TouchWalker: Real-Time Avatar Locomotion from Touchscreen Finger Walking

Geuntae Park[1], Jiwon Yi[1], Taehyun Rhee[2], Kwanguk Kim[1], Yoonsang Lee[1]

Hanyang University[1] University of Melbourne[2]

Motivation

- On touch screen devices, a **virtual joystick** is commonly used
- May not fully leverage the **expressive potential** of touchscreen interaction
- **Lack of tactile feedback** may limit embodiment & immersion

Our Approach

Exploring Finger Walking because…

- Evokes **foot-ground contact sensation** with bi-finger rhythmic gestures
- Provides rich **temporal & spatial cues** for avatar control
- Enables **embodiment & tactile interaction** on touchscreens

TouchWalker

A **real-time system** that enables expressive **full-body avatar control** through **finger-walking** gestures on a touchscreen

TouchWalker

- Reconceptualize **Finger Walking** as a **continuous & expressive input modality**

- Key Components
    - **TouchWalker-MotionNet:** real-time neural motion generator
    - **TouchWalker-UI:** per-frame avatar control UI on touchscreens

- Evaluation
    - **User study** comparing **TouchWalker** vs **virtual joystick baseline**

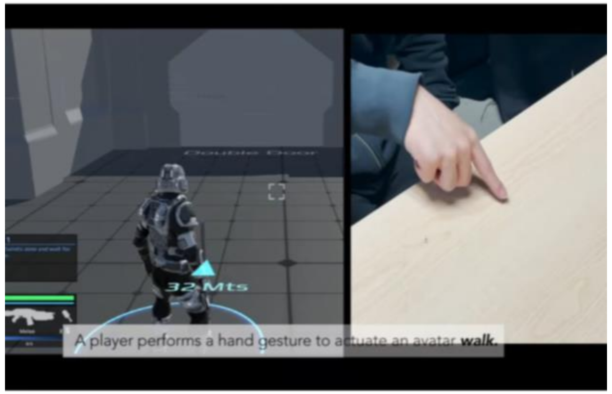
Related Work

Gesture & Touch-based Locomotion

- Mostly rely on **symbolic gestures** or **finger walking**
- Limited to the playback of **predefined motions**

→ Lack of **real-time, responsive motion generation**

[Lockwood and Singh. 2012]



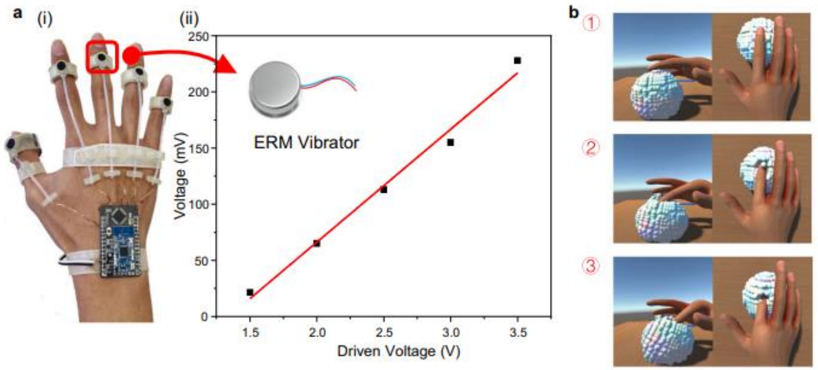
[Hung et al. 2022]

[Hung et al. 2024]

Related Work

Embodied Interaction

- Prior work shows **body-based** & **tactile interactions**
- Enhance **immersion** and **engagement** in VR



[Sun et al. 2022]

