

Switch-JustDance: Benchmarking Whole-Body Motion Tracking Controllers Using a Commercial Console Game

Supplementary Material

A. Implementation Details for Real-Time Streaming of Smoothed gameplay (Smo)

To emulate the online gameplay experience of *Just Dance*, we deploy a streaming mode for Smoothed Gameplay (Smo). In this mode, a live MoCap module extracts human motion directly from the game screen online and continuously transmits the retargeted robot references to the motion tracking controller. To mitigate delay and throughput limitations induced by a computationally heavy GVHMR inference, we adopt a multi-threaded architecture in which a secondary thread interpolates robot frames produced by the retargeting module.

The main thread executes the entire perception and re-targeting pipeline, which consists of the Streaming Module (frame capture), the MoCap Module (GVHMR), and the Retarget Module (GMR). Due to the heavy computation required by GVHMR, this pipeline generates an updated target robot pose roughly every 200ms (5Hz).

A separate interpolation thread smooths the discrete target poses produced by the main thread and sends them to the robot controller via UDP. Since tracking a 5Hz target pose would result in a jerky motion, this thread maintains a buffer of consecutive keyframes (f_{k-1} and f_k). Once two frames are available, the interpolator computes intermediate frames ($f_{0...n}^i$, integer n is the number of interpolation frames that can be set by the user) using Linear Interpolation (LERP) for positions and Spherical Linear Interpolation (SLERP) for rotations. This process introduces a constant delay of roughly 0.2s to maintain consistency but effectively upsamples the 5Hz input into a smooth 10–15Hz trajectory. This upsampled stream is then fed into the robot controller, enabling stable and continuous motion execution.

B. Broader Embodiment and Generalization

By design, our framework is agnostic to different robot morphologies. In terms of scoring, the Just Dance system is designed for a global population and evaluates motion across diverse body sizes and proportions, making it largely morphology-agnostic. For retargeting, our pipeline builds on the open-sourced GMR framework, which has been successfully applied to multiple humanoid morphologies, including Booster and H1. To further validate generality, we apply our pipeline to a HOVER [15]-controlled H1 robot in simulation, where the robot successfully follows the reference dance trajectory. A representative snapshot is shown in Fig. B1.

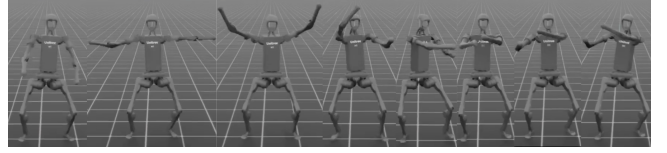


Figure B1. Snapshot of our pipeline being applied to a HOVER [15]-controlled H1 robot in simulation, dancing to *Old Town Road*; from 0:35 to 0:39.

Song	Game Edition	Difficulty	Length (s)
Old Town Road	JustDance 2020	Easy (lvl 1)	161
Heart Of Glass	JustDance 1	Easy (lvl 2)	216
Unstoppable	JustDance 2025 Edition	Easy (lvl 2)	204
Padam Padam	JustDance 2025 Edition	Hard (lvl 3)	149
Pink Venom	JustDance 2025 Edition	Hard (lvl 3)	178

Table C1. Summary of selected Just Dance tracks with game editions, difficulty, and durations. (lvl #) is the in-game difficulty. we group lvl 1–2 as Easy and 3–4 as Hard.

C. Song Selection for Benchmark Evaluation

For the benchmark, we select songs using two criteria: (1) the absence of hand–floor interaction and (2) minimal in-place turning. Songs that involve dancers placing their hands on the ground are excluded, as the Joy-Con controller is mounted on the robot’s hand, making such interactions infeasible. We additionally exclude songs that require significant in-place turning. Unlike human dancers who wear soft, rounded footwear, the robot’s feet are not designed for continuous pivoting, and these motions demand excessive ankle torque. In practice, we observed that in-place turning gestures are frequently ignored due to hardware limitations.

Applying these criteria, we select five songs for our benchmark, as summarized in Table C1.

D. IRB and Human Subjects Approval

This research involves human-motion data extracted from gameplay footage. All procedures involving human data were reviewed and approved by the corresponding Institutional Review Board (IRB). The study was conducted in accordance with applicable ethical guidelines for human subjects research. Consent was obtained where required by the approved protocol.