End-to-end Model-based Gait Recognition using Synchronized Multi-view Pose Constraint ——Supplementary material——

Xiang Li[1] Yasushi Makihara[1] Chi Xu[1] Yasushi Yagi[1]
[1] Osaka University, Osaka, Japan
{li, makihara, xu, yagi}@am.sanken.osaka-u.ac.jp

Abstract

The supplementary material shows some additional pose visualization results of the proposed method. The experimental results of individual pose and shape features for the baseline ModelGait and the proposed method on both OU-MVLP and CASIA-B are provided as well.

1. Pose visualization

We show the pose visualization results of additional ten view input images with almost the same phase on OU-MVLP in Fig. 1. From all views' results, the proposed method could well handle various view variations and successfully estimate almost similar 3D and 2D poses in the unified human-centered coordinate.

2. Experiments

OU-MVLP. Table 1 shows the recognition results of individual pose and shape features for the baseline ModelGait [1] and the proposed method on OU-MVLP. Because the probe sets of OU-MVLP contain some subjects that are not in the gallery, we provide both results with and without non-enrolled probes for rank-1 identification rates. From the results, the proposed method shows large improvements on the pose feature for almost all probe views. In total, the mean accuracy of all probe views of the proposed method gets 5.7% higher rank-1 identification rate and 0.15% lower EER than the baseline ModelGait. This indicates the effectiveness of the proposed synchronized multi-view pose constraint in solving the view covariate. On the other hand, the proposed method shows subtle improvements on the shape feature because the proposed synchronized multi-view pose constraint has little effects on it.

CASIA-B. Table 2 shows the recognition results of individual pose and shape features for the baseline ModelGait [1] and the proposed method on CASIA-B. For three different settings (i.e., NM, BG and CL), the proposed method was trained on three different training set including the corresponding probe and gallery sets. Similar to OU-MVLP, the proposed method shows large improvements on the pose feature for most of the probe views. In total, the mean accuracy of all probe views of the proposed method gets 1.9%, 4.1% and 4.1% higher rank-1 identification rate for NM, BG and CL, respectively. On the other hand, the proposed method shows almost the same performance as

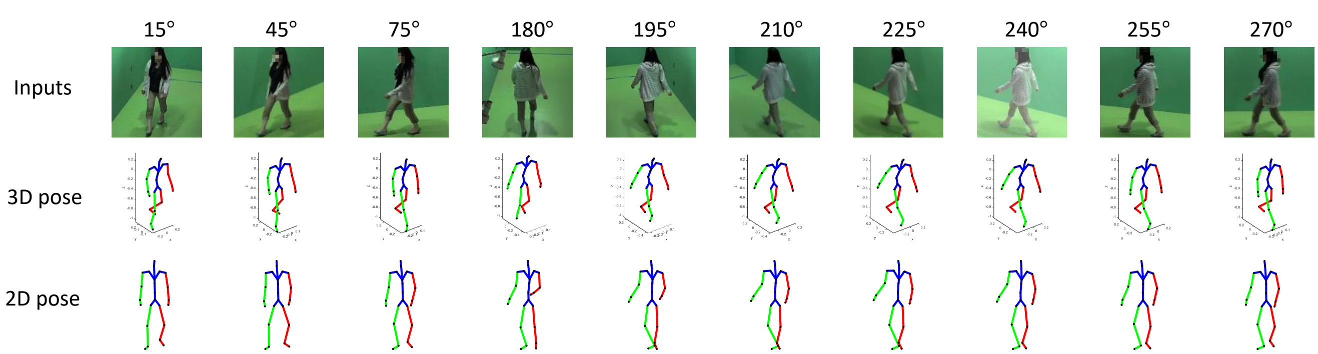

Figure 1. Examples of pose estimation results for the proposed method from OU-MVLP. The rest ten views in almost the same phase are given.

Table 1. Rank-1 rates and EERs of the baseline ModelGait [1] and the proposed method for each probe view averaged over the 14 gallery views on OU-MVLP, where the identical view is excluded. “-” means not provided. The Rank-1 rates show the results without non-enrolled probes, while the Rank-1(*) rates show the results with non-enrolled probes. Bold indicates the best accuracy.

