Temporal Super Resolution
from a Single Quasi-Periodic Image Sequence
Based on Phase Registration

Yasushi Makihara, Atsushi Mori, and Yasushi Yagi

Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan

Abstract. This paper describes a method for temporal super resolution from a single quasi-periodic image sequence. A so-called reconstruction-based method is applied to construct a one period image sequence with high frame-rate based on phase registration data in sub-frame order among multiple periods of the image sequence. First, the periodic image sequence to be reconstructed is expressed as a manifold in the parametric eigenspace of the phase. Given an input image sequence, phase registration and manifold reconstruction are alternately executed iteratively within an energy minimization framework that considers data fitness and the smoothness of both the manifold and the phase evolution. The energy minimization problem is solved through three-step coarse-to-fine procedures to avoid local minima. The experiments using both simulated and real data confirm the realization of temporal super resolution from a single image sequence.

1 Introduction

Image super resolution [1] is one of the fundamental image processing techniques for increasing spatial resolution. Methods for super resolution fall mainly into two categories: (1) reconstruction-based methods (RBM) using multiple images registered at sub-pixelorder [2] [3] [4], and (2) example-based methods using correspondence between low and high resolution image patches from training sets [5] [6] [7]. Furthermore, Glasner et al. [8] proposed a sophisticated framework, combining the two approaches by using the patch recurrence within and across scales of a single image. Most of the existing methods, however, are applicable to static scenes only.

Several methods deal with dynamic scenes in the context of near-real-time super resolution [9], taking advantage of hot-air optical turbulence [10] and image super resolution in the presence of motion debluring [11] [12]. Nevertheless, these methods still focus on spatial super resolution, and fail to address temporal super resolution.

Contrary to the above methods, Shechtman et al. [13] proposed a space-time super resolution by combining information from multiple low-resolution video sequences of the same dynamic scene. This method falls into the so-called RBMs and hence, requires multiple video cameras to obtain multiple image sequences.

Inspired by the effective use of patch recurrence [8], it is noticeable that temporal recurrence in a single sequence can be used for temporal super resolution. In particular, given a periodic image sequence, multiple periods of subsequences with sub-frame order *phase displacement* can be used for the temporal super resolution. If a sequence can be segmented into period-based subsequences composed of more than a few frames, a spatio-temporal sequence registration [14] method is applicable. On the contrary, in cases where each subsequence contains only a few frames or a single frame, period segmentation and the spatio-temporal sequence registration may be error prone, since the method discards useful cues of inter-period phase continuity. Moreover, if the period fluctuates due to a non-uniform sampling rate or to target motion fluctuation appearing in human periodic actions (e.g., gait), phase registration becomes much more difficult.

Therefore, we propose a method to reconstruct a one period image sequence at a high frame-rate from a single low frame-rate quasi-periodic image sequence. Instead of segmenting the whole sequence into multiple subsequences, we solve the *phase assignment* problem for the whole sequence of frames to take advantage of the inter-period phase continuity. Then, a one period image sequence is expressed as a *manifold in phase-parametric eigenspace* and the manifold is reconstructed based on the registered phase information. These two processes are repeated in an energy minimization framework by taking into consideration the following three aspects: (1) data fitness, (2) smoothness of the manifold, and (3) smoothness of the phase evolution.

2   Related work

Temporal interpolation: In addition to spatial interpolation, temporal interpolation has been developed for video editing (e.g., retiming [15]) and video compression based on motion compensation methods [16]. In particular, morphingapproaches [17] [18] based on optical flow [19] are regarded as some of the most promising methods for temporal interpolation, with these ideas having been extended to view and time interpolation [20] and spatio-temporal interpolation from two sequences, one with high resolution and low frame rate and the other with low resolution and high frame rate [21]. Temporal interpolation does not, however, work well in cases where motion between two frames is relatively large, in other words, where the sampling interval is relatively long in fairly low frame-rate video.
Geometric fitting: Our manifold reconstruction step is similar to the fitting of geometric primitives, such as lines, ellipses, or conics [22], [23], in terms of parameter estimation from a series of sampling points. In geometric fitting, global parameters are used for the primitive representation (e.g., 5 parameters for an ellipse) and sampling points are treated independently. On the other hand, piecewise local parameters are used for manifold representation and sampling points are dependent on each other due to the phase evolution smoothness as described in the following sections.

