ResiHMR: Residual-Limb Aware Single-Image 3D Human Mesh Recovery for Individuals with Limb Loss

Supplementary Material

1. Broader Impact

Our work aims to advance accessibility and disability inclusion in computer vision by enabling anatomically valid 3D human modeling for individuals with limb loss, an underrepresented population in existing HMR datasets [3–6] and methods [2, 7, 8].

By explicitly reconstructing residual-limb geometry and adapting kinematic structure to limb-loss topology, ResiHMR offers a step toward more equitable and accessible 3D human modeling for applications that extend beyond graphics and pose estimation. The potential positive outcomes extend beyond graphics or pose estimation. Accurate residual-limb geometry is relevant to prosthetic design, rehabilitation assessment, parasport classification, and human-centered AI technologies that must account for bodies with varied physical structures. Our method offers a technical foundation that may support these domains when applied responsibly and in collaboration with clinical or biomechanics experts.

At the same time, modeling real people, particularly individuals with limb loss, requires careful ethical consideration. Residual-limb data may be sensitive, and any downstream clinical or biomechanical use should involve appropriate oversight, domain experts, and consent. Our method does not infer medical diagnoses nor replace clinical assessment; instead, it provides a technical capability that may complement future accessible technologies when used responsibly.

Our work also highlights the broader need for inclusive datasets, evaluation protocols, and modeling assumptions in human-centric AI. We hope it encourages the community to consider disability inclusion as a core component of fair and responsible human modeling, ultimately contributing to AI systems that serve a wider and more diverse range of users.

2. Importance of Residual-Limb Modeling

From the perspective of parasport science, clinical biomechanics, and disability-movement analysis, incorporating the residual limb as an explicit component of the human body model is not merely beneficial, it is essential. The residual limb is an active, load-bearing, and dynamically expressive structure that influences posture, balance control, segment coordination, socket alignment, compensatory strategies, and whole-body movement organization. Ignoring or collapsing this structure, as in methods that hallucinate intact limbs (e.g., HSMR [7]) or truncate geometry to the nearest joint (e.g., AJAHR [1]), fundamentally distorts the underlying anatomy and removes information that is critical for understanding how amputee athletes actually move. Residual-limb length, orientation, and stump-surface geometry directly shape gait kinematics, limb-loading asymmetries, and upper and lower body coordination factors central to parasport classification, performance evaluation, and prosthetic design.

In contrast, ResiHMR reconstructs the residual limb as a first-class anatomical entity, producing continuous termination points and watertight stump surfaces that correspond to the individual's true morphology. This explicit treatment enables anatomically faithful 3D meshes that retain clinically meaningful geometry rather than simplified joint-level surrogates. As a result, ResiHMR provides not only more accurate reconstructions but also scientifically interpretable models that align with real-world biomechanics. This capability is indispensable for inclusive human-pose research and positions our work as the first step toward anatomically grounded and clinically relevant 3D mesh recovery for individuals with limb loss.

3. AJAHR vs. ResiHMR

Amputated Joint Aware HMR (AJAHR) [1] introduces BPAC-Net, which classifies amputated body regions from images and 2D keypoints, and an AJAHR-Tokenizer, a VQ-VAE–style pose tokenizer trained on large-scale intact-body pose datasets together with a synthetic amputee dataset (A3D). In A3D, amputations are generated by modifying SMPL pose parameters: the pose of an amputated joint and all its descendants is zeroed, and the SMPL skinning function collapses distal vertices toward the parent joint. This representation preserves the original SMPL kinematic tree and restricts amputations to occur strictly at existing joint locations. Consequently, AJAHR cannot localize anatomical residual-limb endpoints and produces no stump surface. Instead, the distal mesh simply converges to the nearest intact joint. The supervision in AJAHR, including its amputee training data, therefore reflects only joint-level amputations and lacks ground-truth residual-limb geometry. As a result, while AJAHR can recognize which joint is amputated, the recovered mesh still terminates at the SMPL joint and cannot represent the true residual-limb length, shape, or surface continuity. These limitations constrain AJAHR's applicability to downstream tasks that require explicit residual-limb geometry, such as prosthetic alignment, stump-volume estimation, residuum–socket interaction analysis, or biome-

