Habitus: Boosting Mobile Immersive Content Delivery through Full-body Pose Tracking and Multipath Networking

Anlan Zhang\textsuperscript{1}\* Chendong Wang\textsuperscript{4} Yuming Hu\textsuperscript{2} Ahmad Hassan\textsuperscript{1}\*
Zejun Zhang\textsuperscript{1}\* Bo Han\textsuperscript{3} Feng Qian\textsuperscript{1}\* Shichang Xu\textsuperscript{5}

\textsuperscript{1}University of Southern California \textsuperscript{2}University of Minnesota, Twin Cities
\textsuperscript{3}George Mason University \textsuperscript{4}University of Wisconsin-Madison \textsuperscript{5}Google

April, 2024
Immersive Content is Everywhere

• 3-DoF (degree-of-freedom) to 6-DoF motion
  • x, y, z, yall, pitch, roll
• Bandwidth-intensive
  • Hard to deliver through common wireless links (e.g., 802.11ac)


Media sources:
[3] https://www.youtube.com/watch?v=aO3TAke7_MI
Omnidirectional vs. mmWave Radio

Networking Spectrum Bands

- **Omnidirectional**
  - Low Bandwidth
  - Slow signal attenuation
  - Less vulnerable to blockage

- **mmWave**
  - High Bandwidth
  - Fast signal attenuation
  - More vulnerable to blockage

---

Promising for streaming bw-hungry immersive content!

Image sources:
[1] https://www.androidauthority.com/sub-6ghz-vs-mmwave-5g-3051246/
Case Study: Volumetric Video Streaming over mmWave

- 802.11ad (60 GHz) vs. 802.11ac (5 GHz)
- Test app: [ViVo, MobiCom’20]
- Impact on QoE (quality-of-experience)
  - Video quality +113%, stall +502%
Existing Systems Using mmWave

- Improving the PHY layer, e.g., SpaceBeam [MobiSys’21]
- Enhancing line-of-sight (LoS), e.g., VIVE Wireless Adapter [2]
- Using specialized device, e.g., MoVR [NSDI’17]

Existing Systems Using mmWave

- Improving the PHY layer, e.g., SpaceBeam [MobiSys’21]
- Enhancing line-of-sight (LoS), e.g., VIVE Wireless Adapter [2]
- Using specialized device, e.g., MoVR [NSDI’17]

We need a robust and easy-to-deploy solution!

Our Solution: Habitus

- Multipath networking over heterogeneous links
  - Omnidirectional (e.g. 802.11ac) + mmWave (e.g., 802.11ad) radios

- Actively predicting the fluctuating mmWave throughput
  - Under constant motion of the viewer
Challenges

• Predicting mmWave throughput w/ ML models
  • How to improve the accuracy under constant motion

• Applying offline pre-trained ML-based prediction models at runtime
  • How to react to unseen changes deviating from training data

• Heterogeneous (omnidirectional + mmWave) links
  • How to do multipath scheduling
Basic ML-based mmWave Throughput Prediction

- History network measurements $\rightarrow$ future network condition (e.g., throughput)
Motion-enhanced mmWave Throughput Prediction

• Insight 1: mmWave throughput is correlated with 6-DoF motions
  • Also validated by previous work: Lumos5G [IMC’20], Aggarwal et al. [PAM’21]

Full-body Pose Guided mmWave Throughput Prediction

- Insight 2: **Spatial correlation** among body parts during human motion [1, 2]
  - Example: hand holding controller moves → head movement → throughput changes
  - Example: leg moves (e.g., viewer turns left) → head rotation → throughput changes

- Tracking **full-body pose** can improve mmWave throughput prediction
  - Body pose: a set of 3D key points

**Full-body Pose Guided mmWave Throughput Prediction**

- Data collection at 4 locations w/ 802.11ac/ad APs + a stereo camera (for tracking pose)
  - 3 viewers (1.6m, 1.7m, 1.8m / 1 Female, 2 Males) exercise 10 motion patterns
    - Collect: ① 802.11ac/ad throughput & signal strength, ② 6DoF head motion, ③ full-body pose
  - More details are in the paper

<table>
<thead>
<tr>
<th>Patterns</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>The user stands in the center of the room, turning around in a clockwise direction.</td>
</tr>
<tr>
<td>S2</td>
<td>The user stands in the center of the room, turning around in a counterclockwise direction.</td>
</tr>
<tr>
<td>S3</td>
<td>The user walks around in a clockwise direction.</td>
</tr>
<tr>
<td>S4</td>
<td>The user walks around in a counterclockwise direction in a normal speed.</td>
</tr>
<tr>
<td>S5</td>
<td>The same as S4, but in a slow speed.</td>
</tr>
<tr>
<td>S6</td>
<td>The same as S4, but in a fast speed.</td>
</tr>
<tr>
<td>S7</td>
<td>A chair occupies the front place of the access point. The user walks around in a counterclockwise direction.</td>
</tr>
<tr>
<td>S8</td>
<td>The same as S3, but the user does not change the orientation of his/her head.</td>
</tr>
<tr>
<td>S9</td>
<td>The same as S4, but the user does not change the orientation of his/her head.</td>
</tr>
<tr>
<td>S10</td>
<td>The user walks around following the walking trace in S7, but there is no chair.</td>
</tr>
</tbody>
</table>

*Table 1: User motion patterns.*
**Full-body Pose Guided mmWave Throughput Prediction**

- Prediction target: mmWave throughput in the next 1 second
  - MAE (mean absolute error) of model w/o and w/ pose
    - Seq2Seq model
    - Other models: GBDT (gradient boosting decision tree), MLP, RNN in our paper
  - w/ Pose: 5% - 29% MAE reduction

