovers SMPL body motion from RGB frames, and the beginner sequence is summarized against the stroke-specific expert reference. The visual panels show RGB event frames, a contact-frame mesh overlay, clean recovered-body views, and the resulting feature-level profile; quantitative features are computed from the full smoothed sequence.

each with 270 frames, 17 COCO-format 3D joints, 90 fps timing, and calibration metadata [16]. Each sequence is centered at maximum reconstructed wrist speed and summarized in a $\pm$45-frame window. For 12 right-view RGB videos, HMR2 and the independent SportsPose skeleton use identical 90 fps frame indices and windows. A separate 90-to-18/10 fps trajectory-resampling check measures sensitivity to THETIS-like temporal sampling.

Release. Code, reviewed events, derived features, and environment records are available in our public repository. Licensed source media, body-model assets, and third-party checkpoints remain under their original terms.

4 Method: Reference-Based Motion Profiling

HMR2 returns dense model-coordinate body motion, while stroke review needs named variables, contact-relative timing, and a reference scale that puts degrees, frames, and speed summaries into one profile. TennisProfile makes this conversion in three stages. Event-reviewed recovery aligns each sequence to forward-swing start, ball contact, and swing end. Expert-reference profiling converts the aligned trajectories into an 11-feature scalar profile and a five-phase joint-time map. Reliability-aware interpretation then attaches review markers and feature-specific interpretation levels. The scalar profile answers which variables depart most from the reference, while the phase map locates those departures within preparation, contact, and follow-through.

4.1 From Recovered Joints to Stroke Variables

Body-line signals. Groundstroke mechanics are commonly described through whole-body rotation and segmental timing, so the recovered joints are first reduced to pelvis and shoulder body-line orientations. At frame f, pelvis and shoulder lines are defined by $p_f = J^R_{\text{hip},f} - J^L_{\text{hip},f}$ and $s_f = J^R_{\text{shoulder},f} - J^L_{\text{shoulder},f}$, respectively. Projection onto the recovered SMPL horizontal x-z plane gives signed orientations $\theta^P_f = \text{atan2}(p_{f,z}, p_{f,x})$ and $\theta^S_f = \text{atan2}(s_{f,z}, s_{f,x})$. The orientation sequences are unwrapped, smoothed, and differentiated to obtain angular speed [7]. Filtering uses each clip's native frame rate with cutoff $\min(6\text{ Hz}, 0.45 \times \text{fps})$; short intervals use the highest order supported by standard padding. Reported ranges, angular-speed magnitudes, and phase-map angle changes are invariant to left–right axis sign.

Stroke-level features. The scalar profile compresses dense framewise trajectories into variables that answer three review questions: when the segments move fastest, how much body-line rotation accumulates, and how large peak movement speed is around contact. The timing group contains three fastest-motion times and two lags: shoulder–pelvis and wrist–shoulder. The rotation-range group contains pelvis, shoulder-line, and shoulder-relative-to-pelvis rotation ranges. The peak-speed group contains pelvis and shoulder-line peak angular speed and higher-speed-wrist peak linear speed. Timing is reported in frames and milliseconds relative to ball contact; negative values precede contact. These 11 features follow variables used in groundstroke biomechanics [1, 3, 6, 20, 21]. Table 1 summarizes their computation and interpretation level.

Body-line rotation captures global trunk mechanics, while model-space wrist speed adds distal upper-limb evidence. TennisProfile separates frame-to-frame wrist displacement from the predicted global camera component and selects the wrist with the larger peak displacement in the reviewed interval. Racket handedness comes from the independent visual audit; the higher-speed side remains a trajectory-defined descriptor.

Five-phase joint-time map. Knowing which variable differs is only part of stroke review; the timing of that departure is equally important. Absolute stroke durations vary across subjects, so the joint-time map anchors all strokes at ball contact and normalizes pre-contact and post-contact intervals separately. Let f_s, f_c, and f_e denote forward-swing start, ball contact, and swing end. Each frame f is mapped to

$$\rho(f) = \begin{cases} -\dfrac{f_c - f}{f_c - f_s}, & f_s \le f \le f_c, \\ \dfrac{f - f_c}{f_e - f_c}, & f_c < f \le f_e, \end{cases} \tag{1}$$

so that the reviewed interval spans $[-1, 1]$ with contact at zero. Five fixed windows partition this coordinate into early pre-contact $[-1, -0.6]$, late pre-contact

Table 1: Movement features used for expert reference profiling. Timing is reported relative to ball contact, with negative values before contact and positive values after contact.