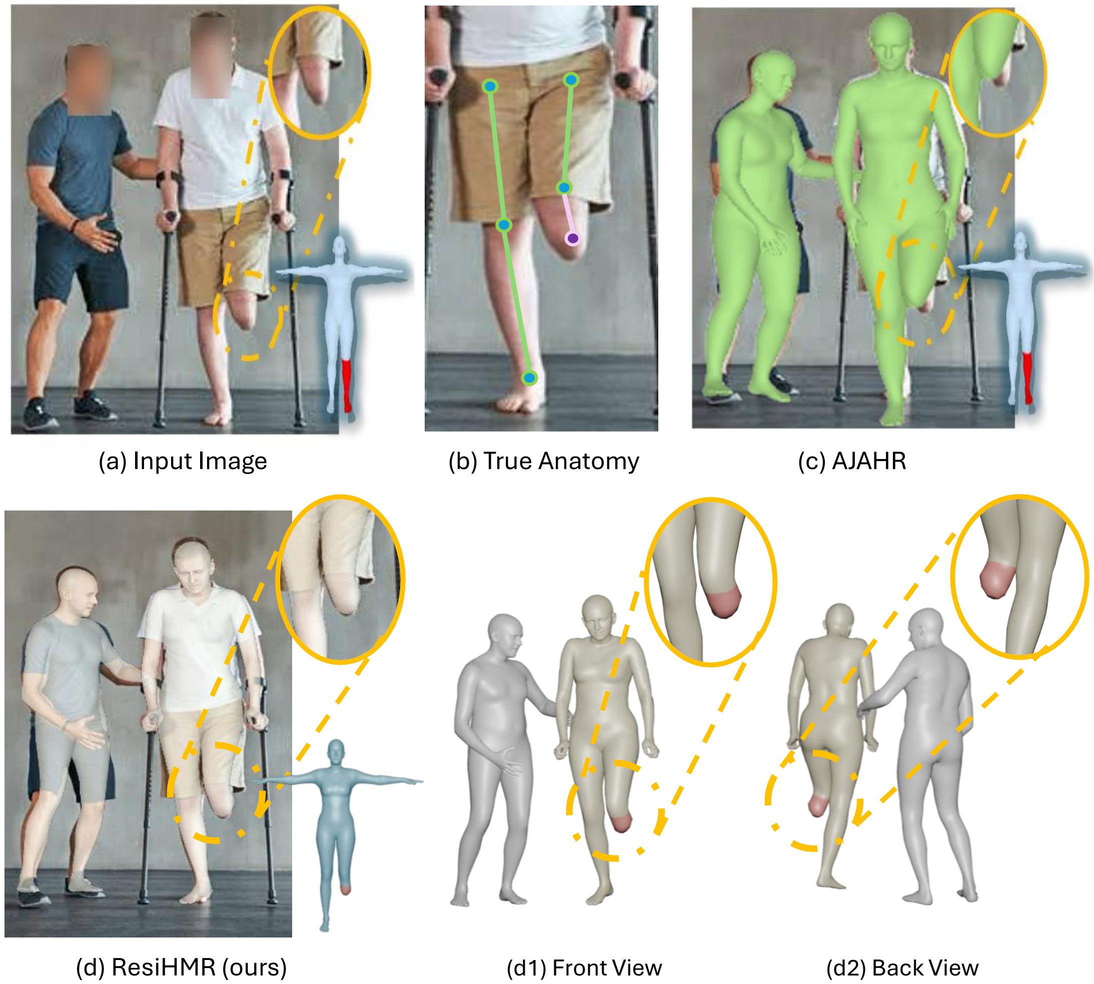

Figure 1. **A visual comparison of AJAHR and ResiHMR (ours).** (a,c) are copied from the AJAHR paper [1], where limb loss is represented by collapsing SMPL vertices toward the parent joint, resulting in joint-level truncation. (b) shows expert-verified evidence and an additional real image of the same individual revealing a clear below-knee residual limb, indicating that joint-level collapse fails to match the true anatomy. (d) shows ResiHMR reconstructions, including a normalized T-pose, with the residual limb highlighted in red (d1,d2). ResiHMR explicitly estimates the anatomical stump surface rather than collapsing geometry, producing a more realistic and clinically interpretable limb termination.

chanical evaluation, where continuous stump surfaces and accurate termination points are required.

3.1. Qualitative Comparison

Up to the date of submission, the official AJAHR implementation had not been released, which prevents us from performing a controlled quantitative comparison. For this reason, AJAHR is not included in our experimental evaluation. Instead, we provide a qualitative examination by running ResiHMR on the same input image used in the AJAHR paper and comparing the outputs. As shown in Figure 1, the AJAHR results reproduced from the original paper exhibit a joint-level collapse pattern: the predicted mesh contains no residual limb below the left knee, and the left knee itself is collapsed to the proximal joint. To verify the true anatomy of the target individual, we consulted a disability and human-movement expert and also located publicly available images of the same person performing a different action. These independent sources clearly confirm that the subject does have a residual limb below the left knee. Therefore, we include panel (b) in Figure 1 to illustrate the correct anatomy, which AJAHR fails to represent. In contrast, ResiHMR reconstructs an anatomically coherent below-knee residual limb, producing a realistic, water-