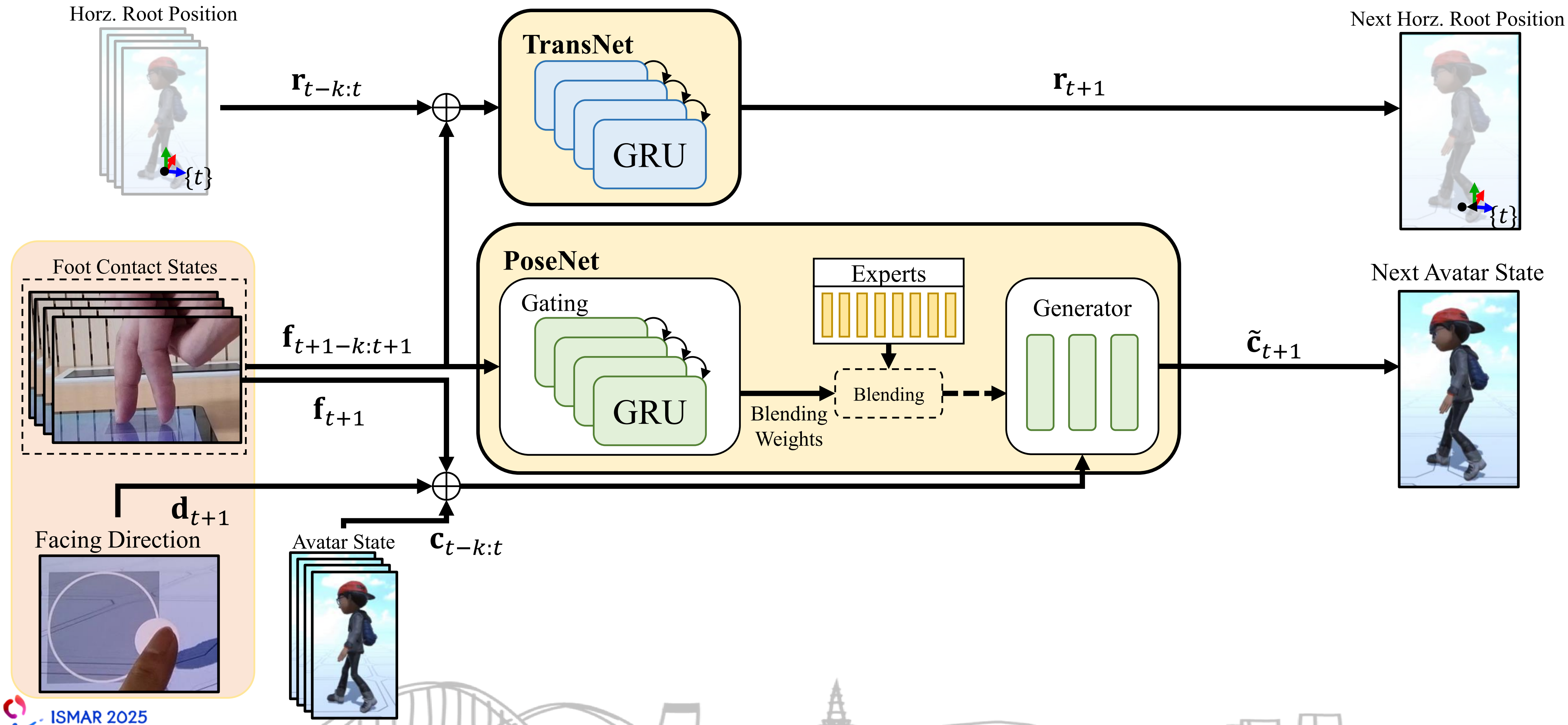
[Li et al. 2024]

