Gait Recognition using Periodic Temporal Super Resolution for Low Frame-rate Videos

Naoki Akae, Yasushi Makihara, and Yasushi Yagi
Osaka University
{akae,makihara,yagi}@am.sanken.osaka-u.ac.jp

Abstract

This paper describes a method of gait recognition where both a gallery and a probe are based on low frame-rate videos. The sparsity of phases (stances) per gait period makes it much harder to match the gait using existing gait recognition algorithms. Consequently, we introduce a super resolution technique to generate a high frame-rate periodic image sequence as a preprocess to matching. First, the initial phase for each frame is estimated based on an exemplar of a high frame-rate gait image sequence. Images between a pair of adjacent frames sorted by the estimated phases are then filled using a morphing technique to avoid ghosting effects. Next, a manifold of the periodic gait image sequence is reconstructed based on the estimated phase and morphed images. Finally, the phase estimation and manifold reconstruction are iterated to generate better high frame-rate images in the energy minimization framework. Experiments with real data on 100 subjects demonstrate the effectiveness of the proposed method particularly for low frame-rate videos of less than 5 fps.

1. Introduction

Gait recognition is one of the behavioral biometrics which has attracted much attention because it can be applied to uncooperative subjects and/or subjects at a distance from cameras. In fact, a burglar's gait captured by a CCTV camera was accepted as evidence in a UK court [27].

Although gait recognition has promising properties, it still suffers from many types of intra-class variations such as observation views [10][2][28][7], walking speeds [11][25], clothing [6], and elapsed times [21][13]. In addition to these difficulties, gait recognition for *low frame-rate* videos has remained an important and challenging problem because (1) digital videos captured by CCTV cameras are often recorded at quite low frame-rates (e.g., 1 to 6 fps) due to limitations of the communication bandwidth and/or storage size, and (2) the sparsity of phases (stances) per one gait period significantly degrades the existing gait recognition algorithms.

Despite its importance, only a few studies have tackled gait recognition with low frame-rates. Mori et al. [14] proposed a matching method between a low frame-rate probe (input/query) sequence and a normal frame-rate gallery (enrollment) sequence based on period-based phase synchronization. Although the problem setting was effective for several cases of criminal investigation, it was naturally not applicable to matching between dual low frame-rate videos such as cross-camera person matching.

In such dual low frame-rate problem settings, temporal interpolation is one possible approach. Al-Huseiny et al. [1] proposed the level-set approach to temporal interpolation and applied it to gait recognition for low frame-rate videos. The temporal interpolation works well for videos with a middle range of frame-rates (e.g., more than 6 fps), it fails, however, for quite low frame-rate videos (e.g., 1 or 2 fps).

Another possible approach is temporal super resolution. Shechtman et al. [22] proposed a space-time super resolution by combining information from multiple low-resolution video sequences of the same dynamic scene. The multiple camera setting is, however, not always available in actual scenes.

Furthermore, methods of temporal super resolution from a single video were also proposed. Shimano et al. [23] proposed an example-based temporal super resolution using self similarity and Makihara et al. [9] proposed a reconstruction-based method using multiple-period observation and phase registration within a single video.

The initial contribution of this paper is to improve gait recognition performance for low frame-rate videos by introducing periodic temporal super resolution [9]. Once an input to a low frame-rate video has been subjected to temporal super resolution, existing gait features [8][5][10][21] are successfully applied.

On the other hand, it was reported that periodic temporal super resolution suffers from errors of initial phase estimation using linear phase evolution prior, which typically oc-

978-1-4577-1359-0/11/$26.00 ©2011 IEEE

curred due to the *wagon wheel effect* in quite low frame-rate cases (e.g., 1 or 2 fps) [9]. Therefore, the second contribution of this paper is the improvement of initial phase estimation using an exemplar high frame-rate gait image sequence. Although the subject of the exemplar gait image sequence is different from those of a probe and a gallery, the exemplar sequence still makes the initial phase estimation much more robust than only linear phase evolution prior.

Moreover, if sufficient images are not available as inputs and hence the observed gait stances are still sparse after temporal super resolution, the reconstructed periodic image sequence often contains alpha blending-like *ghosts* because the method [9] interpolates it in image intensity based eigen space. Therefore, the third contribution of this paper is to improve the quality of the reconstructed periodic image sequence by incorporating morphing techniques to fill the gait stances between a pair of adjacent images sorted by estimated phases. Note that the proposed method fills images between a pair of adjacent frames of relatively similar gait stances after a sorting phase, while the existing method [1] fills images between a pair of adjacent frames of the largely different gait stances in low frame-rate videos.

