

Gait Energy Image Reconstruction from Degraded Gait Cycle Using Deep Learning

Maryam Babaee^{1,2}, Linwei Li¹, and Gerhard Rigoll¹

¹ Chair of Human-Machine Communication, Technical University of Munich, Theresienstrasse 90, Munich, Germany

² maryam.babaee@tum.de

Abstract. Gait energy image (GEI) is considered as an effective gait representation for gait-based human identification. In gait recognition, normally, GEI is computed from one full gait cycle. However in many circumstances, such a full gait cycle might not be available due to occlusion. Thus, the GEI is not complete, giving a rise to degrading gait identification rate. In this paper, we address this issue by proposing a novel method to reconstruct a complete GEI from a few frames of gait cycle. To do so, we propose a deep learning-based approach to transform incomplete GEI to the corresponding complete GEI obtained from a full gait cycle. More precisely, this transformation is done gradually by training several fully convolutional networks independently and then combining these as a uniform model. Experimental results on a large public gait dataset, namely OULP demonstrate the validity of the proposed method for gait identification when dealing with very incomplete gait cycles.

Keywords: Gait Energy Image · Gait Recognition · Deep Learning

1 Introduction

Gait (the style of natural walking) [2] can be used to identify individuals at a distance, when other biometric features such as face and fingerprint might not be available. Here, a sequence of images showing a person walking is analyzed as input data. Since the natural walking of a person is periodic, it is sufficient to consider only one period (gait cycle) from the whole sequence. The gait cycle is defined as the time interval between the exact same repetitive events of walking (any position of the foot during walking can be regarded as the starting point of the gait cycle).

The main assumption in many gait-based human identification techniques [1, 5] is that a full gait cycle of individuals is available, which is a strong assumption in video surveillance applications, where occlusion occurs a lot and a person might be observed in only a few frames. From a full gait cycle, a simple and effective gait representation, namely gait energy image (GEI) [4] can be computed, which is the average of silhouette images of a walking person (see Fig. 1). This standard gait feature has been widely used alone or in combination with other features for gait recognition.

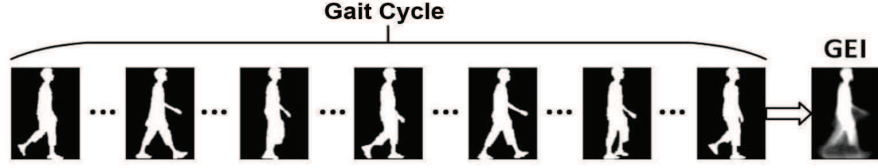


Fig. 1. GEI is computed by averaging gait silhouettes over one gait cycle.

We propose a deep-based method to reconstruct a GEI from a few frames, i.e., incomplete gait cycle. More specifically, having only a few frames of a full gait cycle, we first generate an incomplete GEI. Next, we train a uniform fully convolutional neural network (FCN) which gets the computed incomplete GEI (average of a few frames) as input and outputs the reconstructed complete GEI. The conducted experiments confirm that this network can successfully reconstruct a complete GEI.

2 Method

In our method, the complete GEI reconstruction is done in a progressive way (i.e. various types of incomplete GEIs are gradually converted to the complete GEI). To this end, we propose an incremental GEI reconstruction approach using ten FCNs that each single FCN enhances the quality of input incomplete GEI (IC-GEI). Since the gait cycle length depends on the frame rate and is different from a dataset to another dataset, we consider the partial transformations every 10% of the gait cycle length. The first FCN transforms a GEI generated from one frame to the GEI corresponds to the consecutive 10% of the gait cycle. Similarly, the other FCNs enhance their incomplete GEI by predicting the information of the following 10% of gait cycle. The structures of the all FCNs are the same, but they are trained on different types of GEI in terms of the number of frames and the starting frame in a gait cycle. Each FCN like an auto-encoder consists of two parts; the encoder (convolutional) and the decoder (deconvolutional) part.

After training the ten FCNs, their last convolutional hidden layers are combined to have one uniform model named ITCNet (see Fig. 2). The input of the ITCNet could be any type of IC-GEI, and the target is the corresponding complete GEI. The convolutional hidden layer i maps an mf -GEI (composed of m frames) to the corresponding nf -GEI, where $n - m = 0.1 * T$, and T is the gait cycle length. For example, if the aim is to transform a 1f-GEI to the complete GEI and the full gait cycle is 30 frames long, the input is first mapped to 3f-GEI, then to 6f-GEI by passing through the first two hidden layers, and so on. All generated incomplete GEIs are used to fine tune the ITCNet. In this way, this end-to-end network can transform any type of IC-GEI to corresponding complete GEI, without the need of any prior knowledge about the type of the input IC-GEI.

