Gait Recognition Using 3-D Human Body Shape Inference
Paper Review 04/29

Bo Han, Chen

National Yang Ming Chiao Tung University, Taiwan
bhchen312551074.cs12@nycu.edu.tw

April 29, 2024

Info

- Title: Gait Recognition Using 3-D Human Body Shape Inference [1]
- Authors: Zhu, Haidong and Zheng, Zhaoheng and Nevatia, Ram
- Conference: IEEE/CVF Winter Conference on Applications of Computer Vision
- Year: 2023

Motivation

- Gait recognition based on 2-D images is difficult.
- Camera positions, carried objects, and clothing, etc.

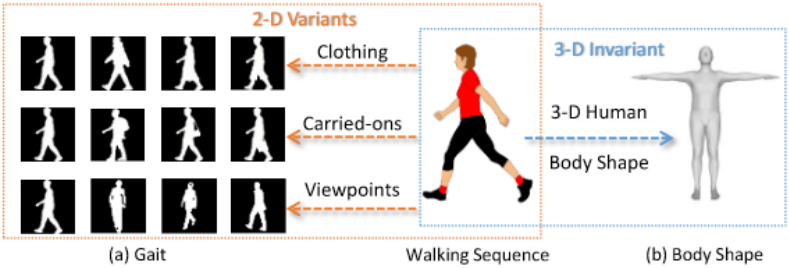

Figure 1: Variance of 2-D silhouette images

Proposed Method

- 3-D body shape inference from silhouettes sequences
- Knowledge distillation for body shape

Proposed Method

Gait Feature Extraction

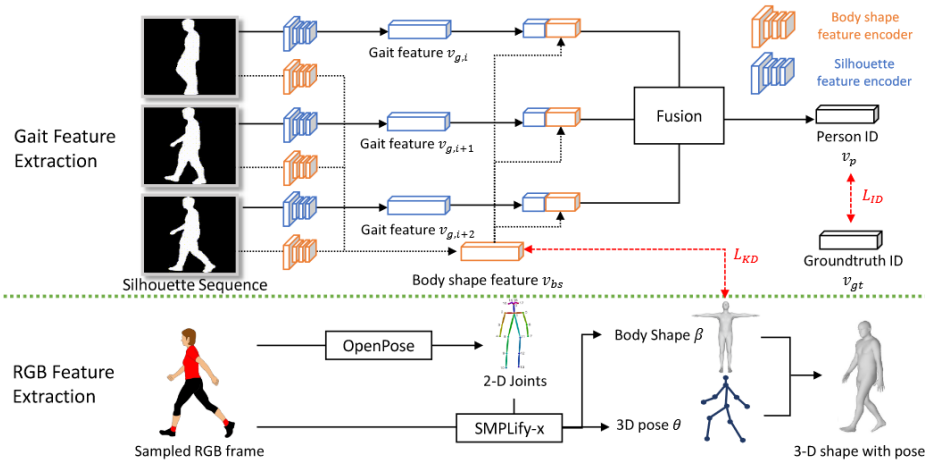

Figure 2: Overview of the proposed method

- sequence of silhouettes
- feature weights of different frame (12.5%)

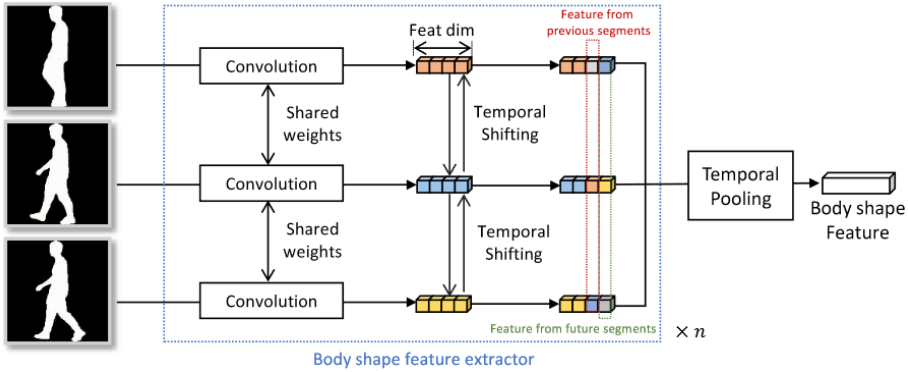

Figure 3: Body Shape Feature Encoder

- SMPLify-X
- Time consumption - select single RGB image

$$
\begin{aligned}
L_{KD} =& \mathbb{E}_{q(v_{br}, v_{bs}|C=1)}[\log\ h(v_{br}, v_{bs})] \\
&+ \mathbb{E}_{q(v_{br}, v_{bs}|C=0)}[\log(1 - h(v_{br}, v_{bs}))] \\
h(s,t) =& \frac{\exp(f_1(v_{bs})^T \cdot f_2(v_{br}))}{\exp(f_1(v_{bs})^T \cdot f_2(v_{br})) + \frac{N}{M}}
\end{aligned}
\tag{1}
$$