Research Gap

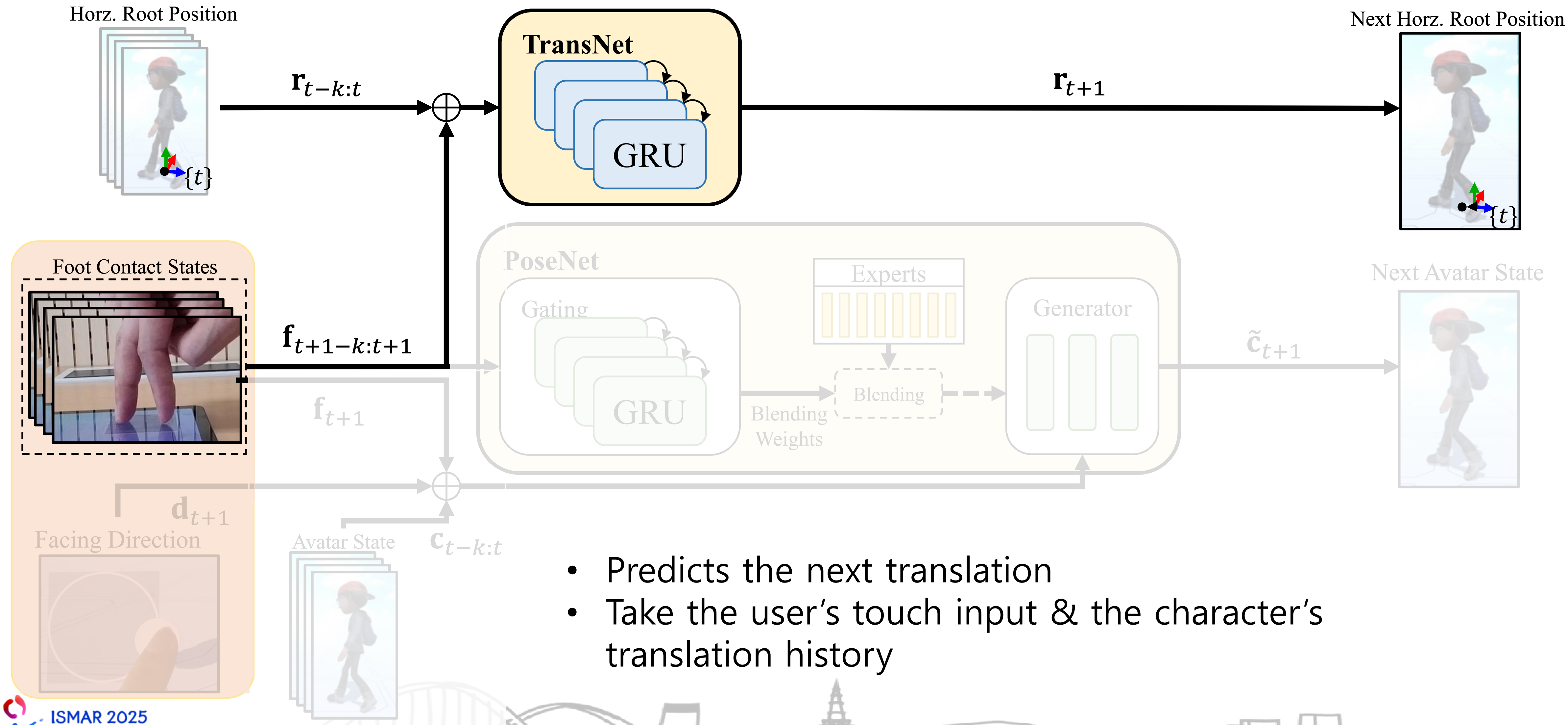
- No prior work combines **touch-based finger walking** with **real-time full-body motion generation**


