

Supplementary material for 3D Human Pose Estimation with Muscles

A Technical appendix

A.1 Human model

We assume a rigid multibody dynamics model of a human with $N_k = 18$ joints – *pelvis, lumbar joint, thoracic joint, neck, scapulas, shoulders, elbows, wrists, hips, knees, and ankles*. The pelvis is set as the root, with 3 rotational and 3 translational degrees of freedom (DoFs). The scapulas have 2 DoFs, corresponding to depression/elevation and protraction/retraction. The elbows have 2 DoFs, corresponding to flexion/extension and forearm pronation/supination. The wrists have 2 DoFs, corresponding to flexion/extension and ulnar/radial deviation. The knees have 1 DoF, corresponding to flexion/extension. All remaining joints have 3 DoFs, for a total of 47 DoFs. We selected this configuration as it aligns best with existing biomechanics models that we implemented, such as for anthropometrics estimation [16] and MTGs [53, 25].

Anthropometrics estimation. We predict the human’s anthropometrics by combining our regressed residuals \mathcal{E} with initial estimates based on literature values $\bar{\mathcal{A}}$ scaled by our predicted human dimensions. Specifically, we want to predict $\mathcal{A} = \cup_k \{m_k, I_{0,k}, \chi_k\}$, where m_k is the **mass** of segment k , with $I_{0,k}$ as its **inertia** tensor at zero rotation with scaling matrix Λ_k , and χ_k as its local **CoM** position relative to its segment length. For the remainder of the section, we assume relevant units to be in seconds, radians, meters, kilograms, Newtons.

From the predicted β and Eq. (2), we can compute the human’s volume and all segment lengths L_k . We further compute the human’s initial bodymass estimate \hat{M} as its volume multiplied by a constant density of 985 kg/m^3 . For segment k , let $s_{L,k} = L_k/H$ be its segment length relative to height, and $s_{m,k} = m_k/M$ be its mass relative to bodymass. Let “bar” ($\bar{\cdot}$) denote the human values measured by Dumas *et al.* in [16]. We set our initial estimates as $\bar{\mathcal{A}}$ scaled by $s_{L,k}/\bar{s}_{L,k}$:

$$\{M, s'_{m,k}, \Lambda_k, \chi_k\} = \{\hat{M}, \bar{s}_{m,k} \frac{s_{L,k}}{\bar{s}_{L,k}}, \bar{\Lambda}_k, \bar{\chi}_k\} + \mathcal{E} \quad (19)$$

$$m_k = s_{m,k} M, \text{ where } s_{m,k} = \frac{s'_{m,k}}{\sum_j s'_{m,j}} \quad (20)$$

$$I_{0,k} = m_k L_k^2 \Lambda_k \quad (21)$$

Lastly, we compute the **body weight** W of the human in Newtons, with $g = 9.8 \text{ m/s}^2$ as

$$W = g \sum_k m_k \quad (22)$$

A.2 GRFM Model

Let $\mathcal{F} = [\mathbf{F}, \mathbf{M}]^\top$ be the ground reaction forces and moments (GRFM) applied at the CoM of a foot in global cartesian coordinates. Let $\mathbf{F} = [F_X, F_Y, F_Z]^\top$ where Y is the vertical direction, and $\mathbf{z} = [z_x, z_y, z_z]^\top$ be the center of pressure (CoP) in the foot’s local coordinates where x is along the length of the foot, such that

$$\mathbf{M} = R_{ankle}^0 \mathbf{z} \times \mathbf{F} \quad (23)$$

where R_k is joint k ’s local rotation matrix, and $R_k^0 = R_{p(k)}^0 R_k$ describes the chain of rotational transformations from the world frame to its local frame. We use lower case x, y, z to denote the foot’s local coordinates, and its dimensions $\{l_l, l_w, l_h\}$ as shown in Fig. 6.

We predict the force in the vertical direction scaled by body weight $F_Y^W = F_Y/W$, and CoP along the foot scaled by foot length $z_x^l = z_x/l_l$, from initial estimates based on the foot’s kinematics Ψ and linear coefficients η . Furthermore, let μ be the **coefficient of friction**, initialized at 0.8. With our regressed residuals δ , and binary contact \mathbf{c} , we infer:

$$\{F_Y^W, z_x^l, \mu\} = \{\eta_{FY} \Psi, \eta_{zx} \Psi, 0.8\} + \delta_{\{Y,l,\mu\}} \quad (24)$$

$$F_Y = F_Y^W \cdot \text{body weight} \quad (25)$$

$$z_x = z_x^l \cdot l_l \quad (26)$$

Specifically, $\Psi = [1, P_{ankle,Y}, P_{oppAnkle,Y}, \dot{P}_{ankle,Y}, \ddot{P}_{ankle,Y}, q_{ankle,z}, \dot{q}_{ankle,z}, \ddot{q}_{ankle,z}]$ includes the ankle's linear kinematics in the direction opposite of gravity, and angular kinematics corresponding to plantar/dorsiflexion. Linear coefficients were fitted on the forceplate data in [75], with $\eta_{FY} = [0.3116, 3.1785, -2.2963, 0.4151, 0.0088, 0.3374, -0.1206, -0.0089]$ and $\eta_{zx} = [0.68996, -3.1508, 0.5925, 0.21997, 0.0035, 0.18502, -0.03311, -0.00212]$.