Feature group	Reported features	Computation and interpretation
Timing	Pelvis, shoulder-line, and higher-speed-wrist fastest-motion time; shoulder–pelvis and wrist–shoulder gaps	Peak frame of absolute pelvis/shoulder angular speed or higher-speed-wrist linear speed. Values locate segmental sequencing around ball contact.
Rotation range	Pelvis rotation range; shoulder-line rotation range; shoulder-relative-to-pelvis rotation range	Projected pelvis and shoulder lines are unwrapped, smoothed, and summarized over the reviewed stroke interval. Values describe accumulated body-line rotation.
Peak speed	Pelvis and shoulder-line peak angular speed; higher-speed-wrist peak linear speed	Pelvis/shoulder entries are angular-speed magnitudes from angle derivatives. The wrist entry is frame-to-frame HMR2 model-space displacement after removing predicted global camera motion. Values summarize contact-window movement intensity.

$(-0.6, -0.2]$, contact $(-0.2, 0.2]$, early follow-through $(0.2, 0.6]$, and late follow-through $(0.6, 1]$. The resulting matrix has six movement rows (pelvis speed, shoulder speed, wrist speed, pelvis rotation, shoulder-line rotation, and shoulder-relative-to-pelvis rotation) and five phase columns. Speed signals are averaged within each phase. Rotation signals first convert each frame to its absolute angular change from the first reviewed frame, then average these values within phase. Beginner z-scores use the stroke-type-specific expert phase reference.

4.2 Expert Reference Profiling

Raw feature values mix degrees, frames, milliseconds, and model-space speeds. Following the reference-population logic used in movement-profile indices [5,27], expert reference profiling converts these quantities into a common stroke-specific scale. For each stroke type, $\mathcal{E}_{ref}$ defines the reference distribution. Given a stroke with feature value v, we compute

$$z = \frac{v - \mu_{\text{expert}}}{\sigma_{\text{expert}}}. \tag{2}$$

The profile also records expert min–max range membership. The final output lists the three features with the largest absolute standardized deviations for that stroke. The full ranked profile remains available for inspection.

A companion rotation-range score summarizes the pelvis, shoulder-line, and shoulder-relative-to-pelvis rotation-range z-scores:

$$S_{\text{range}} = \sqrt{\frac{z_{\text{pelvis}}^2 + z_{\text{shoulder}}^2 + z_{\text{relative}}^2}{3}}. \tag{3}$$

The score is expressed in standard-deviation units and summarizes reference closeness while the named feature profile preserves component-level information. Its cross-source behavior is evaluated in Section 5.7.

4.3 Reliability Checks and Interpretation Scope

Alongside the z-score profile, the output records the temporal conditions that shape video review. Each profile carries two automatic review markers. A marker is raised when fewer than four frames are available before ball contact or when the fastest body motion falls at the edge of the reviewed interval. These markers expose the conditions behind the feature summary and direct review to the relevant frames. Feature-specific interpretation levels are assigned from the cross-source consistency analysis in Section 5.7.

5 Evaluation

5.1 Evaluation Objectives

The evaluation follows the three parts of the protocol. Expert-reference characterization and held-out testing assess reference consistency beyond the construction set. Beginner–expert and phase-level comparisons locate departures in named variables and stroke phases. Stability and SportsPose checks examine event, smoothing, reference-set, pose-source, and sampling-rate sensitivity. Mann–Whitney U results are accompanied by Benjamini–Hochberg (BH) adjusted values and effect sizes.