---

**Seq2Seq model w/ Encoder-Decoder Architecture**

- Encoder state
-Historical features

- mmWave throughput, signal strength
- 6DoF head motion (X, Y, Z, Yaw, Pitch, Roll)
- Body pose

- Predictions: mmWave throughput

---

**Chart:**

- MAE (Mbps)
- 17% reduction
Full-body Pose Guided mmWave Throughput Prediction

- Prediction target: mmWave throughput in the next 1 second
  - MAE (mean absolute error) of model w/o and w/ pose
    - Seq2Seq model
    - Other models: GBDT (gradient boosting decision tree), MLP, RNN in our paper
  - w/ Pose: 5% - 29% MAE reduction

We are not done yet!
Impact of Unseen Changes

- Use a pre-trained model at runtime: performance drop
  - Reason: cannot adapt to changes deviating from training data, e.g.,
    - C1: location change
    - C2: user change
    - C3: motion pattern change
    - C4: static environmental change (e.g., a chair being moved)
    - C5: dynamic environmental change (e.g., a walk-by spectator)
  - Intuition: There is fundamental knowledge learned by models
    - Physical property of mmWave
    - Throughput distribution in certain positions
  - Intuition: The pre-trained model has different sensitivities to these changes
    - The changes that reshape the fundamental knowledge have larger impact
Impact of Unseen Changes

• Systematically quantify model sensitivity to the changes
  • Apply pre-trained models to
    • New location/user/motion pattern
    • Manually created static/dynamic environmental changes

![Static environmental change](image1)

Change

Static environmental change

![Dynamic environmental change](image2)

Dynamic environmental change

Smartphone

Camera & AP
Impact of Unseen Changes

- Systematically quantify model sensitivity to the changes
  - C1: location change
  - C2: user change
  - C3: motion pattern change
  - C4: static environmental change
  - C5: dynamic environmental change

- Impact of the changes on MAE
Impact of Unseen Changes

• Systematically quantify model sensitivity to the changes
  • C1: location change
  • C2: user change
  • C3: motion pattern change
  • C4: static environmental change
  • C5: dynamic environmental change

• Impact of the changes on MAE

Impact of Unseen Changes

Large impact

Small impact

Median ▲ Mean

Normalized MAE Increase (%)
Handle Unseen Changes

- Our solution: **Transfer Learning (TL)**
  - Key assumption: there is **invariant learned knowledge** before & after a change
  - Benefit: adapt model to the change much faster than training a new model from scratch
Handle Unseen Changes

• Our solution: **Transfer Learning (TL)**
  • C1: location change
  • C2: user change
  • C3: motion pattern change
  • C4: static environmental change
  • C5: dynamic environmental change

• More details are in the paper

Vision-based NLoS detection

![Diagram](image.png)
Multipath Scheduling

- Multipath: omnidirectional radio + mmWave
  - Prioritize omnidirectional radio
  - Opportunistically use mmWave
- Trend-aware scheduling
  - Conservatively or aggressively using mmWave
- See paper for details
Holistic View of Habitus

• mmWave throughput prediction
  • Enhanced by tracking full-body pose
  • React to unseen changes
    • Online/Offline transfer learning
    • NLoS detection

• Multipath networking
  • Omnidirectional radio + mmWave
  • Trend-aware scheduling
Habitus Prototype

- Habitus is a general framework for immersive apps
- Implementation w/ commodity HW/SW
  - Challenges, e.g., accurate throughput estimation

Hardware: ROG Phone II, KCXGHYI VR Headset, Netgear Nighthawk X10 AP, ZED 2i Camera
Software: Linux iw, libpcap, Google ARCore, PyTorch & Torchscript, Zed OpenPose

Integrate Habitus to ViVo [MobiCom’20]: only changing 47 LoC
Case Study: Volumetric Video Streaming

- Left: trace-driven emulation
- Right: user trial (N=12)
Case Study: Volumetric Video Streaming

- Left: trace-driven emulation
- Right: user trial (N=12)
- \textit{ad}: only use 802.11ad (mmWave) w/o prediction

![Graph showing Average Quality versus Stall](image)
Case Study: Volumetric Video Streaming

- Left: trace-driven emulation
- Right: user trial (N=12)

- **ad**: only use 802.11ad (mmWave) w/o prediction
- **Simple**: ac + ad w/o prediction
Case Study: Volumetric Video Streaming

- Left: trace-driven emulation
- Right: user trial (N=12)
  - ad: only use 802.11ad (mmWave) w/o prediction
  - Simple: ac + ad w/o prediction
  - Pro: Habitus ac + ad w/ prediction (6DoF features only)
Case Study: Volumetric Video Streaming

- Left: trace-driven emulation
- Right: user trial (N=12)

- **ad**: only use 802.11ad (mmWave) w/o prediction
- **Simple**: ac + ad w/o prediction
- **Pro**: Habitus ac + ad w/ prediction (6DoF features only)
- **Full**: Habitus ac + ad w/ prediction (6DoF + full-body Pose features)

Habitus (Pro, Full) considerably outperforms baseline approaches
Using full-body pose (Full) further boosts the QoE
Find more evaluation in our paper
Summary

• Challenge of high-quality immersive content delivery over mmWave

• The design of Habitus
  • Multipath scheduling over omnidirectional radio and mmWave
  • Full-body pose guided mmWave throughput prediction
    • Handle unseen changes

• QoE improvement of Habitus demonstrated by trace-driven emulation & user trial
  • We release our dataset and the source code for data collection
    • ① 802.11ac/ad throughput & signal strength, ② 6DoF head motion, ③ full-body pose
  • See our paper for the links