The remaining δ terms are scaling factors between -1 and 1 to ensure the values in the other directions are physically possible (i.e. $F_X^2 + F_Z^2 \leq \mu^2 F_Y^2$ and z is within the foot's dimensions)

$$F_X = \delta_X \mu F_Y, \quad F_Z = \delta_Z \sqrt{\mu^2 F_Y^2 - F_X^2} \quad (27)$$

$$z_y = -|\delta_h l_h|, \quad z_z = \delta_s (l_w/2) \quad (28)$$

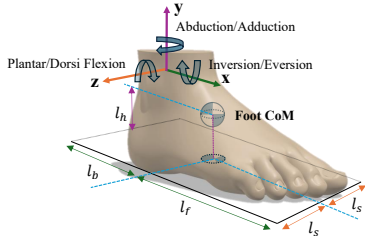


Figure 6: Foot local coordinate system and dimensions, with length $l_l = l_f + l_b$, width $l_w = 2l_s$, and CoM height l_h .

A.3 Muscle torque generators

We compute MTG torque τ_{MTG} using the equations below that are parameterized by the γ coefficients that can be found in the tables of [56, 55]. For each joint rotational DoF $q \in \mathbf{q}_{[6:]}$, with angular velocity \dot{q} , let muscle signal $\alpha \in [0, 1]$ represent the joint's corresponding activation level for this DoF, we separate its τ_{MTG} into active torque generation τ_{active} and passive impedance $\tau_{passive}$

$$\tau_{MTG} = \tau_{active} + \tau_{passive} \quad (29)$$

We compute the **active torque** as

$$\tau_{active} = \alpha \tau_\omega \tau_\theta \tau_0 \quad (30)$$

where τ_ω models the **active-torque-angular-speed relationship** [72, 69] and is parameterized as a piecewise function with coefficients $\gamma_{1:3}$.

$$\tau_\omega(\dot{q}) = \mathbb{1}_{\dot{q} < 0} \left(\frac{(1 - \gamma_1)|\omega_{max}| - (\gamma_2 + 1)\gamma_1\gamma_3\dot{q}}{(1 - \gamma_1)|\omega_{max}| + (\gamma_2 + 1)\gamma_1\dot{q}} \right) + \mathbb{1}_{\dot{q} \geq 0} \left(\frac{|\omega_{max}| - \dot{q}}{|\omega_{max}| + \gamma_2\dot{q}} \right) \quad (31)$$

The peak velocity ω_{max} for each joint we use the values from [56, 55]. The coefficient γ_1 is the ratio of the maximum eccentric isokinetic torque over the maximum isometric torque [72, 15], γ_2 is the slope of the eccentric and concentric functions when the angular velocity is zero [72], and γ_3 is a shape factor that influences the curvature of the hyperbola in the torque-velocity concentric relationship [4].

τ_θ models the **active-torque-angle relationship** [21, 34] and is represented by the non-negative portion of a polynomial (32) with coefficients $\gamma_{4:6}$

$$\tau_\theta(q) = (\gamma_4 + \gamma_5 q + \gamma_6 q^2)_+ \quad (32)$$

τ_0 is the **peak isokinetic torque** that controls peak MTG output at zero joint velocity, which can be measured via dynamometry.

$\tau_{passive}$ is the **passive torque** [1] of a joint that arises when the surrounding muscles, tendons, and ligaments are strained and intensifies near anatomical joint limits [1, 24, 85]. A joint's viscous damping and nonlinear stiffness are commonly described by a double exponential function [83]

$$\tau_{passive} = \gamma_{10e}^{-\gamma_{11}(q - q_{min})} - \gamma_{12}e^{\gamma_{13}(q - q_{max})} - \gamma_{14}\omega \quad (33)$$

where γ_{10-14} are passive coefficients from [50] and γ_{11} is the rotational damping linear coefficient [82] to reflect viscoelasticity. This encourages the joint to move within its **range of motion** (RoM), as a large restoring torque is produced otherwise.

A.4 Inverse dynamics ¹

We compute τ_q using Lagrange's equations derived from d'Alembert's Principle of virtual work

$$\tau_q = \mathfrak{M}\ddot{\mathbf{q}} + \mathfrak{C} - \mathfrak{F} \quad (34)$$

We can write the terms on the right hand side as:

$$\mathfrak{M} = \sum_k J_k^\top \mathcal{M}_k J_k, \quad (35)$$

$$\mathfrak{C} = \sum_k (J_k^\top \mathcal{M}_k \dot{J}_k + J_k^\top \begin{bmatrix} 0 & 0 \\ 0 & [J_{\Omega,k} \dot{\mathbf{q}}]_s \end{bmatrix} \mathcal{M}_k J_k) \dot{\mathbf{q}} \quad (36)$$

$$\mathfrak{F} = J_{LFoot}^\top \mathcal{F}_{LFoot} + J_{RFoot}^\top \mathcal{F}_{RFoot} \quad (37)$$

where \mathbf{I}_3 is the identity matrix, $([\cdot]_s)$ denotes the skew-symmetric form, and

$$\mathcal{M}_k = \begin{bmatrix} m_k \mathbf{I}_3 & 0 \\ 0 & R_k^0 I_{0,k} (R_k^0)^\top \end{bmatrix} \quad (38)$$