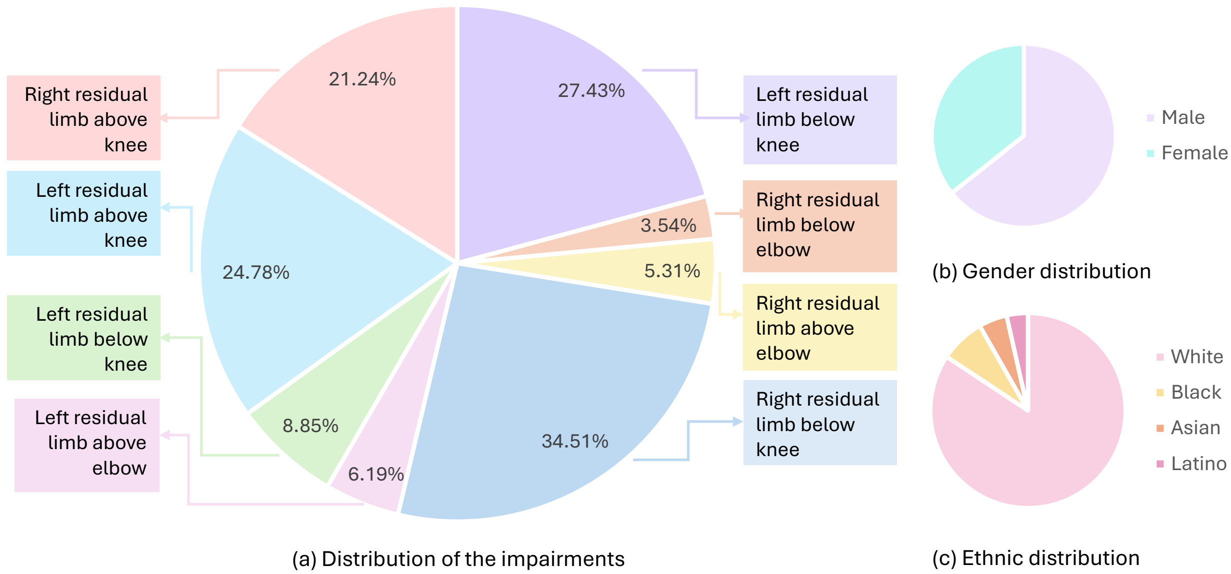

Figure 2. **Key Statistics of our LDPose-LimbLoss Evaluation Dataset.** (a) Distribution of residual-limb types, covering upper- and lower-limb amputations across both sides of the body. (b) Gender distribution. (c) Ethnic composition of the participants. Together, these statistics demonstrate the dataset's demographic and impairment diversity, providing a representative foundation for benchmarking residual-limb–aware 2D/3D human pose and mesh reconstruction. .

tight stump surface as shown in Figure 1(d), with the reconstructed residual limb highlighted in red in (d1) and (d2). This side-by-side comparison underscores the fundamental difference between AJAHR's joint-level truncation and our explicit residual-limb estimation, demonstrating the necessity of anatomically grounded modeling for amputee subjects.

3.2. Conceptual Differences

ResiHMR is designed from a fundamentally different perspective and addresses limitations that AJAHR does not resolve.

Joint-level vs. residual-limb level modeling. AJAHR is *amputated-joint aware*: it predicts whether specific SMPL joints (and their descendants) are amputated and zeroes their rotations, leaving the limb to terminate at the nearest intact SMPL joint. [1] ResiHMR is *residual-limb aware*: it explicitly models residual-limb endpoints in 2D (via LDPose [9] keypoints) and lifts them into 3D as continuous points along the kinematic chain, rather than snapping them to the nearest SMPL joint. This allows residual limbs to end between canonical joints (e.g., mid-thigh, mid-shank), which is typical for real-world amputations.

Topology collapse vs. explicit stump geometry. In AJAHR, the SMPL topology is never altered: the skeleton and mesh connectivity remain intact, and amputations are simulated by collapsing vertices toward the parent joint. This produces a visually amputated region but not a resolved stump surface anatomically. ResiHMR, by contrast, performs a *geometry-based limb termination*: it removes distal vertices, identifies a local boundary ring near the optimized residual endpoint, and reconstructs a smooth, convex, watertight stump surface. As a result, the residual limb becomes an actual geometric structure in the mesh, rather than an implicit collapse of vertices.

Learning regression vs. optimization objective. AJAHR is a learned, token-based regression model, which uses a VQ-VAE tokenizer and a transformer-based HMR architecture trained end-to-end on large-scale synthetic and real data. [1] Its amputation representation is baked into the SMPL pose parameters and the BPAC-Net classifier. ResiHMR is an optimization-based framework built on SMPLify-X, augmented with a Residual Anchor–Factor Optimization (RAFO) stage. RAFO explicitly optimizes the 3D anchor joint locations and continuous residual-length factors under residual 2D keypoint supervision and anthropometric length priors, enabling anatomically meaningful residual-limb endpoints even without 3D stump ground truth.

Supervision signals. AJAHR relies on synthetic amputee 3D annotations generated from the SMPL joint-zeroing pipeline; its supervision is therefore aligned with the intact SMPL kinematic tree and joint positions. [1] ResiHMR instead leverages *residual-limb endpoints* detector developed from LDPose, then uses RAFO plus geometry-based reconstruction to infer stump geometry consistent with these 2D cues and anthropometric segment ratios. This makes ResiHMR more suitable for scenarios where accurate residual-limb localization matters (e.g., residual-length estimation, socket-region analysis), even if only 2D annotations are available.

