Supplementary Material of "Lagrange Motion Analysis and View Embeddings for Improved Gait Recognition"

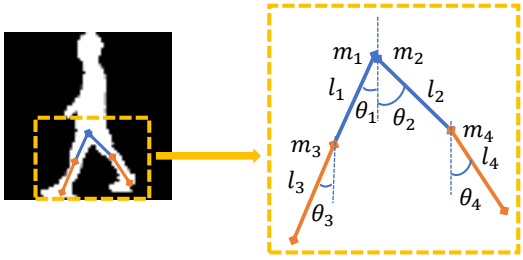

Figure 1. Analysis diagram of pedestrian walks.

A. Simplified version of Lagrange Motion Analysis

Assuming that the human thighs and legs are rigid and model them mechanically, the length and mass of the two thighs and two legs are denoted as l_1, l_2, m_1, m_2 and l_3, l_4, m_3, m_4 respectively. θ_i represents the angle between them and vertical lines. Also, the human body is assumed to move forward with a small distance x.

The translational kinetic energy(TKE) of the whole system is

$$T_{TKE} = \frac{1}{2}(m_1 + m_2 + m_3 + m_4)(\frac{\mathrm{d}x}{\mathrm{d}t})^2. \tag{1}$$

The rotational kinetic energy(RKE) of the four parts is similar, we take the part which length and mass are l_1 and m_1 as example:

$$\begin{aligned} T_{RKE,1} &= \frac{1}{2}\int_0^{l1} \frac{m_1}{l_1} z^2 (\frac{\mathrm{d}\theta_1}{\mathrm{d}t})^2 \mathrm{d}z \\ &= \frac{1}{2}\frac{m_1}{l_1}\frac{l_1^3}{3}(\frac{\mathrm{d}\theta_1}{\mathrm{d}t})^2 \\ &= \frac{1}{6}m_1 l_1^2 (\frac{\mathrm{d}\theta_1}{\mathrm{d}t})^2 \end{aligned} \tag{2}$$

Then we can obtain the kinetic energy T as:

$$\begin{aligned} T = &\frac{1}{2}(m_1 + m_2 + m_3 + m_4)(\frac{\mathrm{d}x}{\mathrm{d}t})^2 + \frac{1}{6}(m_1 l_1^2 (\frac{\mathrm{d}\theta_1}{\mathrm{d}t})^2 \\ &+ m_2 l_1^2 (\frac{\mathrm{d}\theta_2}{\mathrm{d}t})^2 + m_3 l_1^2 (\frac{\mathrm{d}\theta_3}{\mathrm{d}t})^2 + m_4 l_1^2 (\frac{\mathrm{d}\theta_4}{\mathrm{d}t})^2) \end{aligned}, \tag{3}$$

and the potential energy V as:

$$\begin{aligned} V = &-\frac{1}{2}m_1 g l_1 \cos\theta_1 - m_3 g(l_1 \cos\theta_1 + \frac{l_3}{2}\cos\theta_3) \\ &- \frac{1}{2}m_2 g l_2 \cos\theta_2 - m_4 g(l_2 \cos\theta_2 + \frac{l_4}{2}\cos\theta_4) \end{aligned}, \tag{4}$$

We calculate the L=T-V as:

$$\begin{aligned} L = &\frac{1}{2}(m_1 + m_2 + m_3 + m_4)(\frac{\mathrm{d}x}{\mathrm{d}t})^2 + \frac{1}{6}(m_1 l_1^2 (\frac{\mathrm{d}\theta_1}{\mathrm{d}t})^2 \\ &+ m_2 l_1^2 (\frac{\mathrm{d}\theta_2}{\mathrm{d}t})^2 + m_3 l_1^2 (\frac{\mathrm{d}\theta_3}{\mathrm{d}t})^2 + m_4 l_1^2 (\frac{\mathrm{d}\theta_4}{\mathrm{d}t})^2) \\ &+ \frac{1}{2}m_1 g l_1 \cos\theta_1 + m_3 g(l_1 \cos\theta_1 + \frac{l_3}{2}\cos\theta_3) \\ &+ \frac{1}{2}m_2 g l_2 \cos\theta_2 + m_4 g(l_2 \cos\theta_2 + \frac{l_4}{2}\cos\theta_4) \end{aligned}, \tag{5}$$

