

Supplementary Materials for “It Takes Two: Accurate Gait Recognition in the Wild via Cross-granularity Alignment”

Anonymous Authors

1 INTRODUCTION

In this supplementary material, we first visualize the distribution of features. Next, we provide some detailed exemplar results. At last, several potential future works are represented.

2 DISTRIBUTION VISUALIZATION

To verify the effectiveness of our proposed XGait method from more perspectives, we visualize the feature distribution on the Gait3D dataset [3]. For convenience, we randomly sample seven IDs from the test set in Gait3D and utilize the tSNE [2] to draw the feature distribution of all sequences corresponding to these sampled IDs.

Figure 1 (a) illustrates that the feature distribution based solely on the silhouette is highly scattered. In contrast, by replacing silhouette with parsing, the feature distribution becomes more compact, as shown in Figure 1 (b). This demonstrates that parsing has higher information entropy, which enables the model to learn more discriminative gait features.

Figures 1 (c) to (d) demonstrate that combining silhouette and parsing, along with the incremental integration of our proposed Global Cross-granularity Module (GCM) and Part Cross-granularity Module (PCM), can enhance the distinctiveness. Furthermore, our XGait shows superior discriminative ability, demonstrating the effectiveness of our method.

3 EXEMPLAR RESULTS

In this section, we provide some exemplar results of XGait (Only Silhouette), XGait (Only Parsing), and XGait (Ours). We conduct exemplar visualization experiments on the Gait3D dataset [3]. For convenience, we sample 24 consecutive frames from each sequence. The top row depicts the query sequence with blue bounding boxes. The subsequent three rows exhibit the top-3 gallery results ranked by their similarity to the query sequence. The correct results are marked in green bounding boxes, while incorrect results are marked in red bounding boxes.

From Figure 2 and Figure 3, we can observe that our XGait helps the model achieve more correct matches by integrating silhouette and parsing. This demonstrates that these two appearance representations, i.e., silhouette and parsing, have mutual information, and our XGait effectively explores their complementarity to enhance performance.

Furthermore, Figure 3 reveals that different individuals with similar body shapes, viewpoints, and walking postures can cause serious interference. This is an important challenge for appearance-based methods.

4 FUTURE WORK

Although our XGait method proposed in this paper achieves State-Of-The-Art (SOTA) performance on two challenging gait datasets, i.e., the real-world dataset Gait3D [3] and the cloth-changing dataset CCPG [1], there are several directions worthy of further study:

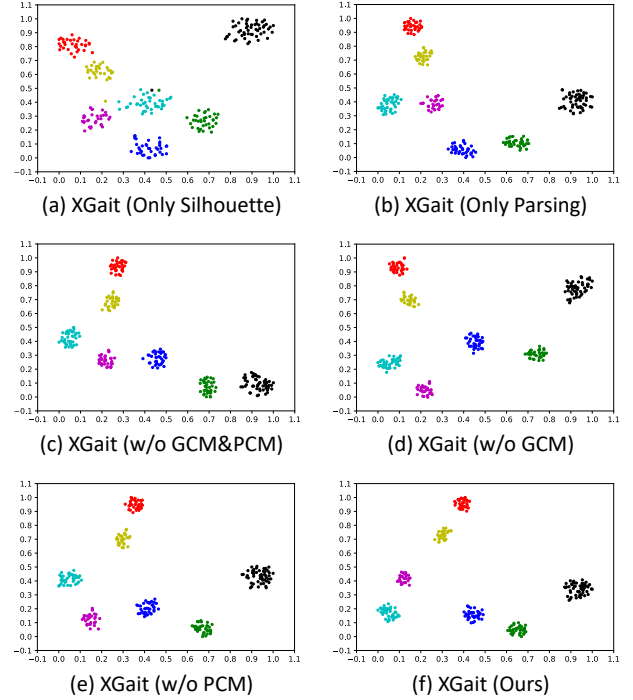


Figure 1: The visualization of feature distributions. (a) XGait (Only Silhouette), (b) XGait (Only Parsing), (c) XGait (w/o GCM&PCM), (d) XGait (w/o GCM), (e) XGait (w/o PCM), and (f) XGait (Ours). Samples of the same color belong to the same person. (Best viewed in color.)

- One direction is to explore more efficient integration of various gait representations, including silhouette, parsing, 2D skeleton, 3D skeleton, 3D SMPL&Mesh, point cloud, etc., to further improve the performance of gait recognition technology in real-world scenarios.
- Additionally, how to mitigate the interference from different pedestrians with similar appearances when utilizing informative appearance representations, as discussed in Section 3.
- Lastly, investigating more efficient methods for temporal modeling to address challenges posed by irregular walking speeds and routes in real-world scenarios is also a promising direction for further research.

REFERENCES

- [1] Weijia Li, Saihui Hou, Chunjie Zhang, Chunshui Cao, Xu Liu, Yongzhen Huang, and Yao Zhao. 2023. An In-Depth Exploration of Person Re-Identification and Gait Recognition in Cloth-Changing Conditions. In *CVPR*. 13824–13833.
- [2] Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *JMLR* 9, 11 (2008).
- [3] Jinkai Zheng, Xinchun Liu, Wu Liu, Lingxiao He, Chenggang Yan, and Tao Mei. 2022. Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. In *CVPR*. 20228–20237.