5.2 Expert Reference Ranges

The expert references show two stroke organizations rather than one generic groundstroke template. In the one-handed-backhand reference, pelvis and wrist fastest-motion times fall after contact, while shoulder-line motion remains near contact; in contrast, the flat-forehand reference shifts pelvis and wrist peaks before contact and again keeps shoulder-line motion near contact. The mean offsets were 1.2, 2.2, and -0.7 frames for backhand pelvis, wrist, and shoulder-line peaks, compared with -3.2, -1.8, and 0.5 frames for forehands. This timing split matches

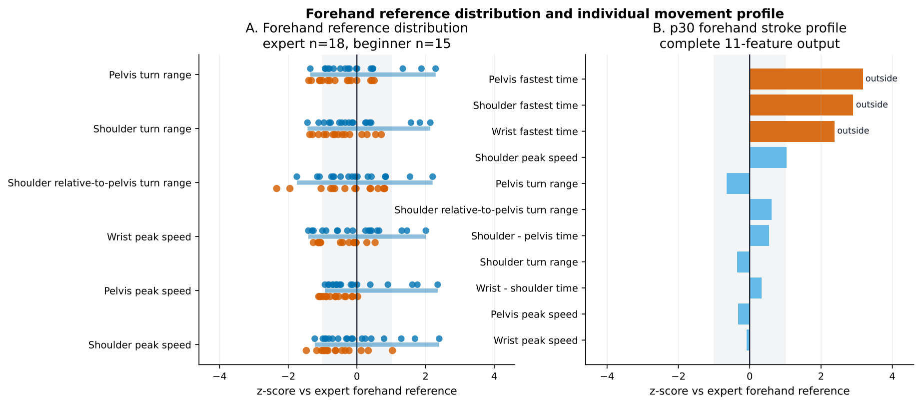

Fig. 2: Forehand reference distribution and individual movement profile. Left: group-level z-score distribution for six selected flat-forehand features, computed from 18 THETIS expert-labeled reference strokes and 15 beginner strokes. Blue points show expert reference trials, orange points show beginner trials, and the gray band marks expert mean ± 1 SD. Right: p30's flat forehand represented as the complete 11-feature profile against the same forehand reference; orange bars fall outside the expert minimum–maximum range, and blue bars remain inside it. Bar color encodes range membership; the top-three ranking is computed independently from absolute z-scores.

the mechanics of the two strokes: the flat-forehand reference concentrates recovered motion before contact, whereas the one-handed-backhand reference carries more peak motion through contact into early follow-through.

Rotation range reinforces the same stroke-specific separation. Flat forehands show larger pelvis, shoulder-line, and shoulder-relative-to-pelvis ranges than backhands by roughly 1.3–1.6×, with mean ranges of 77.9°, 128.2°, and 56.3°, compared with 57.9°, 83.6°, and 35.2°. Because timing and range both change with stroke type, TennisProfile builds separate references for forehands and backhands. A pooled-stroke baseline reduces the beginner–held-out mean-deviation gap from 0.27 to 0.16 SD and the outside-range gap from 9.7 to 3.0 percentage points (supplement), supporting this design.

Figure 2 applies the flat-forehand reference to p30. The shoulder and higher-speed wrist reach their fastest motion later than the expert forehand pattern, producing a forward-swing segmental-timing lag in the profile. This example illustrates how the reference converts raw timing offsets into a reviewable movement node with more detail than a global skill label.

5.3 Beginner–Expert Group Differences

Across beginner strokes, group separation is driven mainly by reduced movement speed along the forward-swing chain, not by a wholesale change in rotation amplitude. The pattern is clearest in forehands: pelvis peak angular speed and higher-speed-wrist peak linear speed fall below the expert mean in nearly every beginner subject, with shoulder angular speed following the same direction in most cases. Backhands show a similar reduction at the pelvis and wrist, although shoulder angular speed is less uniform. Rotation-range values rarely leave the expert range, which places the beginner–expert contrast in contact-directed movement intensity rather than in the overall amount of body-line rotation.

Subject-level tests support this biomechanical pattern. In flat forehands, 14 of 15 beginners show lower pelvis peak angular speed (one-sided binomial BH $q = 0.005$), and 13 of 15 show lower higher-speed-wrist peak linear speed ($q = 0.020$). In backhands, both variables are lower in 12 of 15 beginners ($q = 0.039$). The smallest U-test values occur for forehand pelvis speed and backhand wrist speed (uncorrected $p = 0.041$ and 0.024; Cliff's $\delta = -0.42$ and -0.47). No U-test remains below 0.05 after within-stroke BH correction (minimum $q = 0.452$ and 0.193); the full table is included in the release.