3   Problem setting

3.1   Assumptions

Several assumptions are used in this paper to allow us to focus only on temporal super resolution based on a phase registration framework. The first assumption is that an image sequence is spatially registered in advance. This assumption is basically true in several scenes, such as those with periodic sign language such as a waving hand, periodic actions at the same position such as jumping jacks or walking on a treadmill, or a rotating object such as a fan. Even in other scenes with periodic actions such as walking, skipping, and running on the ground, spatial registration is possible once the target object has been accurately tracked to some extent.

A further assumption is that motion blurs are negligible. Although this assumption may not be true in scenes with fast moving objects, there are still several possible situations in which low frame-rate image sequences with less motion blur are stored. For example, consider the situation where a CCTV camera captures a person walking on the street. In this case, the image sequence is often stored at a low frame-rate due to limited communication band width or storage size, and besides motion blur is not significant since walking motion is relatively slow compared with normal shutter speeds. Therefore, gait recognition in low frame-rate video is one of the typical applications of the proposed method.

3.2   Quasi-periodic image sequence

Next, we define 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. Then, two non-dimensional time parameters, *phase s* and *relative phase $\tilde{s}$*, are introduced as

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

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

where s_P is a phase evolution function and $\lfloor \rfloor$ is a floor function. Now, the periodic image sequence is 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 image sequence is composed of N^{in} discretely observed images $\boldsymbol{X}^{in} = \{\boldsymbol{x}_i^{in}\}(i = 0, \ldots, N^{in} - 1)$. In cases where the scene is completely periodic and where the frame-rate f is completely constant during

the input image sequence, the input image sequence $\boldsymbol{X}^{in}$, time sequence $\boldsymbol{t} = \{t_i\}$, and phase sequence $\boldsymbol{s}_P = \{s_{P,i}\}$ are denoted as

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

$$t_i = t_0 + \frac{i}{f} \tag{6}$$

$$s_{P,i} = s_P(t_i) = s_0 + \frac{1}{fP}. \tag{7}$$

This assumption is, however, often violated due to fluctuations in the frame-rate or in the timing of the periodic motion. Hence, the input image sequence is degraded to a *quasi-periodic image sequence* $\boldsymbol{X}_Q^{in} = \{\boldsymbol{x}_{Q,i}^{in}\}$ defined as

$$\boldsymbol{x}_{Q,i}^{in} = \boldsymbol{x}_s(s_{Q,i}) \tag{8}$$

$$s_{Q,i} = s_i + \Delta s_i, \tag{9}$$

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 $\boldsymbol{X}_Q^{in}$. Analogous to the spatial super resolution case, note that the phase sequence estimation process and periodic manifold estimation correspond to an image registration process and super-resolution process, respectively. In addition, the problem setting falls into the so-called reconstruction-based methods using multiple observations, although a single sequence is used in this problem setting. This is because *multiple periods* in the single sequence serve as multiple observations.

3.3   Manifold representation

The periodic manifold $\boldsymbol{x}_s$ is represented by a parametric eigenspace method [24]. This is motivated by the property of the parametric eigenspace method that a phase for an input image can be estimated by projection to the manifold, which plays quite a significant role in the phase registration stage. More specifically, we use a cubic N-spline function parameterized by the phase in the eigenspace. 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 $\boldsymbol{a}_{j,k}^{sp} \in \mathbb{R}^M$ and 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 $A^{sp} = [A_0^{sp\,T}, \ldots, A_{N^{cp}-1}^{sp}{}^T]^T \in \mathbb{R}^{4N^{cp} \times M}$, respectively. Then, an interpolated point $\hat{\boldsymbol{y}}(\tilde{s})$ in the eigenspace for a relative phase $\tilde{s}$ at the jth interval is expressed as

$$\hat{\boldsymbol{y}}(\tilde{s}) = A^{sp\,T} \boldsymbol{w}(\tilde{s}) \tag{10}$$

$$\boldsymbol{w}(\tilde{s}) = [0, \ldots, 0, 1, w, w^2, w^3, 0, \ldots, 0]^T \tag{11}$$

$$w = \frac{\tilde{s} - s_j^{cp}}{s_{j+1}^{cp} - s_j^{cp}}, \ (s_j^{cp} \leq \tilde{s} \leq s_{j+1}^{cp}) \tag{12}$$

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, the relation 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

$$CA^{sp} = DY^{cp} \tag{13}$$

$$C = \begin{bmatrix} C_1 & C_2 & O & \cdots \\ \ddots & \ddots & \ddots & \ddots \\ C_2 & O & \cdots & C_1 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 6 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \end{bmatrix} \tag{14}$$

$$D = \begin{bmatrix} D_1 & D_2 & O & \cdots \\ \ddots & \ddots & \ddots & \ddots \\ D_2 & O & \cdots & D_1 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{15}$$