LOBBY
STAGE1
STAGE2
STAGE3

TouchWalker-MotionNet

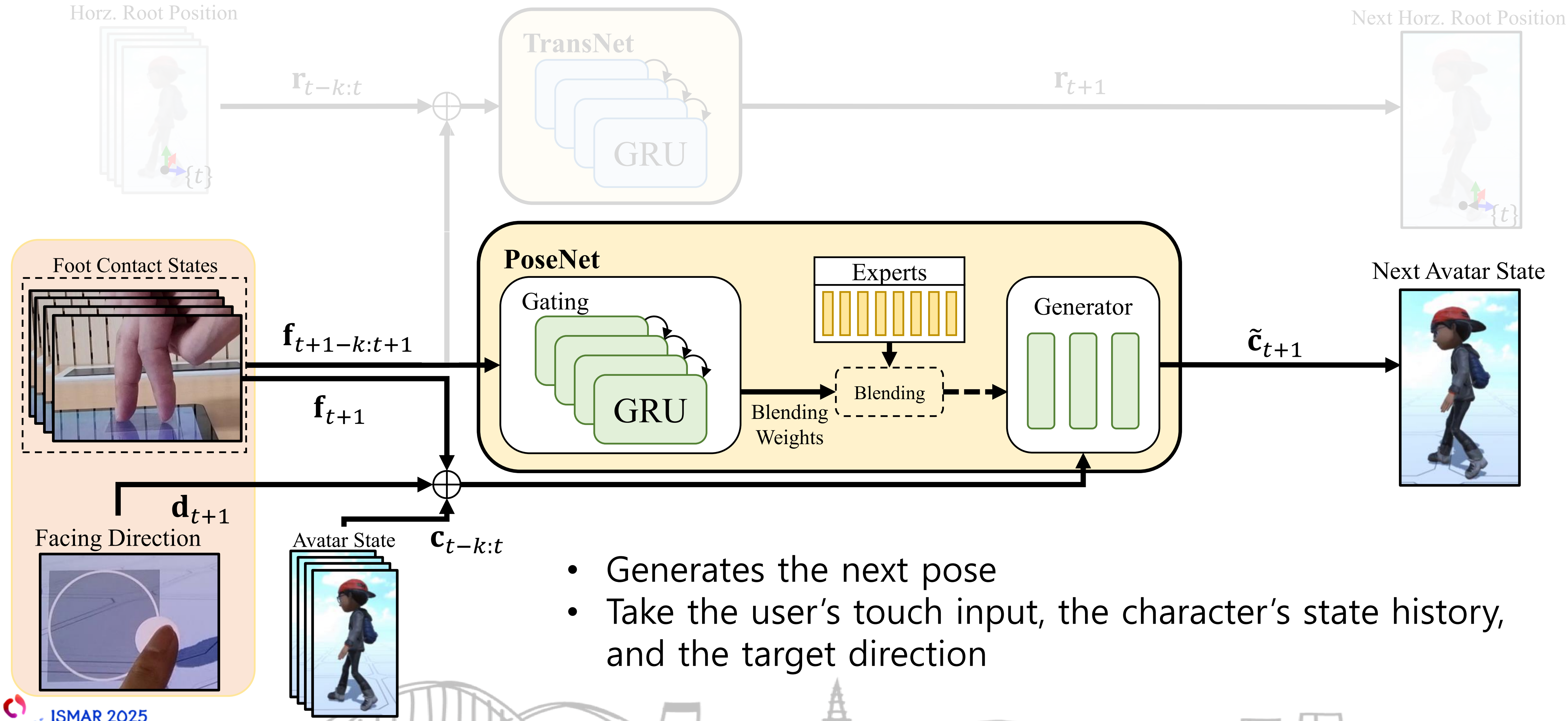


TouchWalker-MotionNet



- Predicts the next translation
- Take the user's touch input & the character's translation history

TouchWalker-MotionNet



- Generates the next pose
- Take the user's touch input, the character's state history, and the target direction

Loss Function

$$\boldsymbol{L} = \boldsymbol{w_1} \cdot \boldsymbol{L_{rec}} + \boldsymbol{w_2} \cdot \boldsymbol{L_{\mathrm{FK}}} + \boldsymbol{w_3} \cdot \boldsymbol{L_{dir}} + \boldsymbol{w_4} \cdot \boldsymbol{L_{ct}} + \boldsymbol{w_5} \cdot \boldsymbol{L_{\mathrm{ct}\_trnas}}$$

pose　　　　Foot contact

$\boldsymbol{L_{rec}}$: Reconstruction loss

$\boldsymbol{L_{FK}}$: Forward Kinematics (FK) loss

$\boldsymbol{L_{dir}}$: Direction loss

$\boldsymbol{L_{ct}}$: Contact loss

$\boldsymbol{L_{ct\_trans}}$: Contact transform loss

*Details in paper

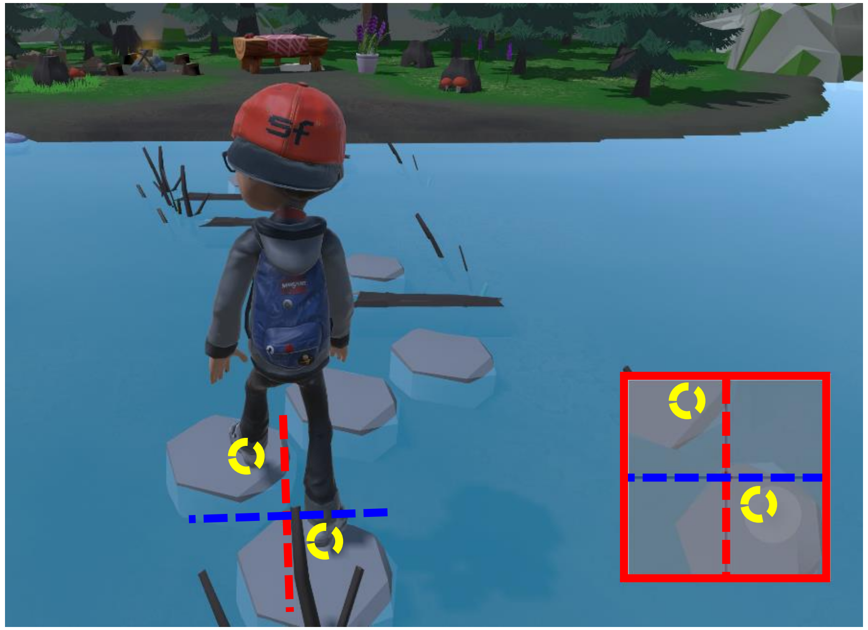
Training

- Dataset: LaFAN1
  - Input
    - Contacted foot state (calc. by height threshold)
  - Output
    - Next root position relative to the current character coordination
    - 6D rotation representation (relative to the parent joint)

