

Gait Recognition using Period-based Phase Synchronization for Low Frame-rate Videos

Atsushi Mori, Yasushi Makihara, Yasushi Yagi
The Institute of Scientific and Industrial Research
Osaka University
Ibaraki, Osaka, Japan
 Email: {mori, makihara, yagi}@am.sanken.osaka-u.ac.jp

Abstract—This paper proposes a method for period-based gait trajectory matching in the eigenspace using phase synchronization for low frame-rate videos. First, a gait period is detected by maximizing the normalized autocorrelation of the gait silhouette sequence for the temporal axis. Next, a gait silhouette sequence is expressed as a trajectory in the eigenspace and the gait phase is synchronized by time stretching and time shifting of the trajectory based on the detected period. In addition, multiple period-based matching results are integrated via statistical procedures for more robust matching in the presence of fluctuations among gait sequences. Results of experiments conducted with 185 subjects to evaluate the performance of the gait verification with various spatial and temporal resolutions, demonstrate the effectiveness of the proposed method.

Keywords—gait recognition; low frame rate; phase synchronization; gait period; PCA

I. INTRODUCTION

There is a growing necessity in modern society for recognition of individuals in many situations, such as in surveillance systems and for access control. For personal recognition, many biometric-based authentication methods have been proposed using a wide variety of cues, such as fingerprints, finger or hand vein, voiceprint, iris, face, handwriting, and gait. Of these, gait recognition has recently gained considerable attention as a promising cue for surveillance systems, due to the ability of ascertaining identity from a distance.

Many approaches to gait recognition have been proposed, including direct matching of gait silhouette sequences [1], gait silhouette matching in the eigenspace [2], spatio-temporal gait volume-based feature matching [3][4], and gait period-based feature matching [5][6].

Of these approaches, gait period-based feature matching [5][6] has the potential to adapt to gait fluctuations by statistically integrating multiple period results, whereas sequence-based methods [2] cannot adapt to such gait fluctuations. In terms of performance evaluation, most of the existing methods were tested on image sequences with standard spatial resolution (e.g., 128-pixel silhouette height) and standard temporal resolution (e.g., 30 fps).



Figure 1. Gait silhouette sequence at 10 fps

On the other hand, videos captured by CCTV cameras in the street are usually of low quality both in terms of spatial and temporal resolution due to limitations on communication bandwidth and storage devices. Therefore, a drop in performance is expected if the existing methods are applied to the CCTV videos. For example, in gait period-based methods [5][6] false gait period detection may occur as gait silhouette images are sparse on the temporal axis.

This paper proposes a method for gait recognition using period-based phase synchronization for low frame-rate videos. The proposed method adapts to gait fluctuations by using a gait period as the matching unit, and performs matching by phase synchronization taking into account false gait period detection due to the low frame-rate videos. In addition, using the period as the matching unit enables robust matching by removing outliers for gait fluctuations [5][6][1]. By means of experiments, we evaluate the performance of the proposed methods with respect to both spatial and temporal resolution, whereas Mowbray et al. [7] refer only to the effect of spatial resolution on recognition performance.

II. GAIT FEATURE EXTRACTION

A. Extraction of Gait Silhouette Sequence

The first step involves extracting gait silhouette images, by exploiting background subtraction. The next step includes scaling and registration of the extracted silhouette images. We scale the silhouette images to ensure that the height of the region is a predefined size, and the aspect ratio of the region is maintained. Finally, we produce an 88×128 pixel image as shown in Fig. 1.

B. Gait Period Detection

Gait is a periodic motion of two steps, and the gait period is detected by maximizing the normalized autocorrelation of the gait silhouette sequence for the temporal axis. We

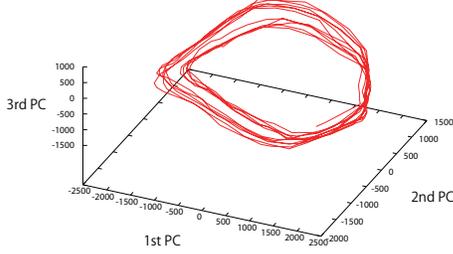


Figure 2. Gait silhouette sequence in PCA space

calculate the normalized autocorrelation as

$$C(N) = \frac{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n)g(x,y,n+N)}{\sqrt{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n)^2} \sqrt{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n+N)^2}} \quad (1)$$

$$T(N) = N_{total} - N - 1 \quad (2)$$

where $C(N)$ is the autocorrelation for an N -frame shift, $g(x,y,n)$ is the silhouette value at position (x,y) in the n -th frame, and N_{total} is the total number of frames in the sequence. We set the domain of the natural gait period to be $[0.83, 1.3]$ sec empirically and convert it into a frame-based domain for each frame rate.