Hence, the spline parameter matrix A^{sp} is linearly solved as $A^{sp} = C^{-1}DY^{cp}$ given the control points Y^{cp}. This also indicates that interpolation $\hat{\boldsymbol{y}}(\tilde{s})$ for a relative phase $\tilde{s}$ is obtained by Eqs. (10) and (15) once the control points Y^{cp} have been given. Then, the reconstruction problem of the periodic manifold $\boldsymbol{x}_s(s; Y^{cp})$ can be replaced by an estimation problem of the control points Y^{cp} as discussed in the following sections.

4 Energy minimization framework

4.1 Energy function

Similar to the spatial super resolution case, we adopt an energy minimization approach. First, suppose that an input quasi-periodic image sequence in the eigenspace is expressed as $Y_Q^{in} = \{\boldsymbol{y}_{Q,i}^{in}\}$ 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. (11), and then projection onto the periodic manifold by the ith phase $s_{Q,i}$ is

$$\hat{\boldsymbol{Y}}(Y^{cp}, s_{Q,i}) = A^{spT}\boldsymbol{w}(s_{Q,i}) = Y^{cpT}(C^{-1}D)^T\boldsymbol{w}(s_{Q,i}). \tag{16}$$

The energy function is 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 the eigenspace, and (3)

smoothness of the phase evolution $\boldsymbol{s}_Q$. The actual form of the function is

$$
\begin{aligned}
E(Y^{cp}, \boldsymbol{s}_Q) = {} & \frac{1}{N^{in}} \sum_{i=0}^{N^{in}-1} ||Y^{cpT}(C^{-1}D)^T \boldsymbol{w}(s_{Q,i}) - \boldsymbol{y}_{Q,i}^{in}||^2 \\
& + \lambda_m \frac{1}{N^{cp}} \int_0^1 \left|\left| \frac{d^2 \boldsymbol{y}_s(s; Y^{cp})}{ds^2} \right|\right|^2 ds \\
& + \lambda_s \frac{1}{N^{in}} \sum_{i=1}^{N^{in}-1} \left(s_{Q,i+1} - s_{Q,i} - \frac{1}{P'} \right)^2 ,
\end{aligned}
\tag{17}
$$

where the first, second, and third terms are the data term, and the smoothness terms for the periodic manifold and the phase evolution, respectively, and P' is a global period in the frame domain, which is defined as the product of the frame-rate f and period P in the time domain, that is, $P' = fP$. Since the integration in the second term is calculated in advance with respect to the domain of the interpolation ratio w for each interval, it is rearranged as a quadratic form of Y^{cp} as

$$
\int_0^1 \left|\left| \frac{d^2 \boldsymbol{y}_s(s; Y^{cp})}{ds^2} \right|\right|^2 ds = Y^{cpT}(C^{-1}D)^T B(C^{-1}D)Y^{cp}
\tag{18}
$$

$$
B = \begin{bmatrix} B_{sub} & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & B_{sub} \end{bmatrix}, \quad B_{sub} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 6 \\ 0 & 0 & 6 & 12 \end{bmatrix}
\tag{19}
$$

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, that 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, three-step coarse-to-fine solutions are provided in the following sections.

4.2   Solution by linear approximation

In the first step, the phase sequence $\boldsymbol{s}_Q$ is limited to a completely periodic domain, that is, linear phase evolution $\boldsymbol{s}_P$. Because the initial phase s_0 is not significant in this problem setting, it is set to zero without loss of generality. Hence, the objective function in the first step is rewritten as

$$
E_{init}(P) = E(Y_P^{cp}, \boldsymbol{s}_P)
\tag{20}
$$

$$
Y_P^{cp} = \arg \min_{Y^{cp}} E(Y^{cp}, \boldsymbol{s}_P)
\tag{21}
$$

$$
s_{P,i} = \frac{i}{fP}.
\tag{22}
$$

Note that manifold control points Y_P^{cp} for each hypothesis of period P are linearly solved as described in 4.1. On the other hand, the objective function $E_{init}(P)$ has many local minima, so a gradient descent method from multiple initial solutions is applied. Optimal solutions of the period, phase and manifold control points in the periodic-domain are denoted as P^*, $\boldsymbol{s}_{P^*}$, and $Y_{P^*}^{cp}$, respectively.