- Training time: 3 hours (Approximately)
- Device
  - CPU: AMD Ryzen 7 2700X
  - GPU: Nvidia GeForce RTX 3090

TouchWalker-UI

- Touch region displays the ground beneath the avatar
- Touched position → relative to the avatar's position & orientation on the ground
- Vertical line: avatar's forward direction
- Horizontal line: for lateral alignment

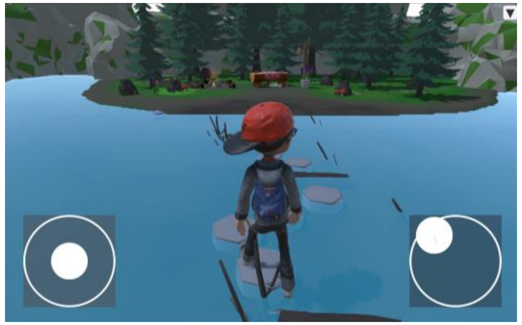


User Study

Conditions

- **TouchWalker (TW):**
  - TouchWalker-UI + facing direction joystick

- **Virtual Joystick (VJ)** *(baseline)* :
  - Dual-analog virtual joystick



Hypotheses

- **H1:** TouchWalker → **higher embodiment, enjoyment, and immersion** then VJ
- **H2:** Differences btwn. TW & VJ may depend on participants' **prior joystick experience**

Task Design

Task 1 (Multi-stage Navigation)

- A task of passing through a course with 5 stages

- **Stage 1 (Hole Avoidance):** Holes scattered along the path
- **Stage 2 (Speed Control):** Strong wind, speed-limit part
- **Stage 3 (Narrow Path):** Narrow and winding path
- **Stage 4 (Foot Buttons):** Five foot buttons
- **Stage 5 (Rolling Obstacles):** Rolling ball obstacles

Task Design

Task 2 (Stepping Stones)

- Cross the river without touching the water
- Stones, wooden planks

Participant

- Participants: 14 (9 males, 5 females)
- Age Range: 18~39 years

- Post-task Questionnaire: 11 items
  - Embodiment Questionnaire (EQ)
  - Intrinsic Motivation Inventory (IMI)
  - Immersive Experience Questionnaire (IEQ)

- Device: Android tablet (Galaxy Tab S7+)