3.3. Why Joint-Level Truncation Is Insufficient for Downstream Biomechanics

Because AJAHR represents amputations through joint-level collapse, implemented by drawing distal vertices toward the parent SMPL joint, the resulting mesh lacks an explicit stump surface and does not provide a measurable residual-limb length along the limb axis. This joint-based truncation preserves the intact SMPL kinematic structure but removes the anatomical information required to characterize the geometry of the residual limb.

As discussed in Section 2, explicit modeling of the residual limb is essential from the standpoint of parasport science, clinical biomechanics, and disability-movement analysis. The residual limb constitutes a load-bearing and dynamically active segment that contributes to balance control, inter-segment coordination, joint kinetics, compensatory strategies, and prosthetic socket alignment. Without a geometric representation of the residual limb, including its termination point, surface continuity, and cross-sectional shape, these biomechanical factors cannot be meaningfully assessed. Consequently, the joint-level representation adopted by AJAHR is unsuitable for downstream applications that rely on physical stump geometry, such as residual-limb volume estimation, residuum–socket interaction modeling, and load-transfer analysis.

In contrast, ResiHMR explicitly estimates residual-limb endpoints and continuous residual-length factors, and reconstructs a watertight stump surface tailored to the individual. This yields anatomically interpretable quantities, including segment length, local cross-sections along the stump, and overall stump geometry, enabling analyses relevant to prosthetic design, rehabilitation, classification research, and broader inclusive biomechanics.

4. ResiHMR with other HMR Methods

ResiHMR is designed as a backbone-agnostic extension that can be integrated with existing HMR pipelines. As illustrated in Figure 5, we first obtain an initial SMPL-X estimate (camera, pose, and shape) from an off-the-shelf HMR method, which can be either optimization-based (e.g., SMPLify-X) or regression-based (e.g., HSMR). This initialization provides a coarse full-body reconstruction under the intact-limb assumption.

Given the input image and corresponding 2D keypoints (including residual-limb endpoints), we then apply our Residual Anchor-Factor Optimization to adapt the kinematic structure by refining anchor joints and residual-limb proportions. This step corrects the mismatch between intact-limb priors and amputated anatomy. Subsequently, Residual Limb Reconstruction replaces the invalid distal geometry with a smooth, watertight stump surface.

This modular design enables ResiHMR to operate as a plug-in on top of existing HMR systems, extending them from intact-body reconstruction to anatomically coherent modeling of individuals with limb loss. For more details, please visit our project page:

(a) Input Image (b) ResiHMR Mesh on Image

Figure 3. **Limitation: Residual-limb surface shape approximation.** Although ResiHMR accurately localizes the residual-limb endpoint, the reconstructed stump surface adopts a smooth, convex prior that does not fully reflect the subject's true residual-limb contour (yellow). The intact limb is reconstructed faithfully (green), indicating that the discrepancy arises from limited instance-specific stump shape cues rather than errors in body alignment. This limitation reflects the absence of residual-limb 3D training data and highlights a natural growth direction toward modeling individualized residual-limb geometry.

5. LDPose-LimbLoss Evaluation Dataset Statistics

The LDPose-LimbLoss Evaluation Dataset is designed exclusively for evaluation. It is not used for training, pre-training, optimization, hyper-parameter tuning, or any other component of ResiHMR. This ensures strict separation between training data and evaluation data, eliminating the possibility of implicit supervision or data leakage. All quantitative and qualitative results reported in the main paper are therefore unbiased by this dataset.

We apply rigorous selection criteria to ensure the dataset provides a reliable and fair basis for evaluating amputee mesh reconstruction. We include only images in which the majority of the subject's body is visible, and where residual-