The remainder of this paper is organized as follows. Section 2 addresses related work on gait recognition and temporal super resolution, and Section 3 describes the preprocessing for silhouette-based gait recognition. Section 4 addresses the proposed method of periodic temporal super resolution. Section 5 describes experiments using real gait image sequences and Section 6 discusses the limitations of the proposed method. Finally, Section 7 gives the conclusions and identifies future research.

2. Related work

Period-based gait feature: Because a gait is a type of periodic motion, period-based gait features have been widely used for gait recognition. The most well-known one is the averaged silhouette [8] (also called GEI [5]) which is obtained by simply averaging gait silhouette images over one gait period. In addition, gait features based on frequency analysis are also proposed such as amplitude spectra [10][29], Fourier descriptors [15][30], and Gabor features [26]. These period-based gait features are computed from multiple images within one gait period, and hence they deteriorate significantly when the number of images within one gait period reduces to low levels, as in low frame-rate videos.

Frame-based gait feature: A frame-based gait feature such as a raw gait silhouette image is used in conjunction with phase synchronization using phase shift and/or stretching in the linear domain [21][17][14] or using Dynamic Time Warping (DTW) or Hidden Markov Model (HMM) in the non-linear domain [4][24]. The frame-based gait features themselves do not deteriorate even in the case of low frame-rate videos. They suffer, however, from a phase synchronization step in the case where both a gallery and a probe are low frame-rate videos because corresponding phases (stances) between the gallery and a probe are often missing due to the sparsity of the frame-based gait features per gait period.

Figure 1. Normalized gait silhouette sequence

Temporal super resolution: Approaches to temporal super resolution fall into two groups: example-based approaches and reconstruction-based approaches. The example-based methods typically try to generate a high frame-rate video from a single low frame-rate video based on a number of exemplar pairs of spatio-temporal patches in both the high frame-rate and low frame-rate videos [23]. The reconstruction-based method typically tries to generate a high frame-rate video from multiple low frame-rate videos based on spatio-temporal registration [22]. On the other hand, in the case of a periodic or quasi-periodic image sequence, multiple periods of image sub-sequences with sub-frame-order differences are available for reconstruction-based periodic temporal super resolution from a single low frame-rate video [9]. In such periodic temporal super resolution, stable phase estimation for each frame is key to the generation of accurate high frame-rate periodic image sequences.

3. Preprocessing

In this section, a preprocessing part of our silhouette-based gait recognition is briefly addressed. First, given a gait image sequence, a silhouette for each frame is extracted as a foreground region by background subtraction-based graph-cut segmentation [12]. Next, a normalized gait silhouette sequence of pre-determined size is generated by image size-normalization and registration based on the height and the center of gravity of each silhouette as shown in Fig. 1.

Principal Component Analysis (PCA) is then applied to the normalized gait silhouette sequence and each normalized gait silhouette is projected into PCA space for dimension reduction. The dimension of the PCA space is decided so that information loss is less than 1%, in other words, the cumulative contribution rate is over 99%. Note that the normalized gait silhouette sequence is expressed as a *manifold* parameterized by *phase* in parametric eigen space [16]. 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.

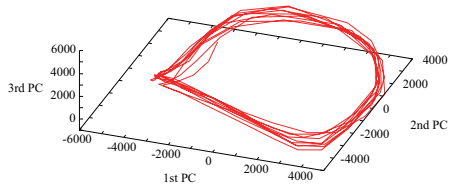


Figure 2. Gait silhouette sequence in PCA space

4. Periodic temporal super resolution

4.1. Definition of quasi-periodic image sequence

In this section, we define an input gait silhouette sequence as a quasi-periodic image sequence. An image drawn from the periodic image sequence at time t is denoted by the vector-form $\boldsymbol{x}(t)$, which satisfies

$$\boldsymbol{x}(t+P) = \boldsymbol{x}(t) \ \ \forall t, \tag{1}$$

where P is a period. Two non-dimensional time parameters, *phase* s and *relative phase* $\tilde{s}$, are then introduced as

$$s = s_P(t) = \frac{t}{P} \tag{2}$$

$$\tilde{s} = s - \lfloor s \rfloor, \tag{3}$$

where $s_P(\cdot)$ is a phase evolution function and $\lfloor \cdot \rfloor$ is a floor function. The periodic image sequence is now represented in the phase domain as

$$\boldsymbol{x}_s(s) = \boldsymbol{x}(s_P^{-1}(s)). \tag{4}$$

Note that the periodic image sequence constructs a manifold with respect to the relative phase $\tilde{s} \in [0,1]$ and which satisfies $\boldsymbol{x}_s(1) = \boldsymbol{x}_s(0)$.