According to the Lagrange's Equation:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \frac{\mathrm{d}x}{\mathrm{d}t}}) - \frac{\partial L}{\partial x} &= Q_0 \\ \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \frac{\mathrm{d}\theta_1}{\mathrm{d}t}}) - \frac{\partial L}{\partial \theta_1} &= Q_1 \\ \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \frac{\mathrm{d}\theta_2}{\mathrm{d}t}}) - \frac{\partial L}{\partial \theta_2} &= Q_2, \\ \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \frac{\mathrm{d}\theta_3}{\mathrm{d}t}}) - \frac{\partial L}{\partial \theta_3} &= Q_3 \\ \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \frac{\mathrm{d}\theta_4}{\mathrm{d}t}}) - \frac{\partial L}{\partial \theta_4} &= Q_4 \end{aligned} \tag{6}$$

The final system can be formulated as:

$$\begin{aligned} (m_1 + m_2 + m_3 + m_4)\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} &= Q_0 \\ \frac{1}{3}m_1 l_1^2 \frac{\mathrm{d}^2\theta_1}{\mathrm{d}t^2} - \frac{1}{2}(m_1 + m_3) g l_1 \sin\theta_1 \frac{\mathrm{d}\theta_1}{\mathrm{d}t} &= Q_1 \\ \frac{1}{3}m_2 l_2^2 \frac{\mathrm{d}^2\theta_2}{\mathrm{d}t^2} - \frac{1}{2}(m_2 + m_4) g l_2 \sin\theta_2 \frac{\mathrm{d}\theta_2}{\mathrm{d}t} &= Q_2, \\ \frac{1}{3}m_3 l_3^2 \frac{\mathrm{d}^2\theta_3}{\mathrm{d}t^2} - \frac{1}{2}m_3 g l_3 \sin\theta_3 \frac{\mathrm{d}\theta_3}{\mathrm{d}t} &= Q_3 \\ \frac{1}{3}m_4 l_4^2 \frac{\mathrm{d}^2\theta_4}{\mathrm{d}t^2} - \frac{1}{2}m_4 g l_4 \sin\theta_4 \frac{\mathrm{d}\theta_4}{\mathrm{d}t} &= Q_4 \end{aligned} \tag{7}$$

where Q_0, Q_1, Q_2, Q_3, Q_4 are generalized force, including the force from human muscles and resistance. These force is the essence of pedestrian and they change gradually and continuously in a gait cycle.

B. Comprehensive analysis of Lagrange Motion

When using silhouettes, simplified version of the analysis is sufficient to represent human motion. However, this analysis ignores the linkage between the legs and thighs and the influence of the overall velocity of the human body on the position of the two legs.

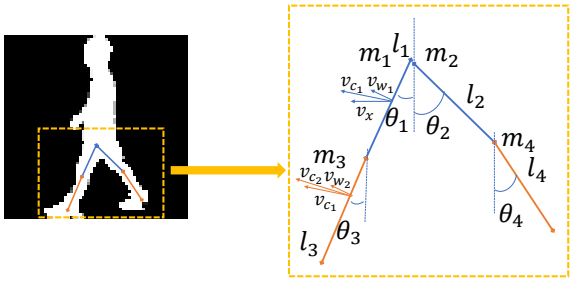

Figure 2. Analysis diagram of pedestrian walks.

Considering a more general situation, that is, the human body is not always in the center of the image, and the image can well reflect the displacement of the human body.

As shown in Figure 2, we analyze and combine the velocity of each part of legs. Considering the horizontal velocity, we change the center of rotation to the midpoint of each part. The translational velocity is $v_x = \frac{\mathrm{d}x}{\mathrm{d}t}$ and rotational velocity is $v_{w_1} = \frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{l_1}{2}$. Then we decompose the velocity into XY two dimensions and synthesize the resultant as:

$$v_{c_1}^2 = \frac{\mathrm{d}x}{\mathrm{d}t}^2 + (\frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{l_1}{2})^2 + \frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_1 cos\theta_1. \quad (8)$$