To deal with potential energy, we offset the root acceleration in the direction of gravity by -9.8 m/s^2 . Jacobian matrix J is the mapping from the generalized space to the global Cartesian coordinates, such that for linear and angular velocities $\mathbf{V}_k, \mathbf{\Omega}_k$ in global Cartesian coordinates, we have:

$$J_k \dot{\mathbf{q}} = \begin{bmatrix} J_{V,k} \\ J_{\Omega,k} \end{bmatrix} \dot{\mathbf{q}} = \begin{bmatrix} \mathbf{V}_k \\ \mathbf{\Omega}_k \end{bmatrix} \quad (39)$$

J can be computed analytically using a recursive algorithm such as in [17]. For segment k , we define its parent segment $p(k)$ as its neighboring segment that is closer to the root. Other than the root, each segment has one and only one parent. We define k 's children $ch(k)$ as its neighboring segments further away from the root. Let $\mathbf{r}_{a \rightarrow b}$ denote the 3D displacement from point a to b . For segment k , with linear velocity \mathbf{V}_k at its CoM and linear velocity \mathbf{V}_k^{joint} at its corresponding joint, we have

$$\mathbf{V}_k^{joint} = \mathbf{V}_{p(k)} + \mathbf{\Omega}_{p(k)} \times \mathbf{r}_{p(k) \rightarrow k^{joint}} \Rightarrow J_{V,k}^{joint} = J_{V,p(k)} - [\mathbf{r}_{p(k) \rightarrow k^{joint}}] J_{\Omega,p(k)} \quad (40)$$

and the velocity at the CoM of segment k becomes:

$$\mathbf{V}_k = \mathbf{V}_k^{joint} + \mathbf{\Omega}_k \times \mathbf{r}_{k^{joint} \rightarrow k} \Rightarrow J_{V,k} = J_{V,p(k)} - [\mathbf{r}_{p(k) \rightarrow k^{joint}}] J_{\Omega,p(k)} - [\mathbf{r}_{k^{joint} \rightarrow k}] J_{\Omega,k} \quad (41)$$

From (41), we can compute the time derivative recursively as:

$$\dot{J}_{V,k} = \dot{J}_{V,p(k)} - [\dot{\mathbf{r}}_{p(k) \rightarrow k^{joint}}] \dot{J}_{\Omega,p(k)} - [\mathbf{r}_{k^{joint} \rightarrow k}] \dot{J}_{\Omega,k} \quad (42)$$

The global angular velocity of k in skew symmetric form is:

$$[\mathbf{\Omega}_k] = \dot{R}_k^0 (R_k^0)^\top = (\dot{R}_{p(k)}^0 R_k^0) (R_{p(k)}^0 R_k^0)^\top \quad (43)$$

$$= \dots = \dot{R}_{p(k)}^0 (R_{p(k)}^0)^\top + R_{p(k)}^0 (\dot{R}_k^0 R_k^\top) (R_{p(k)}^0)^\top \quad (44)$$

$$= [\mathbf{\Omega}_{p(k)}] + R_{p(k)}^0 [\boldsymbol{\omega}_k] (R_{p(k)}^0)^\top \quad (\because [\boldsymbol{\omega}_k] = \dot{R}_k^0 R_k^\top) \quad (45)$$

$$\Rightarrow \mathbf{\Omega}_k = \mathbf{\Omega}_{p(k)} + R_{p(k)}^0 \boldsymbol{\omega}_k \quad (\because [Ab] = A[b]A^\top) \quad (46)$$

To avoid confusion of notation, we also write joint k 's rotation $\boldsymbol{\theta}_k \triangleq \mathbf{q}_k$, i.e. we have generalized coordinates $\mathbf{q} = \begin{bmatrix} \mathbf{X}_0 \\ \boldsymbol{\theta} \end{bmatrix} \in \mathbb{R}^{N_{Dof} \times 1}$ where \mathbf{X}_0 is the global root translation, $\boldsymbol{\theta}_0$ is the global root rotation, $\boldsymbol{\theta}_k$ describes the local rotation of segment k relative to its parent $p(k)$, and $\boldsymbol{\theta}^\top = [\boldsymbol{\theta}_0^\top \ \boldsymbol{\theta}_1^\top \ \dots \ \boldsymbol{\theta}_{N_k}^\top]$. Let $J_{\omega,k}$ be the local Jacobian such that $\boldsymbol{\omega}_k = J_{\omega,k} \dot{\boldsymbol{\theta}}_k$. We can compute $\mathbf{\Omega}_k$ recursively:

$$\mathbf{\Omega}_k = \mathbf{\Omega}_{p(k)} + R_{p(k)}^0 J_{\omega,k} \dot{\boldsymbol{\theta}}_k \quad (47)$$

$$= 0 + J_{\omega,0} \dot{\boldsymbol{\theta}}_0 + \dots + R_{p(p(k))}^0 J_{\omega,p(k)} \dot{\boldsymbol{\theta}}_{p(k)} + R_{p(k)}^0 J_{\omega,k} \dot{\boldsymbol{\theta}}_k \quad (48)$$

$$\triangleq J_{\Omega,k} \dot{\mathbf{q}} \quad (49)$$

¹Derivations in this section are based on C. Karen Liu and Sumit Jain's multibody dynamics notes: https://fab.cba.mit.edu/classes/865.18/design/optimization/dynamics_1.pdf.