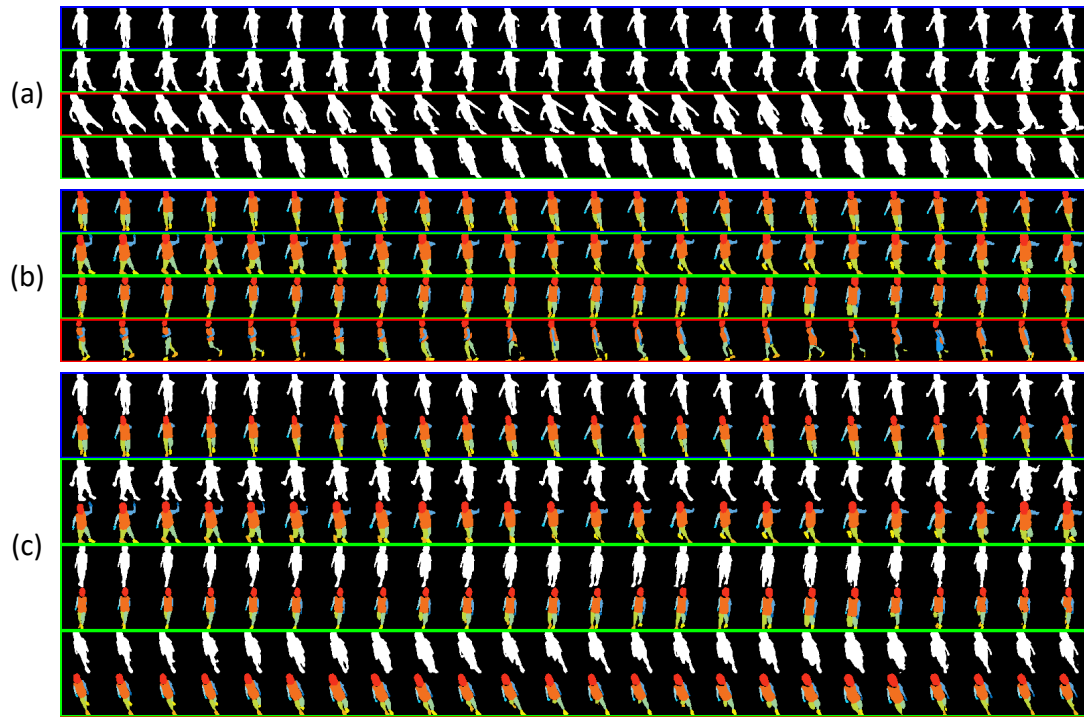


Figure 2: Exemplar results of (a) XGait (Only Silhouette), (b) XGait (Only Parsing), and (c) XGait (Ours). This case shows that utilizing both silhouette and parsing helps the model match more correct results. (Best viewed in color.)

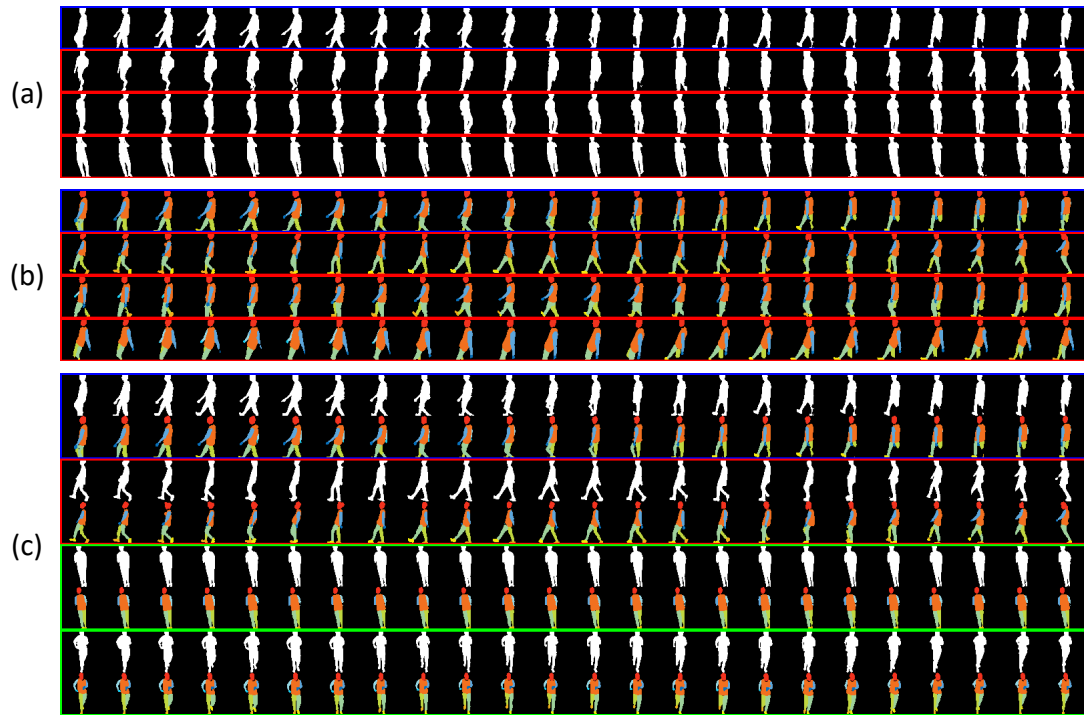


Figure 3: Exemplar results of (a) XGait (Only Silhouette), (b) XGait (Only Parsing), and (c) XGait (Ours). This case shows that similar body shapes and poses are important challenges in gait recognition. (Best viewed in color.)