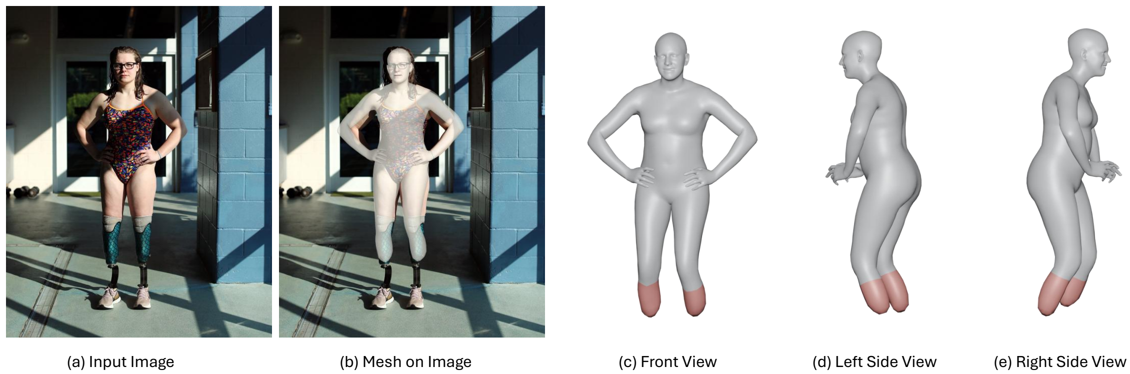

Figure 4. **Limitation: Depth ambiguity from monocular input.** Although the mesh aligns well with the input image in the front view (a–c), the side views (d–e) reveal depth ambiguities inherent to single-image reconstruction. Specifically, the arms are positioned in front of the torso rather than resting on the hips, and the upper body exhibits forward flexion that is not visible from the frontal perspective. These discrepancies arise from the absence of 3D residual-limb ground truth and the reliance on 2D annotations and healthy-body priors. Future 3D datasets with explicit residual-limb geometry may help resolve such ambiguities.

limb regions are either clearly observable or can be unambiguously inferred. Images with heavy occlusion, severe cropping, or ambiguous limb-termination cues are removed. For partially occluded subjects, missing body regions are estimated from nearby visible context to maintain anatomical completeness. We also minimize unrelated visual distractors such as cluttered backgrounds, large prosthetic occlusions, or external objects that may interfere with pose estimation. These curation steps ensure that evaluation accuracy reflects the algorithm's performance rather than noise from missing anatomy or extreme occlusion.

The final dataset contains 255 curated images that span a broad variety of residual-limb configurations, amputation levels, residual-limb lengths, activities, genders, and ethnic backgrounds. Each image is annotated with 17 standard 2D body keypoints and 8 additional residual-limb endpoints. To further facilitate optimization-based mesh fitting, we provide a per-person segmentation mask that isolates the human body region and reduces the influence of clothing, hair, prosthetic devices, equipment, and background objects. These annotations create a clean and controlled setting for evaluating residual-limb aware pose and mesh reconstruction models.

All annotations follow the LDPose protocol with extensions specific to limb-loss anatomy. Annotators were given detailed guidelines on identifying residual-limb endpoints, stump boundaries, and termination cues. Cases with uncertain or ambiguous stump locations were independently reviewed by multiple annotators and resolved through consensus. When multiple images of the same subject were available, cross-image consistency checks were performed to ensure stable residual-limb endpoint placement. This multi-stage quality control procedure ensures that the evaluation labels are precise, consistent, and anatomically meaningful.

The dataset includes demographic and impairment-level diversity, which reduces bias and improves generalizability. Finally, precise residual-limb annotations and segmentation masks ensure that reconstructed meshes can be evaluated with respect to anatomically accurate ground-truth cues.

Figure 2 presents the full statistics of the LDPose-LimbLoss Evaluation Dataset. The dataset includes 255 images distributed across a wide range of amputation types, genders, and ethnic backgrounds:

Residual-limb types (Figure 2 (a)). The dataset spans upper- and lower-limb amputations, covering above- and below-elbow and above- and below-knee cases on both sides of the body. This diversity supports evaluation across a broad set of real-world amputation configurations.

Gender distribution (Figure 2 (b)). Both male and female subjects are included, enabling analyses that account for gender-related anthropometric variation.

Ethnic composition (Figure 2 (c)). Participants represent multiple ethnic backgrounds, providing variation in body shape, skin tone, and appearance critical for inclusive pose and mesh reconstruction.

Together, these statistics highlight the *demographic* and *impairment-level diversity* of the dataset, establishing a representative benchmark for residual-limb aware 2D/3D human pose and mesh reconstruction.

6. Discussion

Although ResiHMR advances anatomically grounded 3D modeling for individuals with limb loss, several limitations remain that reflect the current state of available data and the inherent ambiguity of single-image reconstruction rather than deficiencies of the method.