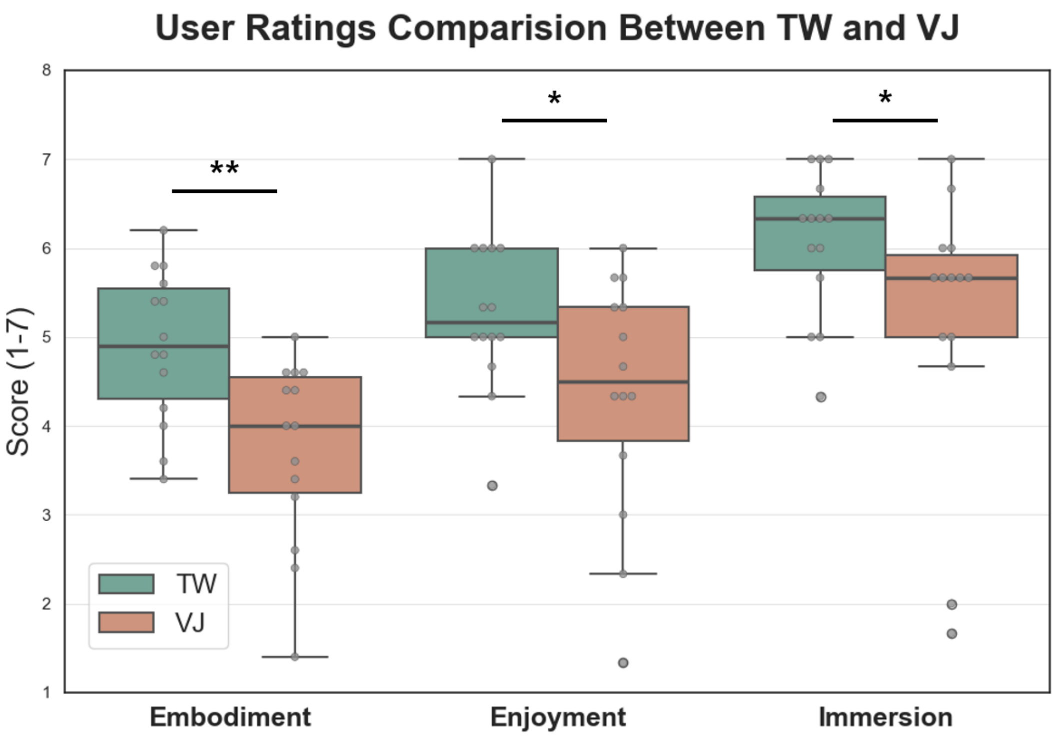
Subjective Evaluation

H1: *TouchWalker → **higher embodiment, enjoyment, and immersion** then VJ*

Findings

- **TW > VJ (significantly)** in all three measures (Embodiment, Enjoyment, Immersion)
- TW shows higher medians & narrower ranges.

→ H1 supported



Subjective Evaluation

H2: *Differences btwn. TW & VJ may depend on participants' **prior joystick experience***

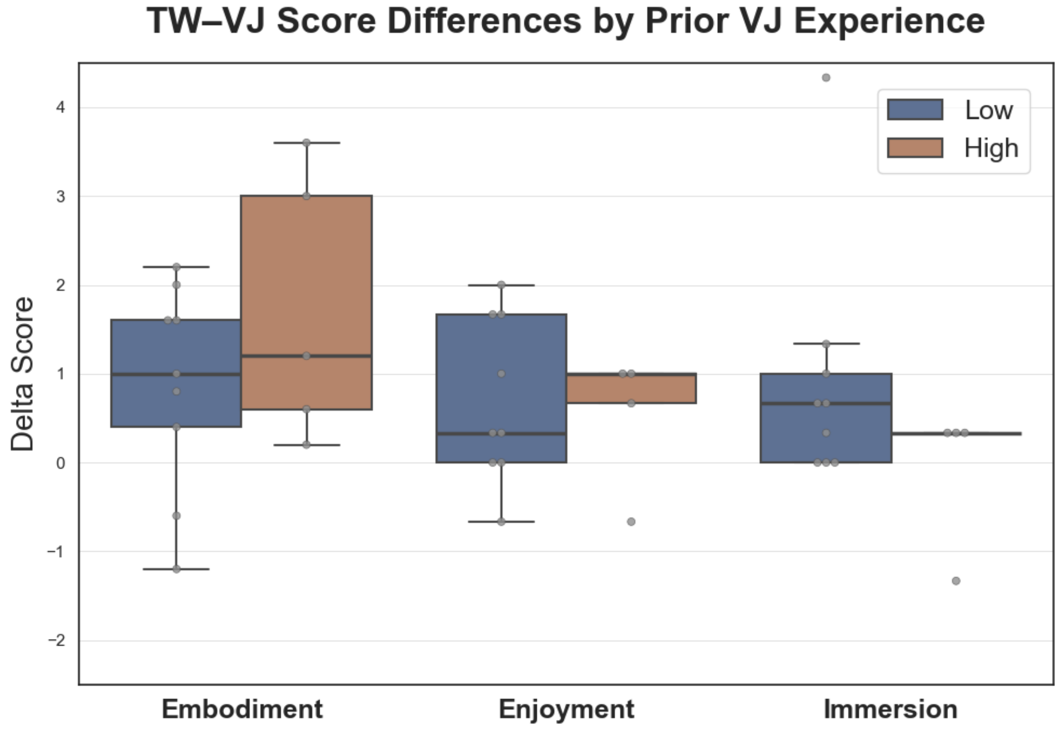
Method

- Participants split into **Low (1~3)** vs. **High (4~5)** familiarity groups

Findings

- No significant differences between groups
- Effect sizes small across all measures