4.3   Dynamic programming solution

In the second step, in order to extend from a periodic domain to a quasi-periodic domain, continuous Dynamic Programming (DP) [25] is applied within a so-called corridor, that is, the neighboring search area from the periodic-domain phase solution $\boldsymbol{s}_{P^*}$ as $R_i^{cdr} = \{s | s_{P^*,i} - s^{cdr} \leq s \leq s_{P^*,i} + s^{cdr}\}$ under fixed manifold control points $Y_{P^*}^{cp}$. 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-phase interpolation coincides with the jth control point $\boldsymbol{y}_{P^*,j}^{cp}$ and that the phase difference between the jth and kth phases is equal to $(j-k)/N^{cp}$. Next, a cumulative cost and the optimal transition from a previous step at the ith input frame and jth phase are 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

$$c(0,j) = ||\boldsymbol{y}_{P^*,j}^{cp} - \boldsymbol{y}_0^{in}||^2, \quad \forall j \in R_0^{cdr} \tag{23}$$

2. Update cumulative cost and transition path

$$p(i,j) = \arg\min_{k \in R_{i-1}^{cdr}} \{c(i{-}1,k) + \lambda g_{P^*}(j,k)\} \tag{24}$$

$$c(i,j) = c(i{-}1,p(i,j)) + \lambda g_{P^*}(j,p(i,j)) \\ + ||\boldsymbol{y}_{P^*,j}^{cp} - \boldsymbol{y}_i^{in}||^2, \forall j \in R_i^{cdr} \tag{25}$$

$$g_{P^*}(j,k) = \left(\frac{\min\{|j-k|, |j-k+N^{cp}|\}}{N^{cp}} - \frac{1}{fP^*} \right)^2 \tag{26}$$

3. Optimize the terminal phase

$$p^*(N^{in}-1) = \arg\min_j c(N^{in}-1,j), \quad j \in R_{N^{in}-1}^{cdr} \tag{27}$$

4. Back track

$$p^*(i-1) = p(i,p^*(i)), \quad \forall\, 1 \leq i \leq N^{cp}-1 \tag{28}$$

Finally, the optimal phase $\boldsymbol{s}_{DP}^*$ is set based on the optimal path $\{p^*(i)\}$ and the manifold control points are updated as

$$Y_{DP}^{cp} = \arg\min_{Y^{cp}} E(Y^{cp}, \boldsymbol{s}_{DP}^*) \tag{29}$$

Strictly speaking, one can bypass the DP procedure and proceed to the following step where there is a sufficiently small fluctuation in phase from the periodic solution. On the other hand, in cases with a substantial fluctuation in phase, the following step may result in local optima only and therefore, the DP procedure is still essential for finding the global optimum.

4.4 Iterative solution by quadratic approximation

A crucial procedure in the third step is quadratic approximation of the data term with respect to the phase within a narrow search range from the DP solution. First, the DP solution, $\boldsymbol{s}^*_{DP}$ and Y^{cp}_{DP}, is set as the initial solution for this step, $\boldsymbol{s}^0_Q$ and $Y^{cp,0}$, respectively. Then, the narrow search region of the phase at the rth iteration is set as $R^r_i = \{s | s^{r-1}_{Q,i} - s^{tol} \leq s \leq s^{r-1}_{Q,i} + s^{tol}\}$. The data term minimum for the ith phase $s_{Q,i}$ is found by the Newton method within the intervals that include at least a part of the search region R^r_i. After the phase minimizing the data term has been obtained as $s^{data,r\,*}_{Q,i}$, the ith data term in Eq. (17) is approximated by a Taylor expansion up to the second-order terms as

$$\hat{E}^{data,r}_i(s_{Q,i}) = E^{data,r}_i(s^{data,r\,*}_{Q,i}) + \left.\frac{dE^{data,r}_i}{ds_{Q,i}}\right|_{s_{Q,i}=s^{data,r\,*}_{Q,i}} (s_{Q,i} - s^{data,r\,*}_{Q,i})$$

$$+ \frac{1}{2}\left.\frac{d^2E^{data,r}_i}{ds^2_{Q,i}}\right|_{s_{Q,i}=s^{data,r\,*}_{Q,i}} (s_{Q,i} - s^{data,r\,*}_{Q,i})^2. \tag{30}$$

