Cross-Subject and Cross-Modal Transfer for Generalized Abnormal Gait Pattern Recognition – Supplementary Materials

Xiao Gu, Yao Guo, Fani Deligianni, Benny Lo, and Guang-Zhong Yang

I. Architecture Details

The architectures details implemented in our study are given in Table I. The E_s, E_p & D are the same for both the cross-subject single-modal E_s^s, E_p^s & D^s and cross-subject cross-modal E_s^c, E_p^c & D^c parts, although the first layer and the last layer might be slightly different according to the specific modality and specific task. The input size of C_p are based on the output of E_p.

TABLE I
Architecture Details

Network	Layer	Parameters	in/out
E_s	Conv + LReLU	k=8, s=2	J_x/48
	Conv + LReLU	k=8, s=2	48/96
	Conv + LReLU	k=8, s=2	128/128
	Conv + LReLU	k=8, s=2	128/128
	Conv + LReLU	k=8, s=2	128/128
E_p	Conv + LReLU	k=8, s=2	J_x/48
	Conv + LReLU	k=8, s=2	48/96
	Conv + LReLU	k=8, s=2	128/128
	Conv + LReLU	k=8, s=2	128/128
	Conv + LReLU	k=8, s=2	128/64
D	Up + Conv + Avg + LReLU	k=7, s=1	192/128
	Up + Conv + Avg + LReLU	k=7, s=1	128/128
	Up + Conv + Avg + LReLU	k=7, s=1	128/128
	Up + Conv + Avg + LReLU	k=7, s=1	128/96
	Up + Conv	k=7, s=1	96/J_x or J_z
C_p	Linear+LReLU	-	In/32
	Linear	-	32/5

* Conv: 1D convolution with reflected padding;
♯ LReLU: LeaklyReLU with $\alpha = 0.2$;
‡ Up: Upsampling
§ Avg: Average Pooling

II. Experimental Details

A. Abnormal Gait Types

In our abnormal gait database[1], 18 healthy volunteers (16 male and 2 female) were recruited and instructed to walk normally and imitate four pathological gait patterns (i.e., toe-in, toe-out, supination, and pronation), following the settings of previous simulation based works [1], [2], [3]. The details of the data are listed in Table II.

Among pathological gait categories as shown in Fig. 1, the supination refers to the outward roll of ankle joint while the pronation indicates the inward rotation. The toe-in and toe-out are the inward and outward of the foot forward direction during walking [4]. To avoid exaggerated imitation, specialized correction insoles were placed inside the shoes to resemble supination and pronation characteristics.

[1]This experiment was approved by Imperial College Research Ethics Committee under Reference No. 18IC4915.

TABLE II
Dataset Details

Subj No.	Gender	#	# of Each Gait Pattern N	S	P	I	O	Gait Modal Mocap	RGBD	EMG
1	F	335	56	54	64	71	90	✓	✓	✓
2	M	182	31	32	37	36	46	✓	✓	✓
3	M	271	50	51	55	56	59	✓	✓	✓
4	M	119	27	29	25	22	16	✓	✓	✓
5	M	175	29	43	35	34	34	✓	✓	✓
6	M	266	44	66	61	35	60	✓	✓	✗
7	M	267	29	65	64	51	58	✓	✓	✓
8	M	117	24	22	22	23	26	✓	✓	✓
9	M	158	22	31	39	35	31	✓	✓	✓
10	M	230	39	38	38	48	67	✓	✓	✓
11	M	271	62	55	54	58	42	✓	✓	✓
12	M	263	41	54	59	54	55	✓	✗	✓
13	M	125	23	24	24	29	25	✓	✓	✓
14	M	127	20	23	25	17	42	✓	✓	✓
15	M	179	24	49	36	30	40	✓	✓	✓
16	M	212	40	42	43	42	45	✓	✗	✓
17	F	143	21	31	34	24	33	✓	✓	✗
18	M	154	27	32	31	32	32	✓	✓	✓
Total	-	3602	610	741	746	704	801	18	16	16

Fig. 1. Visualization of different gait types. Supination refers to the outer rotation of ankle while pronation the inner rotation. Toe-in modification refers to the walking with toes pointing in, while toe-out pointing out.

The simulation of different gait styles on healthy subjects has been widely adopted in previous studies to facilitate the validation of pathological gait recognition or assessment. The pronation/supination refers to inwarded/outwarded ankle positions. To simulate both, Mahmood *et al.* [1] used self-designed wedge-shaped foot insoles to help subjects imitate eversion and inversion foot deficiencies. The settings of using specialized shoes to simulate pronated/supinated feet are similar to the works of Xu *et al.* [3] and Soares *et al.* [5].

