

# A new gait-based identification method using local Gauss maps

Hazem El-Alfy<sup>1,2</sup>, Ikuhisa Mitsugami<sup>1</sup> and Yasushi Yagi<sup>1</sup>

<sup>1</sup> The Institute of Scientific and Industrial Research, Osaka University, Japan

<sup>2</sup> On sabbatical leave from Alexandria University, Egypt  
{hazem,mitsugami,yagi}@am.sanken.osaka-u.ac.jp

**Abstract.** We propose a new descriptor for human identification based on gait. The current and most prevailing trend in gait representation revolves around encoding body shapes as silhouettes averaged over gait cycles. Our method, however, captures geometric properties of the silhouettes boundaries. Namely, we evaluate contour curvatures locally using Gauss maps. This results in an improved shape representation, as contrasted to average silhouettes. In addition, our approach does not require prior training. We thoroughly demonstrate the superiority of our method in gait-based human identification compared to state-of-the-art approaches. We use the OU-ISIR Large Population dataset, with over 4000 subjects captured at different viewing angles, to provide statistically reliable results.

## 1 Introduction

Human identification using gait is gaining significant attention by computer vision researchers and crime prevention practitioners alike. The reason is that gait can serve as a biometric evidence for determining identities without the cooperation of the subjects. In addition, gait information can be collected at a distance. This is in contrast to traditional biometric methods, such as fingerprinting, iris recognition or even face recognition which are both invasive and require subjects' cooperation. Unfortunately, collecting biometric information at a distance comes at the expense of performance degradation. Moreover, several parameters affect the accuracy of identification using gait such as variation of viewing angle [1–3], walking speeds [4, 5], types of cloth [6] and walking surface [7], to mention a few.

Given the previously mentioned challenges, a considerable amount of research has been carried out in the area of gait analysis over the past two decades. The first decade focused on developing new techniques and devising appropriate descriptors. Among the first features used were contour signals [8], image sequence correlation [9], self similarity plots [10, 11] and unwrapped contour signals [12, 13]. The first decade concludes with the development of a simple, yet efficient and accurate descriptor, namely the average silhouette [14]. This latter approach is currently widespread and represent the state-of-the-art in period based gait features. It has paved the way for the development of several variants such as

Gait Energy Image [15], Frequency Domain Feature [2], Gait Entropy Image [16], Gait Flow Image [16] and Chrono-Gait [17].

Later approaches employed statistical tools, such as discriminant analysis, to enhance the identification rates of prior methods [18]. More recently, gait analysis methods have made use of new video capturing technology. With the development of affordable 3D capturing devices such as Kinect cameras, some approaches went about generalizing 2D methods to include depth information [19, 20]. Lately, an approach that combines several existing gait features, using different fusion techniques, was presented in [21].

### 1.1 Motivation and intuition

Our objective is to further investigate the geometry of the silhouettes. They contain more shape information than that derived from their average over a cycle. In particular, the local curvatures of a silhouette contour encode the body shape of a subject more robustly than the mere positions of the boundary pixels. Our motivation is the following: mild variations in the silhouette appearance resulting from, say, small gait fluctuations, will cause the average shape size or location to vary. On the other hand, the *curvature* of the body’s outline will only change slightly.

We achieve our goal by introducing a novel gait representation, *histograms of boundary normal vectors*, in order to compute the curvature of body contours. Histograms of surface normal vectors have recently been used in object recognition in still depth images [22]. Unit vectors normal to a surface are linked to its curvature through *Gauss maps* [23, 24]. We will further elaborate on Gauss maps in section 3. For now, we talk about the intuition of using normal vectors. Generally speaking, two vectors are equal when their magnitude and direction are equal. If we fix the normal vectors magnitudes to unity, then “parallel” contours will have the same histograms of normal vectors. This is a desirable property that makes our descriptor robust to small changes in body dimensions, such as those resulting from minor gait fluctuations or slight weight gain or loss, for instance. However, different contours may have similar normal vectors histograms as we will highlight in section 3. This leads us to use Gauss maps locally on small pieces of the contour.

### 1.2 Contributions

Our main contributions in this work are the following:

- We propose a novel gait descriptor that encodes body shape more robustly than existing methods and is less affected by subtle changes in appearance. Our feature is computed only on silhouettes contours and thus is more efficient in terms of storage requirements compared to prevalent approaches that use the entire silhouette area.
- We validate our approach achieves state-of-the-art performance in person recognition using the world’s largest gait database. With over 4,000 subjects,

a fair subject gender representation, a wide age range and a variable viewing angle, this dataset guarantees a statistically reliable performance evaluation.

The rest of this paper proceeds as follows. Section 2 reviews related work. Section 3 introduces the mathematical background necessary to explain how our feature works. Section 4 summarizes the process we follow to compute our descriptor. In section 5, we describe the large population dataset used in our experiments then show our experimental results in section 6. We conclude the paper with section 7.