5.4 Key Phase-Level Deviations

Placed on a normalized stroke timeline, the same speed-centered difference becomes a phase-level movement pattern. Anchoring each stroke at ball contact allows the five-phase map to compare preparation, contact, and follow-through despite differences in absolute duration. Figure 3 reports this joint-time map for six movement features and both stroke types. For forehands, the strongest negative deviations in pelvis and shoulder angular speed, together with higher-speed-wrist linear speed, cluster around contact and early follow-through; the main departure is therefore localized to the contact-centered speed window. Backhands exhibit a distinct kinematic divergence: lower pelvis angular speed after contact is paired with elevated pelvis rotation before contact and elevated shoulder-line rotation through follow-through.

These phase-localized patterns connect the scalar feature profile to reviewable movement nodes. In forehands, the speed deficit around contact marks the frames where lower forward-swing movement intensity is most visible in video review. In backhands, elevated pelvis rotation before contact indicates earlier accumulation of body-line rotation because rotation is measured as angular change from the reviewed start frame; in stroke review, this corresponds to an earlier pelvis-opening pattern before contact. Elevated shoulder-line rotation during follow-through identifies a later rotation node. The scalar profile ranks which variables differ most; the joint-time map turns that ranking into a phase-specific review plan.

5.5 Held-Out Expert Check

After the reference statistics are fixed, the six expert-labeled strokes in $\mathcal{E}_{hold}$ test whether the reference extends beyond the subjects used to build it. This check compares unseen expert-labeled strokes with beginner strokes under the same reference scale. Across all features and strokes, $\mathcal{E}_{hold}$ has a mean absolute deviation of 0.49 standard deviations from the reference, compared with 0.76

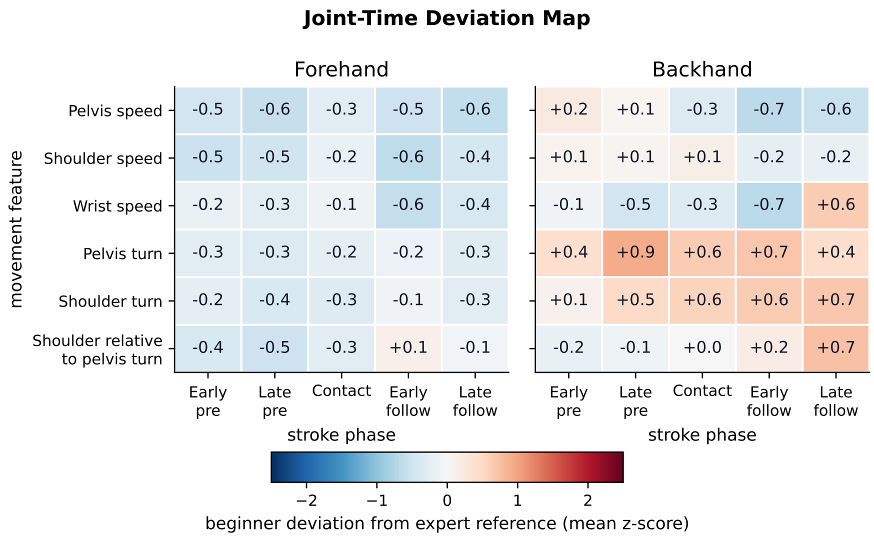

Fig. 3: Five-phase joint-time deviation map for 18 THETIS expert reference subjects and 15 beginner subjects. Stroke windows are normalized around ball contact and divided into early pre-contact, late pre-contact, contact window, early follow-through, and late follow-through phases. Each cell reports the beginner-group mean z-score relative to the stroke-specific expert phase reference; blue indicates lower beginner values and red indicates higher beginner values.

for $\mathcal{B}$. No $\mathcal{E}_{hold}$ feature values fall outside the expert minimum–maximum range, compared with 9.7% for $\mathcal{B}$.

The separation also holds within each stroke type, which matters because the reference is stroke-specific. For backhands, $\mathcal{E}_{hold}$ shows a mean absolute deviation of 0.51, compared with 0.77 for $\mathcal{B}$. For flat forehands, $\mathcal{E}_{hold}$ shows a mean absolute deviation of 0.47, compared with 0.76 for $\mathcal{B}$. These stroke-specific results show that the reference captures expert-group motion beyond $\mathcal{E}_{ref}$.