Let \mathcal{P}_k denote the set of all ancestors of k and itself ($k \in \mathcal{P}_k$), we split $J_{\Omega,k}$ into $N_k + 1$ blocks of size 3×3 :

$$J_{\Omega,k} = \begin{bmatrix} 0_{3 \times 3} & J_{\omega,0} & \mathbb{1}_{1 \in \mathcal{P}_k} R_{p(1)}^0 J_{\omega,1} & \dots & \mathbb{1}_{N_k \in \mathcal{P}_k} R_{p(K)}^0 J_{\omega,N_k} \end{bmatrix} \in \mathbb{R}^{3 \times N_{D \circ F}} \quad (50)$$

For segment k , if we represent rotation $\theta_k = [\theta_{k,1} \ \theta_{k,2} \ \theta_{k,3}] \triangleq [\alpha \ \beta \ \gamma] \in \mathbb{R}^3$ using 3 Euler angles, removing subscript k for notation simplicity, we have

$$[\omega] = \dot{R}R^\top = \sum_i \frac{\partial R}{\partial \theta_i} R^\top \dot{\theta}_i = \frac{\partial R}{\partial \alpha} R^\top \dot{\alpha} + \frac{\partial R}{\partial \beta} R^\top \dot{\beta} + \frac{\partial R}{\partial \gamma} R^\top \dot{\gamma} \quad (51)$$

and it remains to compute \mathbf{J} 's s.t.

$$J_\omega \triangleq [\mathbf{J}_1 \ \mathbf{J}_2 \ \mathbf{J}_3], \quad \text{where } [\mathbf{J}_1] = \frac{\partial R}{\partial \alpha} R^\top, \quad [\mathbf{J}_2] = \frac{\partial R}{\partial \beta} R^\top, \quad [\mathbf{J}_3] = \frac{\partial R}{\partial \gamma} R^\top \quad (52)$$

$$\text{which satisfies } [\omega] = \sum_i [\mathbf{J}_i] \dot{\theta}_i \quad \text{and } \omega = J_\omega \dot{\theta} \quad (53)$$

Finally, let segment $l \in \mathcal{P}_k$ be an ancestor of k and denote $\{J_{\Omega,k}\}_l \triangleq R_{p(l)}^0 J_{\omega,l}$ as the $(l+2)$ -th 3×3 block in $J_{\Omega,k}$ from (50), we can compute its time derivative as:

$$\{\dot{J}_{\Omega,k}\}_l = \dot{R}_{p(l)}^0 J_{\omega,l} + R_{p(l)}^0 \dot{J}_{\omega,l}, \quad \text{with } \dot{J}_{\omega,l} = \sum_{l,i} \frac{\partial J_{\omega,l}}{\partial \theta_{l,i}} \dot{\theta}_{l,i} \quad (54)$$

following (52), it remains to compute $\dot{\mathbf{J}}$'s s.t.

$$\dot{J}_{\omega,l} \triangleq [\dot{\mathbf{J}}_{l,1} \ \dot{\mathbf{J}}_{l,2} \ \dot{\mathbf{J}}_{l,3}], \quad \text{where } \dot{\mathbf{J}}_{l,j} = \sum_{l,i} \frac{\partial \mathbf{J}_{l,j}}{\partial \theta_{l,i}} \dot{\theta}_{l,i} \quad (55)$$

A.5 Neural network

We trained a transformer encoder consisting of 8 layers with a latent dimension of 256, using total loss \mathcal{L}_{total} with weights $\lambda_{kin} = [0.5, 10, 1000, 1, 20, 1000]$ and $\lambda_{dyn} = [100, 100, 20]$. We trained on AMASS with a sequence input length of 16 frames, after removing sequences containing non foot-ground contact, with contact labels from [87], for 25 epochs. We used the AdamW optimizer [46] with a weight decay of 10^{-4} and an initial learning rate of 10^{-4} that decreases by 20% every 5 epochs. The entire process can be trained in about 12 hours on a single Titan Xp GPU.

B Additional discussion

Limitations. Overall, we introduced MusclePose as a prototype for human pose estimation at a more comprehensive musculoskeletal level. While this current version employs basic physics models, more complex and SOTA models could be swapped in and explored. For one, our experiments were trained on foot-ground contact labels from [87], but could be extended to full body-ground contact cases using precomputed pseudo contact labels, and employing a full body GRFM model. The SMPL kinematic model can also be replaced by more biomechanically accurate ones such as SKEL [80, 33]. Furthermore, we are working on collecting force-annotated video data from athletes, to better assess physics pose estimators.

Societal impact. Our research hopes to create accessible biomechanical analyses for all, such as for amateur athletes and rehab patients to track their dynamics from a monocular video (*e.g.* a smartphone camera), without the need for expensive and intrusive physical sensors, that may only be available in specialized labs. We foresee minimal negative societal impact from this work.