On the other hand, an input gait silhouette sequence is composed of N^{in} discretely observed images $\boldsymbol{X}^{in} = \{\boldsymbol{x}_i^{in}\}(i = 0, \ldots, N^{in}-1)$ which are denoted as

$$\boldsymbol{x}_i^{in} = \boldsymbol{x}_s(s_{P,i}) \tag{5}$$

$$s_{P,i} = s_0 + \tfrac{1}{fP}, \tag{6}$$

where $\boldsymbol{s}_P = \{s_{P,i}\}$ is a phase sequence for the input image sequence, s_0 is the phase for the first frame, and f is the frame-rate. Moreover, the input image sequence is often degraded to a *quasi-periodic image sequence* $\boldsymbol{X}_Q^{in} = \{\boldsymbol{x}_{Q,i}^{in}\}$ due to fluctuations of the sampling rate of the network camera and/or gait motion itself and it is defined as

$$\boldsymbol{x}_{Q,i}^{in} = \boldsymbol{x}_s(s_{Q,i}) \tag{7}$$

$$s_{Q,i} = s_{P,i} + \Delta s_i, \tag{8}$$

where $\boldsymbol{s}_Q = \{s_{Q,i}\}$ is a quasi-periodic phase sequence.

In summary, the problem setting can be stated as a simultaneous estimation problem of a periodic manifold $\boldsymbol{x}_s$ and a phase sequence $\boldsymbol{s}_Q$ from an input quasi-periodic image sequence $\boldsymbol{X}_Q^{in}$.

4.2. Manifold representation

The periodic manifold $\boldsymbol{x}_s$ is represented by a parametric eigenspace method [16]. More specifically, a cubic N-spline function parameterized by the phase in the eigenspace is exploited. Let us consider N^{cp} control points $\{\boldsymbol{y}_j^{cp}\}$ in the M dimensional eigenspace accompanied by corresponding phases $\{s_j^{cp}(= j/N^{cp})\}$, $(j = 0, \ldots, N^{cp}-1)$. Next, a spline parameter vector for a k power-term coefficient at the jth interval $[s_j^{cp}, s_{j+1}^{cp}]$ is denoted as $\boldsymbol{a}_{j,k}^{sp} \in \mathbb{R}^M (k = 0,1,2,3)$. Subsequently, a submatrix A_j^{sp} at the jth interval and a total spline matrix A^{sp} are defined as $A_j^{sp} = [\boldsymbol{a}_{j,0}^{sp}, \ldots, \boldsymbol{a}_{j,3}^{sp}]^T \in \mathbb{R}^{4\times M}$ and $A^{sp} = [A_0^{spT}, \ldots, A_{N^{cp}-1}^{spT}]^T \in \mathbb{R}^{4N^{cp}\times M}$, respectively. An interpolated point $\hat{\boldsymbol{y}}(\tilde{s})$ in the eigenspace for a relative phase $\tilde{s}$ at the jth interval is then expressed as

$$\hat{\boldsymbol{y}}(\tilde{s}) = A^{spT}\boldsymbol{w}(\tilde{s}) \tag{9}$$

$$\boldsymbol{w}(\tilde{s}) = [0, \ldots, 0, 1, w, w^2, w^3, 0, \ldots, 0]^T \tag{10}$$

$$w = \frac{\tilde{s} - s_j^{cp}}{s_{j+1}^{cp} - s_j^{cp}}, \ (s_j^{cp} \le \tilde{s} \le s_{j+1}^{cp}) \tag{11}$$

where $\boldsymbol{w}(\tilde{s})$ is an interpolation coefficient vector whose components from $4j$ to $(4j+3)$ are $[1, w, w^2, w^3]$, and w is the interpolation ratio between the control points.

On the other hand, it is well known that the relationship between a control points matrix $Y^{cp} = [\boldsymbol{y}_0^{cp}, \ldots, \boldsymbol{y}_{N^{cp}-1}^{cp}]^T$ and a spline parameter matrix A^{sp} is derived from the C2-continuous boundary conditions as [9]

$$A^{sp} = DY^{cp}, \tag{12}$$

where $D \in \mathbb{R}^{4N^{cp}\times N^{cp}}$ is a coefficient matrix. Hence, once the control points Y^{cp} have been given, interpolation $\hat{\boldsymbol{y}}(\tilde{s})$ for a relative phase $\tilde{s}$ is obtained using Eqs. (9) and (12). The reconstruction problem of the periodic manifold $\boldsymbol{x}_s(s; Y^{cp})$ can then be replaced by an estimation problem of the control points Y^{cp} as discussed in the following sections.