Figure 4: Contrastive Representation Distillation

Experiments

Dataset

- CASIA-B
  - 20% of sequence for data distillation
  - three scenario
    - normal walking
    - carrying bags
    - different clothing
- OUMVLP
  - lack of RGB video sequence
  - using shape encoder pre-trained on CASIA-B

Experiments

Results

Probe	Stats	HBS	Method				Avg. Change
			GaitSet [6]	GaitPart [10]	GLN [15]	GaitGL [23]	
NM #5-6	Mean (↑)	✗	95.6	96.4	96.5	97.3	***+0.3***
		✓	96.0	96.6	96.8	97.5	
		Δ	*+0.3*	*+0.2*	*+0.3*	*+0.2*	
	STD. (↓)	✗	3.4	2.8	2.8	1.7	***-0.1***
		✓	3.2	3.0	2.4	1.6	
		Δ	*-0.2*	*+0.2*	*-0.4*	*-0.1*	
BG #1-2	Mean (↑)	✗	90.1	91.2	93.1	94.4	***+0.4***
		✓	90.5	91.7	93.6	94.8	
		Δ	*+0.4*	*+0.5*	*+0.5*	*+0.4*	
	STD. (↓)	✗	4.6	4.2	3.3	2.7	***-0.3***
		✓	4.2	3.5	3.2	2.7	
		Δ	*-0.4*	*-0.7*	*-0.1*	*0.0*	
CL #1-2	Mean (↑)	✗	75.1	78.6	81.5	83.5	***+1.1***
		✓	76.0	79.7	83.2	84.1	
		Δ	*+0.9*	*+1.1*	*+1.7*	*+0.6*	
	STD. (↓)	✗	6.2	5.8	5.5	8.0	***-0.4***
		✓	5.9	5.6	5.5	7.0	
		Δ	*-0.3*	*-0.2*	*0.0*	*-1.0*	

Experiments

Results

Method	Camera Positions														Mean
	0°	15°	30°	45°	60°	75°	90°	180°	195°	210°	225°	240°	255°	270°	
GEINet [35]	23.2	38.1	48.0	51.8	47.5	48.1	43.8	27.3	37.9	46.8	49.9	45.9	45.7	41.0	42.5
GaitSet [6]	79.2	87.7	89.9	90.1	87.9	88.6	87.7	81.7	86.4	89.0	89.2	87.2	87.7	86.2	87.0
GaitPart [10]	82.8	89.2	90.9	91.0	89.7	89.9	89.3	85.1	87.7	90.0	90.1	89.0	89.0	88.1	88.7
GaitGL [23]	84.2	89.8	91.3	**91.7**	90.8	**91.0**	90.4	88.1	88.2	**90.5**	**90.5**	89.5	**89.7**	88.8	89.6
GaitSet-HBS	79.0	87.9	90.4	90.6	88.4	89.2	88.4	82.3	87.1	89.6	89.6	87.7	88.4	86.9	87.5
GaitPart-HBS	82.4	89.1	91.1	91.3	89.8	90.2	89.7	84.8	88.0	90.3	90.3	89.2	89.4	88.4	88.9
GaitGL-HBS	**84.7**	**90.2**	**91.4**	**91.7**	**90.9**	**91.0**	**90.5**	**88.4**	**88.7**	**90.5**	**90.5**	**89.6**	89.6	**88.9**	**89.8**

Figure 5: OUMVLP Results

Experiments

Results

- criteria
  - performance
  - viewpoints
  - appearance variance
- novel viewpoints

Conclusion

- Relay on exist model
- Quality of RGB video sequence
- Using skeleton for gait recognition?

References

[1] Haidong Zhu, Zhaoheng Zheng, and Ram Nevatia. "Gait recognition using 3-d human body shape inference". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2023, pp. 909–918.

Thanks for Listening

Q & A