## 2 Related work

Gait recognition has been studied extensively over the past two decades. Major approaches to gait representation broadly fit into two categories: model-based and image- or appearance-based (model-free). In model-based approaches, the observed human body parts are fit to a human body model. The work of Johansson using moving light displays [25] is considered to be one of the earliest model-based human gait recognition approaches. More recent approaches include the work by Bobick and Johnson [26] in which they use a three-linked model to fit the torso, leg lengths and strides. Yam *et al.* [27] extract joint angle sequences of legs and fit them to a pendulum-like model. Urtasun and Fua [28] employ a 3D model of links, and Yang *et al.* [29] exploit a 3D human model with cylindrical links. Model-based approaches have the advantage of being unaffected by variations in body shape. Unfortunately, they suffer from high costs of fitting images to models in addition to errors involved in the fitting process. Recently, as the technology of depth cameras evolved, Kumar and Babu [30] used Kinect cameras to capture 3D joint locations and hence build a more accurate skeleton. In that case, the covariance of joint location sequences is used as a feature. The same approach is used by Hussein *et al.* [31] in the wider area of motion recognition. This technology is still recent, and unavailable in most public surveillance systems. Moreover, large scale datasets have yet to be prepared. For all previous reasons, appearance-based approaches are widely spread and more commonly used.

Appearance-based approaches extract gait features directly from 2D images. The current trend is to use side-view silhouettes since those capture most variability in gait motion. No models need to be fit here. Again, this approach is further divided into two subgroups based on the gait features used: frame-based and period-based gait features. Frame-based gait features are matched frame by frame. For a successful evaluation, a synchronization step has to precede the matching step. Sequences have to be aligned (or be in phase) at a preprocessing stage. This approach was more frequently used a decade ago. Philips *et al.* [32] propose a direct silhouette sequence matching as a baseline method. Wang *et al.* [12] exhaustively search the phase shift with the minimal classification distance. Murase and Sakai [9] employ a parametric eigenspace to represent periodic gait silhouette sequences. Liu *et al.* [18] propose a gait dynamics normalization by using a population hidden Markov model to match two silhouettes at the same

phase. Beyond the raw silhouette sequences, Cuntoor *et al.* [33] project the silhouettes into a width vector and Liu *et al.* [34] project it into a frieze pattern. The main disadvantage of frame-based image-based approaches is that frame-to-frame matching is very sensitive to noise and to slight phase shifts, specially when low frame rates are used. On top of that, phase synchronization has often been a time consuming process.

Finally, period-based gait features are evaluated by integrating individual frame features over a given period, commonly a computed gait cycle (two stances). This approach makes the extracted gait features more robust to noise and slight phase fluctuations when contrasted with frame-based features. Thus, it is no surprise to find such features commonly used in most current approaches. One of the earliest robust approaches still used frequently until now is the gait energy image (GEI) by Han and Bhanu [15] or average silhouette [14]. It is computed by averaging silhouettes values, pixel by pixel, over an entire gait cycle period. Given the periodic nature of gait motion, Makihara *et al.* [2] compute, pixel by pixel, amplitude spectra of lower frequency components elements in what they call the frequency domain feature (FDF). Other approaches have also been derived as variants of the GEI method [16, 35, 17]. Besides, other features have also been proposed, such as self-similarity plots [10, 11], Gabor filter based feature [36], local auto-correlations of spatio-temporal gradients [37, 38] and histograms of oriented gradients (HOG) [39, 40]. Recently, motion and shape features were represented and learned separately, eventually combining both for identification [41, 42].

### 3 Gauss maps

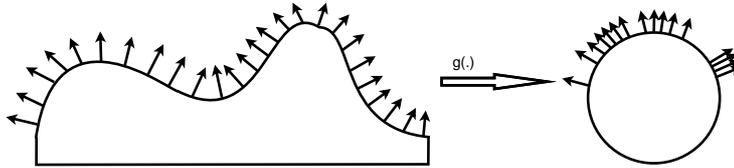
The key tool in estimating the curvature of a silhouette contour is to evaluate its Gauss map. In this section, we introduce this notion, highlighting the necessary mathematical background. A detailed account of the differential geometry of curved surfaces is being referred to in [23].

Without loss of generality, a *Gauss map* is a mapping (function)  $g$  from a surface  $M \in \mathbb{R}^3$  to the unit sphere  $\mathbb{S}^2$  such that  $g$  associates to each point  $p \in M$  the unit vector  $\mathbf{n}_p$  normal to  $M$  at  $p$ .

$$\begin{aligned} g : M &\rightarrow \mathbb{S}^2 \\ p &\mapsto \mathbf{n}_p \end{aligned} \tag{1}$$

A Gauss map measures how curved a surface is. Suppose a surface  $M$  is a (flat) plane. All the normal vectors to  $M$ , i.e. the image of  $g$ , are parallel to each other, thus there is no variation between them. On the other hand, the normal vectors to an “overly” curved surface vary greatly from point to point. For that reason, it is reasonable to use the mapping  $g$  to investigate the curvature of the surface.