4.3. Energy minimization framework

Similar to the other temporal super resolution approaches [22][23], we adopt an energy minimization approach. First, suppose that an input quasi-periodic image sequence in the eigenspace is expressed as $\boldsymbol{Y}_Q^{in} = \{\boldsymbol{y}_{Q,i}^{in}\}$ and recall that the accompanying phase sequence $\boldsymbol{s}_Q = \{s_{Q,i}\}$ is unknown. Subsequently, the interpolation coefficient vector for the ith phase $s_{Q,i}$ is defined as $\boldsymbol{w}(s_{Q,i})$, in the same way as Eq. (10), and then approximation by the periodic manifold with the ith phase $s_{Q,i}$ is

$$\hat{\boldsymbol{y}}(Y^{cp}, s_{Q,i}) = A^{spT}\boldsymbol{w}(s_{Q,i}) = Y^{cpT}D^T\boldsymbol{w}(s_{Q,i}). \tag{13}$$

The energy function is then constructed by considering the following three aspects: (1) data fitness between the interpolation $\hat{\boldsymbol{y}}(Y^{cp}, s_{Q,i})$ and the input $\boldsymbol{y}_{Q,i}^{in}$, (2) smoothness

of the periodic manifold $\boldsymbol{y}_s(s;Y^{cp})$ in eigenspace, and (3) smoothness of the phase evolution $\boldsymbol{s}_Q$ based on the linear phase evolution prior. The actual form of the function is

$$\begin{aligned} E(Y^{cp},\boldsymbol{s}_Q) &= \frac{1}{N^{in}}\sum_{i=0}^{N^{in}-1}||Y^{cpT}D^T\boldsymbol{w}(s_{Q,i})-\boldsymbol{y}^{in}_{Q,i}||^2 \\ &+ \lambda_m\frac{1}{N^{cp}}||BY^{cp}||^2 \\ &+ \lambda_s\frac{1}{N^{in}}\sum_{i=1}^{N^{in}-1}\left(s_{Q,i}-s_{Q,i-1}-\frac{1}{P'}\right)^2, \quad (14) \end{aligned}$$

where the first, second, and third terms are the data term, the smoothness terms for the periodic manifold and the phase evolution. B is a matrix for the manifold curvature, and $P'(=fP)$ is a global period in the frame domain derived from the linear phase evolution prior.

We can see that the objective function $E(Y^{cp},\boldsymbol{s}_Q)$ is a quadratic form with respect to the manifold control points Y^{cp} and, therefore, a linear solution of the manifold control points Y^{cp} is provided under the fixed phase $\boldsymbol{s}_Q$. On the other hand, the phase $\boldsymbol{s}_Q$ is a complex form, since the spline curves are switched piecewise based on phase $\boldsymbol{s}_Q$ and the interpolation ratio w appears as a sixth-order polynomial in the data term. To solve this highly nonlinear optimization problem, iterative solutions are employed in the following sections.

4.4. Initial phase estimation

In the initial phase estimation step, instead of using linear phase evolution prior [9], continuous Dynamic Programming (DP) [18] is applied directly between an input low frame-rate image sequence Y^{in}_Q and an exemplar high frame-rate image sequence which is expressed as a set of manifold control points $Y^{ex}=\{\boldsymbol{y}^{ex}_j\}(j=0,\ldots,N^{cp}-1)$. Note that the only data term in the energy function (14) is considered in this step because the manifold smoothness term is independent of phase and the phase smoothness term based on linear phase evolution prior is negligible.

First, the phase space is quantized by the same interval as that of the manifold control points $(1/N^{cp})$. This means that the jth quantized phase coincides with the jth control point $\boldsymbol{y}^{ex}_j$. Next, a set of phases which can be transited from a previous step to jth quantized phase is defined as R_j by considering the frame-rate of the input image sequence and gait period range. A cumulative cost and the optimal transition path from a previous step at the ith input frame and jth exemplar control point are then denoted as $c(i,j)$ and $p(i,j)$, respectively, and the optimal phase is provided by the following DP steps.