→ **H2 not supported**



⇒ TW outperformed VJ, regardless of prior joystick experience

Objective Evaluation

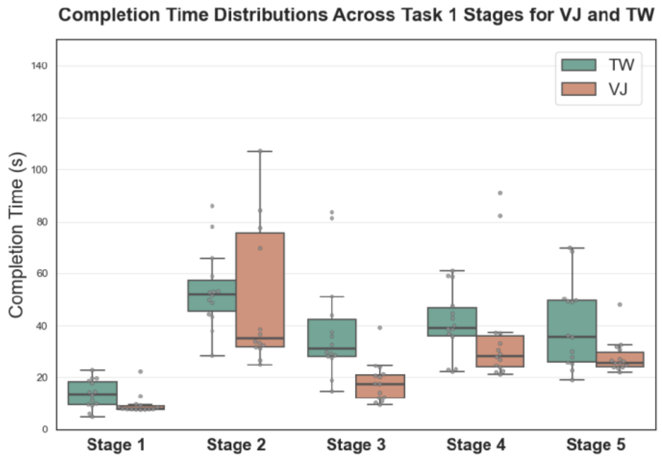
Completion Time

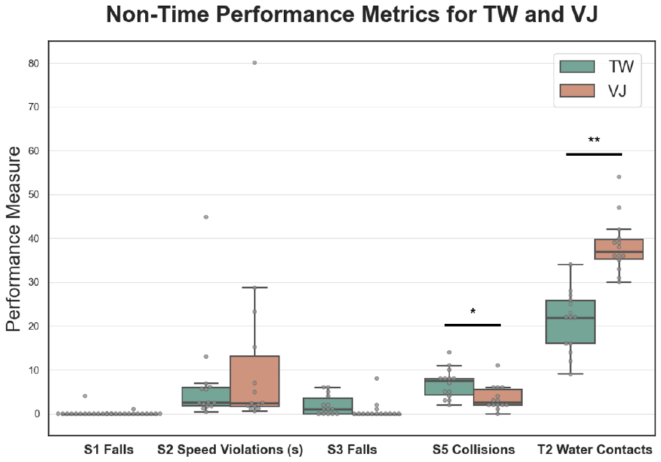
- **VJ faster** in most stages
  (Hole Avoidance, Narrow Path, Rolling Obstacles)
- **TW more consistent** in
  dynamic speed control & foot-button tasks

Non-Time Performance

- **VJ fewer collisions** in fast-paced tasks
- **TW fewer water contacts** in precise stepping tasks

VJ → faster in quick / spatially constrained tasks
TW → more consistent & precise in foot-based tasks





User Feedbacks Summary

Embodiment

- Enhanced by **rhythmic alignment** & **direct mapping** to footsteps
- Reduced when avatar did **not respond as intended**

Enjoyment

- Driven by **novelty** & **fun of walking gesture**
- Reduced by **input difficulty** or **unfamiliarity**

User Feedbacks Summary

Immersion

- Felt more **immersive** & **engaging**
- Requires **continuous focus** & **physical coordination**

Additional Insights

- Mixed views on **intuitiveness** & **learnability**
- TW praised for **precision**, VJ for **ease of control**
- Suggested use cases: **puzzle, stealth, platformer games**

Discussion

- **Subjective Experience:**
  - TW > VJ in embodiment, enjoyment, immersion (supporting H1)
    → Due to **foot–ground mapping + real-time motion generation**
- **Joystick Familiarity:**
  - No significant effect (H2 not supported)
- **Task-specific Trade-offs:**
  - **TW** more consistent & precise in deliberate control (e.g., stepping, button tasks)
  - **VJ** faster in obstacle-heavy & spatially constrained tasks
- **Future Potential:**
  - Best suited for **puzzle, adventure, stealth** genres.
  - Needs improvement in **directional responsiveness** & reducing **input effort / fatigue**