C Datasets and code

While our code is not yet ready for open source, our method and experiments can be reproduced in PyTorch² based on the publicly available datasets and code repositories that we heavily borrowed,

²PyTorch license: <https://github.com/pytorch/pytorch/blob/main/LICENSE>.

listed below. All physics models in this paper can be implemented based on the equations in Sec. 3 and Sec. A

Datasets

- **Training data:**

AMASS [48], url: <https://amass.is.tue.mpg.de>,
license: <https://github.com/nghorbani/amass?tab=License-1-ov-file#readme>

- **Evaluation data:**

Human3.6M [27], url: <http://vision.imar.ro/human3.6m>,
license: <http://vision.imar.ro/human3.6m/eula.php>

3DPW [74], url: <https://virtualhumans.mpi-inf.mpg.de/3DPW/>,
license: <https://virtualhumans.mpi-inf.mpg.de/3DPW/license.html>

PennAction [88], url: <http://dreamdragon.github.io/PennAction/>,
license: <http://dreamdragon.github.io/PennAction/>

Additional qualitative comparison with biomechanics results from [75, 18, 79, 8, 29, 78, 57, 84, 64].

- **Model coefficients used:**

Anthropometrics from [16], url: <https://www.tandfonline.com/doi/full/10.1080/23335432.2015.1112244>,
license: <https://www.tandfonline.com/action/showCopyRight?scroll=top&doi=10.1080%2F23335432.2015.1112244>

GRFM coefficients fitted on [75], url: <https://zenodo.org/records/6457662>,
license: <https://zenodo.org/records/6457662>

MTG coefficients from [56, 55], url: <https://link.springer.com/article/10.1007/s11044-024-10021-5>,
license: <https://link.springer.com/article/10.1007/s11044-024-10021-5>

Code

- **Neural network:** for implementation, as well as training and evaluation, we heavily borrow from:

SMPL [45], url and license: <https://github.com/vchoutas/smplx>.

SPIN [36], url and license: <https://github.com/nkolot/SPIN>.

MDM [70], url: <https://github.com/GuyTevet/motion-diffusion-model>, MIT license.

CLIFF [41], url and license: <https://github.com/huawei-noah/noah-research/tree/master/CLIFF>.

PhysPT [89], url: <https://github.com/zhangy76/PhysPT>, MIT license.

WHAM [67], url: <https://github.com/yohanshin/WHAM>, MIT license.

- **Physics models:** all physics models in this paper can be implemented based on equations in Sec. 3 and Sec. A.

References

- [1] D. E. Anderson, M. L. Madigan, and M. A. Nussbaum. Maximum voluntary joint torque as a function of joint angle and angular velocity: Model development and application to the lower limb. *Journal of Biomechanics*, 40(14):3105–3113, 2007.
- [2] Y. Bengio, J. Louradour, R. Collobert, and J. Weston. Curriculum learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, page 41–48, 2009.
- [3] N. A. Bernshtein. The co-ordination and regulation of movements. 1967.
- [4] C. Brown, W. McNally, and J. McPhee. Optimal control of joint torques using direct collocation to maximize ball carry distance in a golf swing. *Multibody System Dynamics*, 50(3), 2020.
- [5] C. Brown and J. J. McPhee. Predictive forward dynamic simulation of manual wheelchair propulsion on a rolling dynamometer. *Journal of biomechanical engineering*, 2020.
- [6] A. Buchmann, S. Wenzler, L. Welte, and D. Renjewski. The effect of including a mobile arch, toe joint, and joint coupling on predictive neuromuscular simulations of human walking. *Scientific Reports*, 2024.
- [7] Y. Cai, L. Ge, J. Liu, J. Cai, T. J. Cham, J. Yuan, and N. M. Thalmann. Exploiting spatial-temporal relationships for 3D pose estimation via graph convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2019-October, 2019.
- [8] T. Castermans, M. Duvinage, G. Cheron, and T. Dutoit. Towards effective non-invasive brain-computer interfaces dedicated to gait rehabilitation systems. *Brain Sciences*, 4(1):1–48, 2014.
- [9] S.-W. Chen, W.-T. Tang, J.-T. Kung, T.-Y. Hung, W.-H. Lin, Y.-L. Chen, and D. J. Burgee. Comparison of ground reaction force among stride types in baseball pitching. *Sports Biomechanics*, 0(0):1–14, 2024.
- [10] Y. Cheng, B. Yang, B. Wang, and R. T. Tan. 3D human pose estimation using spatio-temporal networks with explicit occlusion training. In *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, 2020.
- [11] H. Ci, C. Wang, X. Ma, and Y. Wang. Optimizing network structure for 3D human pose estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2019-October, 2019.
- [12] D. J. Cleather and A. M. J. Bull. Lower-extremity musculoskeletal geometry affects the calculation of patellofemoral forces in vertical jumping and weightlifting. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224:1073 – 1083, 2010.
- [13] B. Danaei and J. J. McPhee. Model-based acetabular cup orientation optimization based on minimizing the risk of edge-loading and implant impingement following total hip arthroplasty. *Journal of biomechanical engineering*, 2022.
- [14] R. Drillis, R. Contini, and M. Bluestein. Body segment parameters. *Artificial limbs*, 8(1):44–66, 1964.
- [15] G. A. Dudley, R. T. Harris, M. R. Duvoisin, B. M. Hather, and P. Buchanan. Effect of voluntary vs. artificial activation on the relationship of muscle torque to speed. *Journal of Applied Physiology*, 69(6), 1990.
- [16] R. Dumas, L. Chèze, and J. P. Verriest. Adjustments to mcconville et al. and young et al. body segment inertial parameters. *Journal of biomechanics*, 40 3:543–53, 2007.
- [17] R. Featherstone. *Rigid Body Dynamics Algorithms*. 2008.
- [18] C. A. Fukuchi, R. K. Fukuchi, and M. Duarte. A public dataset of overground and treadmill walking kinematics and kinetics in healthy individuals. *PeerJ*, 6, 2018.
- [19] E. Gartner, M. Andriluka, H. Xu, and C. Sminchisescu. Trajectory Optimization for Physics-Based Reconstruction of 3d Human Pose from Monocular Video. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2022-June, 2022.
- [20] E. Gärtner, M. Andriluka, E. Coumans, and C. Sminchisescu. Differentiable dynamics for articulated 3d human motion reconstruction. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2022.
- [21] D. Haering, C. Pontonnier, N. Bideau, G. Nicolas, and G. Dumont. Using Torque-Angle and Torque-Velocity Models to Characterize Elbow Mechanical Function: Modeling and Applied Aspects. *Journal of Biomechanical Engineering*, 141(8), 2019.