The rotation-range profile score gives the same fixed-split ordering: after averaging forehand and backhand scores by subject, $\mathcal{E}_{hold}$ has a mean of 0.52, compared with 0.89 for $\mathcal{B}$. Exhaustive reassignment shows the effect of capture composition: mean-deviation and outside-range ordering is retained in 53.4% and 67.5% of all 1,330 mixed-setting splits, increasing to 80.2% and 90.1% across 455 gym-only splits. The score provides an analogue to reference-index summaries [5, 27]; complete distributions are reported in the supplement.

5.6 Event Review and Stability Checks

Event anchors, smoothing, and reference membership all shape the reported profile, so the stability checks perturb each of them in turn. Figure 4 illustrates the role of event review: the reviewed forward-swing start and swing-end frames

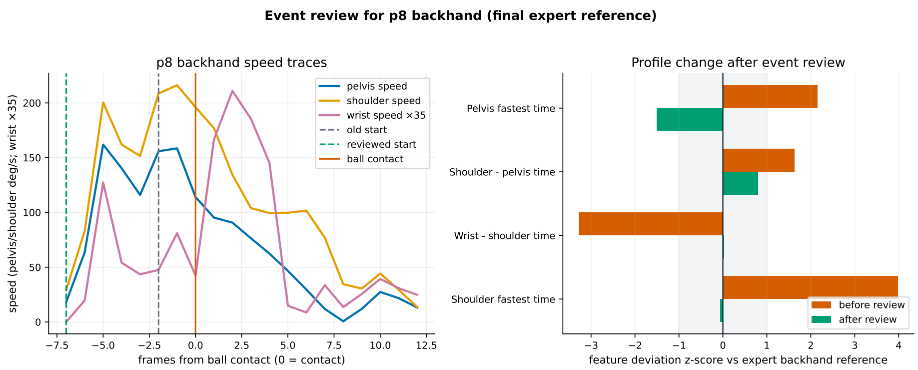

Fig. 4: Effect of reviewing p8 backhand event labels. The left panel shows recovered speed traces aligned to ball contact; the wrist trace is rescaled for visualization. The right panel compares feature-level z-score profiles before and after reviewing the forward-swing start and swing-end frames against the 18-subject expert backhand reference.

Table 2: Profile stability under event, smoothing, and reference perturbations. Top-three overlap is averaged within each row; – indicates that the rotation-range score was not recomputed.

Perturbation	Top-1 retained	Top-3 overlap	Median $\lvert\Delta S_{\mathrm{range}}\rvert$ (SD)
Contact ±1/±2 frames	77.5%	86.9%	–
Swing end ±2/±4 frames	90.8%	95.6%	0.016
5 Hz maximum cutoff	80.0%	92.2%	0.008
4 Hz maximum cutoff	70.0%	82.2%	0.016
Leave-one-expert-out	86.7%	94.4%	–

align the feature summary with the inspected stroke motion and direct attention to the frames that define the reported profile.

Table 2 summarizes perturbations of event labels, smoothing, and reference-set membership. The key pattern is that the ranked set of three is more stable than the single top-ranked feature. Event-label perturbations preserve high top-three overlap, which motivates reporting a ranked profile instead of relying on a single-feature output.

Algorithmic and reference-set adjustments show the same pattern. A 5 Hz maximum cutoff retains 92.2% top-three overlap, and the more conservative 4 Hz setting still retains 82.2% overlap while changing the rotation-range score by 0.016 SD. Leave-one-expert-out references retain 94.4% top-three overlap. Together, the checks favor the ranked top-three summary and the rotation-range score as the stable outputs for video review.

Table 3: Matched 90 fps SportsPose–HMR2 consistency across 12 RGB trials from four subjects; rotation range pools three body-line features.

Feature group	Matched-video result	Interpretation in this study
Rotation range	13.9°; 15.4% relative difference	Lowest matched relative discrepancy in this check
Peak speed	43.2% relative difference	Same-source reviewed cue
Fastest-motion timing	66.7 ms, about 6 frames at 90 fps	Same-source timing cue

5.7 External Consistency Check on SportsPose

SportsPose introduces a different skeleton, camera geometry, and frame rate while keeping the comparison at the level of derived movement features. Across all 20 SportsPose sequences, the external-source protocol in Section 3 yields finite features with interior anchors. The matched RGB subset then measures how feature groups compare against an independent 3D skeleton on the same frame windows.