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

$$\boldsymbol{s}^{r\,*}_Q = \arg\min_{\boldsymbol{s}_Q}\left\{\frac{1}{N^{in}}\sum_{i=0}^{N^{in}-1}\hat{E}^{data,r}_i(s_{Q,i}) + \lambda_s\frac{1}{N^{in}}\sum_{i=1}^{N^{in}-1}\left(s_{Q,i+1}-s_{Q,i}-\frac{1}{fP^*}\right)^2\right\} \tag{31}$$

$$\text{s.t.} \quad s^{r-1}_{Q,i} - s^{tol} \leq s_{Q,i} \leq s^{r-1}_{Q,i} + s^{tol} \tag{32}$$

$$s_{Q,i+1} \geq s_{Q,i}, \tag{33}$$

where Eqs. (32) and (33) 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 one and is solved by the active set method. Then the manifold control points at the rth iteration are updated in the same way as in the previous steps.

$$Y^{cp,r} = \arg\min_{Y^{cp}} E(Y^{cp}, \boldsymbol{s}^{r\,*}_Q) \tag{34}$$

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

5 Experiments

5.1 Experimental setup

In these experiments, the proposed temporal super resolution method is applied to low frame-rate quasi-periodic image sequences. We used two different image sequences: the first is CG data of a conical pendulum viewed from an oblique
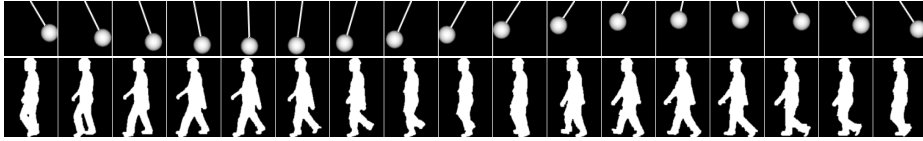
Fig. 1: Periodic image sequences of conical pendulum (top) and gait (bottom) used in the experiments.

Table 1: Image sequence properties.

Image sequence	Image size [pixel]	Original frame-rate [fps]	Period [sec]
Conical pendulum	64×64	- (Arbitrary)	1.17
Gait	88×128	60	1.17

direction (Fig. 1(top)), while the other is real data in the form of a silhouette sequence of a person walking on a treadmill (Fig. 1(bottom)). Regarding the real data, size normalization and spatial registration are done in advance. Moreover, images comprising one period are manually extracted as a subsequence and a completely periodic image sequence is constructed by repeating this subsequence. The image sizes, original frame rates, and periods of the two image sequences are listed in Table 1.

Then, low frame-rate periodic image sequences are down sampled from the original image sequences. In addition, quasi-periodic image sequences are produced by randomly changing the sampling interval up to a predefined phase noise level from the linear phase evolution. Figure 2 shows examples of quasi-periodic image sequences of the conical pendulum at 6 fps and 1 fps, respectively.

Finally, the parameters for the energy minimization framework are determined experimentally as follows. PCA is exploited for eigenspace projection and the information loss rate is set to 1%, in other words, the cumulative contribution ratios of the eigen values is 99%. The number of manifold control points is 100, and the search range for the dynamic programming s^{cdr} and quadratic approximation s^{tol} are 0.25 and 0.02, respectively. The smoothness term coefficients λ_z and λ_s are set to 50.0 and 1.0, respectively, with λ_z reduced by half, down to a minimum of 1.0, in the iterative process of quadratic approximation.

5.2   Refinement process

In this subsection, to explain the refinement process throughout the three steps, i.e., Linear Approximation (LA), Dynamic Programming (DP), and Quadratic Approximation (QA), we focus on a specific example of the conical pendulum image sequence, with frame-rate, number of frames, and phase noise set to 3 fps, 67 frames, and 20%, respectively.

First, phase estimation errors in the Ground Truth (GT) are shown in Fig. 3(a). While the estimated phases in LA deviate widely, since LA cannot absorb
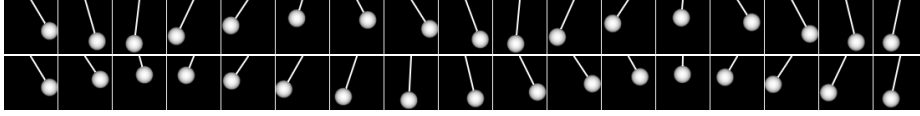
Fig. 2: Examples of input quasi-periodic conical pendulum image sequences with 10% phase noise at 6 fps (top) and 1 fps (bottom).
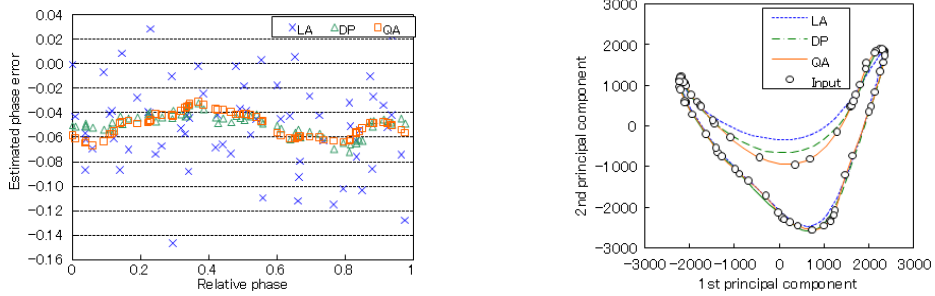