1. Initialize cost matrix

$$c(0,j)=||\boldsymbol{y}^{ex}_j-\boldsymbol{y}^{in}_{Q,0}||^2 \quad (15)$$

2. Update cumulative cost and transition path

$$\begin{aligned} p(i,j) &= \arg\min_{k\in R^j}\{c(i-1,k)\} \quad (16) \\ c(i,j) &= c(i-1,p(i,j))+||\boldsymbol{y}^{ex}_j-\boldsymbol{y}^{in}_{Q,i}||^2 \quad (17) \end{aligned}$$

3. Optimize the terminal phase

$$p^*(N^{in}-1)=\arg\min_j c(N^{in}-1,j) \quad (18)$$

4. Back track

$$p^*(i-1)=p(i,p^*(i)) \quad (19)$$

After the initial phase is estimated as $\boldsymbol{s}^{init}_Q$ as a result of the above DP procedure, the global period P' is computed from averaged phase evolution speed.

4.5. Interpolation between sorted phases

Once an initial phase sequence $\boldsymbol{s}^{init}_Q=\{s^{init}_{Q,i}\}$ is estimated, the input image sequence $\boldsymbol{X}^{in}_Q=\{\boldsymbol{x}^{in}_{Q,i}\}$ is sorted based on the initial phase sequence $\boldsymbol{s}^{init}_Q$ as $\boldsymbol{X}^{sort}_Q=\{\boldsymbol{x}^{sort}_{Q,i}\}$. The intermediate images between each pair of adjacent sorted images $\boldsymbol{x}^{sort}_{Q,i}$ and $\boldsymbol{x}^{sort}_{Q,i+1}$ are then interpolated by the phase quantization interval $(1/N^{cp})$ using the morphing technique [3]. An advantage of morphing-based interpolation is to avoid *ghosts* occurring in the case of interpolation in eigen space. In addition, note that the proposed method interpolates images between a pair of adjacent sorted images, while the existing morphing-based gait recognition approach [1] interpolates images between a pair of adjacent original images whose the elapsed phase is much larger, or in short, whose gait stances are quite different from each other. Finally, the interpolated image sequence $\boldsymbol{X}^{int}_Q=\{\boldsymbol{x}^{int}_{Q,i}\}(i=0,\cdots,N^{int}-1)$ associated with the interpolated sorted phase sequence $\boldsymbol{s}^0_Q=\{s^0_{Q,i}\}$ is fed into the energy minimization framework, and the initial manifold control points $Y^{cp,0}$ under the fixed phase sequence $\boldsymbol{s}^0_Q$ are then easily calculated using a linear solution as described in Section 4.3.

4.6. Phase estimation by quadratic approximation

After the initial solution is obtained, the phase sequence and manifold control points are alternately updated based on the previous solution.

The first step is phase estimation based on the quadratic approximation of the data term with respect to phase. The search domain of the phase at the rth iteration is set based on the previous solution as $R^r_i=\{s|s^{r-1}_{Q,i}-s^{tol}\le s\le s^{r-1}_{Q,i}+s^{tol}\}$. The minimum data term for the ith phase $s_{Q,i}$ is found by the Newton method within the search domain. After phase minimizing, the data term has been obtained as

$s_{Q,i}^{data,r^*}$, and the ith data term in Eq. (14) is approximated by a Taylor expansion up to the second-order term as

$$\hat{E}_i^{data,r}(s_{Q,i}) = E_i^{data,r}(s_{Q,i}^{data,r^*}) + \left.\frac{dE_i^{data,r}}{ds_{Q,i}}\right|_{s_{Q,i}=s_{Q,i}^{data,r^*}} (s_{Q,i} - s_{Q,i}^{data,r^*}) + \frac{1}{2}\left.\frac{d^2E_i^{data,r}}{ds_{Q,i}^2}\right|_{s_{Q,i}=s_{Q,i}^{data,r^*}} (s_{Q,i} - s_{Q,i}^{data,r^*})^2. \quad (20)$$

Now, the total energy function is a quadratic form with respect to the phase $\boldsymbol{s}_Q$ and thus the optimal phase $\boldsymbol{s}_Q$ is given as

$$\boldsymbol{s}_Q^{r\,*} = \arg\min_{\boldsymbol{s}_Q}\left\{\frac{1}{N^{int}}\sum_{i=0}^{N^{int}-1}\hat{E}_i^{data,r}(s_{Q,i}) + \lambda_s\frac{1}{N^{int}}\sum_{i=1}^{N^{int}-1}\left(s_{Q,i+1}-s_{Q,i}-\frac{1}{P'}\right)^2\right\} \quad (21)$$

$$\text{s.t.} \quad s_{Q,i}^{r-1} - s^{tol} \le s_{Q,i} \le s_{Q,i}^{r-1} + s^{tol} \quad (22)$$

$$s_{Q,i+1} \ge s_{Q,i}, \quad (23)$$

where Eqs. (22) and (23) are the lower and upper limit constraints and the monotonically increasing constraints, respectively. As a result, the problem is formulated as a convex quadratic programming problem and is solved by the active set method.