C. Dimension Reduction by PCA

PCA is applied to reduce the dimensions of the gait silhouette. All the gallery sequences are used for PCA training and the number of dimensions is determined so that the cumulative contribution ratio of eigenvalues is more than 90 %.

As a result, the gait silhouette and gait silhouette sequence are expressed as a point and a trajectory in the PCA space, respectively, as illustrated in Fig. 2. Note that the two clusters of trajectories in Fig. 2 correspond to the left and right steps, respectively, and that this derives from the left-right asymmetry of gait motion.

III. MATCHING OF GAIT SILHOUETTE SEQUENCE

A. Assumption on frame rates of probe and gallery

Although the most challenging problem setting occurs when matching two low frame-rate sequences, the problem setting in this paper is relaxed to allow matching between a low frame-rate probe sequence and a normal frame-rate gallery sequence. This problem setting is still realistic in the following example scenario. A CCTV camera captures a suspect in a crime scene as a low frame-rate probe. Then, as a result of the police investigation, a number of possible suspects are identified, and a gait image sequence of each is secretly captured as a normal frame-rate gallery using a conventional video camera. Finally, by matching the sequences the identity of the suspect can be confirmed.

Table I. MATCHING UNIT

Method	Unit	
	Probe	Gallery
Sequence-Period	Sequence	Period
Period-Period	Period	Period
Sequence-Sequence[2]	Sequence	Sequence

B. Matching Unit

Generally speaking, both the gait period and whole sequence are potential matching units. The merit of gait period-based matching is that outliers due to gait fluctuations can be removed from the results by statistical integration of multiple period results [5][6][1], while a demerit is that the relation between gait periods, in other words, the continuity of gait phase between the last frame of the previous period and the first frame of the current period, is discarded. In particular, in very low frame rates, like 1 fps, gait trajectory matching is often degraded to single silhouette matching, as most of the gait periods in such a low frame rate include only a single silhouette.

On the other hand, the merit and demerit of using the whole sequence, are exactly the reverse of those for the gait period, that is, consideration of phase continuity between periods and sensitivity to gait fluctuations. In particular, where whole sequences are used for both the probe and gallery as in [2], gait fluctuations have a deep impact on matching performance.

Thus, by considering a trade-off between gait fluctuations and phase continuity, we implement two methods: sequence-period and period-period as shown in Table I and evaluate the performance of each through experiments. Though a combination of period-sequence is possible, we use the gait period as the gallery gait matching unit, because gait period detection errors in the normal frame-rate gallery are relatively small.

C. Phase Synchronization

In this section, we describe a method for calculating trajectories in eigenspace with phase synchronization of each matching unit. First, we denote the probe and gallery sequences as S_P and S_G , respectively, and their frame rates and true gait periods as f_P and $f_G (> f_P)$ [fps], and P_P and P_G [frame], respectively. Next, the gait phase of S_P in the t_P -th frame and that of S_G in the t_G -th frame are defined as $\theta_P(t_P)$ and $\theta_G(t_G)$, respectively. Because the gait phase is a relative value, we set the initial phase $\theta_P(0)$ of the probe sequence to zero. Thus, each gait phase is expressed as follows

$$\theta_G(t_G) = \theta_G(0) + \frac{t_G}{P_G} \quad (3)$$

$$\theta_P(t_P) = \frac{t_P}{P_P}, \quad (4)$$

where $\theta_G(0)$ is an initial phase of the gallery. The equation $\theta_P(t_P) = \theta_G(t_G)$ is necessary for phase synchronization, a

relation between t_P and t_G is expressed as

$$t_G = at_P + b \quad (5)$$

$$a = \frac{P_G}{P_P}, b = -P_G\theta_G(0), \quad (6)$$

where a and b are the time stretching and time shifting coefficients, respectively. Basically, the detected gait period of the probe $P'_P (= P_P + \Delta P_P)$ and that of the gallery $P'_G (= P_G + \Delta P_G)$ are not always the same as the true gait periods P_P and P_G , owing to the existence of the respective quantization errors, ΔP_P and ΔP_G . In particular, the quantization errors in a low frame-rate probe have a relatively large impact on the time stretching coefficient a . We, therefore, estimate the quantization error ΔP_P up to the gallery frame rate. Note that the previous period-based methods[5][6][1] does not consider such quantization error ΔP_P .