Fig. 3: Phase estimation error on relative phase (left) and manifold in two principal components of eigenspace (right).

non-linear phase noise, those in DP and QA converge within a small range and their deviation is approximately periodic. Although phase error biases from GT (approx. -0.05) are observed in DP and QA, they are not necessarily significant since phase registration up to a relative relation among all the input frames is sufficient for temporal super resolution. In other words, only the standard deviation of the phase estimation errors should be evaluated.

Next, to demonstrate the impact of the phase estimation errors in a more visual way, image sequences sorted by estimated relative phases are shown in Fig. 4. Based on these results, we can see several significant "phase backs" in LA and hence, a sufficiently large manifold smoothness coefficient λ_z is essential for the LA step, otherwise a manifold disturbed by the phase backs is reconstructed. On the other hand, relative phase orders of DP and QA are almost consistent with that of the GT and hence, the smoothness coefficient λ_z can be relaxed.

Finally, manifolds in the two principal components of the eigenspace are shown in Fig. 3(b). We can see that the manifoldsf fitness to the input points improves throughout the three-step phase estimation refinement. In particular, the fitness of QA is the best, since the smoothness coefficient λ_z is relaxed in the iteration process.

5.3   Evaluation

In this section, manifold reconstruction errors and phase estimation errors are evaluated. First, as for manifold reconstruction, a single period of ground truth images are initially projected to a manifold reconstructed from a low frame-rate quasi-periodic input image sequence and then they are back projected to
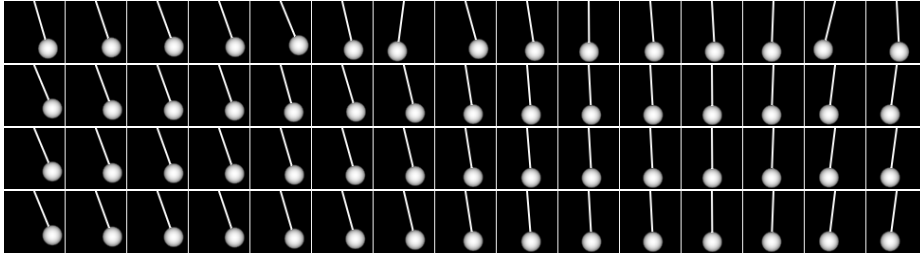
Fig. 4: A section of the conical pendulum image sequences sorted by estimated relative phase. Rows from the top to the bottom indicate LA, DP, QA, and GT, respectively.

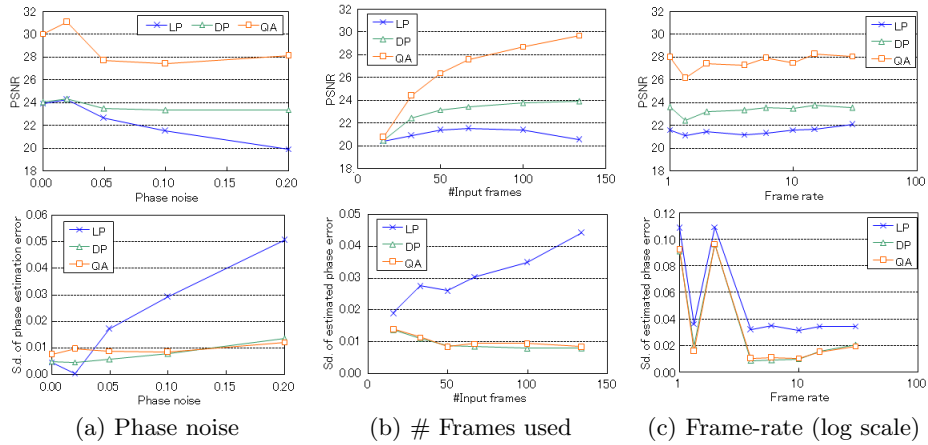(a) Phase noise   (b) # Frames used   (c) Frame-rate (log scale)

Fig. 5: Performance evaluation of manifold PSNR (top row) and phase error SD (bottom row) using gait image sequence. The horizontal axis denotes each factor: (a) phase error, (b) number of frames used, and (c) frame-rate. The vertical axes in the top and bottom rows indicate the PSNR between the reconstructed manifold and ground truth and the SD of the phase estimation error, respectively.

the image domain to calculate Peak Signal to Noise Ratio (PSNR) between the ground truth images and reprojected images. As for phase estimation, errors between the ground truth phase and estimated phases of the input image sequence are calculated and then their Standard Deviation (SD) is evaluated, since their biases are not significant as discussed in the previous section.