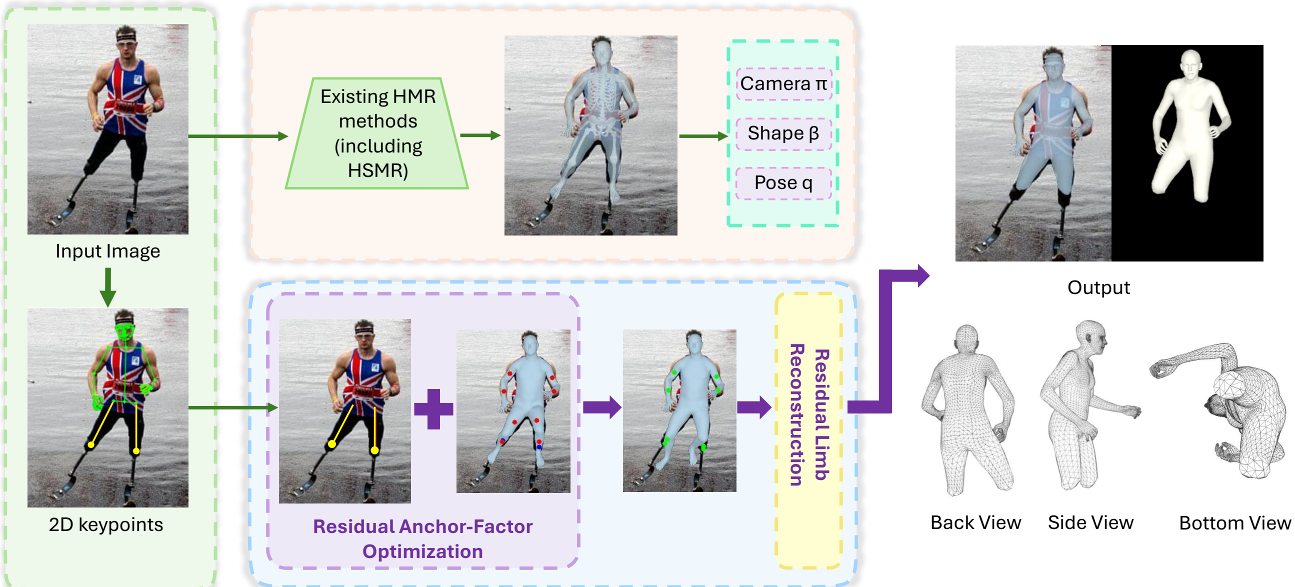

Figure 5. **Overview of our ResiHMR Framework with other existing HMR methods to initialize the SMPL-X/SMPL**, in this case, we have HSMR as the exmaple for the regression-based HMR method.

First, while ResiHMR can reliably estimate continuous residual-limb endpoints for the majority of typical amputation cases, particularly when the residual and intact limbs share broadly similar geometric structure, it cannot yet capture complex limb deformities or highly irregular residual-limb shapes. As illustrated in Figure 3, the reconstructed mesh aligns well with the intact limb but does not fully reproduce the subject's true residual-limb contour. This limitation arises because the model currently infers stump geometry from smooth convex priors rather than detailed, instance-specific surface cues. We view this not as a shortcoming of the framework, but as a natural growth direction: future work incorporating richer shape priors or multi-view information may enable individualized residual-limb surface modeling.

Second, because there is currently no 3D dataset containing ground-truth residual-limb geometry, ResiHMR relies on 2D annotations and anthropometric priors derived from healthy-body datasets. The lack of 3D supervision introduces depth ambiguity that is inherent to single-image reconstruction. As shown in Figure 4, the mesh appears well aligned in the front view, yet the side view reveals discrepancies: the arm is positioned in front of the torso rather than resting on the hip, and the upper body exhibits forward flexion not visible in the frontal image. Such ambiguities reflect limitations of monocular input rather than the optimization framework. The emergence of future 3D datasets containing residual-limb ground truth would likely mitigate these depth-related deviations and enable more accurate modeling of amputee-specific body postures.

These limitations point to several promising future directions, such as integrating multi-view information, incorporating uncertainty-aware optimization, or learning limb-loss–specific 3D priors, while not detracting from the core contribution of ResiHMR, enabling explicit and anatomically grounded modeling for individuals with limb loss.

7. Qualitative Comparison

In this section, we qualitatively compare ResiHMR with state-of-the-art HMR methods on individuals with limb loss from single-image inputs. Representative examples are shown in Figure 6, highlighting the advantages of our approach.

8. ResiHMR Demonstration

In this section, we present additional qualitative results of ResiHMR to further illustrate its behavior across diverse limb-loss configurations, activities, and viewing conditions. Representative examples are shown in Figure 7 and Figure 8, demonstrating the robustness of our residual-limb modeling pipeline and the anatomical coherence of the reconstructed meshes.