- [22] N. Haraguchi, A. Nasr, K. A. Inkol, K. Hase, and J. McPhee. Human and passive lower-limb exoskeleton interaction analysis: Computational study with dynamics simulation using nonlinear model predictive control. *2023 62nd Annual Conference of the Society of Instrument and Control Engineers (SICE)*, pages 844–849, 2023.
- [23] A. V. Hill. The heat of shortening and the dynamic constants of muscle. *Proceedings of The Royal Society B: Biological Sciences*, 126:136–195, 1938.
- [24] P. D. Hoang, R. B. Gorman, G. Todd, S. C. Gandevia, and R. D. Herbert. A new method for measuring passive length-tension properties of human gastrocnemius muscle in vivo. *Journal of Biomechanics*, 38(6):1333–1341, 2005.
- [25] K. A. Inkol, C. Brown, W. McNally, C. Jansen, and J. McPhee. Muscle torque generators in multibody dynamic simulations of optimal sports performance. *Multibody System Dynamics*, 50(4), 2020.
- [26] K. A. Inkol and J. J. McPhee. Using dynamic simulations to estimate the feasible stability region of feet-in-place balance recovery for lower-limb exoskeleton users. *2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, pages 1–6, 2022.
- [27] C. Ionescu, D. Papava, V. Olaru, and C. 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), 2014.
- [28] C. Jansen and J. J. McPhee. Predictive dynamic simulation of olympic track cycling standing start using direct collocation optimal control. *Multibody System Dynamics*, 49:53–70, 2020.
- [29] C. T. John, F. C. Anderson, J. S. Higginson, and S. L. D. and. Stabilisation of walking by intrinsic muscle properties revealed in a three-dimensional muscle-driven simulation. *Computer Methods in Biomechanics and Biomedical Engineering*, 16(4):451–462, 2013.
- [30] S. Johnson and M. Everingham. Clustered pose and nonlinear appearance models for human pose estimation. In *British Machine Vision Conference*, 2010.
- [31] A. Kanazawa, M. J. Black, D. W. Jacobs, and J. Malik. End-to-End Recovery of Human Shape and Pose. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018.
- [32] B. Katz. The relation between force and speed in muscular contraction. *The Journal of Physiology*, 96, 1939.
- [33] M. Keller, K. Werling, S. Shin, S. Delp, S. Pujades, C. K. Liu, and M. J. Black. From skin to skeleton: Towards biomechanically accurate 3D digital humans. *ACM Transaction on Graphics (ToG)*, 42(6):253:1–253:15, Dec. 2023.
- [34] M. A. King, C. Wilson, and M. R. Yeadon. Evaluation of a torque-driven model of jumping for height. *Journal of Applied Biomechanics*, 22(4), 2006.
- [35] M. Kocabas, N. Athanasiou, and M. J. Black. Vibe: Video inference for human body pose and shape estimation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020.
- [36] N. Kolotouros, G. Pavlakos, M. J. Black, and K. Daniilidis. Learning to reconstruct 3d human pose and shape via model-fitting in the loop. In *Proceedings of the International Conference on Computer Vision*, 2019.
- [37] N. Kolotouros, G. Pavlakos, and K. Daniilidis. Convolutional mesh regression for single-image human shape reconstruction. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4496–4505, 2019.
- [38] J. Li, S. Bian, C. Xu, G. Liu, G. Yu, and C. Lu. D & D: Learning Human Dynamics from Dynamic Camera. In *ECCV*, 2022.
- [39] J. Li, C. Xu, Z. Chen, S. Bian, L. Yang, and C. Lu. Hybrik: A hybrid analytical-neural inverse kinematics solution for 3d human pose and shape estimation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021.
- [40] W. Li, H. Liu, H. Tang, P. Wang, and L. Van Gool. MHFormer: Multi-Hypothesis Transformer for 3D Human Pose Estimation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2022-June, 2022.