Under this matched comparison (Table 3), rotation-range features show the strongest agreement: median absolute differences are 13.6° for pelvis, 17.9° for shoulder-line, and 8.5° for shoulder-relative-to-pelvis range. Their pooled median relative difference is 15.4%, compared with 43.2% for peak speed; timing differs by 66.7 ms. Because both sources use the same 90 fps frames, these values do not mix temporal sampling rates. Resampling the independent trajectories to 18/10 fps changes rotation range by 0.4%/0.7%, peak speed by 6.1%/15.8%, and fastest-motion timing by 11.1/22.2 ms (medians).

6 Discussion

The THETIS study indicates that the reference captures reviewable stroke structure beyond its construction subjects. Held-out experts remain close, whereas beginner strokes move farther away (Section 5.5). Beginner deviations concentrate in named, phase-localized patterns, including reduced contact-window rotational speed and earlier pelvis opening in reviewed backhands. The profile therefore links group separation to movement patterns that can be checked in the source video.

The SportsPose check clarifies feature reliability outside the original pose source. Rotation range averages body-line motion across the reviewed interval, whereas peak speed and fastest-motion timing depend on local extrema exposed to frame-level ambiguity, blur, self-occlusion, and smoothing. Two evidence axes therefore emerge: speed features provide stronger within-THETIS group separation, while rotation range has stronger cross-source and sampling-rate support. This hierarchy agrees with CalTennis, whose multi-view benchmark finds joint-angle recovery more reliable than depth and foot-contact estimates [8].

The output form follows from this hierarchy. Event, smoothing, and reference-set perturbations make the single top-ranked feature less stable than the top-three set or the rotation-range score (Section 5.6). The scalar profile ranks deviations, while the five-phase map places them into preparation, contact, and follow-through. Compared with score-based action quality assessment and expert-commentary datasets [4,15,23,28], the result exposes variables and phases for review alongside the original video.

7 Limitations and Future Work

Within reviewed THETIS clips, TennisProfile supports same-source reference behavior and a feature-level reliability hierarchy. Its reference describes the selected cohort rather than a population norm: expert composition and the two indoor capture settings affect exhaustive hold-out results. Event anchors currently come from one annotator's two-pass review, and the SportsPose comparison measures feature agreement rather than marker-based kinematic error. These boundaries motivate three extensions.

Ambient generalization. Future work should test outdoor smartphone capture, wider viewpoints, and larger held-out cohorts; CalTennis offers a testbed for inter-view feature disagreement [8].

Temporal automation. Automatic event alignment and explicit wrist-speed scale normalization would strengthen high-dynamic timing and speed analysis.

Biomechanical dimensionality. Body-line rotation and model-space wrist motion provide reviewable nodes, but finer wrist–racket interaction, forearm rotation, contact mechanics, and injury-risk measures require anatomical models, dedicated labels, or additional sensing. BioPose-style inverse kinematics offers one route toward richer anatomical features [19].

8 Conclusion

TennisProfile frames tennis video analysis as expert-referenced, feature-level motion review rather than a stroke label. On THETIS, the profiles separate held-out expert strokes from beginner strokes and localize the differences into named rotation, timing, and wrist-motion cues. Matched SportsPose checks add an external consistency layer, showing that full-interval body-rotation ranges carry the strongest cross-source support, while peak speed and timing remain reviewed profile cues. By reporting movement deviations together with feature-level evidence, TennisProfile provides a reusable basis for localized motion review from public sports video.