Furthermore, toe-in/toe-out (in-toeing/out-toeing) have been also reported as a common orthopedic problem. Typically, healthy subjects are recruited and asked to mimic these patterns to develop benchmark datasets for training and testing machine learning approaches. In [2], the authors reported the kinematics, dynamics and Electromyography of simulated toe-in, simulated toe-out and neutral gaits. In [6], specialized smart shoe was designed and applied to distinguish the foot

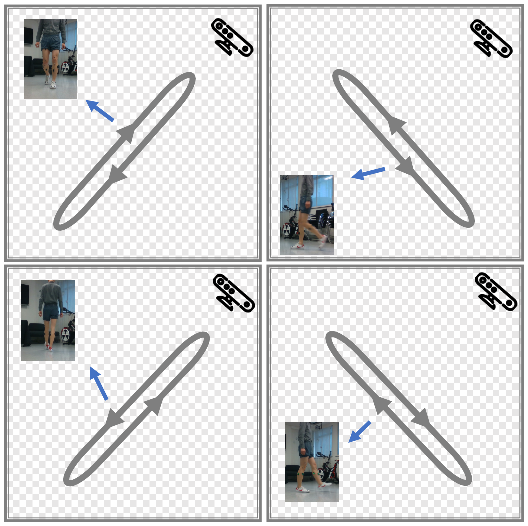

Fig. 2. Experimental settings for walking directions.

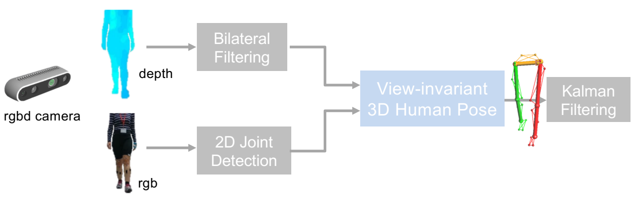

Fig. 3. Extraction of 3D skeleton points based on RGBD data (Image adapted from [4]). Firstly, the 2D joints are extracted from rgb images by using OpenPose [7]. Simultaneously, the associated depth image is denoised by bilateral filtering. Then 3D lower limb pose is estimated from depth data based on the 2D joints derived from associated rgb images. The estimated skeleton is filterd by kalman filter as final output.

progression angles resulted from toe-in/toe-out patterns.

B. Experiment Settings

During the experiment, subjects were instructed to walk roughly along the diagonal lines of around 4 meters inside a multi-camera motion capture laboratory with a RGBD camera fixed at the corner, as shown in the sketch up of Fig. 2. This setting resembles closely an indoor free-living environment and meanwhile, allows us to simultaneously collect Mocap and multi-modal data (i.e., RGBD and EMG).

C. RGBD Skeleton Data

Our framework for estimating 3D skeleton from RGBD data was proposed in our previous work [4] and it is illustrated in Fig. 3. This framework is based on inferring canonical 3d motion data from a mobile rgbd sensor. In our case, the camera was fixed at the corner of the room, whereas the subjects were walking along the diagonal of the room.

Estimating 3d human pose from monocular videos is a hot topic in the computer vision field. Errors and interference during 3d motion capture mostly result from three factors, namely, the 2D pose estimation from rgb, the noises from the depth sensor and partial occlusion of the scene. In order to extract the 3D joint points, we fuse information derived from state-of-the art 2D pose estimation algorithm with depth data [4]. However, depth data are significantly more noisy than rgb and their resolution is poor. Noises involve bias, artefacts and light interference and result in a complex interplay of holes (missing/erroneous depth values), distance dependent biases and instabilities [8]. Similarly, to our previous work, we applied bilateral filtering to denoise the depth data. Subsequently, the 3d positions of the keypoint are extracted from the filtered depth data. To solve the occlusion occurring during walking, we applied Kalman filtering for each joint to further refine the data [4].

References

[1] I. Mahmood, U. Martinez-Hernandez, and A. A. Dehghani-Sanij, “Evaluation of gait transitional phases using neuromechanical outputs and somatosensory inputs in an overground walk,” *Human Movement Science*, vol. 69, p. 102558, 2020.

[2] W. Cui, C. Wang, W. Chen, Y. Guo, Y. Jia, W. Du, and C. Wang, “Effects of toe-out and toe-in gaits on lower-extremity kinematics, dynamics, and electromyography,” *Applied Sciences*, vol. 9, no. 23, p. 5245, 2019.

[3] D. Xu, W. Ouyang, E. Ricci, X. Wang, and N. Sebe, “Learning cross-modal deep representations for robust pedestrian detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5363–5371.

[4] Y. Guo, F. Deligianni, X. Gu, and G.-Z. Yang, “3-d canonical pose estimation and abnormal gait recognition with a single rgb-d camera,” *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3617–3624, 2019.

[5] D. P. Soares, M. P. de Castro, E. Mendes, and L. Machado, “Influence of wedges on lower limbs’ kinematics and net joint moments during healthy elderly gait using principal component analysis,” *Human movement science*, vol. 38, pp. 319–330, 2014.

[6] H. Xia, J. Xu, J. Wang, M. A. Hunt, and P. B. Shull, “Validation of a smart shoe for estimating foot progression angle during walking gait,” *Journal of biomechanics*, vol. 61, pp. 193–198, 2017.

[7] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime multi-person 2d pose estimation using part affinity fields,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7291–7299.

[8] G. Halmetschlager-Funek, M. Suchi, M. Kampel, and M. Vincze, “An empirical evaluation of ten depth cameras: Bias, precision, lateral noise, different lighting conditions and materials, and multiple sensor setups in indoor environments,” *IEEE Robotics & Automation Magazine*, no. 99, pp. 1–1, 2018.