In addition, the initial phase of the gallery $\theta_G(0)$ is unknown, and the time shifting coefficient b also needs to be estimated. Thus, given feature vectors in the eigenspace of the probe and the gallery as $\mathbf{x}_P(t_P)(t_P = 1, \dots, N_P)$ and $\mathbf{x}_G(t_G)(t_G = 1, \dots, P'_G)$, respectively, the time stretching coefficient a and time shifting coefficient b are estimated so as to minimize the distances between the feature vectors as follows

$$d(S_P, S_G) = \arg \min_{a, b} \sum_{t_P=1}^{N_P} \|\mathbf{x}_P(t_P) - \mathbf{x}_G(\lfloor t_G + 0.5 \rfloor)\|_{L_1} \quad (7)$$

$$a \in \left\{ \frac{P'_G}{P'_P - ir} \mid i \in \mathbb{Z}, |i| \leq \frac{1}{2r} \right\} \quad (8)$$

$$b \in \{j \mid j \in \mathbb{Z}, 0 \leq j \leq P_G - 1\}, \quad (9)$$

where $\lfloor \cdot \rfloor$ is the floor function, $\|\cdot\|_{L_1}$ means the L_1 norm, and r is the frame rate ratio ($=f_P/f_G$). As described in the previous section, the gallery feature vector \mathbf{x}_G is expressed as a periodic function $\mathbf{x}_G(t + P_G) = \mathbf{x}_G(t)$, because its matching unit is the gait period.

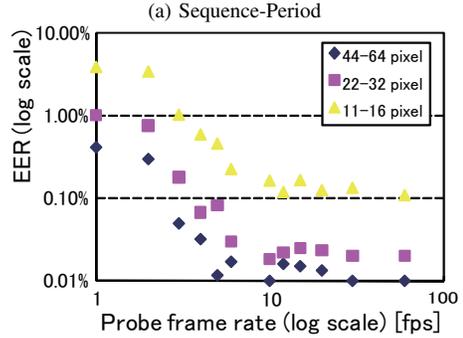
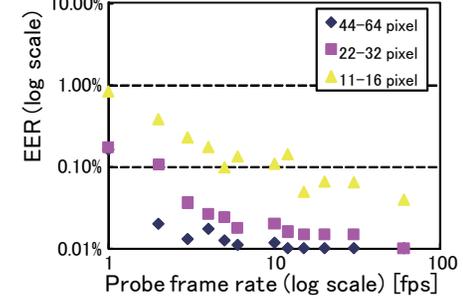
IV. EXPERIMENT

A. Dataset

A total of 370 gait sequences from 185 subjects were used in the experiments. Each subject is requested to walk on a speed-controlled treadmill and two gait sequences are captured: one for the probe, and the other for the gallery. The gait silhouette size, frame rate, and recording time for each sequence are 88×128 pixels, 60 fps, and 6 sec, respectively. Low frame-rate gait sequences are constructed by selecting the gait silhouettes at a specified interval. Low spatial-resolution gait sequences are constructed from the gait silhouettes merely by scaling down the original silhouettes. Table II shows the sets of spatial and temporal resolutions used in this experiment.

Table II. SPATIO-TEMPORAL RESOLUTION OF DATASET

Spatial (pixel)	44 × 64, 22 × 32, 11 × 16
Temporal (fps)	1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60



(a) Sequence-Period
(b) Period-Period
Figure 3. Evaluation on spatial resolution

We also constructed datasets to evaluate the robustness to gait fluctuations. Here, Normalized Auto Correlation (NAC) in Eq. (1) is used as a measure for gait fluctuation, where a higher NAC denotes a stable gait and vice versa. Then, we define the two datasets: DB_{high} comprising 100 subjects with the highest NAC, and DB_{low} comprising 100 subjects with the lowest NAC.

B. Evaluation on Spatial Resolution

We tested the proposed methods, Sequence-Period and Period-Period, on the DB_{high} dataset to investigate spatial resolution. The probe frame rate ranges from 1 to 60 fps, while the gallery frame rate is 60 fps. Sets of spatial resolution used are 44×64 , 22×32 , and 11×16 pixels. The gait verification performance is evaluated using the Equal Error Rate (EER) of FRR and FAR derived from the Receiver Operating Characteristics (ROC) curve [8]. The ROC curve indicates a trade-off between the False Rejection Rate (FRR) and False Acceptance Rate (FAR) when the receiver changes acceptance thresholds.

Results are shown in Fig. 3. In these graphs, the horizontal axis denotes the probe frame rate, while the vertical axis indicates the EER. Note that both axes are expressed in a log scale. We can see that the lower the spatial resolution is, the lower is the performance. In particular, the performance drop associated with the change in resolution from 22×32 pixels to 11×16 pixels is much worse than that from 44×64