References

1. Akutagawa, S., Kojima, T.: Trunk rotation torques through the hip joints during the one- and two-handed backhand tennis strokes. Journal of Sports Sciences **23**(8), 781–793 (Jan 2005). https://doi.org/10.1080/02640410400021609, http://dx.doi.org/10.1080/02640410400021609
2. Aulton, C., Wakili, L., Strafford, B.W., Davids, K., Chiu, C.Y.: The application of deep learning human pose estimation in sport: A systematic review. Sports Medicine - Open **11**(1) (Dec 2025). https://doi.org/10.1186/s40798-025-00953-3, http://dx.doi.org/10.1186/s40798-025-00953-3
3. Bahamonde, R.E., Knudson, D.: Kinetics of the upper extremity in the open and square stance tennis forehand. Journal of Science and Medicine in Sport **6**(1), 88–101 (Mar 2003). https://doi.org/10.1016/S1440-2440(03)80012-9, http://dx.doi.org/10.1016/S1440-2440(03)80012-9
4. Bai, Y., Zhou, D., Zhang, S., Wang, J., Ding, E., Guan, Y., Long, Y., Wang, J.: Action Quality Assessment with Temporal Parsing Transformer, pp. 422–438. Springer Nature Switzerland (2022). https://doi.org/10.1007/978-3-031-19772-7_25, http://dx.doi.org/10.1007/978-3-031-19772-7_25
5. Baker, R., McGinley, J.L., Schwartz, M.H., Beynon, S., Rozumalski, A., Graham, H.K., Tirosh, O.: The gait profile score and movement analysis profile. Gait & Posture **30**(3), 265–269 (Oct 2009). https://doi.org/10.1016/j.gaitpost.2009.05.020, https://doi.org/10.1016/j.gaitpost.2009.05.020
6. Blackwell, J.R., Cole, K.J.: Wrist kinematics differ in expert and novice tennis players performing the backhand stroke: Implications for tennis elbow. Journal of Biomechanics **27**(5), 509–516 (May 1994). https://doi.org/10.1016/0021-9290(94)90062-0, http://dx.doi.org/10.1016/0021-9290(94)90062-0
7. Crenna, F., Rossi, G.B., Berardengo, M.: Filtering biomechanical signals in movement analysis. Sensors **21**(13), 4580 (Jul 2021). https://doi.org/10.3390/s21134580, https://doi.org/10.3390/s21134580
8. Demler, I., Xie, X., Werner, B., Szczuka, A., Perona, P.: CalTennis: Large multi-view tennis video dataset and benchmark of monocular-to-3d pose estimation (2026), https://arxiv.org/abs/2606.20542
9. Elliott, B.: Biomechanics and tennis. British Journal of Sports Medicine **40**(5), 392–396 (Apr 2006). https://doi.org/10.1136/bjsm.2005.023150, http://dx.doi.org/10.1136/bjsm.2005.023150
10. Emmerson, J., Needham, L., Williams, S., Colyer, S.: Evaluation of a markerless motion capture system for measuring mechanical work during tennis strokes. Journal of Sports Sciences **44**(10), 1348–1360 (2026). https://doi.org/10.1080/02640414.2025.2555096
11. Genevois, C., Reid, M., Rogowski, I., Crespo, M.: Performance factors related to the different tennis backhand groundstrokes: A review. Journal of Sports Science and Medicine **14**(1), 194–202 (2015), https://www.jssm.org/jssm-14-194.xml
12. Goel, S., Pavlakos, G., Rajasegaran, J., Kanazawa, A., Malik, J.: Humans in 4d: Reconstructing and tracking humans with transformers. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 14737–14748. IEEE (Oct 2023). https://doi.org/10.1109/iccv51070.2023.01358, http://dx.doi.org/10.1109/ICCV51070.2023.01358
13. Gourgari, S., Goudelis, G., Karpouzis, K., Kollias, S.: Thetis: Three dimensional tennis shots a human action dataset. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 676–681. IEEE (Jun 2013). https://doi.org/10.1109/cvprw.2013.102, http://dx.doi.org/10.1109/cvprw.2013.102