The next step is the update of the manifold control points. The manifold control points $Y^{cp,r}$ at the rth iteration are computed based on the phase sequence $\boldsymbol{s}_Q^{r\,*}$ at the rth iteration via linear solution.

These procedures are iterated until convergence by gradually relaxing the manifold smoothness constraint so that the manifold fits the data.

5. Experiments

5.1. Dataset

We evaluated the proposed method with real data from the OU-ISIR Gait Database [19]. A total of 200 real gait sequences from 100 subjects were used in the experiments. Each subject was requested to walk on a speed-controlled treadmill and two gait sequences were captured: one for the probe, and the other for the gallery. The normalized gait silhouette size, frame rate, and recording time for each sequence were 88×128 pixels, 60 fps, and 6 sec. Low frame-rate gait sequences were constructed by selecting the gait silhouettes at a specified interval, that is, by down sampling.

Figure 3. Results of temporal super resolution. The first row is an input low frame-rate image sequence (2 fps), and rows two to four are temporal interpolation or super resolution results by *Morph* [1], *TSR (LPEP)* [9], and *TSR (Ex + Morph)*, respectively (every 2 steps or control point). Note that the red and green box images are key shapes for level-set morphing. The bottom row is an original frame-rate (60 fps) image sequence as a reference (every 2 steps).

5.2. Result of periodic temporal super resolution

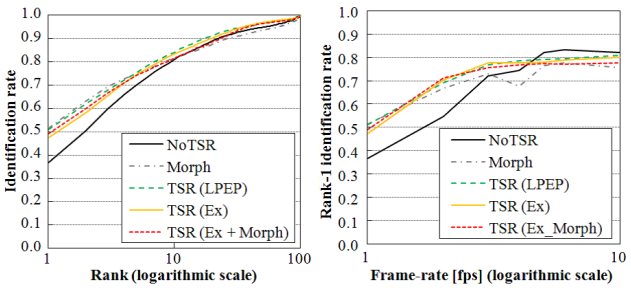
In this section, we showed several results of the periodic temporal super resolution with linear phase evolution prior (*TSR (LPEP)*), with the exemplar image sequence (*TSR (Ex)*), and with the exemplar image sequence and morphing (*TSR (Ex + Morph)*), compared with temporal interpolation using the level-set morphing approach (*Morph*) [1].

Figure 3 shows results for a 2-fps image sequence. We can see that *Morph* (the second row of Fig. 3) cannot successfully recover a high frame-rate image sequence because the two key gait stances are too far apart (a single support phase and the next single support phase) due to the quite low frame-rate for interpolating intermediate gait stances (double support phases). Although *TSR (LPEP)* (the third row of Fig. 3) reconstructs a smooth gait silhouette sequence, it still suffers from backward play, that is, the wagon wheel effect, and also ghosts. On the other hand, *TSR (Ex + Morph)*, (the fourth row of Fig. 3), successfully reconstructs a natural gait silhouette sequence with neither a wagon wheel effect nor ghosting, and hence looks quite similar to the original normal frame-rate image sequence (the bottom row of Fig. 3).

5.3. Result of gait recognition

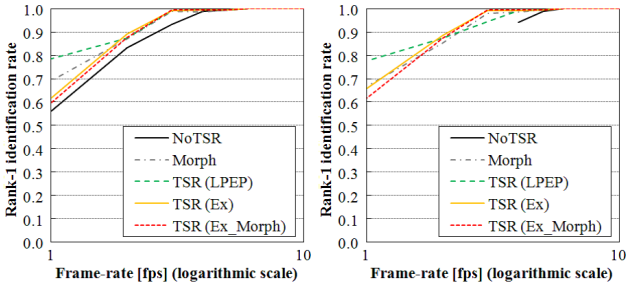
In this section, we evaluate the proposed methods *TSR (LPEP)*, *TSR (Ex)*, and *TSR (Ex + Morph)* in terms of gait recognition performance compared with the approach without temporal super resolution (*NoTSR*) and with temporal interpolation *Morph* [1].