Given a subsurface  $\Omega \subseteq M$ , the *total curvature* of  $\Omega$  is defined to be the area of the image of the Gauss map  $g(\Omega)$ . In modern literature, the curvature of  $M$



**Fig. 1.** Example of a Gauss map from a two-dimensional curve to the unit circle.

at a point is measured by what is called the *Gaussian curvature*  $k$ . It is defined as:

$$k = \lim_{\Omega \rightarrow 0} \frac{\text{area of } g(\Omega)}{\text{area of } \Omega} \quad (2)$$

Now, by looking at  $k$  as a function on  $M$ , we can define the total curvature of  $\Omega$  by:

$$\text{total curvature of } \Omega = \int_{\Omega} k dA \quad (3)$$

where  $dA$  is a surface element on  $M$ .

The total curvature of a surface computed globally cannot always be used to discriminate different surfaces. The reason is that the Gaussian curvature of a surface  $M \in \mathbb{R}^3$  is invariant under *local isometries*. A local isometry of a surface  $M$  is simply a deformation of the surface under which the lengths of *geodesics* (curves of shortest length that lie in  $M$ ) are unchanged. A simple example of isometric surfaces is that of a cylinder and the plane that results from “unwrapping” its surface. Both will have the same total curvature.

For that reason, the Gauss maps are defined locally, i.e. on small pieces of the surface, then local curvatures are computed for each small surface and finally, all those local curvatures are stitched together in a long feature vector. Now, how is that related to gait recognition? We simply reduce one dimension. Instead of computing curvatures of surfaces, we compute them for the contours bounding the silhouettes. In that case, the Gauss map becomes a mapping from a curve to the unit circle as shown in Fig. 1.

## 4 Approach outline

Evaluating Gauss maps of digital images requires first to define a discretization approach. This is realized in the form of a *histogram of normal vectors*. The input to our feature extraction module are sequences of normalized *gait silhouettes volumes*, or GSV in short. Images of subjects undergo a sequence of preprocessing steps, namely segmentation, camera calibration and size normalization, as further presented in section 5.1. We focus on developing the new descriptor and thus use the ready segmented silhouettes. In what follows, we describe our approach to compute the developed feature descriptor.

#### 4.1 Gait cycle detection

Gait is mainly a periodic motion. Hence, capturing representative information requires to determine at least a full cycle, for example motion included between two double support phases that have the same leg on the front. This includes, half way through, a double support phase with the legs inverted. The viewing directions of the sequences used here (and in most literature as well) result in a variation that occurs chiefly in the horizontal direction. For that reason, we define a signal  $m(t)$  that computes the second moments of the body masses around a vertical axis that passes through the center of the silhouettes:

$$m(t) = \sum_x b(x)|x - x_c|^2$$

$$\text{where } b(x) = \begin{cases} 1, & x \in \text{body} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

for all  $x$ -coordinates of pixels in the image;  $x_c$  is the vertical axis location. This signal has its peaks at double support phases (widest silhouettes) and its minima at single support phases. Thus, a full cycle starts at a local maximum, skips the following, then ends at the third one.

The computed moment signals are noisy, highly depending on the quality of the segmented silhouettes (Fig. 2 (a)). Those signals, hence, need to be smoothed first, before they can be used. We use the auto-correlation of  $m(t)$  for that purpose [43]. Auto-correlation signals have the advantage of being smooth and maintain the same cycle length as the original signal. Figure 2 (b) shows the right half of that signal, since it is symmetric. To limit the range of the auto-correlation signal, the moment signals  $m(t)$  are first normalized as follows:

$$\omega(t) = \frac{m(t) - \overline{m(t)}}{\text{range}(m)} \quad (5)$$

where  $\overline{m(t)}$  is the mean value of  $m(t)$  and

$$\text{range}(m) = \max_t \{m(t)\} - \min_t \{m(t)\} \quad (6)$$

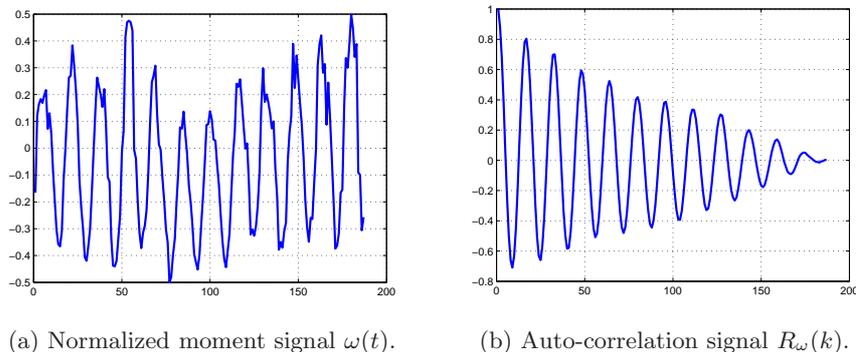
Finally, the auto-correlation of  $\omega(t)$  is computed:

$$R_\omega(k) = \sum_{t=1}^{n-k} \omega(t)\omega(t+k) \quad (7)$$

where  $n$  is the number of samples (frames) in the signal  $\omega(t)$ .