14. Gozlan, Y., Falisse, A., Uhlrich, S., Gatti, A., Black, M., Hicks, J., Delp, S., Chaudhari, A.: Opencapbench: A benchmark to bridge pose estimation and biomechanics. In: 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 4056–4065. IEEE (Feb 2025). https://doi.org/10.1109/wacv61041.2025.00399, http://dx.doi.org/10.1109/WACV61041.2025.00399
15. Grauman, K., Westbury, A., Torresani, L., Kitani, K., Malik, J., Afouras, T., Ashutosh, K., Baiyya, V., Bansal, S., Boote, B., et al.: Ego-exo4d: Understanding skilled human activity from first- and third-person perspectives. arXiv preprint arXiv:2311.18259 (2024)
16. Ingwersen, C.K., Mikkelstrup, C., Jensen, J.N., Hannemose, M.R., Dahl, A.B.: Sportspose: A dynamic 3d sports pose dataset. In: Proceedings of the IEEE/CVF International Workshop on Computer Vision in Sports (2023)
17. Kanazawa, A., Black, M.J., Jacobs, D.W., Malik, J.: End-to-end recovery of human shape and pose. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7122–7131. IEEE (Jun 2018). https://doi.org/10.1109/cvpr.2018.00744, http://dx.doi.org/10.1109/CVPR.2018.00744
18. Kocabas, M., Athanasiou, N., Black, M.J.: VIBE: Video inference for human body pose and shape estimation. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5252–5262. IEEE (Jun 2020). https://doi.org/10.1109/cvpr42600.2020.00530, http://dx.doi.org/10.1109/CVPR42600.2020.00530
19. Koleini, F., Saleem, M.U., Wang, P., Xue, H., Helmy, A., Fenwick, A.: BioPose: Biomechanically-accurate 3d pose estimation from monocular videos. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 6330–6339 (2025). https://doi.org/10.1109/WACV61041.2025.00617, https://doi.org/10.1109/WACV61041.2025.00617
20. Lambrich, J., Muehlbauer, T.: Biomechanical analyses of different serve and groundstroke techniques in tennis: A systematic scoping review. PLOS ONE **18**(8), e0290320 (Aug 2023). https://doi.org/10.1371/journal.pone.0290320, http://dx.doi.org/10.1371/journal.pone.0290320
21. Landlinger, J., Lindinger, S., Stöggl, T., Wagner, H., Müller, E.: Key factors and timing patterns in the tennis forehand of different skill levels. Journal of Sports Science and Medicine **9**(4), 643–651 (2010)
22. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: a skinned multi-person linear model. ACM Transactions on Graphics **34**(6), 1–16 (Oct 2015). https://doi.org/10.1145/2816795.2818013, http://dx.doi.org/10.1145/2816795.2818013
23. Parmar, P., Morris, B.: Action quality assessment across multiple actions. In: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 1468–1476. IEEE (Jan 2019). https://doi.org/10.1109/wacv.2019.00161, http://dx.doi.org/10.1109/wacv.2019.00161
24. Pedro, B., Cabral, S., João, F., Lei, A.M.K., Veloso, A.P.: Kinematic and muscle activation differences between high-performance and intermediate tennis players during the forehand drive. Sensors **26**(7), 2244 (Apr 2026). https://doi.org/10.3390/s26072244, http://dx.doi.org/10.3390/s26072244
25. Reid, M., Elliott, B., Crespo, M.: Mechanics and learning practices associated with the tennis forehand: A review. Journal of Sports Science and Medicine **12**(2), 225–231 (2013), https://pmc.ncbi.nlm.nih.gov/articles/PMC3761830/
26. Sampaio, T., Oliveira, J.P., Marinho, D.A., Neiva, H.P., Morais, J.E.: Applications of machine learning to optimize tennis performance: A systematic review. Applied

Sciences **14**(13), 5517 (Jun 2024). https://doi.org/10.3390/app14135517, http://dx.doi.org/10.3390/app14135517
27. Schwartz, M.H., Rozumalski, A.: The gait deviation index: A new comprehensive index of gait pathology. Gait & Posture **28**(3), 351–357 (Oct 2008). https://doi.org/10.1016/j.gaitpost.2008.05.001, https://doi.org/10.1016/j.gaitpost.2008.05.001
28. Tang, Y., Ni, Z., Zhou, J., Zhang, D., Lu, J., Wu, Y., Zhou, J.: Uncertainty-aware score distribution learning for action quality assessment. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9836–9845. IEEE (Jun 2020). https://doi.org/10.1109/cvpr42600.2020.00986, http://dx.doi.org/10.1109/cvpr42600.2020.00986
29. Uhlrich, S.D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A.S., Hicks, J.L., Delp, S.L.: Opencap: Human movement dynamics from smartphone videos. PLOS Computational Biology **19**(10), e1011462 (Oct 2023). https://doi.org/10.1371/journal.pcbi.1011462, http://dx.doi.org/10.1371/journal.pcbi.1011462