Frame-based gait feature: In these experiments, the method of period-based phase synchronization [14] is employed for matching low frame-rate probe image sequence and super-resoluted or temporally interpolated high frame-rate gallery image sequence, where the periodic gallery image sequences are directly matched to the probe sequence



(a) CMC curve for 1-fps image sequences (b) Rank-1 identification rate for each frame-rate

Figure 4. CMC curve and rank-1 identification rate for frame-based gait feature



(a) Averaged silhouette (b) Frequency-domain feature

Figure 5. Rank-1 identification rate for each period-based gait feature

so as to synchronize the phase by shifting the frame. The performance measure used is Cumulative Matching Characteristics (CMC) curve [20] and rank-1 identification rate.

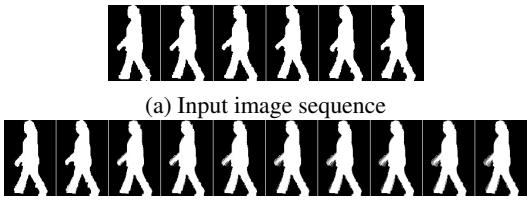
The result is shown in Fig. 4. While all the approaches are degraded as the frame-rate decreases, either of the proposed temporal super resolution methods still achieves the best performance for less than 5 fps.

Period-based gait feature: The gait features used are the most popular averaged silhouette [8][5] and the frequency-domain feature [10]. The performance measure used is the rank-1 identification rate. As for *NoTSR*, the frequency-domain feature cannot be extracted at less than 4 fps because of the sampling theorem.

The results for each gait feature are shown in Fig. 5. While all the approaches achieved a 100 % identification rate at more than 5 fps, either of the periodic temporal super resolution approaches achieved a better performance at less than 5 fps

6. Limitations

While the proposed method solved the wagon wheel effect which the linear phase evolution prior-based method [9] suffers from, the troublesome *stroboscopic effect (temporal aliasing)* still remains. The *stroboscopic effect* typically occurs when the sampling interval coincides with the period of a moving object. In such cases, the observed image sequence appears to be almost standing still because the observed images are always the same even though the object is actually moving periodically as shown in Fig. 6. Any reconstruction-based approach from a single video cannot solve the stroboscopic effect on its own. Therefore a combined framework of the reconstruction-based and example-based temporal super resolution techniques needs to be studied in the future.



(a) Input image sequence

(b) Temporal super resolution result (every 5 control points)

Figure 6. Stroboscopic effect

7. Conclusions

This paper described gait recognition for low frame-rate videos. We first improved gait recognition for low frame-rate videos by incorporating periodic temporal super resolution from a single image sequence. Second, we solved the troublesome wagon wheel problem which the existing periodic temporal super resolution approach suffered from by introducing stable initial phase estimation based on an exemplar high frame-rate video. Third, images between a pair of adjacent images sorted by estimated phases are interpolated by a morphing technique to avoid alpha blending-like *ghosting*. Experimental results using a real gait image sequence showed the effectiveness of the proposed method in terms of not only recognition performance but also the visual quality of temporal super resolution.

In the future we propose to address the problem of the stroboscopic effect by combining the example-based and reconstruction-based approaches.

Acknowledgement

This work was partially supported by Grant-in-Aid for Scientific Research (S) 21220003 and ”R&D Program for Implementation of Anti-Crime and Anti-Terrorism Technologies for a Safe and Secure Society”, Strategic Funds for the Promotion of Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

References

[1] M. S. Al-Huseiny, S. Mahmoodi, and M. S. Nixon. Gait learning-based regenerative model: A level set approach. In *The 20th Int. Conf. on Pattern Recognition*, pages 2644–2647, Istanbul, Turkey, Aug. 2010. 1, 2, 4, 5

[2] C. BenAbdelkader. *Gait as a Biometric For Person Identification in Video*. PhD thesis, Maryland Univ., 2002. 1

[3] D. Cohen-Or, D. Levin, and A. Solomivici. Three-dimensional distance field metamorphosis. *ACM Trans. Graphics*, 17(2):116–141, 1998. 4

[4] N. Cuntoor, A. Kale, and R. Chellappa. Combining multiple evidences for gait recognition. In *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, volume 3, pages 33–36, 2003. 2

[5] J. Han and B. Bhanu. Individual recognition using gait energy image. *Trans. on Pattern Analysis and Machine Intelligence*, 28(2):316– 322, 2006. 1, 2, 6

[6] M. A. Hossain, Y. Makihara, J. Wang, and Y. Yagi. Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control. *Pattern Recognition*, 43(6):2281–2291, Jun. 2010. 1

[7] A. Kale, A. Roy-Chowdhury, and R. Chellappa. Towards a view invariant gait recognition algorithm. In *Proc. of IEEE Conf. on Advanced Video and Signal Based Surveillance*, pages 143–150, 2003. 1