#### 4.2 Histograms of normal vectors

This is where the main computation of our proposed descriptor takes place. Since the total curvature of the silhouette boundary is defined to be the area of the



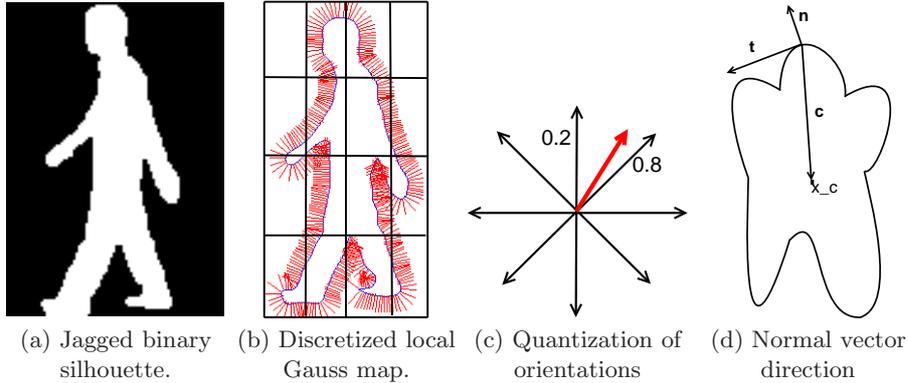
**Fig. 2.** Detecting the cycle length using second moments of silhouettes.

image of the Gauss map, we compute a discretized version of the Gauss map by accumulating the boundary normal vectors into a histogram. First, we extract the silhouette contour by a simple border following algorithm. Depending on the segmentation results, there might be discontinuities in the contour. Those are fixed by stitching contour segments which extremities are within a tolerable proximity. Images of silhouettes contours typically contain lots of noise, which affects the quality of the estimated Gauss map (Fig. 3 (a)). We overcome this artifact by smoothing them, using cubic spline interpolation, before computing the normal vectors.

It is very important to determine the orientation (sign) of the normal vectors to point outside the body of the silhouette. As shown in Fig. 3 (d), we first select a counterclockwise contour orientation by evaluating the cross product of a vector  $\mathbf{t}$  tangent to a convex region (we choose the head) with a vector  $\mathbf{c}$  pointing to the centroid of the silhouette. If the value is positive, the tangent vector points in the counterclockwise direction, otherwise, we flip the contour orientation. The direction of the normal vectors  $\mathbf{n}$  can now be determined by computing their cross-product with the tangent vector  $\mathbf{t}$  at the same contour point. A positive value indicates an outward pointing direction, otherwise, the normal vector is flipped.

The quantization of normal vectors orientations is done by assigning them to the two closest histogram bins, with linear weighting (Fig. 3 (c)). As a simple example, given an 8-bin histogram with  $45^\circ$ -wide bins and an orientation of  $54^\circ$ , and since  $54/45 = 1.2$ , that orientation will be assigned to bins 1 and 2 with respective weights 0.8 and 0.2. This reduces the quantization error that is involved with approaches that assign the orientation to a single bin [38].

As indicated earlier, local Gauss maps have the desirable property of uniquely identifying boundary segments. We approximate locality by dividing the silhouettes (actually their bounding boxes) into a regular grid as in Fig. 3 (b). Within each grid box, we compute the histograms of normal vectors separately. The feature descriptor is finally formed by concatenating the histograms of all grid



**Fig. 3.** Some steps involved in computing the histograms of normal vectors.

boxes into a single vector. We call it the histogram of normal vectors, or HoNV, in short.

### 4.3 Feature matching

We propose two variants to our approach: a frame-based one and a period-based one. In the frame-based variant, we choose a subset of representative key frames to match their feature vectors, one-to-one, to feature vectors of key frames of other sequences. First, key frames are selected by downsampling the number of frames in a full gait cycle to some fixed value for all sequences. Then, a long feature vector is formed by joining the vectors for individual key frames. As for the period-based variant, we compute the histograms for each frame as previously described and then average them over the period. Following the discussion in section 2, we expect the period-based approach to perform better, specially in cases where the frame rate is low, such as young children sequences. We will further investigate the difference between both approaches in the experimental results section.

Feature matching is then performed by evaluating, for all pairs of sequences (gallery and probe), a simple and efficient Euclidean distance metric. Given a probe feature vector  $\mathbf{P}_i$  and a gallery feature vector  $\mathbf{G}_j$ , the distance between them is computed as:

$$D_{i,j} = \|\mathbf{P}_i - \mathbf{G}_j\|_2 \quad (8)$$

In the dataset we are using, only one single sequence is captured per subject. Therefore, learning based methods, such as statistical discriminant analysis tools cannot be applied here. It is definitely expected, as demonstrated elsewhere, that such approaches will improve performance in case datasets with more training samples per class are used.



(a) 55° subsequence. (b) 65° subsequence. (c) 75° subsequence. (d) 85° subsequence.

**Fig. 4.** Typical frames from a gait sequence that illustrate the four different viewing angles.

## 5 Large Population dataset

We have evaluated our approach using the OU-ISIR Large Population dataset [44]. The main significance of using this dataset lies in the large number of subjects it contains. This provides statistically reliable results that cannot be verified with other available smaller datasets. In all, it has 4,007 subjects. However, in our experiments, we have used subsets of sizes ranging from 3,141 to 3,706 subjects. This is by far larger than any other available dataset we know of.