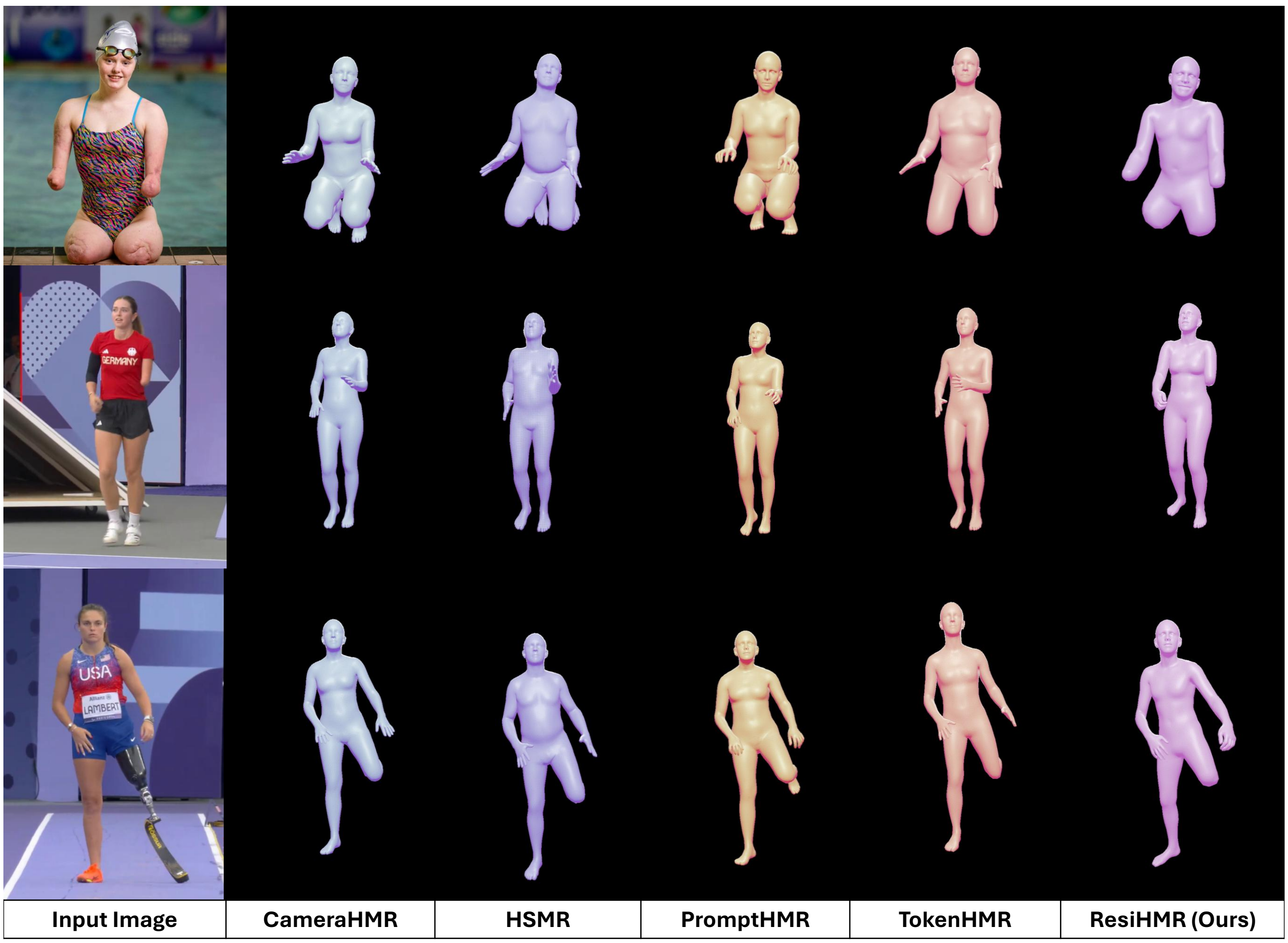

Figure 6. **Qualitative Comparison of ResiHMR with SOTA HMR methods.** Please see this in GIF format in project page ResiHMR.