The above two items are evaluated with respect to three factors: (1) phase noise, (2) number of frames used, and (3) frame-rate as shown in Fig. 5. The baseline settings for the three factors are 20%, 67 frames, and 3 fps, respectively.

As for the phase noise (Fig. 5(a)), although LA achieves the lowest phase error SD in the case of lower phase noise, both the manifold PSNR and the phase error SD in LA deteriorate in proportion to an increase in phase noise. On the contrary, such degradations in DP and QA are suppressed within a certain range.

Regarding the number of frames used (Fig. 5(b)), the performance naturally improves as the number of frames increases in DP and QA. On the other hand, the performance in LP does not improve in proportion to the frames used. This is because LP does not solve non-linear phase noise, and besides, an increase in the phase-error input points in the eigenspace makes the manifold reconstruction worse.

Regarding frame-rate (Fig. 5(c)), the manifold reconstruction PSNRs improve slightly as the frame-rate increases. On the other hand, critical phase errors are observed in an extremely low frame-rate (1 fps and 2 fps) and this is discussed further in the next section.

6   Discussion

In this section, prerequisites and limitations of the proposed temporal super resolution are discussed. In cases where a periodic image sequence is observed at a coarse sampling interval compared with its period, we need to consider two main issues: (1) the *stroboscopic effect* (*temporal aliasing*), and (2) the *wagon wheel effect* as reported in [13].

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 standing still because the observed images are always the same even though the object is actually moving periodically.

From this observation, we can introduce a theoretical upper bound of the temporal super resolution from a single period image sequence. Intuitively speaking, temporal super resolution is impossible when each phase in one period is exactly the same as the corresponding phase in the other periods, that is, the period in the frame domain is not a sub-frame order, but merely a frame order defined as an integer. Let us consider a low frame-rate periodic image sequence and denote its period in the frame domain P' [frame]. Then, assume that the period P' is expressed as a fraction by relative prime numbers, m and n,

$$P' = \frac{m}{n} \in \mathbb{Q}, \; m, n \in \mathbb{N}. \tag{35}$$

Now, if $n = 1$, the period P' is degraded from a fraction to an integer, that is, the period is just a frame order and the temporal super resolution is impossible. On the other hand, if $n > 1$, the period P' is a sub-frame order and hence different phase images can be observed among multiple periods and the temporal super resolution is possible. In summary, the low frame-rate image sequence can be up-converted to a n-times frame-rate, if m input frames are given.

The *wagon wheel effect* typically occurs when the sampling interval is slightly smaller than the period of a moving object as shown in the bottom row of Fig. 2, where the sampling interval and period are 1.0 sec and 1.17 sec, respectively. In this case, the false minimum by backward play competes with the true minimum by forward play, and then the false minimum is often adopted in the LA step as shown in Fig. 6. This is the reason that the critical phase errors occur in the extremely low frame-rate in Fig. 5(c).
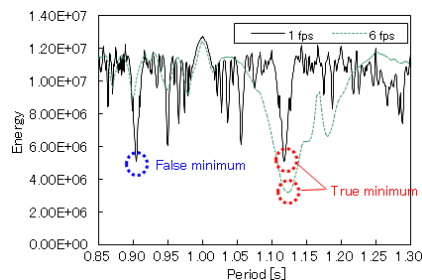
Fig. 6: False period detection by "Wagon wheel effects" in the LA step. The horizontal and vertical axes denote period and energy, respectively. False minimum by backward play (blue broken circle) competes with true minimum by forward play (red broken circle) at 1 fps, while true minimum is obviously the global minimum at 6 fps.

7   Conclusion

This paper described a method for temporal super resolution from a single quasi-periodic image sequence. The temporal super resolution was formulated as simultaneous phase registration and reconstruction of a manifold of the periodic image sequence in a phase parametric eigenspace. An energy minimization framework considering data fitness, and the smoothness of both the manifold and the phase evolution, was introduced and solved through three-step coarse-to-fine procedures to avoid local minima. Experiments using synthesized conical pendulum and real gait silhouette scenes were conducted to evaluate the effects of phase noise, the number of frames used, and frame-rate on the PSNR of the manifold reconstruction and the standard deviation of the phase estimation errors.