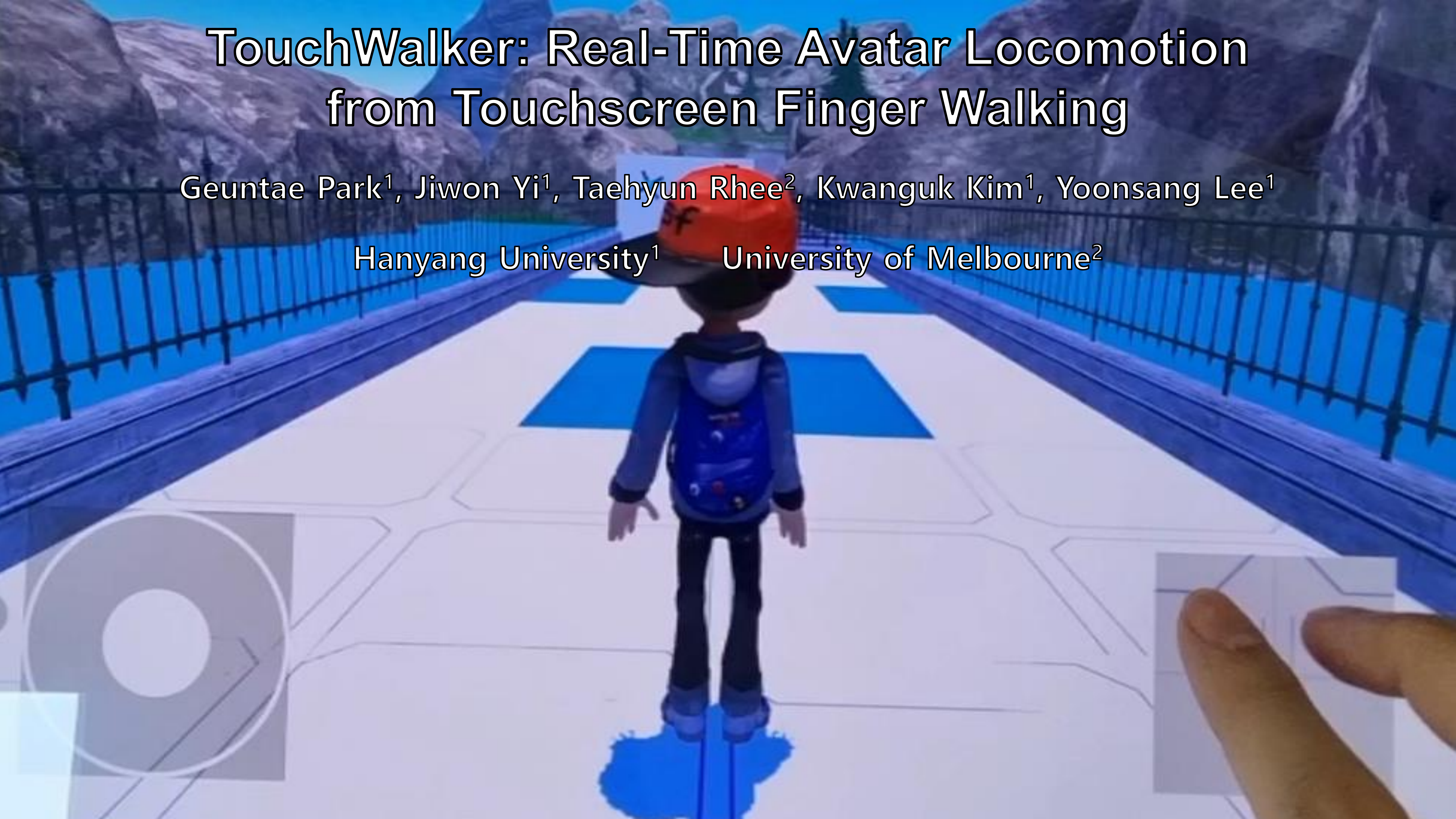
Conclusion

- **TouchWalker**
  - A real-time touchscreen locomotion system
  - Significantly improved **embodiment, enjoyment,** and **immersion**
  - Advantages in **precise** and **deliberate** tasks
- Challenges:
  - **Spatially constrained** or **rapid-response** scenarios
- Promising alternative for **expressive avatar control**

TouchWalker: Real-Time Avatar Locomotion from Touchscreen Finger Walking

Geuntae Park[1], Jiwon Yi[1], Taehyun Rhee[2], Kwanguk Kim[1], Yoonsang Lee[1]

Hanyang University[1] University of Melbourne[2]



Results

Ablation Study

	FCPE ↓ (cm/frame)	FCTE ↓ (%)	MPJPE ↓ (cm/frame)	FS ↓ (cm/frame)
w/o TransNet&GRU	17.23	34.70	20.74	1.66
w/o TransNet	35.70	40.56	66.33	1.56
w/o GRU	24.25	38.55	31.29	**1.40**
w/o L_{rec}	12.20	33.00	25.92	2.50
w/o L_{FK}	37.02	48.12	122.89	3.40
w/o L_{ct}	30.26	34.79	33.66	1.97
w/o $L_{\text{ct\_trans}}$	10.13	26.06	20.19	1.68
Full model	**7.04**	**24.42**	**17.47**	1.68