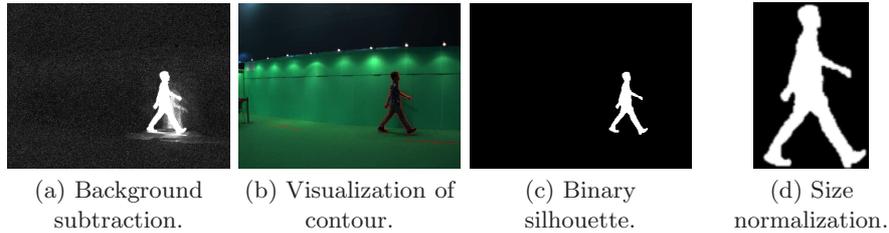
Other than being a large scale gait database, the used dataset balances the gender representation with 2,135 male subjects and 1,872 female subjects. It also has a wide range of ages spanning from 1 to 94 years old subjects. The presence of children, in particular, provides a unique and challenging testing situation, where the number of frames per gait cycle drops significantly. An additional advantage of using this dataset is that silhouette extraction is performed accurately. For the purpose of this paper, the tasks of background subtraction and silhouette segmentation are out of scope. Our aim is to validate our new feature descriptor, separately from errors that may arise from poorly extracted silhouettes.

Moreover, this dataset captures subjects from different viewing angles. Currently available datasets simply present their subjects from a “side view” angle. The subjects in the OU-ISIR dataset, however, are viewed from an angle that is carefully evaluated, gradually changing from about 50° to 90°. Based on that angle, each sequence is split into 4 subsequences, centered at 55, 65, 75 and 85 degrees respectively as illustrated in Fig. 4. This allows us to perform a more rigorous performance evaluation that verifies, first, the robustness of our approach on each view angle separately, then, on entire sequences as well.

### 5.1 Preprocessing

In this section, we briefly note the preliminary stages required to reach the gait silhouettes volumes (GSV) input used by our feature evaluation code.

1. Silhouette extraction: Background subtraction is performed along with a graph-cut approach for segmentation [45]. The contour is visualized to fix any errors manually.



**Fig. 5.** The output of four preprocessing steps performed on the input sequences to extract gait silhouette volumes (GSV).

2. Camera calibration: Intrinsic and extrinsic camera parameters are estimated to correct the distortion and carry out the necessary camera rotation.
3. Image registration: A moving-average filter is applied to the extracted silhouette images then, their sizes are normalized to  $88 \times 128$  pixels.

Some of the previous steps are illustrated in Fig. 5.

## 6 Experimental results

We present here the results of evaluating our new feature descriptor on the OU-ISIR Large Population gait dataset [44]. This dataset is divided into five subsets, based on viewing angle, A-55, A-65, A-75, A-85 and A-All, where the last one contains full sequences with the viewing angle gradually varying from roughly 50 to 90 degrees. The number of subjects in each subset is shown in Tab. 1. We implemented the proposed method using Matlab without any code optimization. Under an Intel Core i7 processor running at 3.5 GHz, our code processes about 60 frames per second, the equivalent of two gait cycles per second given the used dataset.

**Table 1.** Number of subject in each subset of the OU-ISIR dataset.

Dataset	A-55	A-65	A-75	A-85	A-All
No. of subjects	3706	3770	3751	3249	3141

As mentioned earlier in section 4.3, we have developed a frame-based variant and a period-based variant. In addition, several parameters, such as the number of histogram bins and the number of grid boxes can be tuned to enhance the performance. In this section, we compare the variants of our approach and study the effect of varying some parameters on the results. Finally, we compare it to other recent methods.

## 6.1 Evaluation of descriptor parameters

We consider first the effect of changing the number of grid boxes on the performance. A grid with one box corresponds to a global Gauss map, computed for the entire silhouettes. As the number of boxes increases, we have more and more local maps. We illustrate the verification performance using Receiver Operating Characteristic (ROC) curves. This is a common tool used in biometric applications. It denotes the trade-off between false rejection rate (FRR) and false acceptance rate (FAR) as the acceptance threshold is being varied. We also use Cumulative Match Characteristic (CMC) curves to illustrate the identification performance. This tool estimates the probability that a correct match is observed within the top  $k$  “closest” matches.

We use the A-85 subset as a study case, employ the period-based approach and fix the number of histogram bins to 16. The effect of changing the number of grid boxes (1, 16, 36, 64) is illustrated in Fig. 6. As expected, the performance of the global Gauss map (1 box) is the lowest. Next, we vary the number of histogram bins (8, 16, 32) for a fixed grid of 36 boxes. The results are shown in Fig. 7. Further performance enhancements are achievable but cannot be shown at the used scale. Instead, we use the Equal Error Rate (EER) value from the ROC curve, that is the value at which the false acceptance and rejection rates (FAR and FRR) are equal. We also use the rank-1 and rank-5 identification rates as representative measures from the CMC curve. For the  $88 \times 128$  pixel silhouettes that we use, best performance (EER = 0.013, rank-1 id. = 93.0%, rank-5 id. = 96.7%) is obtained using 32-bin histograms and a grid of 169 ( $13 \times 13$ ) boxes.