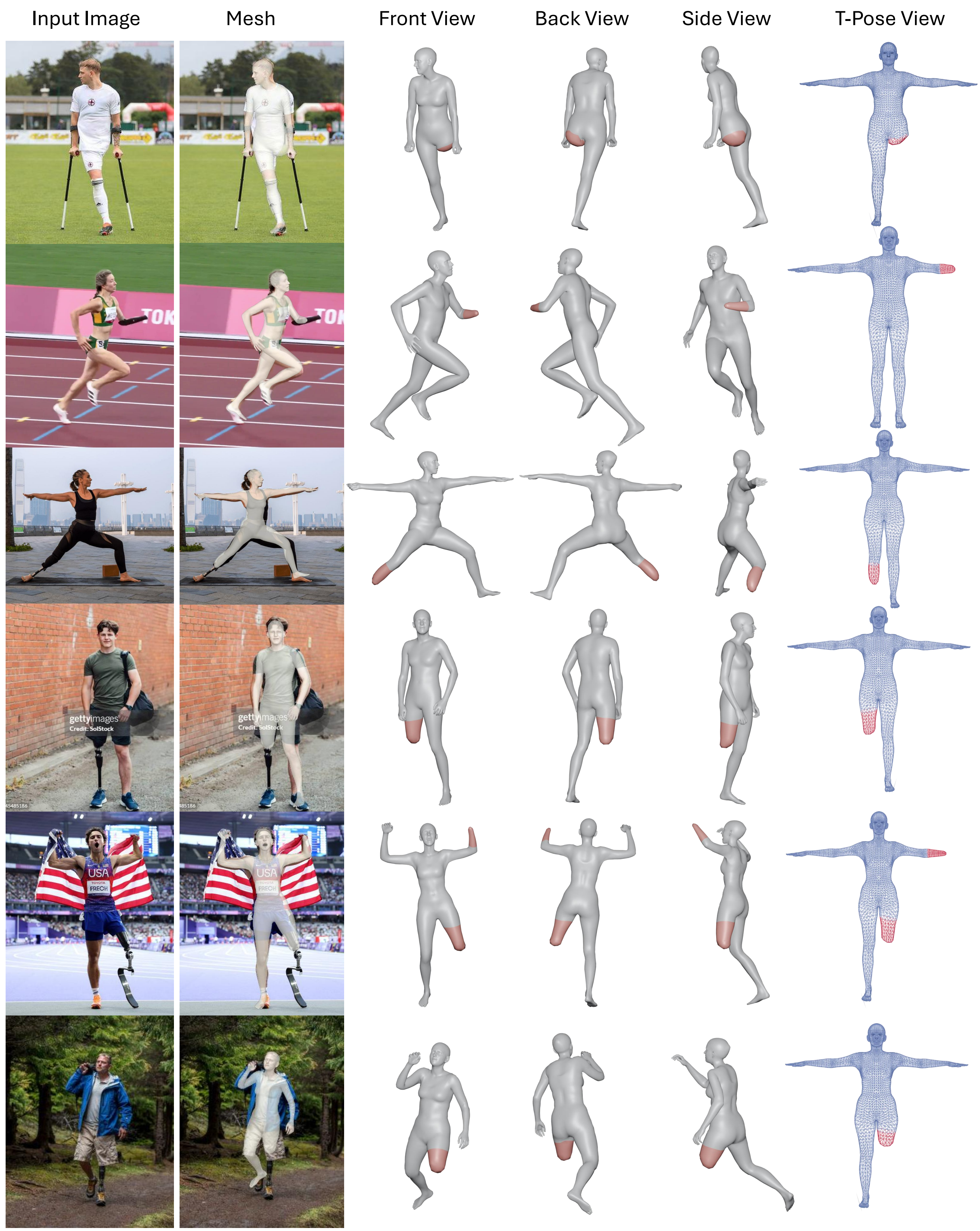

Figure 7. **More Qualitative Evaluation of ResiHMR.** For each input example, we show: (a) the input image, (b) the overlay of SMPL-X in the input view, (c) front view, (d) back view, (e) side view, (f) T-Pose view with model output $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\psi}, \mathbf{R}, \mathbf{t}, \lambda_r\}$ and $\mathbf{m}_r = \{ m_k \mid \mathbf{v}_k = (x_k, y_k, z_k, m_k) \in \mathbf{V}_r \}$, and (c) (d) (e) and (f) all have residual limb being highlighted in red.

Figure 8. **More Qualitative Evaluation of ResiHMR.** For each input example, we show: (a) the input image, (b) the overlay of SMPL-X in the input view, (c) front view, (d) back view, (e) side view, (f) T-Pose view with model output $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\psi}, \mathbf{R}, \mathbf{t}, \lambda_r\}$ and $\mathbf{m}_r = \{\, m_k \mid \mathbf{v}_k = (x_k, y_k, z_k, m_k) \in \mathbf{V}_r \,\}$, and (c) (d) (e) and (f) all have residual limb being highlighted in red.

References

[1] Hyunjin Cho, Giyun Choi, and Jongwon Choi. Ajahr: Amputated joint aware 3d human mesh recovery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7925–7935, 2025. 1, 2, 3

[2] Shubham Goel, Georgios Pavlakos, Jathushan Rajasegaran, Angjoo Kanazawa, and Jitendra Malik. Humans in 4d: Reconstructing and tracking humans with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14783–14794, 2023. 1

[3] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1325–1339, 2014. 1

[4] Naureen Mahmood, Nima Ghorbani, Nikolaus F Troje, Gerard Pons-Moll, and Michael J Black. Amass: Archive of motion capture as surface shapes. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 5442–5451, 2019.

[5] Shashank Tripathi, Lea Müller, Chun-Hao P. Huang, Taheri Omid, Michael J. Black, and Dimitrios Tzionas. 3D human pose estimation via intuitive physics. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4713–4725, 2023.

[6] Keenon Werling, Nicholas A Bianco, Michael Raitor, Jon Stingel, Jennifer L Hicks, Steven H Collins, Scott L Delp, and C Karen Liu. Addbiomechanics: Automating model scaling, inverse kinematics, and inverse dynamics from human motion data through sequential optimization. *Plos one*, 18(11): e0295152, 2023. 1

[7] Yan Xia, Xiaowei Zhou, Etienne Vouga, Qixing Huang, and Georgios Pavlakos. Reconstructing humans with a biomechanically accurate skeleton. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 5355–5365, 2025. 1

[8] Yuliang Xiu, Jinlong Yang, Xu Cao, Dimitrios Tzionas, and Michael J Black. Econ: Explicit clothed humans optimized via normal integration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 512–523, 2023. 1

[9] Jiaying Ying, Heming Du, Kaihao Zhang, Lincheng Li, and Xin Yu. Ldpose: Towards inclusive human pose estimation for limb-deficient individuals in the wild. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9865–9875, 2025. 3