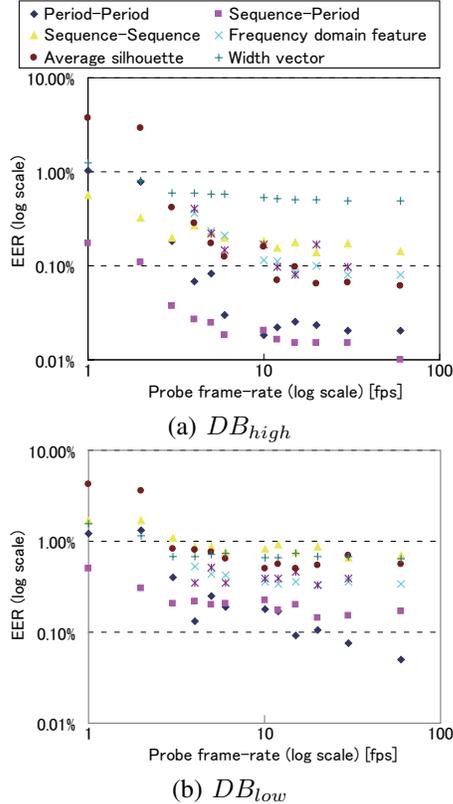


Figure 4. Evaluation on temporal resolution

pixels to 22×32 pixels. Therefore, we conclude that scaling down the spatial resolution to 22×32 pixels has a relatively low impact on the performance decline.

C. Evaluation on Temporal Resolution

We tested the two proposed methods and four related works, namely Sequence-Sequence [2], Averaged silhouette [6], Frequency-domain feature [5], Sequence-Period with Width Vector [9], on the DB_{high} and DB_{low} datasets. Probe frame rates range from 1 to 60 fps, while the gallery frame rate is set to 60 fps. The spatial resolution is fixed at 22×32 pixels.

Results are shown in Fig. 4 where the meaning of each axis is the same as that in Fig. 3. We can see that the proposed Sequence-Period method achieves the best performance for almost all temporal resolutions in DB_{high} . Reflecting on the performance of Sequence-Period, it is clear that the gait fluctuation in the dataset and the loss of phase continuity in the probe sequences are negligible. On the other hand, in DB_{low} , Period-Period achieves the best performance for the higher temporal resolutions, while Sequence-Period achieves the best performance for lower resolutions. This is why at higher temporal resolutions, Period-Period is effective due to gait fluctuations, while at the lower ones, Sequence-Period is effective due to discarding phase continuity. This tendency is consistent with the trade-off described in III-B.

On the other hand, even if we select the optimal method for each frame rate, the performance for DB_{low} is, on average, much worse than that for DB_{high} . This indicates that there is room for improvement in terms of adapting to gait fluctuations.

V. CONCLUSION

This paper describes gait recognition using period-based phase synchronization with low frame-rate probes and a normal frame-rate gallery. By using the gait period as the matching unit and adopting the minimum distance of each unit's distance, the proposed methods can adapt to gait fluctuations. Phase synchronization including time shifting and time stretching is applied by considering quantization errors of the gait period detection in low frame-rate videos. Based on the results of experiments with 185 subjects, the proposed methods outperform existing methods for all frame rates. Future works include the following:

- Experimentation with noisy data;
- More robust matching of fluctuating gait sequences; and
- Matching in cases where both the probes and galleries are low frame-rate videos.

ACKNOWLEDGMENT

This work was supported by Grant-in-Aid for Scientific Research(S) 21220003.

REFERENCES

- [1] S. Sarkar, J. Phillips, Z. Liu, I. Vega, P. Grother, and K. Bowyer, "The humanid gait challenge problem: Data sets, performance, and analysis," *Trans. of Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, 2005.
- [2] H. Murase and R. Sakai, *Moving object recognition in eigenspace representation: gait analysis and lip reading*. Elsevier Science Inc., feb 1996, vol. 17.
- [3] Y. Ohara, R. Sagawa, T. Echigo, and Y. Yagi, "Gait volume: Spatio-temporal analysis of walking," in *Proc. of the 5th Workshop on Omnidirectional Vision, Camera Networks and Non-classical cameras*, 2004, pp. 79–90.
- [4] S. Niyogi and E. Adelson, "Analyzing and recognizing walking figures in xyt," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1994, pp. 469–474.
- [5] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, *Gait Recognition Using a View Transformation Model in the Frequency Domain*. Springer Berlin Heidelberg, jul 2006, vol. 3953.
- [6] Z. Liu and S. Sarkar, "Simplest representation yet for gait recognition: Averaged silhouette," in *Proc. of the 17th Int. Conf. on Pattern Recognition*, vol. 1, Aug. 2004, pp. 211–214.
- [7] S. Mowbray and M. Nixon, *Automatic Gait Recognition via Fourier Descriptors of Deformable Objects*. Springer Berlin / Heidelberg, 2003, vol. 2688.
- [8] P. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The feret evaluation methodology for face-recognition algorithms," *Trans. of Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.
- [9] A. Kale, N. Cuntoor, B. Yegnanarayana, A. Rajagopalan, and R. Chellappa, *Gait Analysis for Human Identification*. Springer Berlin / Heidelberg, 2003, vol. 2688.