According to the Equation 8, the synthetic velocity of the lower leg can be formulated as:

$$\begin{aligned} v_{c_3}^2 &= (\frac{\mathrm{d}x}{\mathrm{d}t})^2 + (\frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{l_1}{2})^2 + \frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_1 cos\theta_1 + (\frac{\mathrm{d}\theta_3}{\mathrm{d}t}\frac{l_3}{2})^2 \\ &+ \sqrt{(\frac{\mathrm{d}x}{\mathrm{d}t})^2 + (\frac{\mathrm{d}\theta_1}{\mathrm{d}t})^2 + \frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_1 cos\theta_1}\frac{\mathrm{d}\theta_1}{\mathrm{d}t}l_3 cos(\theta_1 - \theta_3) \end{aligned}. \quad (9)$$

When people are walking, the horizontal velocity is much greater than the rotation velocity of the legs, so we ignore the quadratic power of the rotation speed which is expressed as $(\frac{\mathrm{d}\theta_i}{\mathrm{d}t})^2$.

The calculation results of $v_{c_2}^2$ and $v_{c_4}^2$ is similar to $v_{c_1}^2$ and $v_{c_3}^2$. Then the Kinetic energy can be formulated as:

$$\begin{aligned} T &= \frac{1}{2}(m_1 + m_2 + m_3 + m_4)(\frac{\mathrm{d}x}{\mathrm{d}t})^2 \\ &+ \frac{m_1}{2}\frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_1 cos\theta_1 + \frac{m_2}{2}\frac{\mathrm{d}\theta_2}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_2 cos\theta_2 + \\ &+ \frac{m_3}{2}\frac{\mathrm{d}\theta_3}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_3 cos\theta_3 + \frac{m_4}{2}\frac{\mathrm{d}\theta_4}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_4 cos\theta_4 \\ &+ \frac{m_3}{2}\sqrt{(\frac{\mathrm{d}x}{\mathrm{d}t})^2 + \frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_1 cos\theta_1}\frac{\mathrm{d}\theta_1}{\mathrm{d}t}l_3 cos(\theta_1 - \theta_3) \\ &+ \frac{m_4}{2}\sqrt{(\frac{\mathrm{d}x}{\mathrm{d}t})^2 + \frac{\mathrm{d}\theta_2}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_2 cos\theta_2}\frac{\mathrm{d}\theta_2}{\mathrm{d}t}l_4 cos(\theta_2 - \theta_4) \end{aligned}. \quad (10)$$

Combining Equation 10 and Equation 4, $L = T - V$ can be formulated as:

$$\begin{aligned} L &= \frac{1}{2}(m_1 + m_2 + m_3 + m_4)(\frac{\mathrm{d}x}{\mathrm{d}t})^2 \\ &+ \frac{m_1}{2}\frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_1 cos\theta_1 + \frac{m_2}{2}\frac{\mathrm{d}\theta_2}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_2 cos\theta_2 + \\ &+ \frac{m_3}{2}\frac{\mathrm{d}\theta_3}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_3 cos\theta_3 + \frac{m_4}{2}\frac{\mathrm{d}\theta_4}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_4 cos\theta_4 \\ &+ \frac{m_3}{2}\sqrt{(\frac{\mathrm{d}x}{\mathrm{d}t})^2 + \frac{\mathrm{d}\theta_1}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_1 cos\theta_1}\frac{\mathrm{d}\theta_1}{\mathrm{d}t}l_3 cos(\theta_1 - \theta_3) \\ &+ \frac{m_4}{2}\sqrt{(\frac{\mathrm{d}x}{\mathrm{d}t})^2 + \frac{\mathrm{d}\theta_2}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}l_2 cos\theta_2}\frac{\mathrm{d}\theta_2}{\mathrm{d}t}l_4 cos(\theta_2 - \theta_4) \\ &+ \frac{1}{2}m_1 g l_1 \cos\theta_1 + m_3 g(l_1 \cos\theta_1 + \frac{l_3}{2}\cos\theta_3) \\ &+ \frac{1}{2}m_2 g l_2 \cos\theta_2 + m_4 g(l_2 \cos\theta_2 + \frac{l_4}{2}\cos\theta_4) \end{aligned} \quad (11)$$

Then bring Equation 11 into Equation 6, the final system can be obtained. Since the form of final result is too complex, we do not list the result here. But it can also be concluded that we need second-order information to describe the system.

C. Experiments on Outdoor-Gait dataset

The effect of our module on OU-MVLP dataset [6] is not very significant. We think this is because there are no samples with large changes of appearance such as changing clothes and backpacks in OU-MVLP dataset [6]. Therefore, we do experiments on Outdoor-gait dataset [5].