## 6.2 Evaluation of descriptor variants

The results shown earlier are for the period-based approach, as pointed out. We now evaluate the performance of the frame-based approach. Matching key frames, as required with this approach, necessitates the alignment of the phases of the gallery and probe sequences as much as possible. We first used the cycle boundaries provided with the dataset in order to align the viewpoint. This caused the sequences to be out of phase, and severely deteriorated the performance. The average EER is 0.38 and the identification rates are 14.9% for rank-1 and 19.6% for rank-5.

Next, we extracted the cycles using our own code. Even though the phases are now synchronized, the view angles are different between probe and gallery sequences. That is due to the fact that the view angle gradually changes over the walking course, as mentioned in the dataset description. We still notice a significant performance boost (EER = 0.079, rank-1 id. = 55.4%, rank-5 id. = 67.6%), which means that the frame-based approach is more robust to view changes than it is to phase shift. However, the overall results are unacceptable when compared to the period-based approach. Thus, we will not pursue the use of this approach further here since it is unsuitable, at least for this dataset.

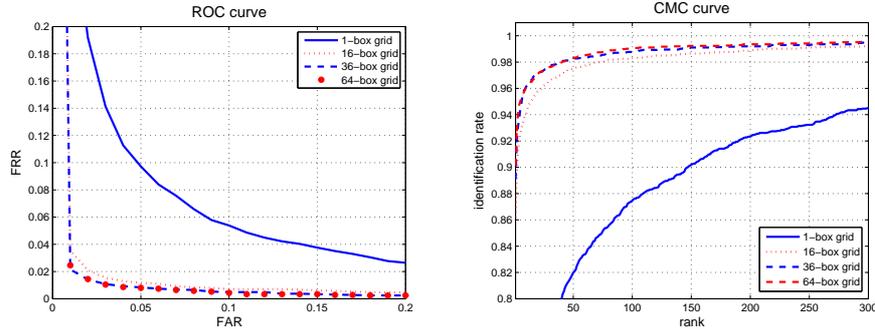


Fig. 6. Performance comparison for varying the number of grid boxes.

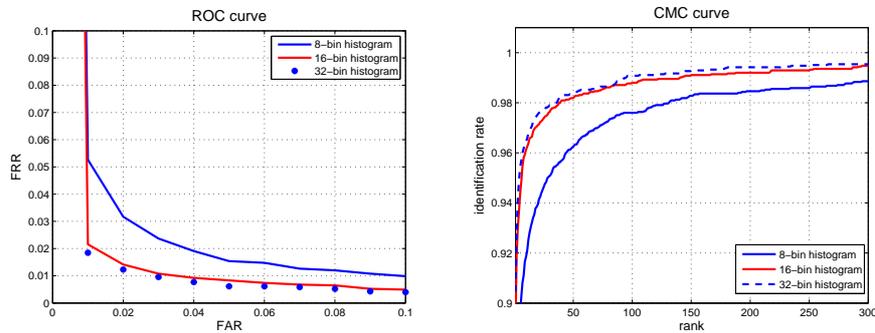
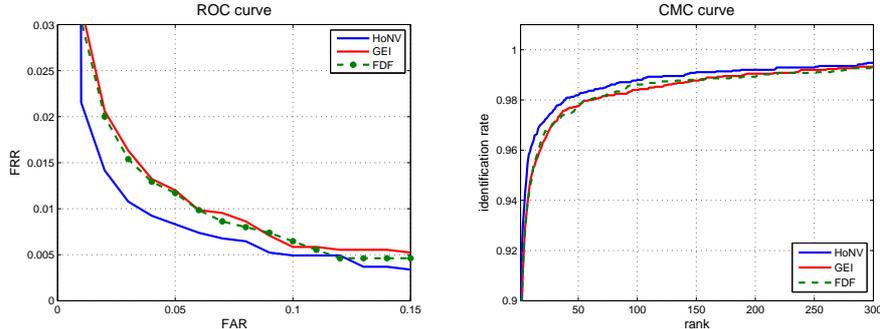


Fig. 7. Performance comparison for varying the number of histogram bins.

### 6.3 Comparison against other methods

Finally, we compare our feature descriptor with state-of-the-art approaches. In [44], five methods are tested on the same dataset we are using here. The gait energy image (GEI) and the frequency domain feature (FDF) approaches exhibit the best performance of all the tested methods. First, we compare our method against those two. The significant parts of the ROC and CMC curves (A-85 subset as above) are shown in Fig. 8. At these high identification rates, visualizing the performance improvements becomes significantly more difficult, and so we will use the EER metric, the rank-1 and the rank-5 identification rates as we did earlier. We also compare our method to the average of histograms of oriented gradients (HOG). That is a very recent approach [40] that employs HOG, a powerful human detection feature [46]. We implement the method according to the details in [40] and [46] and test it on the OU-ISIR dataset. We show all results in Tab. 2. Our method is abbreviated as HoNV for histogram of normal vectors. The best performance out of GEI and FDF is recorded in the GEI column. Which method performs best depends on the used dataset. Finally, the performance of the average HOG technique can be found in the HOG column. It



**Fig. 8.** Performance comparison versus other approaches.

is evident that our approach supersedes all other methods on almost all values of EER and on all values of ranking.