- [41] Z. Li, J. Liu, Z. Zhang, S. Xu, and Y. Yan. Cliff: Carrying location information in full frames into human pose and shape estimation. In *European Conference on Computer Vision*, 2022.
- [42] Z. Li, J. Sedlar, J. Carpentier, I. Laptev, N. Mansard, and J. Sivic. Estimating 3D motion and forces of person-object interactions from monocular video. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2019-June, 2019.
- [43] K. Lin, L. Wang, and Z. Liu. Mesh graphormer. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12919–12928, 2021.
- [44] D. M. Lindsay and J. F. Horton. Trunk rotation strength and endurance in healthy normals and elite male golfers with and without low back pain. *North American journal of sports physical therapy : NAJSPT*, 1 2:80–9, 2006.
- [45] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black. SMPL: A skinned multi-person linear model. In *ACM Transactions on Graphics*, volume 34, 2015.
- [46] I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2017.
- [47] T. Luan, Y. Wang, J. Zhang, Z. Wang, Z. Zhou, and Y. Qiao. Pc-hmr: Pose calibration for 3d human mesh recovery from 2d images/videos. In *AAAI Conference on Artificial Intelligence*, 2021.
- [48] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. Black. AMASS: Archive of motion capture as surface shapes. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2019-October, 2019.
- [49] J. Martinez, R. Hossain, J. Romero, and J. J. Little. A Simple Yet Effective Baseline for 3d Human Pose Estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2017-October, 2017.
- [50] W. McNally and J. McPhee. Dynamic Optimization of the Golf Swing Using a Six Degree-of-Freedom Biomechanical Model. 2018.
- [51] W. J. McNally and J. J. McPhee. Dynamic optimization of the golf swing using a six degree-of-freedom biomechanical model. 2018.
- [52] D. Mehta, H. Rhodin, D. Casas, P. Fua, O. Sotnychenko, W. Xu, and C. Theobalt. Monocular 3D human pose estimation in the wild using improved CNN supervision. In *Proceedings - 2017 International Conference on 3D Vision, 3DV 2017*, 2018.
- [53] M. Millard, T. K. Uchida, A. Seth, and S. L. Delp. Flexing computational muscle: modeling and simulation of musculotendon dynamics. *Journal of biomechanical engineering*, 135 2:021005, 2013.
- [54] G. Moon and K. M. Lee. I2l-meshnet: Image-to-voxel prediction network for accurate 3d human pose and mesh estimation from a single rgb image. *ArXiv*, abs/2008.03713, 2020.
- [55] A. Nasr, A. Hashemi, and J. McPhee. Scalable musculoskeletal model for dynamic simulations of upper body movement. *Computer Methods in Biomechanics and Biomedical Engineering*, 2023.
- [56] A. Nasr and J. McPhee. Scalable musculoskeletal model for dynamic simulations of lower body movement. *Computer methods in biomechanics and biomedical engineering*, pages 1–27, 2024.
- [57] S. Nesbit. Development of a full-body biomechanical model of the golf swing. *International Journal of Modelling and Simulation*, 27(4):392–404, 2007.
- [58] A. Newell, P. Hu, L. Lipson, S. R. Richter, and V. Koltun. Comotion: Concurrent multi-person 3d motion. In *ICLR*, 2025.
- [59] G. Pavlakos, X. Zhou, K. G. Derpanis, and K. Daniilidis. Coarse-to-fine volumetric prediction for single-image 3D human pose. In *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, volume 2017-January, 2017.
- [60] D. Pavlo, C. Feichtenhofer, D. Grangier, and M. Auli. 3D human pose estimation in video with temporal convolutions and semi-supervised training. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2019-June, 2019.
- [61] X. B. Peng, P. Abbeel, S. Levine, and M. Van De Panne. DeepMimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions on Graphics*, 37(4), 2018.