Outdoor-Gait dataset contains 138 subjects. It is similar to CASIA-B dataset [8] since gait sequences of each subjects are collected under NM/BG/CL three conditions.

We reduce the parameters of GaitGL [3] to adapt to this dataset as the baseline. In detail, we change the first 3D convolution layer to two 2D convolution layers and remove the last partial 3D convolution layer of the network.

The results are shown in Table 1, our method outperform other SOTA methods and the baseline which verifys the effectiveness of our second-order motion extraction module. It is noteworthy that we can achieve good performance when only the extracted motion features are used for recognition. This is in line with the expectation, because there is only data of 90° in this dataset, which contains the most obvious motion information.

D. Comparison with model-based methods on CASIA-B dataset

We try to use only the extracted motion features for identity recognition. Since the motion feature does not contain apparent information, which is consistent with the model-based method, we compare it with the two best model-based methods on CASIA-B dataset [8].

It can be seen from Table 2, our method can obtain good performance with only second-order motion feature com-

Table 1. Rank-1 accuracy on Outdoor-Gait dataset(%).

Gallery	NM			CL			BG			Mean
Probe	NM	CL	BG	NM	CL	BG	NM	CL	BG	
GEI+PCA [1]	85.0	29.2	38.9	29.2	86.7	22.9	30.7	20.8	92.5	48.5
GEI-Net [4]	93.2	55.8	59.2	45.9	93.7	36.7	44.2	27.5	96.6	61.4
GaitNet [5]	96.9	60.2	**89.1**	58.7	97.3	55.4	92.0	59.7	97.1	78.5
Baseline	98.3	78.9	88.1	70.8	98.1	68.5	69.1	61.2	**99.7**	81.3
Baseline+Motion	**98.3**	**79.4**	88.6	67.9	98.3	**69.7**	**71.3**	**63.8**	99.5	**81.9**
Motion only	97.8	76.5	87.4	**68.4**	**98.5**	67.3	68.8	63.6	99.5	80.9

Table 2. Rank-1 accuracy on CASIA-B dataset(%) compared with model-based methods.

Method	Probe			Mean
	NM	BG	CL	
PoseGait [2]	68.7	44.5	36.0	49.7
GaitGraph [7]	**87.7**	**74.8**	**66.3**	**76.3**
Motion only	86.3	70.5	57.3	71.4

pared with PoseGait [2] and GaitGraph [7]. It should be pointed out that our method does not need additional information (pose annotation and pose estimation model), therefore the result can show the powerful the proposed second-order motion extraction module is.

References

[1] W. Liang, T. Tan, H. Ning, and W. Hu. Silhouette analysis-based gait recognition for human identification. *Pattern Analysis Machine Intelligence IEEE Transactions on*, 25(12):1505–1518, 2003. 3

[2] Rijun Liao, Shiqi Yu, Weizhi An, and Yongzhen Huang. A model-based gait recognition method with body pose and human prior knowledge. *Pattern Recognition*, 98:107069, 2020. 3

[3] Beibei Lin, Shunli Zhang, Xin Yu, Zedong Chu, and Haikun Zhang. Learning effective representations from global and local features for cross-view gait recognition. *arXiv preprint arXiv:2011.01461*, 2020. 2

[4] K. Shiraga, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi. Geinet: View-invariant gait recognition using a convolutional neural network. In *International Conference on Biometrics*, 2016. 3

[5] C. Song, Y. Huang, Y. Huang, N. Jia, and L. Wang. Gaitnet: An end-to-end network for gait based human identification. *Pattern Recognition*, 96:106988–, 2019. 2, 3

[6] Noriko Takemura, Yasushi Makihara, Daigo Muramatsu, Tomio Echigo, and Yasushi Yagi. Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. *IPSJ TCVA*, 10(1):4, 2018. 2

[7] Torben Teepe, Ali Khan, Johannes Gilg, Fabian Herzog, Stefan Hörmann, and Gerhard Rigoll. Gaitgraph: Graph convolutional network for skeleton-based gait recognition. *arXiv preprint arXiv:2101.11228*, 2021. 3

[8] Shiqi Yu, Daoliang Tan, and Tieniu Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 4, pages 441–444. IEEE, 2006. 2