**Table 2.** Comparison of the performance of our approach (HoNV) versus GEI / FDF and HOG methods.

Dataset	EER [ $\times 10^{-2}$ ]			Rank-1 [%]			Rank-5 [%]		
	HoNV	HOG	GEI	HoNV	HOG	GEI	HoNV	HOG	GEI
A-55	1.70	<b>1.59</b>	2.06	<b>91.6</b>	90.4	84.7	<b>95.1</b>	95.0	92.4
A-65	<b>1.43</b>	1.54	1.83	<b>92.1</b>	90.9	86.6	<b>95.4</b>	95.1	92.8
A-75	<b>1.31</b>	1.70	1.97	<b>93.3</b>	91.5	86.9	<b>96.4</b>	95.3	92.9
A-85	<b>1.32</b>	1.78	2.00	<b>93.0</b>	89.9	85.9	<b>96.7</b>	95.0	92.8
A-All	<b>0.58</b>	0.81	1.13	<b>97.5</b>	96.2	94.2	<b>98.7</b>	98.0	97.1

## 7 Conclusion

We presented a new feature descriptor for human identification using gait: histograms of normal vectors. We developed the method, explained the theory of Gauss maps on which it relies and demonstrated how it can exploit boundary curvature information. We verified that the new feature encodes enhanced shape information when compared with state-of-the-art approaches. Our superior results are validated through rigorous experiments using the world’s largest gait database. In addition, our method does not require any training which makes it also valid when a small number of sequences per subject is available.

We are currently working on several improvements to the presented approach. When a large training sample is available, we are studying to which extent some statistical tools, such as discriminant analysis, can enhance the identification results. The main drawback of our approach is that it relies on the quality of

the extracted silhouettes. In that regard, we will carry out further comparisons with other contour-based approaches. On a different note, recent work [21] has suggested that combining gait features can improve performance. Thus, we would like to study how does fusing our method with other approaches improve the discriminative power of the combined method.

**Acknowledgement.** This work was supported in part by the JST CREST “Behavior Understanding based on Intention-Gait Model” project.

## References

1. Kusakunniran, W., Wu, Q., Zhang, J., Li, H.: Gait recognition under various viewing angles based on correlated motion regression. *Circuits and Systems for Video Technology, IEEE Transactions on* **22** (2012) 966–980
2. Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T., Yagi, Y.: Gait recognition using a view transformation model in the frequency domain. In: *Computer Vision—ECCV 2006*. Springer (2006) 151–163
3. BenAbdelkader, C., Cutler, R., Davis, L.: View-invariant estimation of height and stride for gait recognition. In: *Biometric Authentication*. Springer (2002) 155–167
4. Tsuji, A., Makihara, Y., Yagi, Y.: Silhouette transformation based on walking speed for gait identification. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE* (2010) 717–722
5. Kusakunniran, W., Wu, Q., Zhang, J., Li, H.: Speed-invariant gait recognition based on procrustes shape analysis using higher-order shape configuration. In: *Image Processing (ICIP), 2011 18th IEEE International Conference on, IEEE* (2011) 545–548
6. Hossain, A., Makihara, Y., Wang, J., Yagi, Y., et al.: Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control. *Pattern Recognition* **43** (2010) 2281–2291
7. Sarkar, S., Phillips, P.J., Liu, Z., Vega, I.R., Grother, P., Bowyer, K.W.: The humanid gait challenge problem: Data sets, performance, and analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **27** (2005) 162–177
8. Niyogi, S.A., Adelson, E.H.: Analyzing and recognizing walking figures in xyt. In: *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR’94., 1994 IEEE Computer Society Conference on, IEEE* (1994) 469–474
9. Murase, H., Sakai, R.: Moving object recognition in eigenspace representation: gait analysis and lip reading. *Pattern recognition letters* **17** (1996) 155–162
10. BenAbdelkader, C., Cutler, R., Nanda, H., Davis, L.: Eigengait: Motion-based recognition of people using image self-similarity. In: *Audio-and Video-Based Biometric Person Authentication, Springer* (2001) 284–294
11. BenAbdelkader, C., Cutler, R., Davis, L.: Motion-based recognition of people in eigengait space. In: *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on, IEEE* (2002) 267–272
12. Wang, L., Hu, W., Tan, T.: A new attempt to gait-based human identification. In: *Pattern Recognition, 2002. Proceedings. 16th International Conference on. Volume 1., IEEE* (2002) 115–118
13. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **25** (2003) 1505–1518