- [62] E. D. Pieri, M. Lund, A. Gopalakrishnan, K. P. Rasmussen, D. E. Lunn, and S. J. Ferguson. Refining muscle geometry and wrapping in the tlem 2 model for improved hip contact force prediction. *PLoS ONE*, 13, 2018.
- [63] D. Rempe, L. J. Guibas, A. Hertzmann, B. Russell, R. Villegas, and J. Yang. Contact and Human Dynamics from Monocular Video. In *19th ACM SIGGRAPH / Eurographics Symposium on Computer Animation 2020, SCA 2020 - Showcases*, 2020.
- [64] M. Sabick, M. Torry, Y.-K. Kim, and R. Hawkins. Humeral torque in professional baseball pitchers. *The American journal of sports medicine*, 32:892–8, 07 2004.
- [65] S. Shimada, V. Golyanik, W. Xu, P. Pérez, and C. Theobalt. Neural monocular 3D human motion capture with physical awareness. *ACM Transactions on Graphics*, 40(4), 2021.
- [66] S. Shimada, V. Golyanik, W. Xu, and C. Theobalt. Phys Cap. *ACM Transactions on Graphics*, 39(6), 2020.
- [67] S. Shin, J. Kim, E. Halilaj, and M. J. Black. Wham: Reconstructing world-grounded humans with accurate 3d motion. In *CVPR*, 2024.
- [68] T. Siebert, C. Rode, W. Herzog, O. Till, and R. Blickhan. Nonlinearities make a difference: comparison of two common hill-type models with real muscle. *Biological Cybernetics*, 98:133–143, 2008.
- [69] E. J. Sprigings. Simulation of the force enhancement phenomenon in muscle. *Computers in Biology and Medicine*, 16(6), 1986.
- [70] G. Tevet, S. Raab, B. Gordon, Y. Shafir, A. H. Bermanno, and D. Cohen-Or. Human motion diffusion model. *ArXiv*, abs/2209.14916, 2022.
- [71] S. Tripathi, L. Müller, C.-H. P. Huang, O. Taheri, M. J. Black, and D. Tzionas. 3d human pose estimation via intuitive physics. *ArXiv*, abs/2303.18246, 2023.
- [72] A. J. van Soest and M. F. Bobbert. The contribution of muscle properties in the control of explosive movements. *Biological Cybernetics*, 69(3), 1993.
- [73] G. Varol, D. Ceylan, B. C. Russell, J. Yang, E. Yumer, I. Laptev, and C. Schmid. Bodynet: Volumetric inference of 3d human body shapes. *ArXiv*, abs/1804.04875, 2018.
- [74] T. von Marcard, R. Henschel, M. Black, B. Rosenhahn, and G. Pons-Moll. Recovering accurate 3d human pose in the wild using imus and a moving camera. In *European Conference on Computer Vision (ECCV)*, sep 2018.
- [75] H. Wang, A. Basu, G. Durandau, and M. Sartori. Comprehensive Kinetic and EMG Dataset of Daily Locomotion with 6 types of Sensors, May 2022.
- [76] H. Wang, A. Basu, G. Durandau, and M. Sartori. A wearable real-time kinetic measurement sensor setup for human locomotion. *Wearable Technologies*, Feb. 2023.
- [77] J. Wang, S. Yan, Y. Xiong, and D. Lin. Motion Guided 3D Pose Estimation from Videos. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 12358 LNCS, 2020.
- [78] D. A. Winter. *Biomechanics and motor control of human movement*. John Wiley & Sons, Hoboken, NJ, USA, 4th edition, 2009.
- [79] A. R. Wu, F. Dzeladini, T. J. H. Brug, F. Tamburella, N. L. Tagliamonte, E. H. F. van Asseldonk, H. van der Kooij, and A. J. Ijspeert. An adaptive neuromuscular controller for assistive lower-limb exoskeletons: A preliminary study on subjects with spinal cord injury. *Frontiers in Neurorobotics*, Volume 11 - 2017, 2017.
- [80] Y. Xia, X. Zhou, E. Vouga, Q. Huang, and G. Pavlakos. Reconstructing humans with a biomechanically accurate skeleton. In *CVPR*, 2025.
- [81] K. Xie, T. Wang, U. Iqbal, Y. Guo, S. Fidler, and F. Shkurti. Physics-based Human Motion Estimation and Synthesis from Videos. In *Proceedings of the IEEE International Conference on Computer Vision*, 2021.
- [82] G. T. Yamaguchi. *Dynamic Modeling of Musculoskeletal Motion*. 2001.
- [83] G. T. Yamaguchi. *Dynamic modeling of musculoskeletal motion: A vectorized approach for biomechanical analysis in three dimensions*. Springer, Boston, MA, USA, 1 edition, 2006.

- [84] K. Yoichi, E. Sato, and T. Yamaji. Biomechanical analysis of the pitching characteristics of adult amateur baseball pitchers throwing standard and lightweight balls. *Journal of Physical Therapy Science*, 32:816–822, 12 2020.
- [85] Y. S. Yoon and J. M. Mansour. The passive elastic moment at the hip. *Journal of Biomechanics*, 15(12):905–910, 1982.
- [86] Y. Yuan, S.-E. Wei, T. Simon, K. Kitani, and J. Saragih. Simpoe: Simulated character control for 3d human pose estimation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021.
- [87] S. Zhang, B. L. Bhatnagar, Y. Xu, A. Winkler, P. Kadlec, S. Tang, and F. Bogo. Rohm: Robust human motion reconstruction via diffusion. In *CVPR*, 2024.
- [88] W. Zhang, M. Zhu, and K. G. Derpanis. From actemes to action: A strongly-supervised representation for detailed action understanding. In *2013 IEEE International Conference on Computer Vision*, pages 2248–2255, 2013.
- [89] Y. Zhang, J. O. Kephart, Z. Cui, and Q. Ji. Physpt: Physics-aware pretrained transformer for estimating human dynamics from monocular videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2305–2317, June 2024.
- [90] C. Zheng, S. Zhu, M. Mendieta, T. Yang, C. Chen, and Z. Ding. 3D Human Pose Estimation with Spatial and Temporal Transformers. In *Proceedings of the IEEE International Conference on Computer Vision*, 2021.
- [91] X. Zhou, Q. Huang, X. Sun, X. Xue, and Y. Wei. Towards 3D Human Pose Estimation in the Wild: A Weakly-Supervised Approach. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2017-October, 2017.
- [92] Y. Zhou, C. Barnes, J. Lu, J. Yang, and H. Li. On the continuity of rotation representations in neural networks. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019.
- [93] W. Zhu, X. Ma, Z. Liu, L. Liu, W. Wu, and Y. Wang. Learning human motion representations: A unified perspective. In *Proceedings of the International Conference on Computer Vision*, 2023.