[8] Z. Liu and S. Sarkar. Simplest representation yet for gait recognition: Averaged silhouette. In *Proc. of the 17th Int. Conf. on Pattern Recognition*, volume 1, pages 211–214, Aug. 2004. 1, 2, 6

[9] Y. Makihara, A. Mori, and Y. Yagi. Temporal super resolution from a single quasi-periodic image sequence based on phase registration. In *Proc. of the 10th Asian Conf. on Computer Vision*, pages 107–120, Queenstown, New Zealand, Nov. 2010. 1, 2, 3, 4, 5, 6

[10] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi. Gait recognition using a view transformation model in the frequency domain. In *Proc. of the 9th European Conf. on Computer Vision*, pages 151–163, Graz, Austria, May 2006. 1, 2, 6

[11] Y. Makihara, A. Tsuji, and Y. Yagi. Silhouette transformation based on walking speed for gait identification. In *Proc. of the 23rd IEEE Conf. on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, Jun 2010. 1

[12] Y. Makihara and Y. Yagi. Silhouette extraction based on iterative spatio-temporal local color transformation and graph-cut segmentation. In *Proc. of the 19th Int. Conf. on Pattern Recognition*, Tampa, Florida USA, Dec. 2008. 2

[13] D. Matovski, M. Nixon, S. Mahmoodi, and J. Carter. The effect of time on the performance of gait biometrics. In *Proc. of the 4th IEEE Int. Conf. on Biometrics: Theory Applications and Systems*, pages 1–6, Washington D.C., USA, Sep. 2010. 1

[14] A. Mori, Y. Makihara, and Y. Yagi. Gait recognition using period-based phase synchronization for low frame-rate videos. In *Proc. of the 20th Int. Conf. on Pattern Recognition*, pages 2194–2197, Istanbul, Turkey, Aug. 2010. 1, 2, 5

[15] S. Mowbray and M. Nixon. Automatic gait recognition via fourier descriptors of deformable objects. In *Proc. of IEEE Conf. on Advanced Video and Signal Based Surveillance*, pages 566–573, 2003. 2

[16] H. Murase and S. K. Nayar. Parametric eigenspace representation for visual learning and recognition. In *Proc. of SPIE, 2031*, 1993. 2, 3

[17] H. Murase and R. Sakai. Moving object recognition in eigenspace representation: Gait analysis and lip reading. *Pattern Recognition Letters*, 17:155–162, 1996. 2

[18] R. Oka. Spotting method for classification of real world data. *Computer Journal*, 41(8):559–565, 1998. 4

[19] Ou-isir gait database. http://www.am.sanken.osaka-u.ac.jp/GaitDB/index.html. 5

[20] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The feret evaluation methodology for face-recognition algorithms. *Trans. of Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000. 6

[21] 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*, 27(2):162–177, 2005. 1, 2

[22] E. Shechtman, Y. Caspi, and M. Irani. Space-time super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):531–545, Apr. 2005. 1, 2, 3

[23] M. Shimano, T. Okabe, I. Sato, and Y. Sato. Video temporal super-resolution based on self-similarity. In *Proc. of the 10th Asian Conf. on Computer Vision*, volume 1, pages 93–106, Queeenstown, New zealand, Nov. 2010. 1, 2, 3

[24] A. Sunderesan, A. Chowdhury, and R. Chellappa. A hidden markov model based framework for recognition of humans from gait sequences. In *Proc. IEEE Int'l Conf. on Image Processing 2003*, volume 2, pages 93–96, 2003. 2

[25] R. Tanawongsuwan and A. Bobick. A study of human gaits across different speeds. Technical report, Georgia Tech, 2003. 1

[26] D. Tao, X. Li, X. Wu, and S. Maybank. Human carrying status in visual surveillance. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, volume 2, pages 1670–1677, New York, USA, Jun. 2006. 2

[27] How biometrics could change security. http://news.bbc.co.uk/2/hi/programmes/click_online/7702065.stm. 1

[28] Y. Wang, S. Yu, Y. Wang, and T. Tan. Gait recognition based on fusion of multi-view gait sequences. In *Proc. of the IAPR Int. Conf. on Biometrics 2006*, pages 605–611, Jan. 2006. 1

[29] G. Zhao, R. Chen, G. Chen, and H. Li. Amplitude spectrum-based gait recognition. In *Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pages 23–30, 2004. 2

[30] G. Zhao, R. Chen, G. Chen, and H. Li. Recognition of human periodic movements from unstructured information using a motion-based frequency domain approach. *Image and Vision Computing*, 24:795–809, 2006. 2