14. Liu, Z., Sarkar, S.: Simplest representation yet for gait recognition: Averaged silhouette. In: Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. Volume 4., IEEE (2004) 211–214
15. Han, J., Bhanu, B.: Individual recognition using gait energy image. Pattern Analysis and Machine Intelligence, IEEE Transactions on **28** (2006) 316–322
16. Bashir, K., Xiang, T., Gong, S.: Gait recognition using gait entropy image. (2009)
17. Wang, C., Zhang, J., Pu, J., Yuan, X., Wang, L.: Chrono-gait image: a novel temporal template for gait recognition. In: Computer Vision–ECCV 2010. Springer (2010) 257–270
18. Liu, Z., Sarkar, S.: Improved gait recognition by gait dynamics normalization. Pattern Analysis and Machine Intelligence, IEEE Transactions on **28** (2006) 863–876
19. Sivapalan, S., Chen, D., Denman, S., Sridharan, S., Fookes, C.: Gait energy volumes and frontal gait recognition using depth images. In: Biometrics (IJCB), 2011 International Joint Conference on, IEEE (2011) 1–6
20. Hofmann, M., Bachmann, S., Rigoll, G.: 2.5d gait biometrics using the depth gradient histogram energy image. In: Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on, IEEE (2012) 399–403
21. Makihara, Y., Muramatsu, D., Iwama, H., Yagi, Y.: On combining gait features. In: Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, IEEE (2013) 1–8
22. Tang, S., Wang, X., Lv, X., Han, T.X., Keller, J., He, Z., Skubic, M., Lao, S.: Histogram of oriented normal vectors for object recognition with a depth sensor. In: Computer Vision–ACCV 2012. Springer (2013) 525–538
23. Gauss, K.F.: General investigations of curved surfaces of 1827 and 1825, translated with notes and a bibliography by J.C. Morehead and A.M. Hildebeitel. The Princeton University Library (1902)
24. Hazewinkel, M.: Encyclopaedia of mathematics. Volume 13. Springer (2001)
25. Johansson, G.: Visual motion perception. Scientific American (1975)
26. Bobick, A.F., Johnson, A.Y.: Gait recognition using static, activity-specific parameters. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Volume 1., IEEE (2001) 1–423
27. Yam, C., Nixon, M.S., Carter, J.N.: Automated person recognition by walking and running via model-based approaches. Pattern Recognition **37** (2004) 1057–1072
28. Urtasun, R., Fua, P.: 3d tracking for gait characterization and recognition. In: Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on, IEEE (2004) 17–22
29. Yang, H.D., Lee, S.W.: Reconstruction of 3d human body pose for gait recognition. In: Advances in Biometrics. Springer (2005) 619–625
30. Kumar, M., Babu, R.V.: Human gait recognition using depth camera: a covariance based approach. In: Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing, ACM (2012) 20
31. Hussein, M.E., Torki, M., Gawayyed, M.A., El-Saban, M.: Human action recognition using a temporal hierarchy of covariance descriptors on 3d joint locations. In: Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, AAAI Press (2013) 2466–2472
32. Phillips, P.J., Sarkar, S., Robledo, I., Grother, P., Bowyer, K.: The gait identification challenge problem: data sets and baseline algorithm. In: Pattern Recognition, 2002. Proceedings. 16th International Conference on. Volume 1., IEEE (2002) 385–388

33. Cuntoor, N., Kale, A., Chellappa, R.: Combining multiple evidences for gait recognition. In: *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP'03). 2003 IEEE International Conference on*. Volume 3., IEEE (2003) III–33
34. Liu, Y., Collins, R., Tsin, Y.: *Gait sequence analysis using frieze patterns*. Springer (2002)
35. Lam, T.H., Cheung, K.H., Liu, J.N.: Gait flow image: A silhouette-based gait representation for human identification. *Pattern recognition* **44** (2011) 973–987
36. Tao, D., Li, X., Maybank, S.J., Wu, X.: Human carrying status in visual surveillance. In: *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. Volume 2., IEEE (2006) 1670–1677
37. Kobayashi, T., Otsu, N.: A three-way auto-correlation based approach to human identification by gait. In: *IEEE Workshop on Visual Surveillance*. Volume 1., Citeseer (2006) 4
38. Kobayashi, T., Otsu, N.: Three-way auto-correlation approach to motion recognition. *Pattern Recognition Letters* **30** (2009) 212–221
39. Hofmann, M., Rigoll, G.: Improved gait recognition using gradient histogram energy image. In: *19th IEEE International Conference on Image Processing (ICIP), IEEE (2012)* 1389–1392
40. Hofmann, M., Rigoll, G.: Exploiting gradient histograms for gait-based person identification. In: *IEEE International Conference on Image Processing (ICIP)*. (2013) 4171–4175
41. Lombardi, S., Nishino, K., Makihara, Y., Yagi, Y.: Two-point gait: Decoupling gait from body shape. In: *Computer Vision (ICCV), 2013 IEEE International Conference on, IEEE (2013)* 1041–1048
42. Lee, C.S., Elgammal, A.: Style adaptive contour tracking of human gait using explicit manifold models. *Machine Vision and Applications* **23** (2012) 461–478
43. Boulgouris, N.V., Plataniotis, K.N., Hatzinakos, D.: Gait recognition using dynamic time warping. In: *Multimedia Signal Processing, 2004 IEEE 6th Workshop on, IEEE (2004)* 263–266
44. Iwama, H., Okumura, M., Makihara, Y., Yagi, Y.: The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. *Information Forensics and Security, IEEE Transactions on* **7** (2012) 1511–1521
45. Makihara, Y., Yagi, Y.: Silhouette extraction based on iterative spatio-temporal local color transformation and graph-cut segmentation. In: *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on, IEEE (2008)* 1–4
46. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR. Volume 1., IEEE (2005)* 886–893