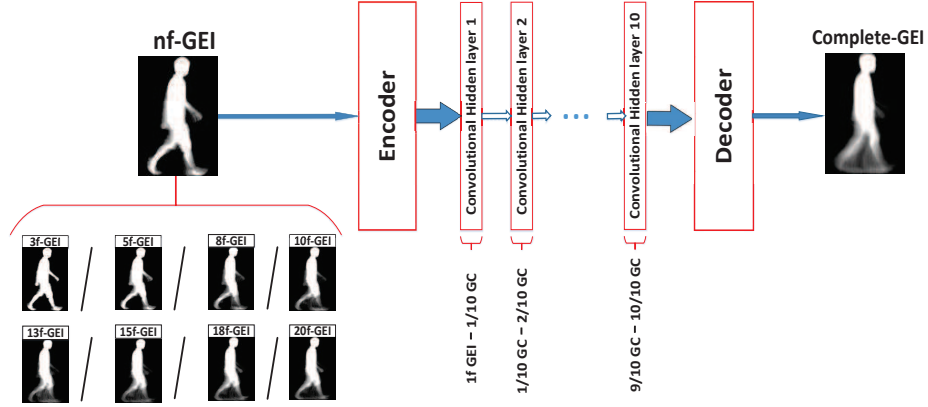


Fig. 2. ITCNet structure for reconstructing a complete GEI from an incomplete GEI.

3 Results

Figure 3 shows some samples of the reconstructed GEIs from only 3 frames for 10 different subjects in the OULP dataset [3]. It can be seen that the reconstructed GEIs recovered by our end-to-end ITCNet are almost similar to the ground truth. Regarding the gait recognition performance, Table 1 presents the rank-1 and rank-5 identification rates for incomplete GEIs and reconstructed GEIs. This experiment was performed on 500 subjects of the OULP dataset. Clearly, the end-to-end ITCNet has greatly improved the identification performance, especially for IC-GEIs generated from smaller partial gait cycle. As the number of frames increases (larger partial gait cycle), the rank-1 and rank-5 identification rates get closer to that of ground truth complete GEI.

Table 1. Comparison of rank-1 and rank-5 identification rates between incomplete GEIs and reconstructed GEIs from different portion of gait cycle. Note that rank-1 and rank-5 for ground truth complete GEI are **86.30%** and **94.24%**, respectively.

Partial Gait Cycle (%)	Rank-1[%]		Rank-5[%]	
	Reconst. GEI	Incomp. GEI	Reconst. GEI	Incomp. GEI
10	53.40	8.42	73.35	12.81
20	65.11	12.70	82.57	17.76
30	71.19	17.09	86.39	24.24
40	77.53	26.50	89.95	35.62
50	80.46	34.63	91.67	47.83
60	82.59	61.96	93.06	77.31
70	84.22	72.40	93.02	85.53
80	85.76	72.94	93.75	86.98
90	86.00	73.06	94.10	86.40

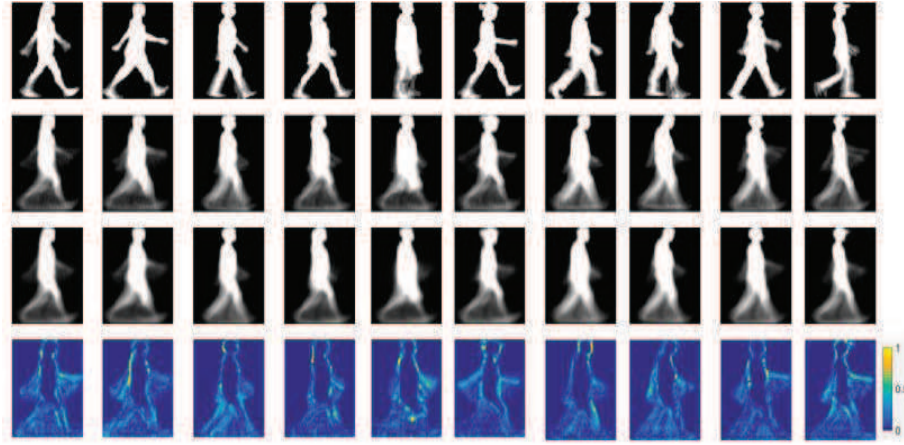


Fig. 3. Best viewed in color. Qualitative results of our GEI reconstruction method; Incomplete GEI generated from 3 frames for different subjects (first row), ground truth GEIs (second row), the corresponding reconstructed GEIs (third row), and the difference images of reconstructed and ground truth GEIs in color scale (last row).

4 Conclusions

We have proposed a fully convolutional neural network for gait energy image (GEI) reconstruction from an incomplete gait cycle. The model could reconstruct a GEI, given an incomplete-GEI which is composed of only a few frames of a gait cycle. Experimental results show that the proposed model can improve recognition rate greatly, particularly when there is only 10% of a gait cycle is available. In future, we will extend this model to an end-to-end model for both gait energy image reconstruction and recognition.

References

1. Babaei, M., Rigoll, G.: View-invariant gait representation using joint bayesian regularized non-negative matrix factorization. In: The IEEE International Conference on Computer Vision (ICCV) (2017)
2. Boyd, J.E., Little, J.: Biometric gait recognition. In: Advanced Studies in Biometrics, pp. 19–42. Springer (2005)
3. Iwama, H., Okumura, M., Makiyara, Y., Yagi, Y.: The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition. IEEE Transaction on Information Forensics and Security pp. 1511–1521 (2012)
4. Man, J., Bhanu, B.: Individual recognition using gait energy image. IEEE Transactions on Pattern Analysis and Machine Intelligence pp. 316–322 (2006)
5. Wolf, T., Babaei, M., Rigoll, G.: Multi-view gait recognition using 3d convolutional neural networks. In: Proceedings of the International Conference on Image Processing (ICIP). pp. 4165–4169. IEEE (2016)