	Methods	Probe view														
		0°	15°	30°	45°	60°	75°	90°	180°	195°	210°	225°	240°	255°	270°	Mean
Rank-1 [%]	ModelGait (pose)	77.4	81.6	84.1	86.4	86.5	87.0	86.1	78.1	82.8	86.0	87.4	86.0	87.3	85.9	84.5
	Ours (pose)	**85.0**	**88.6**	**89.9**	**91.1**	**91.7**	**92.5**	**91.8**	**85.1**	**88.8**	**90.8**	**92.2**	**91.5**	**92.1**	**91.4**	**90.2**
	ModelGait (shape)	92.2	94.4	95.8	96.2	96.4	96.3	95.9	92.6	94.7	95.9	96.3	96.1	96.2	95.8	95.3
	Ours (shape)	**92.8**	**95.1**	**96.2**	**96.4**	**96.5**	**96.5**	**96.3**	**93.0**	**95.4**	**96.4**	**96.5**	**96.6**	**96.5**	**96.2**	**95.7**
Rank-1(*) [%]	ModelGait (pose)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Ours (pose)	79.8	83.6	86.2	85.4	83.2	85.0	86.1	80.0	82.9	84.1	85.6	85.6	85.5	84.9	84.1
	ModelGait (shape)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Ours (shape)	87.1	89.5	90.6	89.6	89.3	89.9	89.7	87.4	89.1	89.3	89.6	90.3	89.5	89.3	89.3
EER [%]	ModelGait (pose)	0.57	0.48	0.48	0.41	0.34	0.30	0.32	0.46	0.45	0.49	0.49	0.36	0.34	0.35	0.42
	Ours (pose)	**0.35**	**0.29**	**0.35**	**0.27**	**0.24**	**0.20**	**0.19**	**0.30**	**0.27**	**0.35**	**0.31**	**0.23**	**0.23**	**0.21**	**0.27**
	ModelGait (shape)	**0.33**	0.30	0.32	0.25	0.20	**0.16**	0.17	**0.28**	0.26	0.32	0.29	**0.21**	**0.18**	0.19	0.25
	Ours (shape)	0.34	**0.29**	**0.29**	**0.24**	**0.19**	**0.16**	**0.16**	**0.28**	**0.25**	**0.30**	**0.26**	**0.21**	0.19	**0.18**	**0.24**

Table 2. Rank-1 rates [%] of the baseline ModelGait [1] and the proposed method on CASIA-B using the first 74 subjects for training. The mean result over all 10 gallery views for each probe view is given, where the identical view is excluded.

Probe	Methods	Probe view											
		0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	Mean
NM #5–6	ModelGait (pose)	87.1	88.3	93.8	**95.4**	92.1	**92.8**	90.5	90.7	88.5	92.4	91.7	91.2
	Ours (pose)	**94.4**	**90.8**	**95.0**	94.3	**93.2**	91.3	**92.5**	**94.2**	**92.0**	**93.5**	**92.6**	**93.1**
	ModelGait (shape)	**97.1**	**97.3**	**98.4**	98.4	**97.4**	98.3	**97.7**	96.2	**96.9**	97.1	97.5	**97.5**
	Ours (shape)	96.8	96.4	**98.4**	**98.9**	96.9	**98.4**	97.6	**96.6**	96.8	**97.5**	**97.9**	**97.5**
BG #1–2	ModelGait (pose)	86.8	81.2	84.6	86.8	84.9	83.0	83.9	82.8	82.1	84.0	83.2	83.9
	Ours (pose)	**90.1**	**86.9**	**88.8**	**92.5**	**88.6**	**86.5**	**87.1**	**86.1**	**86.6**	**88.0**	**86.7**	**88.0**
	ModelGait (shape)	**92.0**	**91.7**	**92.2**	**93.0**	**92.7**	**91.6**	**92.8**	**92.3**	88.4	86.5	83.4	90.6
	Ours (shape)	91.5	91.2	90.7	92.6	92.3	91.2	92.6	90.3	**89.4**	**89.9**	**89.9**	**91.0**
CL #1–2	ModelGait (pose)	**63.0**	62.4	66.3	65.2	61.9	58.2	58.3	59.1	56.8	55.4	55.6	60.2
	Ours (pose)	60.9	**62.7**	**66.8**	**69.8**	**69.2**	**66.9**	**62.8**	**65.4**	**63.6**	**59.8**	**59.1**	**64.3**
	ModelGait (shape)	72.1	74.1	77.2	79.0	77.3	76.7	75.2	76.0	70.1	72.8	**74.8**	75.1
	Ours (shape)	**73.5**	**76.3**	**78.0**	**79.1**	**77.9**	**77.4**	**75.5**	**76.3**	**71.3**	**73.2**	73.4	**75.6**

the baseline on the shape feature.

References

[1] Xiang Li, Yasushi Makihara, Chi Xu, Yasushi Yagi, Shiqi Yu, and Mingwu Ren. End-to-end model-based gait recognition. In *Proceedings of the Asian Conference on Computer Vision (ACCV)*, November 2020. 1, 2