In this paper, we focused mainly on phase registration while ignoring blur effects, and hence, it is necessary, in future, to include the blur effects in the manifold reconstruction framework. Moreover, not only phase fluctuation, but also inter-period image deformation (e.g., view or speed transition in gait scenes) should be considered for real applications.

References

1. van Ouwerkerk, J.: Image super-resolution survey. Image and Vision Computing **24** (2006) 1039–1052
2. Borman, S., Stevenson, R.: Spatial resolution enhancement of low-resolution image sequences: A comprehensive review with directions for future research. Technical report, University of Notre Dame (1998)
3. Irani, M., Peleg, S.: Improving resolution by image registration. Computer Vision, Graphics, and Image Processing **53** (1991) 231–239
4. Tanaka, M., Okutomi, M.: A fast map-based super-resolution algorithm for general motion. In: Proc. of SPIE-IS& T Electronic Imaging 2006, Computational Imaging IV. Volume 6065. (2006) 60651B–1–12

5. Freeman, W., Jones, T., Pasztor, E.: Example-based super-resolution. IEEE Trans. on Computer Graphics and Applications **22** (2002) 56–65
6. Liu, C., Shum, H., Zhang, C.: A two-step approach to hallucinating faces: Global parametric model and local non-parametric model. In: IEEE Int. Conf. on Computer Vision and Pattern Recognition. (2001) 192–198
7. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Trans. Pattern Anal. and Machine Intelligent **24** (2002) 1167–1183
8. Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: Proc. of the 12th Int. Conf. on Computer Vision. (2009)
9. M.Tanaka, M.Okutomi: Near-real-time video-to-video super-resolution. In: Proc of 8th Asian Conference on Computer Vision (ACCV2007), Demo-A. (2007) 1
10. M.Shimizu, S.Yoshimura, M.Tanaka, M.Okutomi: Super-resolution from image sequence under influence of hot-air optical turbulence. In: Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). (2008) 1–8
11. Blake, A., Bascle, B., , Zisserman, A.: Motion deblurring and super-resolution from an image sequence. In: Proc. European Conf. Computer Vision. (1996) 312–320
12. Sezan, M., Patti, A., Tekalp, A.: Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. IEEE Trans. Image Processing **6** (1997) 1064–1076
13. Shechtman, E., Caspi, Y., Irani, M.: Space-time super-resolution. IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005) 531–545
14. Caspi, Y., Irani, M.: Spatio-temporal aignment of sequences. IEEE Trans. Pattern Analysis and Machine Intelligence **24** (2002) 1409–1425
15. REALVIZ2: Retimer. www.realviz.com/products/rt (2000)
16. Wang, Y., Ostermann, J., Zhang, Y.Q.: Video Processing and Communications. Prentice Hall (2002)
17. Beymer, D., Poggio, T.: Image representations for visual learning. Science **272** (1996) 1905–1909
18. Stich, T., Magnor, M.: Image morphing for space-time interpolation. In: SIGGRAPH '07: ACM SIGGRAPH 2007 sketches, New York, NY, USA, ACM (2007) 87
19. LUCAS, B., KANADE, T.: An iterative image registration technique with an application to stereo vision. In: Proc. Seventh International Joint Conf. Artificial Intelligence. (1981) 674–679
20. Stich, T., Linz, C., Albuquerque, G., Magnor, M.: View and time interpolation in image space. Computer Graphics Forum (Proc. Pacific Graphics) **27** (2008)
21. Watanabe, K., Iwai, Y., Nagahara, H., Yachida, M., Suzuki, T.: Video synthesis with high spatio-temporal resolution using spectral fusion. In: Proc. Int. Workshop Multimedia Content Representation, Classification and Security. Volume LNCS4105., Istanbul, Turkey (2006) 683–690
22. Taubin, G.: Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **13** (1991) 1115–1138
23. Kanatani, K.: Ellipse fitting with hyperaccuracy. IEICE Transactions on Information and Systems **E89-D** (2006) 2653–2660
24. Murase, H., Nayar, S.K.: Parametric eigenspace representation for visual learning and recognition. In: Proc. of SPIE. Volume 2031. (1993)
25. Oka, R.: Spotting method for classification of real world data. Computer Journal **41** (1998) 559–565